



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

THIAGO BOTELHO RODRIGUES

**GOVERNANÇA DESCENTRALIZADA DE DADOS COM DATA MESH:
Fundamentos, Desafios e Estratégias Emergentes**

Recife
2025

THIAGO BOTELHO RODRIGUES

**GOVERNANÇA DESCENTRALIZADA DE DADOS COM DATA MESH:
Fundamentos, Desafios e Estratégias Emergentes**

Trabalho de Conclusão de Curso
apresentado ao Centro de Informática
(CIn) da Universidade Federal de
Pernambuco (UFPE), como requisito
parcial para a obtenção do título de
bacharel em Sistemas de Informação.

Orientador (a): Robson do Nascimento Fidalgo

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Rodrigues, Thiago Botelho.

Governança Descentralizada de Dados com Data Mesh: Fundamentos,
Desafios e Estratégias Emergentes / Thiago Botelho Rodrigues. - Recife, 2025.
51 p. : il., tab.

Orientador(a): Robson do Nascimento Fidalgo

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado,
2025.

Inclui referências.

1. Governança de Dados. 2. Descentralização de Dados. 3. Data Mesh. I.
Fidalgo, Robson do Nascimento. (Orientação). II. Título.

000 CDD (22.ed.)

THIAGO BOTELHO RODRIGUES

**GOVERNANÇA DESCENTRALIZADA DE DADOS COM DATA MESH:
Fundamentos, Desafios e Estratégias Emergentes**

Trabalho de Conclusão de Curso
apresentado ao Centro de Informática
(CIn) da Universidade Federal de
Pernambuco (UFPE), como requisito
parcial para a obtenção do título de
bacharel em Sistemas de Informação.

Aprovado em: 03 / 04 / 2025

BANCA EXAMINADORA

Profa. Dr. Robson do Nascimento Fidalgo (Orientador)

Universidade Federal de Pernambuco

Prof. Dr. Vinícius Cardoso Garcia (Examinador Interno)

Universidade Federal de Pernambuco

RESUMO

A governança de dados tem sido tradicionalmente conduzida por arquiteturas centralizadas, como Data Warehouses e Data Lakes, que, apesar de eficientes, apresentam limitações de escalabilidade e flexibilidade. O Data Mesh surge como uma alternativa inovadora, promovendo uma abordagem descentralizada na qual os domínios organizacionais assumem a responsabilidade pela gestão dos seus próprios dados. No entanto, a adoção desse modelo enfrenta desafios significativos, incluindo interoperabilidade, conformidade regulatória e resistência organizacional. Este trabalho identifica os principais desafios, benefícios e estratégias emergentes na implementação do Data Mesh. A análise de 11 estudos publicados entre 2020 e 2025 evidencia que, apesar dos benefícios em escalabilidade e autonomia, a governança descentralizada exige padronização, infraestrutura self-service robusta e mudanças culturais dentro das empresas. Como estratégias para mitigar esses desafios, destacam-se a adoção de catálogos de dados, contratos inteligentes baseados em blockchain e plataformas de governança federada. O estudo conclui que, embora o Data Mesh represente um avanço significativo na governança de dados, sua adoção requer planejamento e evolução contínua para garantir sua eficácia e sustentabilidade a longo prazo.

Palavras-chave: Governança de Dados; Data Mesh; Descentralização; Interoperabilidade; Conformidade Regulatória.

ABSTRACT

Traditional data governance architectures, such as Data Warehouses and Data Lakes, have been effective but face scalability and flexibility limitations. Data Mesh emerges as an innovative alternative, decentralizing data governance by assigning responsibility to organizational domains. However, its adoption presents significant challenges, including interoperability, regulatory compliance, and organizational resistance. This work identifies the main challenges, benefits, and emerging strategies in Data Mesh implementation. Analysis of 11 studies published between 2020 and 2025 reveals that, despite improvements in scalability and autonomy, decentralized governance requires standardization, robust self-service infrastructure, and cultural adaptation. Strategies to mitigate these challenges include the adoption of data catalogs, blockchain-based smart contracts, and federated governance platforms. The study concludes that while Data Mesh represents a significant advancement in data governance, its adoption requires careful planning and continuous evolution to ensure long-term effectiveness and sustainability.

