

Spis treści

Wstęp	13
Podziękowania	15
O niniejszej książce	17
O autorkach	21
CZĘŚĆ I ROZPOCZĘCIE PRZYGODY Z BADANIEM DANYCH	23
1. Czym jest danologia?	25
1.1. Czym jest danologia?	27
1.1.1. Matematyka/statystyka	29
1.1.2. Bazy danych/programowanie	30
1.1.3. Zrozumienie biznesu	32
1.2. Różne rodzaje prac związanych z badaniem danych	33
1.2.1. Analiza	34
1.2.2. Uczenie maszynowe	34
1.2.3. Nauka o podejmowaniu decyzji	35
1.2.4. Pokrewne prace	36
1.3. Wybór swojej drogi	37
1.4. Wywiad z Robertem Changiem, badaczem danych w Airbnb	38
Jak zaczęła się twoja podróż badacza danych?	39
Czego ludzie powinni szukać w pracy związanej z danologią?	39
Jakich umiejętności potrzeba, aby zostać badaczem danych?	40
2. Firmy działające w obszarze danologii	41
2.1. MTC — potężna firma technologiczna	42
2.1.1. Twój zespół —jeden z wielu w MTC	42
2.1.2. Technologia — zaawansowana, ale obecna w firmie w systemie silosowym	43
2.1.3. Zalety i wady pracy w MTC	44
2.2. HandbagLOVE — uznany sprzedawca detaliczny	45
2.2.1. Twój zespół — grupka ludzi walczących o rozwój	46
2.2.2. Twoja technologia — stos technologiczny, który zaczyna się zmieniać	46
2.2.3. Wady i zalety pracy w HandbagLOVE	47
2.3. Seg-Metra — nowo powstały start-up	48

2.3.1. Twój zespół (jaki zespół?)	48
2.3.2. Technologia — najnowsza, wewnętrznie spójna technologia	49
2.3.3. Wady i zalety pracy w Seg-Metra	50
2.4. Videory — udany start-up technologiczny na późnym etapie rozwoju	52
2.4.1. Zespół — wyspecjalizowany, ale z przestrzenią do działania	52
2.4.2. Technologia — próby pokonania kodu odziedziczonego	53
2.4.3. Zalety i wady pracy w Videory	53
2.5. Global Aerospace Dynamics — wielki dostawca rządowy	54
2.5.1. Zespół — danolog w morzu inżynierów	55
2.5.2. Technologia — stara, skostniała i wyposażona w blokadę bezpieczeństwa	56
2.5.3. Zalety i wady pracy w GAD	56
2.6. Podsumowanie	57
2.7. Wywiad z Randym Au, specjalistą Google'a ds. badań ilościowych w zakresie doświadczenia użytkownika (ang. user experience)	58
Czy istnieją duże różnice pomiędzy dużymi a małymi firmami?	58
Czy istnieją różnice w zależności od branży, w jakiej działa dana firma?	59
Jaka jest twoja ostatnia rada dla początkujących badaczy danych?	59
3. Zdobywanie umiejętności	61
3.1. Uzyskiwanie tytułu zawodowego lub stopnia naukowego w zakresie danologii	62
3.1.1. Wybór uczelni	64
3.1.2. Dostanie się do programu akademickiego	66
3.1.3. Podsumowanie kwestii tytułów zawodowych oraz stopni naukowych	68
3.2. Przejście przez intensywny kurs	69
3.2.1. Czego się nauczysz	69
3.2.2. Koszt	71
3.2.3. Wybór programu	71
3.2.4. Podsumowanie intensywnych kursów danologicznych	72
3.3. Zdobyć pracę z zakresu badania danych w swojej firmie	72
3.3.1. Podsumowanie nauki w pracy	75
3.4. Uczenie się na własną rękę	75
3.4.1. Podsumowanie nauki na własną rękę	76
3.5. Dokonanie wyboru	77
3.6. Wywiad z Julią Silge, badaczką danych i programistką w RStudio	78
Zanim zostałeś badaczką danych, pracowałeś na uczelni. Jak zdobyte w środowisku akademickim umiejętności pomogły ci w danologii?	78
Po podjęciu decyzji o zostaniu badaczką danych w jaki sposób zaczęłaś nabywać nowych umiejętności?	79
Czy wiedziałaś, wkraczając do domeny badania danych, jaką pracę chciałabyś wykonywać?	79
Co poleciłabyś osobom, które chciałyby zdobyć umiejętności potrzebne do zostania badaczem danych?	80

