

Data Mesh

**Entregando valor em escala
e orientado a dados**

Zhamak Dehghani

O'REILLY®
Novatec

Authorized Portuguese translation of the English edition of Data Mesh ISBN 9781492092391 © 2022 Zhamak Dehghani. This translation is published and sold by permission of O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

Tradução em português autorizada da edição em inglês da obra Data Mesh ISBN 9781492092391 © 2022 Zhamak Dehghani. Esta tradução é publicada e vendida com a permissão da O'Reilly Media, Inc., detentora de todos os direitos para publicação e venda desta obra.

© Novatec Editora Ltda. [2023].

Editor: Rubens Prates

Tradução: Cibelle Ravaglia

Revisão gramatical: Tássia Carvalho

ISBN do impresso: 978-85-7522-868-5

ISBN do ebook: 978-85-7522-869-2

Histórico de impressões:

Setembro/2023 Primeira edição

Novatec Editora Ltda.

Rua Luís Antônio dos Santos 110

02460-000 – São Paulo, SP – Brasil

Tel.: +55 11 2959-6529

Email: novatec@novatec.com.br

Site: <https://novatec.com.br>

Twitter: twitter.com/novateceditora

Facebook: facebook.com/novatec

LinkedIn: linkedin.com/in/novatec

LIS20230914

Sumário

Apresentação	15
Prefácio	17
Prólogo: imaginando o data mesh	25
Parte I = O que é data mesh?	41
Capítulo 1 = Resumindo o data mesh	43
Resultados	44
Mudanças	44
Princípios	46
Princípio de propriedade de domínio	47
Princípio de dados como produto	47
Princípio de plataforma de autoatendimento de dados	48
Princípio de governança computacional federada	49
Inter-relação dos princípios	50
Visão geral do modelo de data mesh	51
Dados	52
Dados operacionais	52
Dados analíticos	53
Origem	54
Capítulo 2 = Princípio de propriedade de domínio	56
Breve histórico sobre o design orientado a domínio	58
Aplicando o design estratégico do DDD aos dados	59
Arquétipos de dados de domínio	62
Dados de domínio alinhados à fonte	63
Dados agregados de domínio	65
Dados de domínio alinhados ao consumidor	66
Transição para propriedade de domínio	67
Promovendo a propriedade de dados upstream	67
Defina múltiplos modelos conectados	67
Abraça os dados de domínio mais relevantes: não espere uma única fonte de verdade	68
Ocultar os pipelines de dados como implementação interna dos domínios	69
Recapitulando	70

Capítulo 3 = Princípio de dados como produto72

Aplicando a abordagem de product thinking aos dados.....	74
Atributos de baseline de usabilidade de um produto de dados	76
Transição para dados como um produto	86
Inclua a propriedade de produtos de dados em domínios	87
Reformule a nomenclatura para criar mudanças	88
Pense nos dados como um produto, não como um simples ativo	89
Estabeleça a cultura de dados trust-but-verify (confie, mas verifique")	89
Combinar dados e calcular como uma unidade lógica.....	90
Recapitulando	91

Capítulo 4 = Princípio de plataforma de autoatendimento de dados93

Plataforma de data mesh: comparações e diferenças	95
Atendendo equipes autônomas e orientadas a domínio	97
Gerenciando produtos de dados autônomos e interoperáveis	97
Plataforma contínua de capacidades operacionais e analíticas	98
Projetada para uma maioria generalista	99
Favorecendo tecnologias descentralizadas	100
Domínio agnóstico	101
Pensamento da plataforma de data mesh.....	101
Viabiliza que equipes autônomas obtenham valor a partir dos dados.....	104
Troca de valor com produtos de dados autônomos e interoperáveis	106
Acelera a troca de valor ao reduzir a carga cognitiva	107
Scale-out de compartilhamento de dados	109
Apoio à cultura de inovação intrínseca	110
Transição para uma plataforma de autoatendimento de data mesh	111
Projete APIs e protocolos primeiro	111
Prepare-se para a adoção generalista.....	111
Faça um inventário e simplifique as coisas	112
Crie APIs de alto nível para gerenciar produtos de dados	113
Crie experiências, não mecanismos	113
Comece com alicerces mais simples, e a colheita será o progresso	113
Recapitulando	114

Capítulo 5 = Princípio de governança computacional federada 116

Aplique o pensamento sistêmico à governança de data mesh	119
Mantenha o equilíbrio dinâmico entre a autonomia do domínio e a interoperabilidade global.....	120
Adote a topologia dinâmica como estado padrão	125
Utilize a automação e a arquitetura distribuída.....	125
Aplique a federação ao modelo de governança.....	126
Equipe federada.....	128
Valores norteadores	129

Políticas.....	133
Incentivos.....	134
Aplique computação ao modelo de governança	136
Padrões como código	137
Políticas como código	137
Testes automatizados	139
Monitoramento automatizado	139
Transição para governança computacional federada	139
Delegue accountability aos domínios.....	140
Incorpore a implementação de políticas em cada produto de dados	140
Automatize a viabilização e o monitoramento das intervenções	141
Modele as lacunas	141
Meça o efeito de rede	142
Adote a mudança em vez da constância	142
Recapitulando	142

Parte II - Por que data mesh?..... 146

Capítulo 6 - Momento decisivo..... 147

Altas expectativas de dados.....	148
A imensa divisão dos dados.....	151
Escala: a descoberta de uma abordagem nova	153
Além do previsível.....	154
Abordando o platô de retorno	155
Recapitulando	155

