



Universidade Federal do Rio de Janeiro

Escola Politécnica

MBA em Engenharia de Dados
(MBED)

**APLICAÇÃO DE PRINCÍPIOS DE *DATA MESH* EM MODELOS
ORGANIZACIONAIS**

Autor:

Deive Audieres Leal

Orientador:

Manoel Villas Boas Junior, M. Sc.

Coorientador:

Norberto Ribeiro Bellas, M. Sc.

Examinador:

Vinicius Drumond Gonzaga, M. Sc.

**Rio de Janeiro
Agosto de 2023**

Declaração de Autoria e de Direitos

Eu, **Deive Audieres Leal** CPF 061.460.456-75, autor da monografia ***APLICAÇÃO DE PRINCÍPIOS DE DATA MESH EM MODELOS ORGANIZACIONAIS***, subscrevo para os devidos fins, as seguintes informações:

1. O autor declara que o trabalho apresentado na defesa da monografia do curso de Pós-Graduação, Especialização MBA em Engenharia de Dados da Escola Politécnica da UFRJ é de sua autoria, sendo original em forma e conteúdo.
2. Excetuam-se do item 1 eventuais transcrições de texto, figuras, tabelas, conceitos e ideias, que identifiquem claramente a fonte original, explicitando as autorizações obtidas dos respectivos proprietários, quando necessárias.
3. O autor permite que a UFRJ, por um prazo indeterminado, efetue em qualquer mídia de divulgação, a publicação do trabalho acadêmico em sua totalidade, ou em parte. Essa autorização não envolve ônus de qualquer natureza à UFRJ, ou aos seus representantes.
4. O autor declara, ainda, ter a capacidade jurídica para a prática do presente ato, assim como ter conhecimento do teor da presente Declaração, estando ciente das sanções e punições legais, no que tange a cópia parcial, ou total, de obra intelectual, o que se configura como violação do direito autoral previsto no Código Penal Brasileiro no art.184 e art.299, bem como na Lei 9.610.
5. O autor é o único responsável pelo conteúdo apresentado nos trabalhos acadêmicos publicados, não cabendo à UFRJ, aos seus representantes, ou ao(s) orientador(es), qualquer responsabilização/ indenização nesse sentido.
6. Por ser verdade, firmo a presente declaração.

Rio de Janeiro, 05 de agosto de 2023.

Deive Audieres Leal

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Av. Athos da Silveira, 149 - Centro de Tecnologia, Bloco H, sala - 212,
Cidade Universitária Rio de Janeiro – RJ - CEP 21949-900.

Este exemplar é de propriedade Escola Politécnica da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmear ou adotar qualquer forma de arquivamento.

Permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do autor.

DEDICATÓRIA

Dedico este trabalho a todos os professores e colegas que estiveram ao meu lado durante esta jornada de pós-graduação e a quem direta, indiretamente ou poeticamente, me ajudaram a manter a motivação ao longo de todo o processo.

“não lhe prometi o paraíso. e, para evitar enganos, gravei na pele a palavra caos.”

(Dauana Vale)

AGRADECIMENTO

Agradeço aos meus pais, que, embora com as grandes dificuldades da vida, me propiciaram dedicar-me aos estudos e entender que a educação é um bem valioso, e que se com ela a vida é complicada, sem a porta que abre em nossa alma outras vias são abertas para toda sorte de influências negativas e desventuras. Aos professores e colegas de classe, pela troca de ideias e cooperação durante todo o processo acadêmico, muito obrigado pela camaradagem. À minha família, amigos, bichanos e dogs que me deram suporte no cotidiano o meu obrigado.

RESUMO

A abordagem de *Data Mesh* é uma estratégia que visa otimizar a gestão de dados em empresas, descentralizando a responsabilidade pela coleta, armazenamento e análise de dados frente a centralização propiciada pelos *Data Lakes*. Neste trabalho, os quatro princípios do *Data Mesh* são colocados em perspectiva junto aos modelos estruturais organizacionais: funcional, divisional e matricial. A análise desses princípios em consonância com as particularidades de cada estrutura organizacional possibilita uma melhor compreensão de como implementar a filosofia do *Data Mesh*, visto que a desenvolvedora do conceito, Dehghani, não chega a oferecer nenhum ponto de partida na direção da migração em contextos específicos, embora frise a importância de levar em conta o contexto da organização. Desta forma é apresentado uma visão geral dos princípios do *Data Mesh* e dos modelos estruturais, seguido pela análise das diferentes possibilidades de associação entre ambos. Por fim, entende-se que abordar a implementação de acordo com o tipo de modelo organizacional ajuda a inserir a transformação cultural necessária para migração e à descentralização preconizada pelo *Data Mesh*. Por meio dessa exploração, o presente trabalho visa fornecer uma primeira compreensão para a adoção consciente e estratégica do *Data Mesh*, enriquecendo a gestão de dados e tecnologia nas empresas e impulsionando o progresso contínuo na era da transformação digital.

Palavras-Chave: *Data Mesh*, Descentralização de Dados, Democratização de Dados, Engenharia de Dados

ABSTRACT

The Data Mesh approach is a strategy that aims to optimize data management in businesses by decentralizing the responsibility for data collection, storage, and analysis, as opposed to the centralization provided by Data Lakes. In this work, the four principles of the Data Mesh are put into perspective alongside organizational structural models: functional, divisional, and matrix. The analysis of these principles in line with the particularities of each organizational structure allows for a better understanding of how to implement the Data Mesh philosophy, since the concept's developer, Dehghani, does not offer any starting point for migration in specific contexts, although she emphasizes the importance of considering the organization's context. In this way, an overview of the Data Mesh principles and structural models is presented, followed by an analysis of the different possibilities of association between both. Finally, it is understood that approaching implementation according to the type of organizational model helps to insert the cultural transformation necessary for migration and the decentralization advocated by the Data Mesh. Through this exploration, the present work aims to provide a first understanding for the conscious and strategic adoption of the Data Mesh, enriching data and technology management in businesses and driving continuous progress in the era of digital transformation.

Keywords: Data Mesh, Data Decentralization, Data Democratization, Data Engineering.

SIGLAS

API	<i>Application Programming Interface</i>
DevOps	<i>Development Operations</i>
RDBMS	<i>Relational Database Management System</i>
SLA	<i>Service Level Agreements</i>
TI	Tecnologia da Informação

LISTA DE FIGURAS

Figura 2.1	Interoperabilidade no <i>Data Mesh</i>	11
Figura 2.2	Estrutura Funcional	18
Figura 2.3	Estrutura Divisional	19
Figura 2.4	Estrutura Matricial	21

LISTA DE QUADROS

Quadro 2.1	Modelo Centralizado de Governança de Dados vs Modelo <i>Data Mesh</i> de Governança de Dados	16
-------------------	--	----

Sumário

Capítulo 1: Introdução	1
1.1 – Tema	2
1.2 - Justificativa	2
1.3 – Objetivos.....	3
1.4 - Delimitação	3
1.5 – Metodologia.....	4
1.6 – Descrição	4
Capítulo 2: Embasamento Teórico	5
2.1 – Dados no dia a dia	6
2.2 – <i>Big Data</i>	8
2.3 – <i>Data Lake</i>	8
2.4 – <i>Data Mesh</i>	9
2.4.1 – Princípio da propriedade de domínio	11
2.4.2 – Princípio dos dados como um produto	12
2.4.3 – Princípio da plataforma de dados como autosserviço	14
2.4.4 – Princípio da governança computacional federada	15
2.5 – Estruturas Empresariais	17
2.5.1 – Estrutura Funcional	18
2.5.2 – Estrutura Divisional	19
2.5.3 – Estrutura Matricial	20
Capítulo 3: Propostas Tecnológicas	23
3.1 – Princípio da propriedade de domínio nas estruturas	24
3.2 – Princípio dos dados como um produto nas estruturas	25
3.3 – Princípio da plataforma de dados como autosserviço nas estruturas	26
3.4 – Princípio da governança computacional federada nas estruturas	28
Capítulo 4: Resultados Esperados	30
Capítulo 5: Conclusão e Trabalho Futuro	32
5.1 – Conclusão	32
5.2 – Trabalho Futuro	33
Referências Bibliográficas	34

CAPÍTULO 1

Introdução

A rápida transformação digital está redefinindo o cenário empresarial, impelindo as organizações a buscarem novas estratégias e soluções tecnológicas para se adaptarem às mudanças e manterem sua competitividade no mercado. Nesse contexto, o "*Data Mesh*" surge como uma abordagem inovadora e promissora na gestão de dados, proporcionando impactos significativos e resolvendo problemas na centralização de dados em *Data Lakes* no que tange a governança, confiabilidade e disponibilidade. Com o foco na transformação do dado como produto pela área geradora e, portanto, conhecedora do dado em si, agrega-se mais valor, mas sem deixar de lado o viés da democratização.

O termo "*Data Mesh*" faz referência a uma arquitetura de dados distribuída e descentralizada, se opondo às abordagens centralizadas e buscando superar os desafios que são impostos a esse tipo de arquitetura. Essa arquitetura do "*Data Mesh*" permite uma gestão mais eficiente e ágil dos dados, garantindo a qualidade e a velocidade necessárias para impulsionar a inovação e facilitar a tomada de decisões. No entanto, para alcançar o potencial possível com a abordagem do *Data Mesh*, mudanças culturais podem ser necessárias, e compreender como os princípios do "*mesh*" dialogam com diferentes estruturas organizacionais, surge como uma oportunidade e ponto de entrada para a transformação cultural e, assim, auxiliando a organização no caminho de se tornar uma empresa orientada à dados.

Este estudo visa examinar a aplicação dos princípios do *Data Mesh* em diferentes modelos organizacionais, proporcionando uma compreensão mais abrangente de como essa abordagem pode ser adaptada para enfrentar os desafios exclusivos apresentados por cada tipo. Ao analisar as particularidades de cada modelo organizacional em contraste com os princípios do *Data Mesh*, tentaremos identificar as melhores abordagens para que a aplicação da arquitetura seja menos traumática dentro do ambiente empresarial que ainda não tem uma cultura orientada à dados.

Assim, a análise sobre a implementação do *Data Mesh* em diferentes estruturas organizacionais assume relevância ao fornecer uma compreensão valiosa para gestores, líderes e pesquisadores. Ao compreender como essa abordagem pode beneficiar diversas áreas, este

estudo busca fomentar a adoção consciente e estratégica do *Data Mesh*, enriquecendo o campo da gestão de dados e tecnologia e incentivando o avanço contínuo no âmbito corporativo.