Keywords: Data Governance; Data Mesh; Decentralization; Interoperability; Regulatory Compliance.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Protocolo baseado em Kitchenham e Charters	17
Figura 2 –	Fluxo do processo de seleção dos artigos.	24
Gráfico 1 –	Gráfico da quantidade de artigos publicados por ano	27
Gráfico 2 –	Gráfico da distribuição dos artigos nas diferentes bibliotecas digitais utilizadas na pesquisa	27

LISTA DE TABELAS

Tabela 1 – Questões de Pesquisa (Research Questions ou RQ) e Motivações	19
Tabela 2 – Bibliotecas utilizadas	20
Tabela 3 – Critérios de Inclusão	21
Tabela 4 – Critérios de Exclusão	21
Tabela 5 – Critérios de Avaliação de Qualidade	22
Tabela 6 – Síntese dos artigos/estudos analisados	25
Tabela 7 – Distribuição dos artigos/estudos em relação aos critérios de avaliação de qualidade	28

LISTA DE ABREVIATURAS E SIGLAS

AI	Artificial Intelligence
CQ	Critério de Qualidade
GDPR	General Data Protection Regulation
IA	Inteligência Artificial
LGPD	Lei Geral de Proteção de Dados
RBAC	Role-Based Access Control
RQ	Research Question
SGLR	Systematic Gray Literature Review
SLA	Service Level Agreement
SOA	Service-Oriented Architecture

SUMÁRIO

1.	INTRODUÇÃO	11
2.	CONCEITOS BÁSICOS	13
2.1	Arquiteturas de Dados Tradicionais: Data Warehouse, Data Lake e Data Lakehouse	13
2.2	Silos de Dados: Um Obstáculo à Democratização da Informação	14
2.3	Data Mesh: Definição e Justificativa Conceitual	14
2.4	Os Quatro Princípios Fundamentais do Data Mesh	15
2.4.1	<i>Dados como Produto</i>	15
2.4.2	<i>Propriedade Orientada a Domínios</i>	15
2.4.3	<i>Plataforma Self-Service de Dados</i>	16
2.4.4	<i>Governança Computacional Federada</i>	16
3.	METODOLOGIA	17
3.1	Definição das questões de Pesquisa	18
3.2	Estratégia de Busca	19
3.3	Critérios de Inclusão e Exclusão	20
3.4	Critérios de Qualidade	22
3.5	Seleção dos Estudos	23
4.	RESULTADOS	30
4.1	Quais são os principais desafios enfrentados pelas organizações ao implementar a governança descentralizada no Data Mesh?	30
4.1.1	<i>Desafios técnicos de interoperabilidade e qualidade dos dados</i>	30
4.1.2	<i>Desafios organizacionais e ausência de estruturas de governança</i>	31
4.1.3	<i>Segurança da informação e conformidade regulatória</i>	32
4.1.4	<i>Mudança cultural e capacitação dos times</i>	34
4.2	Quais estratégias e práticas emergentes têm sido adotadas para superar os desafios da governança descentralizada no Data Mesh?	34
4.2.1	<i>Padronização da governança e dos metadados</i>	35
4.2.2	<i>Automação da qualidade, rastreabilidade e segurança</i>	36

4.2.3	<i>Plataformas self-service e escalabilidade operacional</i>	38
4.2.4	<i>Estruturas organizacionais híbridas e apoio interdomínio</i>	39
4.2.5	<i>Capacitação e mudança cultural</i>	40
4.3	Quais lacunas e oportunidades de pesquisa ainda existem na literatura sobre governança descentralizada em arquiteturas de Data Mesh?	41
5.	DISCUSSÃO	44
6.	AMEAÇAS À VALIDADE	46
7.	CONCLUSÃO E TRABALHOS FUTUROS	47
	REFERÊNCIAS	

1. INTRODUÇÃO

A explosão do volume de dados nas últimas décadas transformou significativamente a forma como as organizações armazenam, processam e utilizam suas informações estratégicas. As arquiteturas tradicionais de dados, como Data Warehouses e Data Lakes, foram essenciais para a consolidação e análise avançada de grandes volumes de dados, mas hoje enfrentam limitações significativas relacionadas à escalabilidade, integração de diferentes fontes e autonomia dos times organizacionais (Loukiala et al., 2022).

Nesse contexto, surge o Data Mesh como uma proposta inovadora, que rompe com a lógica tradicional centralizada e propõe uma nova forma de organizar a governança de dados. Essa abordagem descentraliza a gestão, atribuindo aos domínios organizacionais a responsabilidade direta pela sua qualidade, documentação e disponibilização. Trata-se de uma mudança paradigmática, em que os dados passam a ser tratados como produtos organizacionais, com maior autonomia local e alinhamento às necessidades específicas de cada área de negócio (Dehghani, 2019; Blohm et al., 2024).

