

Przedmowa	15
Wstęp	17
Prolog. Wyobraź sobie siatkę danych	23
Część I. Czym jest siatka danych?	37
1. Siatka danych w pigułce	39
Rezultaty	39
Transformacje	40
Zasady	42
Zasada własności dziedziny	42
Zasada danych jako produktu	43
Zasada samoobsługowej platformy danych	43
Zasada sfederowanego zarządzania obliczeniowego	44
Wzajemne oddziaływanie zasad	44
Spojrzenie na model siatki danych	45
Dane	46
Dane operacyjne	47
Dane analityczne	47
Pochodzenie	48
2. Zasada własności dziedziny	50
Krótkie wprowadzenie do projektowania dziedziny	51
Zastosowanie projektowania strategicznego DDD do danych	53
Archetypy danych dziedziny	55
Dane dziedziny dostosowane do źródła	56
Zagregowane dane dziedziny	58
Dane dziedziny dostosowane do konsumenta	58

Przechodzenie w kierunku własności dziedziny	59
Przekazuj własność danych upstreamowo	59
Definiuj wiele połączonych modeli	60
Wykorzystaj dane najbardziej odpowiedniej dziedziny:	
nie oczekuj jednego źródła prawdy	60
Ukrywaj potoki danych jako wewnętrzną implementację dziedziny	61
Podsumowanie	62
3. Zasada danych jako produktu	63
Zastosowanie do danych myślenia w kategoriach produktu	65
Podstawowe atrybuty użyteczności produktu danych	67
Przechodzenie do zasady danych jako produktu	75
Uwzględnij własność produktu danych w dziedzinach	75
Przeformułuj nomenklaturę, aby zainicjować zmiany	76
Traktuj dane jako produkt, a nie zwykły zasób	76
Kształtuj kulturę danych opartą na zasadzie „ufaj, ale sprawdzaj”	77
Łącz dane i obliczaj je jako pojedynczą jednostkę logiczną	78
Podsumowanie	78
4. Zasada samoobsługowej platformy danych	80
Porównanie platformy siatki danych z innymi rozwiązaniami	81
Dostosowanie do autonomicznych zespołów opartych na dziedzinach	82
Zarządzanie autonomicznymi i interoperacyjnymi produktami danych	84
Zintegrowana platforma dla funkcjonalności operacyjnych i analitycznych	84
Dostosowanie do generalistów	85
Faworyzowanie technologii zdecentralizowanych	86
Niezależność dziedziny	86
Myślenie w kategoriach platformy siatki danych	87
Umożliwienie autonomicznym zespołom uzyskiwania wartości z danych	88
Wymiana wartości za pomocą autonomicznych	
i interoperacyjnych produktów danych	90
Przyspieszenie wymiany wartości przez zmniejszenie obciążenia poznawczego	91
Pozioame skalowanie udostępniania danych	92
Wspieranie kultury wbudowanych innowacji	93
Przechodzenie na samoobsługową platformę siatki danych	94
Najpierw zaprojektuj interfejsy API i protokoły	94
Przygotuj się na dostosowanie do generalistów	95
Przeprowadź inwentaryzację i dokonaj uproszczenia	95
Utwórz wysokopoziomowe interfejsy API do zarządzania produktami danych	96
Buduj doświadczenia, a nie mechanizmy	96
Zacznij od najprostszych fundamentów, a potem zbieraj plony, aby ewoluować	96
Podsumowanie	97

5. Zasada sfederowanego zarządzania obliczeniowego	98
Zastosowanie myślenia systemowego do zarządzania siatką danych	100
Utrzymywanie dynamicznej równowagi między autonomią dziedzinową a globalną interoperacyjnością	101
Przyjęcie topologii dynamicznej jako stanu domyślnego	105
Wykorzystanie automatyzacji i architektury rozproszonej	105
Zastosowanie federacji do modelu zarządzania	106
Zespół sfederowany	107
Wartości przewodnie	109
Reguły	111
Zachęty	112
Zastosowanie obliczeń do modelu zarządzania	113
Standardy jako kod	114
Reguły jako kod	115
Zautomatyzowane testy	116
Zautomatyzowane monitorowanie	116
Przechodzenie na sfederowane zarządzanie obliczeniowe	116
Delegowanie odpowiedzialności do dziedzin	116
Osadzanie wykonywania reguł w poszczególnych produktach danych	117
Automatyzacja zapewniania możliwości oraz monitorowania zamiast interweniowania	117
Modelowanie luk	118
Pomiar efektu sieciowego	118
Przyjęcie zmian w miejsce stałości	118
Podsumowanie	119

Część II. Dlaczego siatka danych? 121

6. Punkt przegięcia	123
Wielkie oczekiwania wobec danych	124
Wielki podział danych	126
Skala — bliskie spotkania z nowym gatunkiem	127
Nie tylko porządek	128
Zbliżanie się do punktu zwrotu	129
Podsumowanie	130
7. Po punkcie przegięcia	131
Płynne reagowanie na zmiany w złożonym biznesie	132
Wzajemne dostosowanie biznesu, technologii i danych analitycznych	132
Wypełnienie luki między danymi analitycznymi i operacyjnymi	133
Lokalizowanie zmian danych w dziedzinach biznesowych	135
Zmniejszenie przypadkowej złożoności potoków i kopiowania danych	136