1.1 – Tema

O presente trabalho tem como grande área de conhecimento a gestão de dados. O subconjunto de conhecimento sobre o qual o trabalho trata é o *Data Mesh* e a implementação desse tipo de arquitetura nos modelos organizacionais estruturais.

1.2 – Justificativa

A transformação digital tem impulsionado as organizações a adotarem novas abordagens na gestão de dados para enfrentar os desafios impostos pela centralização excessiva em de dados. Embora os *Data Lakes* tenham sido considerados uma solução promissora para o armazenamento de grandes volumes de dados, surgiram problemas significativos, como dados desnecessários, subutilizados pelo negócio, dificuldade na disponibilização e suporte demorado. Diante dessas limitações, a abordagem do *Data Mesh* emerge como uma proposta que busca superar tais obstáculos, promovendo a descentralização dos dados, ao mesmo tempo em que garante a governança e a democratização.

Entretanto, ao adotar o *Data Mesh*, surgem desafios que precisam ser cuidadosamente abordados. Um dos principais desafios é a necessidade de delegar a produção e a disponibilização dos dados a diferentes unidades ou equipes dentro da organização. Isso requer uma mudança cultural significativa.

Além disso, o conceito de *Data Mesh* ainda é relativamente novo e possui poucas referências na literatura acadêmica no Brasil e na indústria, tornando essencial a realização de pesquisas para compreender e avaliar suas implicações práticas. Nesse contexto, a presente pesquisa se propõe a preencher essa lacuna, fornecendo uma análise dos princípios do *Data Mesh* e identificando maneiras de viabilizar sua implementação em diferentes modelos de organização.

Este trabalho visa incentivar a reflexão sobre a importância de considerar a abordagem do *Data Mesh* no contexto da gestão de dados, encorajando as organizações a pensarem no todo e não apenas em uma perspectiva micro quando surgirem oportunidades para a descentralização dos dados. Ao colocar em perspectiva a implementação do *Data Mesh* em organizações

funcionais, divisionais e matriciais, espera-se fornecer visões relevantes para gestores e líderes, permitindo que tomem decisões estratégicas informadas sobre a melhor forma de adotar essa arquitetura e sua cultura.

Ao destacar as vantagens de uma abordagem descentralizada, aliada a uma governança sólida, que são próprias da arquitetura do *Data Mesh*, espera-se com isso estimular a adoção consciente e eficiente do *Data Mesh*, potencializando a inovação, a tomada de decisões ágeis e a competitividade das organizações no ambiente empresarial em constante evolução.

1.3 – Objetivos

Objetivo geral

O objetivo geral do trabalho é explorar a oportunidade que alguns tipos escolhidos de modelos organizacionais empresariais abrem para a implantação do *Data Mesh*.

Objetivo específico

Demonstrar como os princípios do *Data Mesh* quando pensados junto aos modelos organizacionais podem ser a porta de entrada na transformação de organizações para se tornarem orientadas a dados.

1.4 – Delimitação

O presente trabalho foi direcionado a apenas três modelos de estrutura organizacional. Embora ciente de que existem outros, optamos por três por entendermos que são alguns dos mais comuns nas organizações. Vale frisar, que a análise dos outros modelos levaria mais tempo do que dispomos para a execução completa. O estudo também se mostra teórico, sendo possível inúmeras outras análises e perspectivas. Um estudo de caso foi aventado, mas o prazo também foi um delimitador para a execução, sendo então descartado deste trabalho.

Este trabalho visa contribuir para a discussão visto que o *Data Mesh* é uma arquitetura pouco abordada em português, sendo um ponto de partida para análises ou debate com a intenção da aplicação da descentralização dos dados.

1.5 – Metodologia

A partir das definições de cada princípio do *Data Mesh* e dos modelos organizacionais propostos, a saber funcional, divisional e matricial, será colocado em perspectiva como esses princípios e conceitos podem se alinhar ou se distanciar a depender do tipo de abordagem realizada.

1.6 – Descrição

Este trabalho está estruturado da seguinte forma:

No capítulo dois discute-se os principais pontos de diferentes tecnologias. Partindo da ideia de tomada de decisão baseada em dados, passando pelos tipos de dados mais comuns, a saber: estruturados, não-estruturados e semiestruturados. Fala-se sobre o *Data Lake* e dos possíveis problemas que podem vir a acontecer durante sua utilização como centralizador de informações. Apresenta-se o *Data Mesh* e seus princípios e, enfim, as estruturas organizacionais são focalizadas, fechando o capítulo de embasamento teórico.

No capítulo seguinte, são trabalhadas as implementações dos diferentes princípios nos tipos organizacionais. Ressalta-se que todos os tipos estudados são aderentes, ou ao menos tem uma abertura para a implementação do *Data Mesh*. Sendo o principal desafio vencer a cultura no qual a própria empresa pode estar inserida.

Nos capítulos finais, quatro e cinco, são abordados os resultados observados da implementação teórica dos princípios do *Data Mesh* e quais as conclusões podemos observar ao realizar essa abordagem.

CAPÍTULO 2

Embasamento Teórico

Desde que o acesso a dados para suportar a tomada de decisão das empresas se tornou disponível, seu uso foi se expandindo e se tornando cada vez mais estruturado. Através de diversas etapas de desenvolvimento, os dados, assim como seu armazenamento e aplicação, foram ganhando destaque, sendo considerados como o "novo petróleo".

A tomada de decisão orientada a dados é um campo de estudo que se destaca pela capacidade de auxiliar na descoberta de informações úteis a partir de grandes ou complexas bases de dados. Existem vários artigos científicos sobre o assunto, como "Da observação à decisão: métodos de pesquisa e de análise quantitativa e qualitativa de dados" (Freitas; Moscarola, 2002) e "Modelos de tomada de decisão e sua relação com a informação orgânica" (Lousada; Valentim, 2011).

O uso de dados em empresas é um assunto muito discutido atualmente. Um estudo realizado pela McKinsey com 124 empresas de grande e médio porte em diversos setores no Brasil identificou que já existe uma mentalidade baseada em dados em áreas específicas de empresas de alguns setores. A tecnologia da informação (TI) desempenha um papel importante nesse processo, fornecendo dados, informações e conhecimento para as organizações.

Em cerca de dez anos, entre 2010 e 2020, houve uma explosão no volume de dados criados e armazenados, chegando à ordem de 59 zettabytes (Viana, 2021). Isso se deve aos investimentos na aquisição de dados e extração de informações que sustentam o direcionamento mercadológico e executivo. Através da análise e realização de diagnósticos do ambiente, gerentes e tomadores de decisão em geral podem "antecipar o futuro e reduzir riscos e incertezas" (Guimarães; Évora, 2004, p.73) durante o processo decisório.

Para as autoras Guimarães e Évora, para que uma solução proposta para um problema seja eficaz e correta, é fundamental que as informações utilizadas na tomada de decisão sejam precisas, relevantes e oportunas, permitindo uma análise abrangente da situação. A informação é, portanto, a base para tomar decisões acertadas.

A tecnologia da informação (TI), nesse contexto, é uma ferramenta que visa auxiliar na tomada de decisões empresariais de várias maneiras. A TI consiste no uso dos recursos e técnicas da computação com o objetivo de captar, tratar e organizar dados para transformá-los

em conhecimento, compartilhar e analisar informações, automatizar processos e ajudar no gerenciamento e tomada de decisão nas organizações. Com a integração proporcionada pelos sistemas de gestão, softwares personalizados, automação de processos e computação em nuvem, por exemplo, a manipulação e compartilhamento de informações é uma realidade. Isso permite aos líderes e gestores conseguirem um melhor embasamento para as suas decisões.

Partindo dessa premissa sobre a informação, adotando que estamos caminhando cada vez mais em direção a uma sociedade gerida por dados e direcionando para empresas que buscam orientação pelos mesmos para seus negócios e estrutura, veremos como os dados hoje são produzidos e algumas possibilidades de aplicação.

2.1 – Dados no dia a dia

Com o desenvolvimento de novas tecnologias, a miniaturização dos componentes dos produtos e o espalhamento da internet pelos mais diversos ambientes de vivência das pessoas, observou-se no cotidiano o emprego cada vez maior de informações geradas pelo *feedback* desses mesmos produtos. Com isso foi implementado a customização do uso do produto, o desenvolvimento direcionado de novas características e dispositivos, além de se criar uma fonte dispersa e inesgotável de dados.

No contexto de miniaturização de componentes que geram produtos, pode-se falar dos chamados dispositivos vestíveis, os *wearables*.

A wearable computing é o campo da ciência da computação que pesquisa e desenvolve dispositivos computacionais e sensoriais que têm o corpo como suporte e principal fonte de dados. O uso de tecnologias associadas a objetos que já fazem parte do nosso cotidiano possibilita expandir a função dos mesmos a partir da inclusão de um comportamento programável. (Bruno, 2014, p.12)

Difundidos hoje, cita-se os diversos tipos de relógios ou pulseiras que coletam dados dos usuários enquanto praticam atividades físicas ou exercem as funções cotidianas. Os dados podem ser dos mais diferentes tipos: ritmo cardíaco, distância percorrida, pressão e assim por diante. Com isso, faz-se acompanhamento da performance, no caso de atividades desportivas, ou pontos relativos à saúde quando se perfazem para o cotidiano simples do usuário. Estas informações podem então ser sincronizadas com celulares, enviadas através da internet e permitem gerar verdadeiros relatórios dos hábitos diários dos usuários.

Os dados, assim sendo, são uma parte importante do nosso dia a dia e são usados em muitas áreas, incluindo tomada de decisão, resolução de problemas complexos, melhoria da experiência do cliente e aperfeiçoamento dos processos operacionais. A gestão de dados cobre

todo o chamado “ciclo de vida” dos dados, ou seja, desde a sua coleta até o armazenamento de longo prazo, passando por uma série de processamentos de limpeza, curadoria, anotação, indexação, agregação e transformação.

A administração adequada dos dados é crucial para garantir que as informações sejam precisas, confiáveis e acessíveis quando necessário. Isso pode ajudar a tomar decisões mais informadas e resolver problemas complexos de maneira mais eficiente.

Após ciclos de coleta e uso do dado, observou-se que a análise ainda era superficial se comparada ao volume de dados em produção. Mas mais do que o volume, os tipos de dados também são diferentes. Portanto, extrair informações relevantes é um aspecto que desafia os analistas.