Apesar das evidentes vantagens dessa nova abordagem, como maior escalabilidade e flexibilidade operacional, sua adoção envolve desafios técnicos, regulatórios e, especialmente, culturais significativos. Aspectos como interoperabilidade entre domínios distintos, dificuldades na padronização dos dados e a resistência organizacional à mudança são frequentemente citados pela literatura (Dolhopolov et al., 2024; Machado et al., 2022).

Diante desse cenário, o presente trabalho tem como objetivo geral identificar os principais desafios, benefícios e estratégias emergentes relacionados à adoção do modelo descentralizado de governança de dados conhecido como Data Mesh. Para alcançar esse objetivo, foi utilizado o protocolo metodológico proposto por Kitchenham e Charters (2007), garantindo rigor na seleção e análise das publicações mais recentes e relevantes sobre o tema. Este estudo também busca identificar lacunas existentes na literatura atual e propor caminhos para futuras pesquisas. Com isso, espera-se contribuir tanto para o avanço teórico quanto para o apoio prático à implementação dessa abordagem por profissionais e organizações.

A monografia está estruturada da seguinte forma: o Capítulo 2 apresenta os conceitos fundamentais que sustentam o Data Mesh e fornece o embasamento

necessário para a compreensão das questões centrais da pesquisa. O Capítulo 3 descreve detalhadamente a metodologia empregada na condução do mapeamento. Em seguida, o Capítulo 4 responde diretamente às perguntas de pesquisa, apresentando e discutindo os principais resultados encontrados. O Capítulo 5 realiza uma análise crítica desses resultados, destacando implicações teóricas e práticas, enquanto o Capítulo 6 aborda as ameaças à validade deste estudo. Por fim, o Capítulo 7 apresenta as conclusões finais, evidenciando as contribuições deste trabalho, suas limitações e sugestões concretas para futuras pesquisas.

2. CONCEITOS BÁSICOS

2.1 Arquiteturas de Dados Tradicionais: Data Warehouse, Data Lake e Data Lakehouse

As arquiteturas de dados vêm evoluindo ao longo dos anos na tentativa de atender à crescente complexidade das demandas analíticas e ao volume massivo de dados gerados nas organizações. Tradicionalmente, as empresas adotaram Data Warehouses, que consistem em estruturas centralizadas voltadas ao suporte à decisão. Esses sistemas seguem a abordagem de schema-on-write, ou seja, os dados são previamente estruturados e organizados antes de serem carregados no repositório (Inmon, 2005; Sawadogo & Darmont, 2021). O conceito foi consolidado por Inmon (2005), que define o Data Warehouse como uma coleção de dados orientada por assunto, integrada, não volátil e sensível ao tempo, projetada para apoiar processos decisórios.

Com o aumento exponencial da diversidade e volume dos dados, surgiram os Data Lakes, que se diferenciam por permitir o armazenamento de dados em seu estado bruto, sem a necessidade de modelagem prévia. Esse modelo é mais flexível e capaz de lidar com dados estruturados, semiestruturados e não estruturados, adotando o conceito de schema-on-read, no qual a estrutura dos dados é definida apenas no momento da consulta (Sawadogo & Darmont, 2021). Embora os Data Lakes ofereçam maior agilidade no armazenamento e no acesso aos dados, sua efetividade depende fortemente de mecanismos robustos de gerenciamento de metadados e governança, sob pena de se tornarem os chamados “data swamps”, onde os dados se tornam desorganizados e de difícil aproveitamento (Sawadogo & Darmont, 2021; Nargesian et al., 2019).

Mais recentemente, surgiu o conceito de Data Lakehouse, que busca unir os pontos fortes dos modelos anteriores. Essa arquitetura combina o armazenamento bruto e flexível do Data Lake com capacidades analíticas e transacionais típicas do Data Warehouse, oferecendo suporte a workloads analíticos complexos em uma camada unificada (Loukiala et al., 2022). No entanto, apesar dos avanços técnicos, o Data Lakehouse ainda preserva uma estrutura logicamente centralizada, o que pode limitar sua escalabilidade organizacional e dificultar a autonomia dos domínios em contextos que exigem maior descentralização (González-Velázquez et al., 2024).