4. Tworzenie portfolio	81
4.1. Tworzenie projektu	82
4.1.1. Identyfikowanie danych i zadanie pytania	82
4.1.2. Wybór kierunku	85
4.1.3. Wypełnianie pliku README na GitHubie	86
4.2. Założenie bloga	87
4.2.1. Potencjalne tematy	87
4.2.2. Logistyka	88
4.3. Praca nad przykładowymi projektami	90
4.3.1. Freelancerzy zajmujący się badaniem danych	90
4.3.2. Uczenie sieci neuronowej na podstawie kontrowersyjnych tablic rejestracyjnych	91
4.4. Wywiad z Davidem Robinsonem, danologiem	93
Jak zaczęła się twoja przygoda z blogowaniem?	93
Czy działalność publiczna przyniosła ci konkretne korzyści?	93
Czy są ludzie, którzy twoim zdaniem szczególnie skorzystaliby na działalności publicznej?	93
Jak zmieniał się z czasem twój pogląd na wartość działalności publicznej?	94
Jak powstają pomysły na posty dotyczące analizy danych?	94
Jaka jest twoja ostatnia rada dla początkujących i młodszych badaczy danych?	94
Zasoby do rozdziałów 1. - 4.	95
Książki	95
Posty na blogu	96

CZĘŚĆ II ZNALEZIENIE PRACY ZWIĄZANEJ z ANALIZĄ DANYCH 99

5. Poszukiwania — znalezienie odpowiedniej dla siebie pracy	101
5.1. Znalezienie pracy	102
5.1.1. Dekodowanie opisów	103
5.1.2. Szukanie znaków ostrzegawczych	105
5.1.3. Określanie swoich oczekiwań	106
5.1.4. Uczestnictwo w spotkaniach	107
5.1.5. Korzystanie z mediów społecznościowych	109
5.2. Podejmowanie decyzji o tym, o jaką pracę się ubiegać	110
5.3. Wywiad z Jesse Mostipak, rzeczniką ds. rozwoju oprogramowania w Kaggle'u	112
Jakie masz porady co do rozpoczęcia poszukiwania pracy?	112
Jak można zbudować swoją sieć kontaktów?	112
Co robić, gdy brak nam pewności, żeby odpowiadać na oferty pracy z zakresu badania danych?	113
Co byś powiedziała komuś, kto uważa, że „nie spełnia wszystkich wymogów na to stanowisko”?	113

Jaka będzie twoja ostatnia porada dla początkujących badaczy danych?	113
6. Aplikowanie — życiorysy i listy motywacyjne	115
6.1. Życiorys — podstawy	116
6.1.1. Struktura	118
6.1.2. Zanurzenie się w szczegółach dotyczących doświadczenia zawodowego — generowanie treści	123
6.2. Listy motywacyjne — podstawy	125
6.2.1. Struktura	126
6.3. Szycie na miarę	127
6.4. Polecenie	128
6.5. Wywiad z Kristen Kehrer, wykładowczynią i autorką kursów danologicznych	130
Ile razy, według twoich szacunków, przeredagowywałaś swój życiorys?	130
Jakie są częste błędy, które twoim zdaniem ludzie popełniają?	131
Czy dostosowujesz swoje CV do stanowiska, na które aplikujesz?	131
Jakie strategie polecasz do opisanie stanowisk pracy w CV?	131
Jaką masz ostatnią radę dla początkujących danologów?	132
7. Rozmowa kwalifikacyjna — czego należy się spodziewać i jak sobie z tym poradzić	133
7.1. Czego pragną firmy?	134
7.1.1. Proces rozmowy rekrutacyjnej	135
7.2. Krok 1. Wstępna rozmowa telefoniczna	136
7.3. Krok 2. Rozmowa w siedzibie firmy	138
7.3.1. Rozmowa na tematy techniczne	140
7.3.2. Pytania behawioralne	144
7.4. Krok 3. Analiza przypadku	146
7.5. Krok 4. Rozmowa końcowa	148
7.6. Oferta	149
7.7. Wywiad z Ryanem Williamsem, doświadczonym analitykiem decyzyjnym w Starbucksie	150
Co trzeba zrobić, żeby świetnie wypaść na rozmowie o pracę?	150
Jak sobie radzisz w sytuacjach, gdy nie znasz odpowiedzi?	150
Co należy zrobić w przypadku, gdy odpowiedź spotka się z negatywną reakcją?	151
Czego na temat kandydatów nauczyło cię prowadzenie rozmów o pracę?	151
8. Oferta — co można zaakceptować	153
8.1. Proces	154
8.2. Otrzymanie oferty	154
8.3. Negocjacje	156
8.3.1. Co podlega negocjacji?	157
8.3.2. Jak bardzo możesz negocjować	159
8.4. Taktyki negocjacyjne	162