Utrzymanie zwinności w obliczu wzrostu	136
Usunięcie scentralizowanych i monolitycznych wąskich gardeł	137
Zmniejszenie koordynacji potoków danych	137
Zmniejszenie koordynacji zarządzania danymi	138
Zapewnienie autonomii	139
Zwiększenie współczynnika wartości z danych w stosunku do inwestycji	140
Tworzenie warstwy abstrakcji dla złożoności technicznej	141
za pomocą platformy danych	141
Wszechobecne stosowanie myślenia w kategoriach produktowych	141
Przekraczanie granic	141
Podsumowanie	142
8. Przed punktem przejęcia	145
Ewolucja architektur danych analitycznych	145
Pierwsza generacja — architektura hurtowni danych	145
Druga generacja — architektura jeziora danych	146
Trzecia generacja — multimodalna architektura chmury	149
Charakterystyka architektury danych analitycznych	149
Monolityzm	150
Scentralizowana własność danych	154
Wykorzystanie technologii	156
Podsumowanie	159

Część III. Jak zaprojektować architekturę siatki danych? 161

9. Architektura logiczna	163
Dziedzinowe interfejsy udostępniania danych analitycznych	166
Projekt interfejsu operacyjnego	167
Projekt interfejsu danych analitycznych	168
Międzydziedzinowe zależności danych analitycznych	168
Produkt danych jako kwant architektury	169
Komponenty strukturalne produktu danych	170
Interakcje współdzielenia danych przez produkty danych	176
Interfejsy API wykrywania i obserwowalności danych	178
Wielopłaszczyznowa platforma danych	178
Płaszczyzna platformy	179
(Narzędziowa) płaszczyzna infrastruktury danych	180
Płaszczyzna doświadczenia produktu danych	180
Płaszczyzna doświadczenia siatki	181
Przykład	181

Osadzone reguły obliczeniowe	182
Przyczepa produktu danych	183
Kontener obliczeniowy produktu danych	185
Port sterowania	185
Podsumowanie	186
10. Architektura wielopłaszczyznowej platformy danych	188
Projekt platformy oparty na podrózach użytkowników	190
Podróż twórcy produktu danych	192
Rozpoczęcie projektowania, eksplorowanie, wstępne uruchamianie i określanie źródeł	193
Budowanie, testowanie, wdrażanie i uruchamianie	196
Utrzymywanie, ewoluowanie i wycofywanie	199
Podróż konsumenta produktu danych	201
Rozpoczęcie projektowania, eksplorowanie, wstępne uruchamianie i określanie źródeł	202
Budowanie, testowanie, wdrażanie i uruchamianie	204
Monitorowanie, ewoluowanie i wycofywanie	205
Podsumowanie	205

Część IV. Jak zaprojektować architekturę produktu danych? 207

11. Projektowanie produktu danych według afordancji	209
Afordancje produktu danych	210
Charakterystyka architektury produktu danych	213
Projektowanie inspirowane prostotą złożonych systemów adaptacyjnych	214
Zachowanie wynikające z prostych reguł lokalnych	214
Brak centralnej orkiestracji	215
Podsumowanie	215
12. Projektowanie konsumowania, przekształcania i serwowania danych	217
Serwowanie danych	217
Potrzeby użytkowników danych	217
Właściwości projektowania serwowania danych	220
Projektowanie serwowania danych	232
Konsumowanie danych	233
Archetypy źródeł danych	234
Lokalizacja konsumpcji danych	237
Projektowanie konsumowania danych	239
Przekształcanie danych	240
Przekształcanie programowe i nieprogramowe	241
Transformacja oparta na przepływie danych	243

Uczenie maszynowe jako transformacja	243
Transformacja niestacjonarna	244
Projekt transformacji	244
Podsumowanie	246

13. Projektowanie wykrywania, rozumienia i komponowania danych 247

Wykrywanie, rozumienie, obdarzanie zaufaniem i eksplorowanie	247
Rozpocznij wykrywanie od samorejestracji	250
Wykrywanie globalnego identyfikatora URI	250
Zrozumienie modeli semantycznych i składniowych	250
Ustanowienie zaufania za pomocą gwarancji danych	252
Eksplorowanie kształtu danych	255
Nauka na podstawie dokumentacji	256
Wykrywanie, eksplorowanie i rozumienie projektu	256
Komponowanie danych	256
Właściwości projektu konsumowania danych	258
Tradycyjne podejście do kompozycyjności danych	260
Projekt komponowania danych	263
Podsumowanie	266

14. Projektowanie organizowania i obserwowania danych oraz zarządzania nimi 267

Zarządzanie cyklem życia	267
Projektowanie zarządzania cyklem życia	268
Komponenty manifestu produktu danych	269
Zarządzanie danymi	270
Projektowanie zarządzania danymi	270
Normalizacja reguł	272
Integracja danych i reguł	273
Linkowanie reguł	273
Obserwowanie, debugowanie i audytowanie	274
Projektowanie obserwowalności	275
Podsumowanie	278

Część V. Od czego zacząć? 281

15. Strategia i wykonywanie 283

Czy należy przyjąć siatkę danych już dziś?	283
Siatka danych jako element strategii danych	287
Framework wykonywania siatki danych	290
Wykonywanie oparte na biznesie	291
Wykonywanie kompleksowe i iteracyjne	296
Wykonywanie ewolucyjne	296
Podsumowanie	312

16. Organizacja i kultura	313
Zmiana	315
Kultura	317
Wartości	317
Nagroda	320
Motywacje wewnętrzne	320
Motywacje zewnętrzne	321
Struktura	322
Założenia struktury organizacyjnej	322
Wykrywanie granic produktów danych	330
Ludzie	333
Role	333
Rozwój zestawów umiejętności	336
Proces	338
Zmiany kluczowych procesów	339
Podsumowanie	340