Concentrando-se em três grandes grupos, o armazenamento, tratamento e gerenciamento também acabaram por ser diferentes ou exigir abordagens díspares quanto ao manuseio. Daí, então, encaminhamos a discussão dos tipos e características de cada tipo de dado. Sendo eles: dados estruturados, semiestruturados e não-estruturados.

Dados estruturados são aqueles que possuem uma organização bem definida, geralmente em tabelas e colunas com esquemas pré-definidos, facilitando a manipulação e análise dos dados (ESR, 2022). Esta categoria de dados é comumente encontrada em sistemas de gerenciamento de banco de dados relacionais (RDBMS), onde as informações são armazenadas em tabelas, e as relações entre as tabelas são definidas por meio de chaves primárias e secundárias. Os dados estruturados podem ser facilmente consultados e analisados utilizando linguagens e ferramentas de programação específicas, como SQL.

Por outro lado, os dados semiestruturados possuem uma organização parcialmente definida, não tão rígida quanto nos dados estruturados, mas ainda assim com algum grau de estruturação, como hierarquias ou categorias (Simões, 2022). Exemplos de dados semiestruturados incluem arquivos XML, JSON e CSV. Estes dados são geralmente mais flexíveis do que os estruturados, permitindo a inclusão de informações adicionais sem a necessidade de modificar o esquema subjacente.

A terceira categoria, dados não estruturados, refere-se a informações que não possuem uma estrutura predefinida ou organização formal, como documentos de texto, imagens, vídeos e áudio (ESR, 2022). Este tipo de dado é o mais abundante na era da informação e representa um desafio significativo no armazenamento, processamento e análise, devido à sua natureza heterogênea e não hierárquica.

2.2 – *Big Data*

Big Data é um termo que se refere ao grande volume de dados gerados a partir de diversas fontes, que podem ser, como mencionado acima, estruturados, semiestruturados ou não-estruturados. Esses dados são gerados em alta velocidade e apresentam uma grande variedade de formatos.

O conceito de *Big Data* surgiu com o crescimento exponencial da quantidade de dados gerados e armazenados pelas empresas e organizações. Com o aumento da quantidade de dados, tornou-se necessário desenvolver novas tecnologias e metodologias para armazenar, processar e analisar esses dados de maneira eficiente.

O conceito *Big Data* é utilizado para caracterizar os dados que excedem a capacidade de processamento de sistemas de banco de dados convencionais. *Big Data* é muito grande, se move muito rápido, e não se encaixa nas restrições de arquiteturas de banco de dados (Franco; Bessa, 2016)

Um marco importante na história do *Big Data* foi o desenvolvimento do Hadoop, uma plataforma de software open-source utilizada para armazenar dados e executar aplicações em clusters de hardwares comuns (Franco; Bessa, 2016). O Hadoop fornece armazenamento massivo para qualquer tipo de dado, grande poder de processamento e a capacidade de lidar quase ilimitadamente com tarefas e trabalhos ocorrendo ao mesmo tempo.

Uma maneira comum de descrever o *Big Data* é através dos 5 V's: Volume - relacionado à grande quantidade de dados gerados e armazenados; Velocidade - se refere à rapidez com que os dados são gerados e processados; Variedade - se refere aos diferentes tipos e formatos de dados; Veracidade - se refere à confiabilidade e precisão dos dados; Valor - se refere à capacidade dos dados gerarem compreensão para as empresas.

2.3 – *Data Lake*

Conceitualmente, *Data Lake* pode ser definido como um repositório centralizado que permite armazenar todos os dados, independentemente de sua estrutura e formato, facilitando a análise e a busca de conhecimentos relevantes em um cenário de *Big Data*. Além disso, *Data Lakes* possibilitam a realização de análises avançadas e a aplicação de técnicas de aprendizado de máquina e inteligência artificial para gerar visões estratégicas. Trata-se, portanto, de uma arquitetura de armazenamento escalável, capaz de acomodar enorme quantidade de informações em seu formato original e sem a necessidade de padronização prévia.

(...) *Data Lakes* propõem uma maneira de armazenar dados volumosos e de estruturas diversas em seu formato nativo, em um local de armazenamento evolutivo que permite análises posteriores [...]. O conceito de *Data Lake* se opõe ao de *data warehouse*, que produz um banco de dados integrado, altamente estruturado, orientado para decisões e orientado para assuntos, mas tem a desvantagem de dividir os dados em silos¹. (Nogueira; Romdhane, 2018, p.1)

Entre os principais benefícios do *Data Lake*, destacam-se a flexibilidade para armazenar e gerenciar variados tipos de dados, incluindo dados não estruturados e semiestruturados, provenientes de fontes diversas, como redes sociais, dispositivos móveis e sensores. Agilidade na análise e processamento de dados, uma vez que não há necessidade de transformação prévia dos dados em formatos específicos. Facilitação na integração e colaboração entre diferentes áreas de uma organização, permitindo rápido acesso às informações e a tomada de decisões baseada em dados.

Apesar dos benefícios, a implementação de *Data Lake* também apresenta desafios, como dificuldades no gerenciamento de informações, como a padronização dos dados, qualidade e inconsistências, podendo levar ao surgimento de um *Data Swamp* – pântano de dados – um ambiente desorganizado e ineficiente onde os dados são difíceis de serem localizados e aproveitados. A segurança e privacidade dos dados armazenados no *Data Lake* também são preocupações importantes, uma vez que envolvem aspectos como controle de acesso, criptografia e conformidade com regulamentações de proteção de dados. A necessidade de investimentos em infraestrutura, tecnologia e mão de obra especializada é outro aspecto a ser considerado, especialmente para pequenas e médias empresas.

2.4 – *Data Mesh*

O conceito de "*Data Mesh*" emergiu como uma abordagem inovadora para lidar com o gerenciamento de dados em organizações, especialmente aquelas que buscam escalar suas operações de dados para atender às crescentes demandas de análise e tomada de decisões baseada em dados. O termo "*Data Mesh*" foi cunhado por Zhamak Dehghani, uma engenheira de software do ThoughtWorks, em um artigo publicado em 2019, intitulado "How to Move Beyond a Monolithic *Data Lake* to a Distributed *Data Mesh*." Segundo Dehghani "*Data Mesh* é uma abordagem sociotécnica descentralizada para compartilhar, acessar e gerenciar dados analíticos em ambientes complexos e de grande escala, dentro ou entre organizações"² (2022).

¹ Traduzido pelo autor

² Traduzido pelo autor

Antes do *Data Mesh*, muitas organizações centralizavam seus dados em um único local, como o *Data Lake* mencionado anteriormente, onde todos os dados eram armazenados, processados e compartilhados. No entanto, essa abordagem tradicional começou a apresentar desafios à medida que a quantidade de dados crescia exponencialmente, levando a problemas de escalabilidade, governança, colaboração entre equipes e propriedade dos dados.

O conceito de *Data Mesh* propõe uma mudança fundamental no paradigma do gerenciamento de dados. Em vez de centralizar todos os dados em um único local, ele propõe a descentralização dos dados em diferentes unidades de negócio ou "domínios". Cada domínio seria responsável por seus próprios dados e estabeleceria os princípios de governança, qualidade e compartilhamento dos dados.

A ideia por trás do *Data Mesh* é tornar o gerenciamento de dados mais ágil, escalável e orientado por resultados, capacitando equipes individuais a serem proprietárias e responsáveis por seus próprios dados, ao mesmo tempo que fornecem os serviços e ferramentas necessários para facilitar a colaboração e o compartilhamento de informações em toda a organização.

Entre os benefícios do uso do *Data Mesh* destacam-se: Maior escalabilidade - permite que as organizações escalem suas operações de dados sem ter que centralizar todos os dados em um único local; Melhor governança - facilita a governança dos dados, pois a responsabilidade pela propriedade e gerenciamento dos dados é distribuída entre as diferentes unidades de negócio ou domínios; Maior colaboração - facilita a colaboração entre as diferentes unidades de negócio ou domínios, pois os dados são disponibilizados por meio de APIs, Interface de Programação de Aplicação, ou outras ferramentas; Melhor tomada de decisão - propicia a tomada de decisão baseada em dados, pois as equipes individuais são proprietárias e responsáveis por seus próprios dados.

Deghani apresenta quatro princípios que norteiam os objetivos do *Data Mesh*: Princípio da propriedade de domínio - cada domínio é responsável por seus próprios dados e estabelece os princípios de governança, qualidade e compartilhamento dos dados; Princípio dos dados como um produto - os dados são tratados como um produto e são disponibilizados para outras unidades de negócio ou usuários por meio de APIs ou ferramentas próprias; Princípio da plataforma de dados como autosserviço - as unidades de negócio ou usuários são capazes de acessar e processar dados de forma independente, sem a necessidade de intervenção de uma equipe de TI; Princípio da governança computacional federada - a governança dos dados é distribuída entre as diferentes unidades de negócio ou domínios. Os quatro princípios trabalham na interoperabilidade complementando um ao outro e abordando os desafios que possam vir a

surgir de cada um, Figura 2.1. Tendo sido apresentados os princípios, segue-se a conceitualização de cada um.

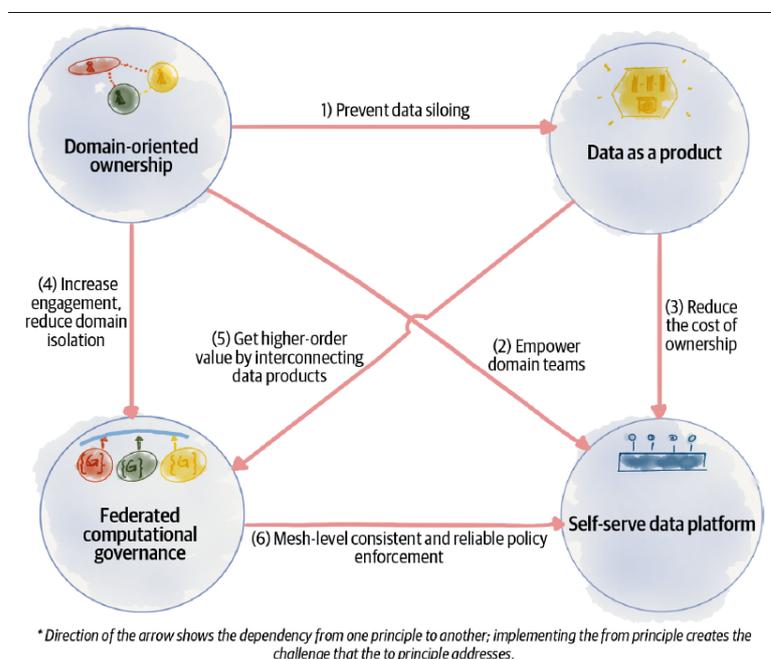


Figura 2.1 – Interoperabilidade no *Data Mesh*.