8.5. Jak wybrać jedną spośród dwóch „dobrych” ofert pracy	163
8.6. Wywiad z Brooke Watson Madubuonwu, starszą badaczką danych w ACLU	165
Co należy wziąć pod uwagę, oprócz wynagrodzenia, gdy rozważa się przyjęcie oferty?	165
Jakie są sposoby na przygotowanie się do negocjacji?	165
Co zrobić, kiedy mamy już jedną ofertę, ale nadal czekamy na inną?	166
Jaka jest twoja ostatnia rada dla początkujących i młodszych badaczy danych?	166
Zasoby do rozdziałów 5. - 8.	167
Książki	167
Posty na blogach i kursy	168

CZĘŚĆ III ZADOMOWIENIE SIĘ W ŚWIECIE BADANIA DANYCH 171

9. Pierwsze miesiące w nowym miejscu pracy 173

9.1. Pierwszy miesiąc	174
9.1.1. Wdrożenie w dużej organizacji — dobrze naoliwiona maszyna	174
9.1.2. Wdrożenie w małej firmie. Jakie wdrożenie?	175
9.1.3. Zrozumienie i ustalenie oczekiwań	175
9.1.4. Znajomość własnych danych	177
9.2. Stawianie się produktywnym pracownikiem	180
9.2.1. Zadawanie pytań	181
9.2.2. Budowanie relacji	182
9.3. Jeśli jesteś pierwszym zatrudnionym badaczem danych	184
9.4. Kiedy praca nie jest tym, co Ci obiecywano	185
9.4.1. Charakter pracy jest okropny	185
9.4.2. Środowisko pracy jest toksyczne	186
9.4.3. Decyzja o odejściu z pracy	187
9.5. Wywiad z Jarvisem Millerem, badaczem danych w firmie Spotify	189
Co cię zaskoczyło w twojej pierwszej pracy naukowej w dziedzinie danologii?	189
Z jakimi problemami się stykałeś?	190
Możesz nam opowiedzieć o jednym ze swoich pierwszych projektów?	190
Jaka byłaby twoja najważniejsza rada na kilka pierwszych miesięcy pracy?	191

10. Przeprowadzanie skutecznej analizy 193

10.1. Wniosek	196
10.2. Plan analizy	198
10.3. Przeprowadzenie analizy	201
10.3.1. Importowanie i czyszczenie danych	201
10.3.2. Eksplorowanie i modelowanie danych	203
10.3.3. Ważne punkty dotyczące eksplorowania i modelowania	205
10.4. Odpowiednia otoczką	209