Capítulo 7 - Após o momento decisivo..... 157

Responder com elegância às mudanças em cenário complexo de negócio.....	158
Alinha negócios, tecnologia e agora dados analíticos	159
Fecha a lacuna entre dados analíticos e operacionais.....	160
Localiza alterações/mudanças de dados em domínios de negócios	162
Reduz a complexidade imprevista de pipelines e de cópia de dados	162
Mantém a agilidade diante do crescimento.	163
Remove gargalos centralizados e monolíticos.....	164
Reduz a coordenação de pipelines de dados	165
Reduz a coordenação da governança de dados	166
Viabiliza a autonomia	167
Aumenta a proporção de valor dos dados em relação ao investimento.....	168
Abstrai a complexidade técnica com uma plataforma de dados	168
Incorpora o product thinking em tudo	169
Ultrapassando fronteiras	169
Recapitulando	170

Capítulo 8 ■ Antes do momento decisivo 173

Evolução das arquiteturas analíticas de dados	173
Primeira geração: arquitetura de data warehouse	174
Segunda geração: arquitetura de data lake	175
Terceira geração: arquitetura de nuvem multimodal	177
Características da arquitetura de dados analíticos	178
Monolíticos	179
Propriedade centralizada de dados	184
Orientado à tecnologia	186
Recapitulando	190

Parte III ■ Como projetar a arquitetura de data mesh 191**Capítulo 9 ■ Arquitetura lógica 193**

Interfaces de compartilhamento de dados analíticos orientadas a domínio	197
Design de interface operacional	198
Design de interface de dados analíticos	198
Dependências de dados analíticos entre domínios	199
Produto de dados como quantum de arquitetura	200
Componentes estruturais de um produto de dados	201
Interações de compartilhamento de dados de produtos de dados	208
APIs de descoberta e de observabilidade de dados	210
Plataforma multiplano de dados	210
Plano de plataforma	212
Plano (utilitário) de infraestrutura de dados	213
Plano de experiência do produto de dados	213
Plano de experiência da rede interconectada	213
Exemplo	214
Políticas computacionais incorporadas	215
Sidecar de produto de dados	217
Contêiner computacional de produto de dados	218
Porta de controle	218
Recapitulando	219

Capítulo 10 ■ Arquitetura da plataforma multiplano de dados 222

Projete uma plataforma impulsionada pelas jornadas do usuário	225
Jornada do desenvolvedor de produtos de dados	227
Concepção, exploração, inicialização e fonte	229
Build, teste, implantação e execução	232
Manutenibilidade, evolução e desativação	236
Jornada do consumidor de produtos de dados	238
Concepção, exploração, inicialização e fonte	241
Build, teste, implantação e execução	241

Manutenibilidade, evolução e desativação	242
Recapitulando	242

Parte IV ■ Como projetar a arquitetura do produto de dados 244

Capítulo 11 ■ Projetando um produto de dados com affordances 246

Affordances do produto de dados	247
Características da arquitetura de produto de dados	251
Design influenciado pela simplicidade de sistemas adaptativos complexos	252
Comportamento emergente a partir de regras locais simples	252
Sem orquestrador central	253
Recapitulando	254

Capítulo 12 ■ Design para consumir, transformar e disponibilizar dados..... 255

Disponibilizando dados	255
Necessidades dos usuários de dados.....	256
Disponibilizando propriedades de design de dados.....	258
Design para disponibilização de dados.....	272
Consumindo dados.....	273
Arquétipos de fontes de dados.....	275
Localidade do consumo de dados.....	278
Design para consumo de dados	280
Transformando dados	282
Transformação programática versus não programática.....	283
Transformação baseada em dataflow	284
ML como transformação.....	286
Transformação de variação temporal	286
Design para transformação	286
Recapitulando	288

Capítulo 13 ■ Design para descoberta, compreensão e composição de dados..... 289

Descobrimo, compreendendo, confiando e explorando	289
Comece a descoberta com autorregistro	292
Descubra o URI global.....	292
Entendendo os modelos semânticos e de sintaxe.....	293
Estabeleça confiança com garantias de dados	295
Explorando a estrutura dos dados.....	298
Obtendo informações com a documentação	299
Descobrimo, explorando e compreendendo o design	300
Compondo dados	301
Consumindo propriedades de design dos dados	303
Abordagens tradicionais para capacidade de composição de dados.....	304
Compondo o design de dados	309
Recapitulando	312

Capítulo 14 = Gerenciamento de design, de governança e de observação de dados 313

Gerenciamento do ciclo de vida	313
Gerenciamento do design do ciclo de vida	314
Componentes do manifesto do produto de dados	315
Governança de dados	316
Design de governança de dados	317
Padronização de políticas	319
Integração de dados e de políticas	321
Vinculação de políticas	321
Observando, debugando e auditando	321
Design de observabilidade	323
Recapitulando	327

Parte V = Primeiros passos 328**Capítulo 15 = Estratégia e execução 330**

Você deve adotar o data mesh hoje?	330
Data mesh como elemento da estratégia de dados	335
Framework de execução do data mesh	339
Execução orientada a negócios	340
Execução integral e iterativa	346
Execução evolutiva	347
Recapitulando	364

Capítulo 16 = Organização e cultura 365

Mudanças	367
Cultura	370
Valores	370
Recompensas	373
Motivações intrínsecas	374
Motivações extrínsecas	375
Estrutura	376
Suposições da estrutura organizacional	376
Identificação das fronteiras do produto de dados	385
Pessoas	389
Papéis	389
Desenvolvimento de conjunto de habilidades	393
Processos	395
Principais mudanças de processo	397
Recapitulando	398

Índice remissivo 401