Fonte: Dehghani, 2022.

2.4.1 – Princípio da propriedade de domínio

O princípio da propriedade de domínio, *Domain Ownership*, é um dos pilares fundamentais do conceito de *Data Mesh*. Ele enfatiza que cada equipe ou unidade de negócio deve ser proprietária dos dados que produz, coleta ou utiliza. Essa abordagem é fundamental para o sucesso da implementação do *Data Mesh*, pois busca descentralizar o gerenciamento de dados e garantir que as equipes que têm o conhecimento especializado sobre um domínio específico sejam responsáveis por seus próprios dados.

Descentralize a propriedade dos dados analíticos para os domínios de negócios mais próximos dos dados, seja a fonte dos dados ou seus principais consumidores. Decomponha os dados (analíticos) de forma lógica e com base no domínio de negócios que eles representam, e gerencie o ciclo de vida dos dados orientados para domínio de forma independente³ (Dehghani, 2022, pos.720-732).

O termo "domínio" refere-se a uma área específica do negócio, como vendas, marketing, logística, recursos humanos, financeiro, entre outros. Cada domínio geralmente possui suas próprias fontes de dados, estruturas de dados e requisitos exclusivos para análises e tomadas de

³ Tradução do autor

decisão. Ao aplicar o princípio da propriedade de domínio, a equipe que opera nesse domínio torna-se responsável por criar, gerenciar e disponibilizar seus próprios dados.

Os benefícios do princípio incluem propriedade e responsabilidade claras, agilidade nas tomadas de decisão, flexibilidade e customização, autonomia para inovação e fomento à colaboração. Ao atribuir a propriedade do domínio aos especialistas de negócios e equipes que atuam nessa área, há uma clara responsabilidade sobre a qualidade, integridade e segurança dos dados. Além disso, as equipes que possuem a propriedade dos dados podem responder rapidamente às necessidades de análise e tomada de decisão do seu domínio específico. Isso reduz a dependência de solicitações a uma equipe centralizada de dados e acelera o processo de obtenção de informações cruciais para a empresa.

Cada domínio pode adaptar a estrutura dos dados de acordo com suas necessidades específicas. Isso permite uma maior customização e flexibilidade, pois diferentes áreas da organização têm requisitos variados de dados. Com a propriedade de dados, as equipes podem explorar novas fontes de dados, implementar novas tecnologias e adotar práticas inovadoras para impulsionar o desenvolvimento e crescimento de seu domínio. Outros domínios podem solicitar dados e, se apropriado, ocorrer a fusão de dados para uma análise mais abrangente.

O princípio da propriedade de domínio coloca a responsabilidade do gerenciamento de dados nas mãos das equipes de negócios que melhor compreendem suas necessidades e contextos específicos. Isso resulta em uma abordagem mais ágil, colaborativa e adaptável para o gerenciamento de dados em toda a organização.

Por fim, podemos recuperar Murray e sua abordagem sobre a propriedade de domínio:

O Princípio da Propriedade de Domínio é a base do design orientado a domínios (D3), uma abordagem para construir aplicativos de software que enfatiza o conhecimento de negócios e os dados. O Princípio da Propriedade de Domínio fornece uma maneira consistente de pensar sobre domínios, sua propriedade e suas características. Isso significa que você pode usá-lo de forma consistente em todos os seus projetos, independentemente de serem totalmente internos ou terceirizados; em qualquer ponto durante o ciclo de vida do seu projeto, não haverá surpresas em relação ao gerenciamento, propriedade e evolução do domínio⁴ (Murray, 2023, pos.450).

2.4.2 – Princípio dos dados como um produto

O princípio dos dados como um produto, *Principle of Data as a Product* no inglês, afirma que os dados devem ser tratados como ativos valiosos dentro de uma organização, em

⁴ Tradução do autor

vez de secundários ou resultados de processos de negócios. Ao adotar esse princípio, as equipes de dados tratam os dados que gerenciam com o mesmo nível de cuidado e atenção dedicado a qualquer produto ou serviço oferecido pela empresa, projetando, desenvolvendo, entregando e oferecendo suporte.

A ideia por trás do princípio dos dados como um produto é que, ao tratar os dados como um produto, as equipes responsáveis pelos dados podem aplicar as mesmas práticas e metodologias usadas no desenvolvimento de produtos para gerenciar e entregar seus dados. Isso pode incluir a definição de requisitos, o design de interfaces de usuário, a implementação de testes automatizados e a entrega contínua de novas versões dos dados.

Dados como produto aplica o pensamento de produto aos dados orientados a domínios para remover atritos de usabilidade e realmente aprimorar a experiência dos usuários de dados - cientistas de dados, analistas de dados, exploradores de dados e qualquer pessoa intermediária. Dados como produto espera que os dados analíticos fornecidos pelos domínios sejam tratados como um produto, e os consumidores desses dados devam ser tratados como clientes⁵ (Dehghani, 2022, pos.1259).

Para entender melhor o princípio dos dados como um produto, é útil comparar a abordagem tradicional de gerenciamento de dados com a abordagem proposta pelo *Data Mesh*. Na abordagem tradicional, os dados são frequentemente vistos como subprodutos dos sistemas e armazenados em um *Data Lake* ou data warehouse centralizado. Os dados são compartilhados entre as equipes de acordo com uma estrutura hierárquica e centralizada, o que pode levar a gargalos, dificuldades de acesso e problemas de governança. Com o princípio, cada equipe de dados trata os dados que ela gerencia como um ativo valioso e independente. Isso significa que a equipe é responsável por todos os aspectos do ciclo de vida dos dados, desde sua captura até sua disponibilização e manutenção contínua.

Os aspectos-chave do princípio dos dados como um produto incluem: Proprietário do produto de dados - cada conjunto de dados tem um proprietário ou uma equipe claramente designada que é responsável por entender as necessidades dos usuários dos dados, garantir a qualidade e governança dos dados e oferecer suporte para o uso correto dos dados; Definição do contrato de serviço - a equipe responsável pelo produto estabelece um contrato claro com os usuários do dado que inclui informações sobre o formato do dado, SLAs (*Service Level Agreements*), limitações de uso e outros detalhes relevantes; Autoatendimento e plataforma de dados - o produto deve ser oferecido por meio de uma plataforma que permite aos usuários descobrirem, acessarem e utilizarem os dados autonomamente; Medição do valor e desempenho - a equipe responsável pelo produto deve medir o valor gerado pelos dados e avaliar o

⁵ Tradução do autor

desempenho do produto para permitir uma melhoria contínua e adaptação às necessidades dos usuários; Orientação ao usuário - o foco do produto é atender às necessidades dos usuários finais para permitir que eles tomem decisões informadas e realizem análises relevantes.

2.4.3 – Princípio da plataforma de dados como autosserviço

Este princípio, *Principle of the Self-Serve Data Platform*, propõe que as unidades de negócio ou usuários tenham a capacidade de acessar e processar dados de forma independente, sem a necessidade de intervenção de uma equipe centralizada de TI. Isso é alcançado por meio da implementação de uma plataforma de dados que permita o autosserviço e que forneça as ferramentas e serviços necessários para que os usuários possam descobrir, acessar e utilizar os dados por conta própria. É uma ferramenta descentralizada e com auto-organização, enfatizando que a responsabilidade da equipe de negócios pelo gerenciamento dos dados que lhe pertencem (Murray, 2023)

A ideia por trás do princípio da plataforma de dados como autosserviço é que, ao fornecer aos usuários as ferramentas e serviços necessários para acessar e processar dados de forma independente, as organizações podem aumentar a agilidade e a eficiência do gerenciamento de dados. Isso pode levar a uma redução nos gargalos e atrasos causados pela dependência de uma equipe centralizada de TI, permitindo que as equipes individuais respondam rapidamente às necessidades de análise e tomada de decisão.

A plataforma de dados como autosserviço capacita as equipes a trabalharem de maneira mais ágil, eficiente e independente, ao mesmo tempo em que mantém a governança e a qualidade dos dados. Algumas características e benefícios importantes do princípio da plataforma de dados como autosserviço incluem o acesso facilitado aos dados, maior autonomia para as equipes, padronização e governança, redução da sobrecarga na equipe central de dados, fomento à inovação e experimentação, feedback contínuo e melhoria e facilitação da adoção de práticas *DevOps* - Desenvolvimento e Operação.

A plataforma fornece uma interface amigável que permite aos usuários descobrirem e acessar facilmente os dados necessários para suas análises e tomadas de decisão. Isso elimina a necessidade de passar por uma equipe centralizada para cada solicitação. Com a plataforma, as equipes têm maior autonomia e controle sobre os dados que possuem ou precisam utilizar. Elas podem selecionar os dados relevantes para seu domínio sem esperar por longos processos de aprovação ou integração. Embora a plataforma seja autosserviço, ela ainda deve garantir a

conformidade com os padrões estabelecidos pela organização. Isso ajuda a manter a consistência e confiabilidade dos dados mesmo com a descentralização.

Ao permitir que as equipes acessem os dados por meio da plataforma, a equipe central fica menos sobrecarregada com solicitações diárias, podendo se concentrar em tarefas mais estratégicas. A plataforma encoraja a inovação e experimentação, permitindo que as equipes combinem e analisem dados para encontrar relações úteis. A plataforma deve ser projetada com base em feedback contínuo dos usuários para melhorar sua usabilidade, recursos e desempenho. Isso garante que ela atenda às necessidades em constante mudança dos usuários. A plataforma muitas vezes segue práticas *DevOps*, permitindo que as equipes itere rapidamente no desenvolvimento, implantação e gerenciamento da plataforma.

A chave para construir a infraestrutura de dados como uma plataforma é (a) não incluir nenhum conceito específico de domínio ou lógica de negócios, mantendo-a agnóstica de domínio, e (b) certificar-se de que a plataforma esconde toda a complexidade subjacente e fornece os componentes da infraestrutura de dados na forma de autosserviço⁶ (Dehghani, 2019).