10.4.1. Finalna prezentacja	210
10.4.2. Kończenie pracy	211
10.5. Wywiad z Hilary Parker, badaczką danych w Stitch Fix	211
W jaki sposób myślenie o innych ludziach pomaga w twoich analizach?	212
Jaką strukturę nadajesz swoim analizom?	212
Jakiego rodzaju korektę przeprowadzasz w ramach wersji ostatecznej?	212
Jak sobie radzisz z osobami proszącymi o wprowadzenie zmian do analizy?	213
11. Wdrażanie modelu do środowiska produkcyjnego	215
11.1. Czym w ogóle jest wdrożenie do środowiska produkcyjnego?	216
11.2. Tworzenie systemu produkcyjnego	218
11.2.1. Gromadzenie danych	219
11.2.2. Budowa modelu	220
11.2.3. Serwowanie modeli przy użyciu interfejsu programowania aplikacji	221
11.2.4. Budowa interfejsu programowania aplikacji	222
11.2.5. Dokumentacja	224
11.2.6. Testowanie	225
11.2.7. Wdrażanie API	225
11.2.8. Testy obciążeniowe	229
11.3. Utrzymanie działającego systemu	229
11.3.1. Monitorowanie systemu	230
11.3.2. Ponowne trenowanie modelu	230
11.3.3. Wprowadzanie zmian	231
11.4. Na zakończenie	232
11.5. Rozmowa z Heather Nolis, inżynierem zajmującym się uczeniem maszynowym w T-Mobile	232
Co oznacza bycie „inżynierem zajmującym się uczeniem maszynowym” w twoim zespole?	232
Jakie to było uczucie, gdy po raz pierwszy wdrażałaś fragment kodu do środowiska produkcyjnego?	233
Jeśli coś pójdzie nie tak w środowisku produkcyjnym, to co wtedy?	233
Jaka będzie twoja ostatnia rada dla danologów pracujących z inżynierami?	234
12. Współpraca z interesariuszami	235
12.1. Typy interesariuszy	236
12.1.1. Interesariusze biznesowi	236
12.1.2. Interesariusze inżynieryjni	237
12.1.3. Korporacyjne kierownictwo	239
12.1.4. Twój menedżer	240
12.2. Współpraca z interesariuszami	240
12.2.1. Zrozumienie celów interesariusza	241
12.2.2. Ciągła komunikacja	243

12.2.3. Bycie konsekwentnym	245
12.3. Ustalanie priorytetów w pracy	247
12.3.1. Praca zarówno innowacyjna, jak i wywierająca wpływ	248
12.3.2. Praca, która nie jest innowacyjna, ale wciąż wywiera wpływ	249
12.3.3. Praca innowacyjna, która jednak nie wywiera żadnego wpływu	249
12.3.4. Praca, która ani nie jest innowacyjna, ani też nie wywiera żadnego wpływu	250
12.4. Uwagi końcowe	251
12.5. Wywiad z Sade Snowden-Akintunde, badaczką danych w Etsy	251
Dlaczego zarządzanie interesariuszami jest ważne?	252
Jak nauczyłaś się zarządzać interesariuszami?	252
Czy miałaś kiedyś problemy z interesariuszem?	252
Co młodzi badacze danych często robią źle?	252
Czy zawsze starasz się wyjaśniać techniczne aspekty danologii?	253
Jaka jest twoja ostatnia porada dla młodszych lub początkujących danologów?	253
Zasoby do rozdziałów 9. - 12.	253
Książki	253
Blogi	255

CZĘŚĆ IV TWÓJ ROZWÓJ W ROLI BADACZA DANYCH 257

13. Kiedy Twój projekt danologiczny kończy się niepowodzeniem 259

13.1. Dlaczego projekty danologiczne kończą się niepowodzeniem	261
13.1.1. Dane nie spełniają Twoich oczekiwań	261
13.1.2. Brak sygnału w zakresie danych	262
13.1.3. Utrata zainteresowania klientów	264
13.2. Zarządzanie ryzykiem	266
13.3. Co możesz zrobić, gdy Twój projekt kończy się niepowodzeniem?	267
13.3.1. Co należy zrobić z projektem?	267
13.3.2. Radzenie sobie z negatywnymi emocjami	269
13.4. Wywiad z Michelle Keim, dyrektorką ds. danologii i uczenia maszynowego w Pluralsight	270
Czy zdarzyło ci się doświadczyć porażki w swojej karierze?	271
Czy potrafisz rozpoznać sygnały ostrzegawcze przed rozpoczęciem projektu?	271
Czy firmy różnią się w sposobie podejścia do niepowodzeń?	271
Skąd wiesz, że projekt, który realizujesz, skończy się niepowodzeniem?	272
Jak pokonać strach przed porażką?	272