2.4.4 – Princípio da governança computacional federada

O princípio da governança computacional federada, *Principle of Federated Computational Governance*, preconiza a descentralização da governança dos dados, distribuindo-a entre as diversas unidades de negócio ou domínios da organização, em vez de centralizá-la em um único departamento. Para alcançar esse objetivo, são implementadas políticas e práticas de governança que se aplicam a toda a empresa, mas que permitem a cada domínio adaptar e personalizar tais políticas de acordo com suas necessidades específicas.

[...] a governança de dados deve ser descentralizada e distribuída por toda a organização, em vez de centralizada. Isso permite maior flexibilidade e adaptabilidade na forma como os dados são gerenciados e acessados, e garante que a propriedade e o controle dos dados permaneçam distribuídos entre as várias equipes e indivíduos dentro de uma organização⁷ (Murray, 2023, pos.828).

Para garantir a eficácia dessa governança distribuída, o princípio incentiva a criação de acordos claros e bem definidos entre os diferentes domínios que produzem, consomem ou compartilham dados. Esses acordos abrangem aspectos como a definição de esquemas de dados, qualidade dos dados, segurança, conformidade com regulamentações e outras políticas relevantes.

⁶ Tradução do autor

⁷ Tradução do autor

A ênfase na colaboração entre os domínios é outro elemento-chave da governança computacional federada. Essa colaboração é incentivada por meio de fóruns de discussão, reuniões regulares ou outras formas de interação, permitindo que os domínios trabalhem juntos para compartilhar dados relevantes e encontrar soluções para desafios comuns.

Além disso, a transparência e visibilidade são princípios importantes na tomada de decisões relacionadas aos dados. Documentar e comunicar as decisões de governança de forma transparente é essencial para garantir que todas as partes envolvidas compreendam as políticas e procedimentos estabelecidos.

Apesar da descentralização, a autonomia dos domínios deve ser exercida com responsabilidade. Isso significa que cada equipe ou unidade de negócio deve agir de acordo com os acordos estabelecidos, considerando o impacto de suas ações no ecossistema de dados mais amplo.

Para assegurar a conformidade com os acordos estabelecidos e a eficácia das políticas de governança, a governança computacional federada requer um processo contínuo de monitoramento. O uso de métricas e indicadores relevantes auxilia na avaliação da qualidade dos dados, no desempenho da governança e na satisfação dos usuários.

Dehghani (2020) aponta as diferenças entre uma governança centralizada em contraste com a governança proposta pelo *Data Mesh* no seguinte quadro:

Quadro 2.1 – Modelo Centralizado de Governança de Dados vs Modelo *Data Mesh* de Governança de Dados.⁸

Aspectos de Governança Centralizada	Aspectos de Governança do <i>Data Mesh</i>
Equipe centralizada	Equipe federada
Responsável pela qualidade dos dados	Responsável por definir como modelar o que constitui qualidade
Responsável pela segurança dos dados	Responsável por definir aspectos da segurança dos dados, como níveis de sensibilidade de dados para a plataforma construir e monitorar automaticamente
Responsável por cumprir a regulamentação	Responsável por definir os requisitos de regulamentação para a plataforma construir e monitorar automaticamente
Custódia centralizada de dados	Custódia federada de dados por domínios
Responsável pela modelagem de dados canônicos globais	Responsável pela modelagem de polissemias - elementos de dados que cruzam as fronteiras de vários domínios
A equipe é independente dos domínios	A equipe é composta por representantes de domínios
Com o objetivo de obter uma estrutura de dados estática bem definida	Com o objetivo de permitir uma operação de malha eficaz, adotando uma topologia de malha continuamente mutável e dinâmica

⁸ Traduzido pelo autor

Aspectos de Governança Centralizada	Aspectos de Governança do <i>Data Mesh</i>
Tecnologia centralizada usada por lago/armazenamento monolítico	Tecnologias de plataforma de autosserviço usadas por cada domínio
Mede o sucesso com base no número ou volume de dados governados (tabelas)	Mede o sucesso com base no efeito de rede - as conexões que representam o consumo de dados na malha
Processo manual com intervenção humana	Processos automatizados implementados pela plataforma
Previne erros	Detecta erros e se recupera por meio do processamento automatizado da plataforma

Fonte: Dehghani, 2020

2.5 – Estruturas Empresariais

A estrutura empresarial é um elemento de extrema importância para o bom funcionamento de qualquer negócio. Ela desempenha um papel crucial ao facilitar o alcance das metas estabelecidas, coordenar as atividades internas e criar um ambiente de trabalho eficiente e produtivo. Em essência, a estrutura organizacional define como a empresa é organizada em unidades menores, como departamentos, cada uma com responsabilidades específicas. A coordenação e colaboração entre essas unidades são fundamentais para garantir o sucesso da empresa como um todo.

A diversidade das organizações em termos de tamanho, setor de atuação, estratégia e cultura torna cada negócio único, e é exatamente por isso que existem diversos tipos de estruturas organizacionais. Cada modelo traz suas próprias vantagens e desvantagens, o que deixa a escolha do mais adequado uma decisão crucial para o futuro da empresa.

Dentre os tipos mais comuns de estruturas organizacionais, destacam-se a estrutura funcional, a estrutura divisional e a estrutura matricial. Ressalta-se que há diversos outros tipos, mas que se opta por apresentar apenas os três citados.

A estrutura organizacional é a base que sustenta o funcionamento da empresa e influencia diretamente sua eficiência e sucesso. A compreensão das diferentes opções disponíveis e a análise criteriosa das vantagens e desvantagens de cada modelo são essenciais para tomar uma decisão adequada, garantindo que a estrutura escolhida esteja alinhada com os objetivos e estratégias da empresa. Uma escolha bem-feita proporciona uma base sólida para o crescimento e a prosperidade do negócio, enquanto uma estrutura mal pensada pode criar desafios desnecessários e dificultar o alcance dos resultados desejados.

Segue-se os principais aspectos relacionados a cada modelo de estrutura organizacional abordados por esse trabalho.

2.5.1 – Estrutura Funcional

A estrutura funcional, figura 2.2, é um dos modelos mais comuns de organização empresarial, em que as unidades são agrupadas de acordo com suas funções específicas, como marketing, vendas, finanças, produção e recursos humanos. “A organização funcional é o tipo de estrutura organizacional que aplica o princípio funcional ou princípio da especialização das funções” (Chiavenato, 2003, p.189). Essa abordagem é especialmente adequada para empresas de porte menor e que operam em mercados mais estáveis, proporcionando uma divisão clara de tarefas e responsabilidades.

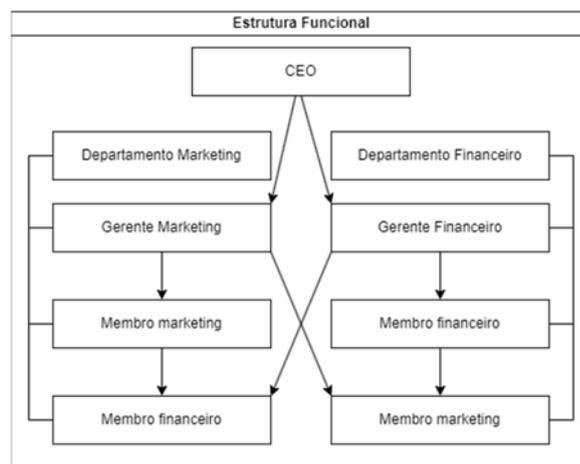


Figura 2.2 – Estrutura Funcional.

Fonte: O autor, 2023.

Uma das principais vantagens da estrutura funcional é a promoção da eficiência ao reunir funcionários com habilidades e conhecimentos similares. A proximidade entre profissionais da mesma área facilita a comunicação e coordenação das atividades, contribuindo para um fluxo mais eficiente de informações e recursos. Além disso, a concentração de colaboradores em suas respectivas especializações permite que eles aprimorem suas habilidades, aumentando a produtividade e a qualidade do trabalho.

Outro aspecto positivo é que a estrutura funcional concede maior controle aos gestores sobre as atividades da empresa. Com os departamentos claramente definidos, os gerentes têm uma visão mais detalhada do desempenho de cada setor e podem tomar decisões com base em informações mais precisas. Isso contribui para uma gestão mais assertiva e ágil, especialmente em situações que demandam respostas rápidas.

Entretanto, a estrutura funcional também possui desvantagens importantes. Sua rigidez pode tornar a adaptação a mudanças mais lenta e desafiadora, sendo problemática em mercados altamente dinâmicos e em constante transformação. O foco centrado nos processos internos

pode levar à negligência das necessidades e expectativas dos clientes, resultando em menor satisfação e fidelidade do público-alvo.

Outro ponto crítico é a falta de estímulo à inovação. A estrutura funcional, ao enfatizar a especialização e a eficiência em processos já estabelecidos, pode desencorajar a busca por novas ideias e práticas disruptivas. Em mercados altamente competitivos, onde a inovação é essencial para se destacar, essa rigidez pode representar uma desvantagem significativa.

Em suma, a estrutura funcional é uma abordagem organizacional que traz benefícios para empresas menores e atuantes em mercados mais estáveis. Sua ênfase na eficiência e especialização favorece a comunicação e o desempenho interno. No entanto, é importante ponderar suas limitações, como a falta de flexibilidade e inovação, que podem afetar a competitividade e a capacidade de adaptação em ambientes de negócios em constante transformação. Portanto, a escolha da estrutura organizacional deve levar em consideração a natureza do negócio, suas metas e o contexto em que a empresa está inserida.

2.5.2 – Estrutura Divisional

A estrutura divisional, figura 2.3, é uma forma de organização empresarial em que a empresa é dividida em unidades autônomas, cada uma responsável por produtos, serviços, regiões ou mercados específicos. Essa abordagem é comumente adotada por empresas grandes e complexas, que necessitam de agilidade e precisão nas tomadas de decisão em diferentes áreas de atuação.

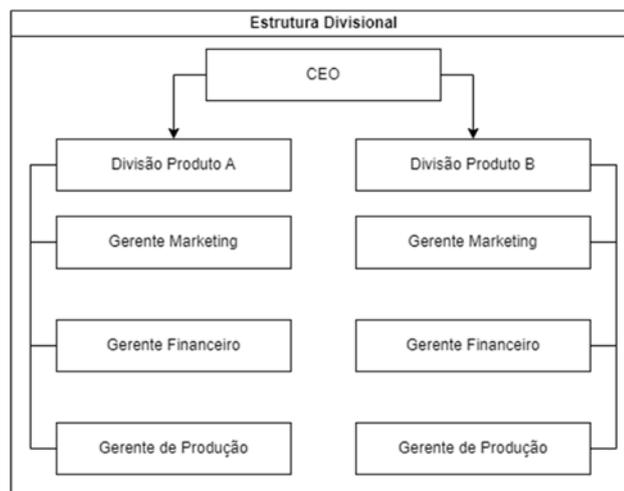


Figura 2.3 – Estrutura Divisional.
Fonte: O autor, 2023.

Cada divisão dentro da estrutura divisional opera de maneira independente, com sua própria equipe de gestão, recursos e orçamento. Essa autonomia permite uma tomada de decisão mais ágil e adaptável às necessidades específicas de cada unidade, possibilitando a empresa a se expandir e crescer de forma rápida e eficiente. A criação de novas divisões para gerenciar operações em novos mercados ou o lançamento de novos produtos se torna mais fácil, evitando mudanças complexas na estrutura organizacional como um todo. “A estrutura divisional agrupa as atividades por unidade em que cada agrupamento tem suas particularidades” (Fraga da Costa; Silva de Souza; Albuquerque Fell, 2012, p.64).

Além disso, a estrutura divisional favorece a concentração nos mercados-alvo e no atendimento personalizado aos clientes. Ao direcionar suas divisões para áreas específicas, a empresa pode dedicar-se de forma mais assertiva às necessidades de seus clientes, o que pode resultar em maior satisfação e fidelidade.

Entretanto, a estrutura divisional também apresenta algumas desvantagens. A complexidade inerente desse modelo pode dificultar a comunicação e coordenação entre as diversas divisões. À medida que a empresa cresce e se torna mais ampla, a troca de informações e o alinhamento estratégico podem se tornar mais desafiadores.

Outra desvantagem é o potencial aumento de custos. Como cada divisão opera de forma autônoma, é necessário um conjunto de recursos gerenciais para cada uma delas. Essa multiplicação de equipes de gestão pode aumentar os custos operacionais da empresa, o que precisa ser cuidadosamente avaliado.

Assim, a estrutura divisional é uma opção adequada para empresas grandes e complexas que enfrentam a necessidade de tomar decisões rápidas e precisas em diferentes áreas de negócios. A autonomia das divisões permite uma adaptação ágil a novos mercados e oportunidades de crescimento.

2.5.3 – Estrutura Matricial

A estrutura matricial é uma abordagem organizacional que reúne características da estrutura funcional e da estrutura divisional. Nesse modelo, os colaboradores possuem dupla subordinação, reportando-se a um gerente funcional, responsável por sua área de especialização, e a um gerente de projeto, responsável pelo projeto específico em que estão envolvidos.

A utilização da estrutura matricial, figura 2.4, é frequentemente adotada por empresas que buscam equilibrar flexibilidade e eficiência. Ela oferece vantagens significativas ao possibilitar que as organizações se adaptem rapidamente às mudanças do mercado e implementem projetos complexos com maior agilidade. Dessa forma, essa estrutura é especialmente adequada para empresas que atuam em setores dinâmicos e competitivos, onde a rapidez na tomada de decisões e a eficácia na execução de projetos são essenciais para o sucesso.

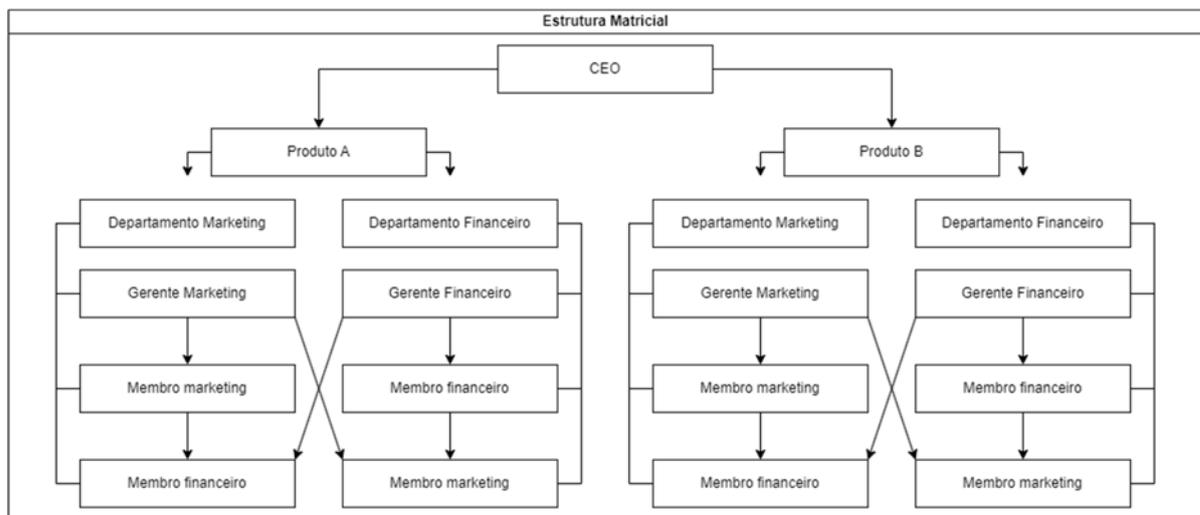


Figura 2.4 – Estrutura Matricial
Fonte: O autor, 2023.

Entretanto, a estrutura matricial também apresenta desafios significativos. O fato de os funcionários responderem a dois chefes diferentes pode gerar conflitos de prioridades e objetivos. Além disso, a natureza complexa dessa abordagem pode levar à confusão entre os colaboradores, especialmente quando precisam lidar com múltiplos projetos simultaneamente.

(...) tendo em vista sua característica de responsabilidade compartilhada, exige nível de confiança mútua e capacidade de improvisação da solução de problemas. Dessa forma, é importante o estudo de liderança dos elementos da alta administração, que têm grande influência em relação ao conflito inevitável desse tipo de departamentalização, que pode ser minimizado, se administrado com eficiência e eficácia (Oliveira, 2004, p.115).

Apesar dos desafios, a estrutura matricial tem se mostrado bastante eficaz em setores específicos, como tecnologia, consultoria e marketing, onde a capacidade de gerenciar projetos complexos e responder rapidamente às demandas do mercado é crucial para o sucesso da empresa. Também é uma escolha comum em empresas que precisam coordenar projetos de grande escala, como construções ou desenvolvimento de novos produtos.

Dessa forma, a estrutura matricial é uma solução organizacional versátil e adaptável, capaz de combinar os pontos fortes das estruturas funcional e divisional. Seu uso é especialmente relevante em setores dinâmicos e complexos, onde a agilidade e a eficiência na execução de projetos são fundamentais.

CAPÍTULO 3

Propostas Tecnológicas

Ao considerar os problemas enfrentados na utilização de *Data Lakes*, tal como o pântano de dados, um ambiente desorganizado e ineficiente onde os dados são difíceis de serem localizados e aproveitados, e a solução preconizada de implantação do *Data Mesh* como substitutivo e contendo uma arquitetura aderente à descentralização da informação, surge a necessidade de explorar as possibilidades de execução de implantação em diferentes cenários.

Ao procurar as propostas de Dehghani (2022) para a implantação verificamos que ela sugere que seja feita de maneira iterativa e incremental, permitindo que a cada iteração o adotante aprenda, refina a estratégia e siga incrementalmente.

Também sugere a adoção de uma função fitness:

Uma função de fitness é uma função objetiva usada para resumir quão próxima uma solução de design prospectiva está de atingir seus objetivos definidos. Na computação evolutiva, a função de fitness determina se um algoritmo melhorou ao longo do tempo⁹ (Dehghani, 2022, pos.7071-7086).

Ainda que a autora dê pistas de como realizar a migração ou implementação da solução, destacando que as estruturas organizacionais legadas, arquitetura e tecnologia devam ser incorporadas no projeto de migração, ela não apresenta como isso pode ser feito, levando em conta as características únicas de cada organização. “Elaborar uma estratégia de migração de um estado específico para o *Data Mesh* está fora do escopo deste livro e requer uma estreita colaboração com uma organização e a compreensão do seu contexto único”¹⁰ (Dehghani, 2022, pos.7158).

Assim, nos esforçamos neste trabalho para trazer uma visão de como isso pode ser realizado em ambientes que detêm alguns tipos de modelos organizacionais. No recorte adotado analisamos os modelos: funcional, divisional e matricial. Tendo sempre como orientação a visão de Dehghani sobre a solução:

O objetivo final do *Data Mesh* é aumentar a capacidade das organizações de utilizar dados para fins analíticos e obter valor de seus dados em escala, alinhada com o

⁹ Tradução do autor

¹⁰ Tradução do autor

crescimento e a complexidade organizacional. Cada um dos princípios do *Data Mesh* contribui para esse resultado de alto nível¹¹ (Dehghani, 2022, pos.7086-7093).

Assim sendo, a análise inicia-se pela separação dos princípios fundamentais do *Data Mesh*, conforme apresentados no capítulo anterior, que serão aplicados em cada uma das características específicas de cada estrutura organizacional. Ao concluir a aplicação dos princípios do *Data Mesh* em cada uma das estruturas organizacionais mencionadas, ter-se-á uma visão mais clara dos pontos positivos e dos desafios específicos em cada contexto. Esse estudo pode servir como ponto de partida para que outras organizações interessadas em implementar *Data Mesh* em seus ambientes, obtenha uma primeira compreensão de como isso pode ser realizado. Sendo consciente de que a compreensão das nuances e particularidades de cada estrutura organizacional é crucial para o sucesso da implantação de estratégias de gerenciamento de dados.

3.1 – Princípio da propriedade de domínio nas estruturas

Com relação à estrutura funcional, nota-se uma proximidade conceitual entre o princípio da propriedade de domínio e as funções específicas desempenhadas por cada departamento. Essa similaridade torna a estrutura funcional uma rota de entrada viável para a aplicação do *Data Mesh* dentro da organização.

Nesse contexto, o princípio da propriedade de domínio do *Data Mesh* alinha-se perfeitamente com as características das funções individuais. Cada função ou departamento se tornaria o proprietário dos dados que produz, sendo responsável por coletar, gerenciar e governar as informações relacionadas às suas atividades específicas. Levando em consideração que estes serão os que mais conhecerão os dados que serão produzidos.