14. Dołączenie do środowiska danologicznego 273

14.1. Rozwijanie własnego portfolio	275
14.1.1. Więcej wpisów na blogu	275
14.1.2. Więcej projektów	276

14.2. Uczestnictwo w konferencjach	277
14.2.1. Radzenie sobie z lękiem społecznym	280
14.3. Wystąpienia	281
14.3.1. Poszukiwanie możliwości wystąpienia	281
14.3.2. Przygotowanie	284
14.4. Przyczynianie się do rozwoju otwartego oprogramowania	285
14.4.1. Wkład w pracę innych ludzi	285
14.4.2. Tworzenie własnego pakietu lub biblioteki	287
14.5. Rozpoznawanie i unikanie wypalenia	288
14.6. Wywiad z Renee Teate, dyrektor ds. danologii w HelioCampus	289
Jakie są główne korzyści płynące z działalności w mediach społecznościowych?	289
Co byś powiedziała ludziom, którzy mówią, że nie mają czasu na angażowanie się w życie społeczności?	290
Czy warto wytwarzać niewielką ilość treści?	290
Stresowałaś się publikacją pierwszego wpisu na blogu lub pierwszym przemówieniem?	290
15. Jak odejść z pracy z wdziękiem	293
15.1. Decyzja o odejściu	294
15.1.1. Podsumuj swoje postępy w nauce	294
15.1.2. Omów swoją sytuację z menedżerem	295
15.2. Jak wygląda poszukiwanie pracy, kiedy masz już za sobą pierwszą pracę	297
15.2.1. Decydowanie o tym, czego chcesz	298
15.2.2. Rozmowa kwalifikacyjna	298
15.3. Znalezienie nowej pracy w trakcie zatrudnienia	300
15.4. Złożenie wypowiedzenia	301
15.4.1. Rozważanie przyjęcia kontroferty	302
15.4.2. Poinformowanie zespołu	303
15.4.3. Ułatwienie przeprowadzenia zmiany	304
15.5. Wywiad z Amandą Casari, menedżerem ds. technicznych w Google'u	305
Po czym poznać, że nadszedł czas na szukanie nowej pracy?	305
Czy kiedykolwiek rozpoczęłaś poszukiwania pracy i zdecydowałaś się jednak pozostać?	306
Widzisz ludzi, którzy zbyt długo wykonują tę samą pracę?	306
Czy można zmienić pracę zbyt szybko?	306
Jaka jest twoja ostatnia rada dla ambitnych i początkujących danologów?	307
16. Wspinanie się po szczeblach kariery	309
16.1. Ścieżka menedżerska	311
16.1.1. Zalety bycia menedżerem	312
16.1.2. Wady bycia menedżerem	312

16.1.3. Jak zostać menedżerem	313
16.2. Ścieżka głównego danologa	315
16.2.1. Zalety bycia głównym danologiem	317
16.2.2. Wady bycia głównym danologiem	318
16.2.3. Jak zostać głównym danologiem	318
16.3. Zostanie niezależnym konsultantem	319
16.3.1. Zalety bycia niezależnym konsultantem	321
16.3.2. Wady bycia niezależnym konsultantem	321
16.3.3. Jak zostać niezależnym konsultantem	322
16.4. Wybór ścieżki	323
16.5. Wywiad z Angelą Bassą, dyrektorką ds. danologii, inżynierii danych i uczenia maszynowego w i Robot	323
Jaka wygląda codzienna praca na stanowisku menedżerskim?	324
Jakie są oznaki tego, że należy przestać być szeregowym pracownikiem?	324
Czy musisz w końcu całkiem przestać wykonywać robotę szeregowego pracownika?	324
Jakiej rady udzieliłabyś komuś, kto chce zostać technicznym liderem, ale jeszcze nie jest na to gotowy?	325
Jaka jest twoja ostatnia rada dla początkujących i młodszych badaczy danych?	325
Zasoby dla rozdziałów 13. -16.	326
Książki	326
Blogi	327
Epilog	329
Dodatek. Pytania z rozmów kwalifikacyjnych	330

oprac. BPK