Já no contexto da estrutura divisional, podemos abordar como cada divisão da organização pode se tornar um domínio de dados, sendo responsável por possuir, gerenciar e governar as informações específicas relacionadas às suas operações. Cada divisão atua de forma autônoma, com sua própria estratégia e objetivos, o que torna a aplicação da propriedade de domínio uma escolha natural. Embora o modelo organizacional possa trabalhar sobre um produto específico, poderemos separá-lo em menores conjuntos já existentes, financeiro por exemplo, direcionando para que se alinhe com o esperado dentro do princípio aqui estudado.

¹¹ Tradução do autor

Ao aplicar a abordagem do *Data Mesh* à estrutura matricial de uma organização, a implantação da propriedade de domínio torna-se uma tarefa desafiadora. Nessa estrutura, a organização opera por meio de equipes multifuncionais, que colaboram em projetos específicos enquanto mantêm suas responsabilidades nas áreas de especialização. Cada equipe multifuncional pode vir a se tornar um potencial domínio de dados, com a responsabilidade de possuir e governar as informações produzidas no âmbito do projeto em que estão envolvidas.

Contudo, a implantação da propriedade de domínio na estrutura matricial requer uma coordenação cuidadosa para garantir a integração e o compartilhamento de informações quando necessário, já que nesse modelo temos dois gerentes diferentes, o que requer um suporte por parte da administração, o que pode levar a grandes embates. Desta forma, os níveis hierárquicos superiores estarem comprometidos com a adoção do *Data Mesh* é de fundamental importância.

3.2 – Princípio dos Dados como um Produto e estruturas

Para lembrar, o Princípio dos Dados como um Produto defende que os dados produzidos pela organização sejam tratados como um produto em si, com requisitos, qualidade e valor a serem entregues aos consumidores internos e externos.

Dessa forma, na estrutura funcional, cada departamento torna-se responsável por produzir e entregar dados de qualidade, alinhados às necessidades e expectativas dos seus usuários. Isso implica em estabelecer uma mentalidade de "produto" em relação aos dados, sendo sempre um desafio na propagação da cultura que, não incomum, entende os dados como resultados de outras atividades e não que possam ter um valor em si.

A implantação do Princípio dos Dados como um Produto requer também a definição de métricas de qualidade e de desempenho dos dados. Cada departamento deve estabelecer padrões e critérios para garantir a confiabilidade, integridade e segurança das informações geradas, tornando os dados mais confiáveis e valiosos para toda a organização.

Ademais, é fundamental que cada departamento funcional considere os feedbacks dos usuários de seus dados. A escuta ativa permite identificar oportunidades de melhoria, ajustar os requisitos dos produtos e garantir que as informações atendam plenamente às demandas dos usuários. Esse retorno deve ser implementado em qualquer projeto que anseie trabalhar o dado como produto, portanto sendo aplicado em todas as estruturas organizacionais.

Aproveitando a mesma abordagem do princípio anterior, no qual cada divisão pode ser pensada como um domínio, o mesmo pode ser aplicado na estrutura divisional. Assim, cada

divisão passa a poder construir seu ou seus dados como produtos. Separadamente, cada divisão é responsável por tratar os dados que produz como um produto, com requisitos claros e qualidade assegurada.

Cada divisão deve passar a atuar como um fornecedor de dados, tratando suas informações como um produto que é oferecido a outras áreas da organização que serão então os clientes desses dados para suas operações e tomada de decisões.

As divisões devem ter suas características respeitadas enquanto domínio, mas propiciando a construção de um número de produtos com base nas ingestões e fontes de dados que nascem nos domínios. Cada produto tem então um proprietário e pode ser fornecido entre as demais divisões da organização.

Por fim, na estrutura matricial, cada equipe multifuncional passa a atuar como um produtor de dados, tratando suas informações como produtos. Sendo os proprietários a própria equipe. Sugere-se que nessa abordagem, e pela possibilidade de uma mesma equipe ter inúmeros projetos, uma grande exploração dos diferentes tipos de dados seja feita, agrupando sobre o mesmo proprietário produtos semelhantes e permitindo que eles, sendo conhecedores próximos à fonte, possam ofertar produtos com maior qualidade e aplicabilidade.

A implantação do Princípio dos Dados como um Produto requer uma mudança de mentalidade e cultura organizacional no que diz respeito ao entendimento de “o que é o dado”. Assim, não bastará observar apenas os modelos organizacionais, independente do observado, e se adaptar, mas será necessária uma ação ativa para que a mudança passe de “dado resultante” para “dado com valor e intencionalmente produzido”, levando a transformações muito mais internas do que se supunha pela organização.

3.3 – Princípio da plataforma de dados como autosserviço nas estruturas

Entende-se que esse seja o princípio que menos sofra influência dos modelos organizacionais. Nele observamos que a construção da plataforma, que será disseminada para utilização entre os diferentes produtores, é realizada por uma equipe dedicada à essa função e que não tem uma relação direta com os outros times.

Porém é o princípio que mais deve demonstrar o comprometimento da organização para a execução da implantação ou migração. A organização deve investir em tecnologias e plataformas de dados que suportem o modelo *self-service*. Isso inclui a implementação de uma plataforma de dados que seja amigável e fácil de usar, com recursos que permitam a descoberta

e o acesso rápido aos conjuntos de dados relevantes. Dehghani sugere que a plataforma seja desenvolvida em arquitetura cloud, pois permitiria uma maior interação para todas as equipes que serão as donas dos produtos ou consumidoras dos que já foram criados e estão disponíveis para acesso.

A comunicação aberta e transparente será fundamental para garantir que todas as equipes compreendam os recursos disponíveis na plataforma de dados *self-service* e as políticas de governança que devem ser seguidas. O compartilhamento de feedback entre as equipes e a liderança da organização permitirá a identificação de oportunidades de melhoria contínua na plataforma de dados, ponto que deve ser fortemente trabalhado nos modelos organizacionais.

Independente do modelo organizacional, as equipes serão incentivadas a se tornarem autossuficientes no acesso e gerenciamento de dados relevantes para suas operações e projetos. Cada equipe será capacitada para usar a plataforma de dados *self-service* dedicada ao seu domínio específico, permitindo que elas tomem decisões sobre os dados sem a necessidade de intervenção constante de uma equipe centralizada de TI ou especialistas em dados.

A implementação é iniciada com o projeto e desenvolvimento da plataforma, que deve ser projetada de forma a atender às necessidades específicas de cada equipe. A ideia é proporcionar autonomia a cada equipe para gerenciar seus próprios dados de maneira eficaz e assim, a plataforma, embora genérica, deve conseguir abarcar os nuances e características relacionadas aos diferentes times.

Um ponto crucial para o sucesso da plataforma de autoatendimento é a capacitação e a possibilidade de fornecer treinamentos e suporte para que as equipes possam utilizar a plataforma de maneira eficiente. Isso garantirá que todos os membros das equipes estejam aptos a coletar, preparar e analisar dados de forma independente, mas adequada. Isso pode incluir a disponibilidade de ferramentas de integração de dados, painéis de controle, APIs e outras funcionalidades que permitam o acesso e compartilhamento de dados de forma rápida e descomplicada.

Para otimizar o uso dos recursos disponíveis na plataforma, é fundamental que as equipes compartilhem conhecimentos, experiências e melhores práticas. Promover a colaboração é essencial para alcançar o máximo potencial da plataforma e garantir que todos estejam alinhados com os mesmos padrões e diretrizes.

A cultura da organização também desempenha um papel importante na implementação do princípio. É essencial promover uma cultura data-driven, incentivando as equipes a tomarem decisões com base em dados e evidências, em vez de intuição ou opiniões subjetivas. A

plataforma de autoatendimento contribui para essa cultura, facilitando o acesso rápido e fácil às informações necessárias para suportar os times e as tomadas de decisão.

3.4 – Princípio da governança computacional federada nas estruturas

Lembrando que esse princípio promove descentralização da governança dos dados, distribuindo-a entre as diversas unidades de negócio ou domínios da organização, em vez de centralizá-la em um único departamento, passamos a tratar de sua possível aplicação em diferentes modelos.

Dessa forma segue-se com a estrutura funcional, assim sendo a organização deve estabelecer conselhos de governança compostos por representantes dos diferentes departamentos funcionais. Esses conselhos serão responsáveis por supervisionar a governança dos dados e processos computacionais em suas respectivas áreas de atuação. Eles devem trabalhar de forma colaborativa, compartilhando conhecimentos e experiências para tomar decisões informadas e alinhadas com os objetivos estratégicos da organização.

Cada conselho de governança deve definir protocolos e diretrizes comuns que orientem as equipes dentro de seus domínios. Esses protocolos podem abordar questões como padronização de metadados, segurança de dados, privacidade, compartilhamento de informações e melhores práticas para análise de dados. A colaboração entre os diferentes conselhos de governança é fundamental para garantir a coerência e a consistência nas decisões tomadas em toda a organização.

Para implementá-lo na divisional, envolver as divisões de forma ativa no processo de governança é pré-requisito, permitindo que elas tenham autonomia para tomar decisões relacionadas ao gerenciamento, análise e compartilhamento de dados. Cada divisão assume a responsabilidade pela gestão dos dados em seu próprio domínio, assim sendo, o conselho de governança de dados é composto por representantes de cada divisão, o que garante a representatividade de todas as partes interessadas nas decisões relacionadas aos dados.

A colaboração entre as divisões é incentivada em todo o processo de governança de dados. Grupos de trabalho, reuniões regulares e fóruns de discussão são estabelecidos para permitir a troca de conhecimentos, experiências e boas práticas entre as equipes. Definir responsabilidades e papéis é fundamental para garantir que cada divisão compreenda suas obrigações no gerenciamento de dados. A atribuição de proprietários de dados em cada divisão torna possível garantir a qualidade e a integridade dos dados dentro de cada domínio.

Chegando enfim à estrutura matricial, para implementar esse princípio, faz-se necessário estabelecer uma governança distribuída para as equipes funcionais e de projeto. Cada equipe será responsável pela gestão dos dados e operações computacionais dentro do seu domínio específico, enquanto uma abordagem colaborativa será adotada para a tomada de decisões em relação aos dados em toda a organização.

A estrutura matricial de organização pode ajudar a garantir que os dados sejam gerenciados de forma consistente em toda a organização, mesmo que os dados sejam usados por diferentes funções ou equipes. Isso pode ser feito através de comitê de governança de dados, como nos outros modelos, que é responsável por supervisionar o gerenciamento de dados em toda a organização.

Ao propor que a governança de dados seja compartilhada entre as diferentes equipes multifuncionais, permite-se que cada uma delas tenha autonomia para definir suas próprias políticas, padrões e práticas de dados, considerando as particularidades e necessidades específicas de seus projetos e iniciativas.

Por fim, é importante lembrar que todos os princípios são interconectados. Portanto, ao aplicá-los, independentemente do tipo de modelo organizacional, é necessário considerar a existência do domínio da propriedade, do dado como produto, do autosserviço e da governança em paralelo com a implantação. Isso levará à descentralização e, como sugerido anteriormente, a uma mudança cultural na organização.

CAPÍTULO 4

Resultados Esperados

Ao analisar as abordagens realizadas no capítulo anterior, depreende-se que a estrutura funcional se alinha bem com o princípio da propriedade de domínio porque cada departamento é responsável por uma função específica dentro da organização. Isso significa que eles têm a melhor compreensão dos dados que produzem e são mais propensos a tomá-los para decisões informadas.

A estrutura divisional também se beneficia da propriedade de domínio porque cada divisão é responsável por uma linha de negócios específica. Dessa maneira o princípio pode ser adotado gerando, como na estrutura funcional, uma melhor compreensão dos dados que precisam para tomar decisões sobre seu negócio e acabam fazendo isso de maneira mais eficaz.

A estrutura matricial é mais desafiadora para implementar a propriedade de domínio porque tem uma hierarquia dupla. Significando que cada equipe é responsável por um departamento e uma linha de negócios. Isso pode tornar difícil determinar quem é o proprietário de dados específicos. A partir dessa deliberação, baseada em muito diálogo e a adoção hierárquica das premissas do *Data Mesh* passamos a ter o vislumbre já da descentralização.

Assim sendo, o princípio da propriedade de domínio, demonstra ser uma excelente abordagem para que as equipes iniciem o processo de descentralização e implementação do *Data Mesh*, levando à granularidade dos departamentos e mantendo a coerência e conhecimento necessário para o desenvolvimento e aplicação dos outros princípios. Aumentando a chance de uma implementação com resultados positivos, embora difíceis de determinar sem um estudo de caso adequado.

O princípio dos dados como um produto pode ser aplicado nas diferentes estruturas organizacionais. Em cada uma delas – funcional, divisional e matricial – os dados são tratados como produtos com proprietários responsáveis por sua qualidade e disponibilização. O feedback dos usuários é valorizado, visando aprimorar continuamente a qualidade dos dados.

A implementação de uma plataforma de autoatendimento de dados é essencial para todos os modelos de organização e deve ser projetada para atender às necessidades específicas das equipes e assim promover uma cultura data-driven. Por sua vez a organização deve investir em treinamento e suporte para garantir que as equipes usem a plataforma de

autoatendimento de dados de forma eficaz. Por fim, a comunicação aberta e transparente é essencial para garantir o sucesso da plataforma de autoatendimento de dados.

Os princípios de governança de dados são interconectados e devem ser considerados em conjunto. A colaboração entre os diferentes conselhos de governança é fundamental para garantir a coerência e a consistência nas decisões tomadas em toda a organização. A definição de padrões e práticas é realizada de forma colaborativa pelo conselho de governança de dados, garantindo consistência, qualidade e segurança em toda a organização e é necessário estabelecer uma governança distribuída para as equipes.

CAPÍTULO 5

Conclusão e Trabalho Futuro

5.1 – Conclusão

Podemos concluir que o *Data Mesh* é uma abordagem emergente para gerenciamento de dados que pode ajudar as organizações a lidar com os desafios de um mundo cada vez mais orientado pela informação.

A implementação do *Data Mesh* pode ser desafiadora, mas oferece oportunidades para aprimorar o gerenciamento e governança dos dados em diferentes estruturas organizacionais. Da mesma maneira é importante compreender as nuances e particularidades de cada estrutura organizacional para o sucesso da implementação do *Data Mesh*. Sendo que a aplicação da propriedade de domínio pode servir de entrada principal para a descentralização em si, enquanto os outros princípios tendem a fortalecer e melhorar os resultados compartilhados.

As implementações do *Data Mesh* têm em comum a promoção de uma cultura orientada a dados, o tratamento dos dados como produtos tangíveis, o estabelecimento de processos e responsabilidades para o gerenciamento de dados e a adoção de uma arquitetura orientada a serviços.

Na perspectiva de observação dos princípios do *Data Mesh* em contraste com os modelos organizacionais abordados, vislumbramos que um princípio acabou se sobressaindo e assim “puxando” a análise, o princípio da propriedade de domínio. Corroborando com Dehghani que previa esse princípio como premissa esperada na execução desse tipo de implantação.

Passamos também a ter um vislumbre de como poderia ser abordada a implementação em contextos organizacionais específicos, algo não abordado pela autora do conceito de *Data Mesh* em seus escritos. Oferecendo um ganho para o pensamento estratégico dos interessados em adotar o *Data Mesh* e colaborando com o objetivo de “aumentar a capacidade das organizações de utilizar dados para fins analíticos e obter valor de seus dados em escala, alinhada com o crescimento e a complexidade organizacional” (Dehghani, 2022, pos.7086-7093) já mencionado anteriormente e novamente destacado aqui.

Portanto, o *Data Mesh* pode ser uma abordagem valiosa para as organizações que buscam melhorar a qualidade dos dados, aumentar a agilidade decisória e aumentar a

competitividade em um ambiente de negócios cada vez mais orientado por dados e trabalhar a estratégia organizacional de implementação é essencial para se obter os melhores resultados possíveis.

5.2 – Trabalho Futuro

Como uma nova etapa do estudo realizado, deve-se analisar algum caso de uso para extração e materialidade das dificuldades pelo qual uma implementação do *Data Mesh* pode passar. Da mesma forma, passa-se a ter uma visão dos benefícios que tal abordagem traz para os negócios e para a cultura orientada a dados de uma organização.

REFERÊNCIAS BIBLIOGRÁFICAS

BRUNO, Natalia Chaves. **Wearables, deficiência intelectual de desenvolvimento e movimentação corporal: um estudo sob a perspectiva do design em parceria com o grupo do IPCEP (instituto de psicologia clínica educacional e profissional)**. 2015. 80 f. Dissertação (Mestrado) - Curso de Departamento de Artes e Design, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2015. Disponível em: <https://www.maxwell.vrac.puc-rio.br/29495/29495.PDF>. Acesso em: 15 ago. 2023.

CHIAVENATO, Idalberto. **Introdução à Teoria Geral da Administração**. 8ª ed. São Paulo: Makron Books, 2004.

DEGHANI, Zhamak. **How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh**. 2019. Disponível em: <https://martinfowler.com/articles/data-monolith-to-mesh.html>. Acesso em: 01 jul. 2023.

DEGHANI, Zhamak. **Data Mesh Principles and Logical Architecture**. 2020. Disponível em: <https://martinfowler.com/articles/data-mesh-principles.html>. Acesso em: 01 jul. 2023.

DEGHANI, Zhamak. **Data Mesh: Delivering Data-Driven Value at Scale**. 1ª edição. Kindle Edition. Sebastopol, CA: O'Reilly Media, 2022.

ESCOLA SUPERIOR DE REDES (ESR). **Dados estruturados, não-estruturados e semiestruturados: diferenças e similaridades**. 2022. Disponível em: <https://esr.rnp.br/ciencia-de-dados/dados-estruturados-nao-estruturados-e-semiestruturados/>. Acesso em: 21 ago. 2023.

FRAGA DA COSTA, Maria Clara, SILVA DA SOUZA, Bruno Silvestre, DE ALBUQUERQUE FELL, André Felipe. **Um estudo da estrutura organizacional e as mudanças organizacionais: proposta de um novo modelo**. NAVUS - Revista de Gestão e Tecnologia. 2012. Disponível em: <https://www.redalyc.org/articulo.oa?id=350450809004>. Acesso em: 01 jul. 2023.

FRANCO, Anderson Larcher; BESSA, Geraldo Magela Almeida. **Aplicabilidade, utilidade e ganhos do Big Data utilizando a ferramenta Hadoop**. Caderno de Estudos em Sistemas de Informação, Juiz de Fora, v. 3, n. 1, 2016. Disponível em: <https://seer.uniacademia.edu.br/index.php/cesi/article/view/817>. Acesso em: 16 ago. 2023.

FREITAS, Henrique; MOSCAROLA, Jean. **Da observação à decisão: métodos de pesquisa e de análise quantitativa e qualitativa de dados**. RAE eletrônica, São Paulo, v. 1, n. 1, jun. 2002. Disponível em: <https://www.scielo.br/j/raeel/a/nzhrSTzq6Tm4K6sbKx4Gcqs/>. Acesso em: 20 jul. 2023.

GUIMARÃES, E. M. P.; ÉVORA, Y. D. M.. **Sistema de informação: instrumento para tomada de decisão no exercício da gerência**. *Ciência da Informação*, v. 33, n. 1, p. 72–80, jan. 2004.

LOUSADA, Mariana; VALENTIM, Marta Lígia Pomim. **Modelos de tomada de decisão e sua relação com a informação orgânica**. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 16, n. 1, p. 147-164, mar. 2011. Disponível em: <https://www.scielo.br/j/pci/a/t8XnmBCGfRGYpm9YQh4CKFh/>. Acesso em: 19 jul. 2023

MURRAY, Brian. *Data Mesh: Reinventing Data Architecture for Decentralized and Autonomous Data Teams*. Kindle Edition. 1ª edição. 2023.

NOGUEIRA, I. D.; ROMDHANE, M. **Modeling Data Lake Metadata with a Data Vault**. 22nd International Database Engineering & Applications Symposium, Villa San Giovanni, p. 253 – 261, Jun 2018. Disponível em: <https://arxiv.org/abs/1807.04035>. Acesso em: 25 de ago de 2023.

OLIVEIRA, D. P. R. **Sistemas, organização e métodos: uma abordagem gerencial**. 14. ed. São Paulo: Atlas, 2004.

SIMÕES, Rachel Albuquerque Manguiera. **A importância dos Dados Estruturados, Não estruturados e Semiestruturados os desafios da sua utilização nas organizações brasileiras**. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Unidade Acadêmica de Educação a Distância e Tecnologia, Universidade Federal Rural de Pernambuco, 2022.

VIANA, Bernardo. **O mar de dados virou um oceano e não para de crescer. Mas nem tudo é aproveitado**. 2021. Disponível em: <https://www.insper.edu.br/noticias/o-mar-de-dados-virou-um-oceano-e-nao-para-de-crescer-mas-nem-tudo-e-aproveitado/>. Acesso em: 13 jun. 2023.