Apesar das contribuições relevantes desses modelos, suas limitações tornam-se cada vez mais evidentes no contexto atual, sobretudo pela centralização na gestão dos dados e pela sobrecarga de equipes especializadas. Diante desse cenário, emergem propostas alternativas que buscam romper com a lógica centralizada. Uma das mais promissoras é o Data Mesh, que será explorado em profundidade nas próximas seções.

2.2 Silos de Dados: Um Obstáculo à Democratização da Informação

Um dos problemas mais recorrentes e prejudiciais às iniciativas de dados nas organizações modernas é a existência dos chamados silos de dados. Esse termo refere-se à segmentação isolada de informações entre diferentes departamentos, sistemas ou equipes, que dificultam a integração, o compartilhamento e a reutilização de dados em escala organizacional (Goedegebuure et al., 2024).

Segundo Blohm et al. (2024), os silos surgem historicamente devido à estrutura funcional das organizações, onde cada área cria seus próprios repositórios de dados com finalidades específicas, sem a preocupação com interoperabilidade ou compartilhamento com outros setores. Isso compromete a visibilidade global dos dados, gera duplicidade de esforços, inconsistência de informações e baixa confiança nos dados corporativos. Dolhopolov et al. (2024) reforçam que os silos não são apenas uma limitação técnica, mas também organizacional e cultural, sendo um reflexo da falta de responsabilidade clara sobre os dados e da ausência de incentivos para colaboração interdepartamental.

2.3 Data Mesh: Definição e Justificativa Conceitual

O conceito de Data Mesh, idealizado por Zhamak Dehghani em 2019, surgiu como uma proposta disruptiva às limitações dos modelos anteriores. Trata-se de uma mudança de paradigma sociotécnica, que envolve não apenas novas tecnologias, mas uma reestruturação organizacional da gestão de dados (Dehghani, 2019). Dehghani (2019) propõe o abandono da visão centralizadora dos dados em prol de um modelo descentralizado e orientado a domínios, que representam unidades organizacionais responsáveis por áreas específicas de negócio, como Vendas, Marketing, Logística e Financeiro. Cada domínio passa a ser responsável

pelos dados que produz e consome, garantindo sua qualidade, documentação e disponibilidade para outros domínios que possam precisar dessas informações. Essa abordagem busca alinhar a gestão de dados à lógica de negócios, promovendo autonomia, agilidade e maior proximidade com os contextos onde os dados são gerados e utilizados.

Blohm et al. (2024) reforçam essa perspectiva ao posicionar o Data Mesh como um conceito que transcende a camada tecnológica, promovendo um modelo de gestão orientado à escala, interoperabilidade e democratização de dados. O foco não está apenas na infraestrutura, mas na transformação organizacional e cultural necessária para que os dados sejam geridos de forma eficiente e estratégica.

2.4 Os Quatro Princípios Fundamentais do Data Mesh

2.4.1 Dados como Produto

A ideia de que “dados são produtos” é o ponto central do Data Mesh. Isso significa que os dados devem ser tratados com o mesmo rigor e atenção que se aplica a qualquer produto entregue ao cliente: com qualidade assegurada, usabilidade clara, segurança, documentação completa, suporte e ciclo de vida definido. Segundo Blohm et al. (2024), os dados devem ser concebidos para gerar valor, sendo descobertos, acessados e reutilizados com facilidade por qualquer equipe interessada. Isso envolve entregar dados com documentação adequada, contratos de acesso, garantias de atualização e governança de uso, práticas inspiradas no desenvolvimento de APIs ou produtos de software. Wider et al. (2023) acrescentam que dados com esse tratamento promovem maior confiança, reuso e agilidade no ecossistema organizacional.

2.4.2 Propriedade Orientada a Domínios

O segundo princípio do Data Mesh é a descentralização da responsabilidade sobre os dados, atribuindo a propriedade de cada conjunto de dados ao domínio organizacional que os gera. Isso implica que, por exemplo, a equipe de operações passa a ser responsável pelos dados que ela própria gera, garantindo sua qualidade, estrutura e disponibilização. Esse modelo resolve um dos principais

gargalos das arquiteturas tradicionais, onde equipes de dados centralizadas são sobrecarregadas e muitas vezes desconhecem o contexto dos dados. De acordo com Bode et al. (2024) e Dolhopolov et al. (2024), esse modelo aumenta a responsabilização, o engajamento com os dados e a capacidade de resposta aos consumidores internos, tornando a governança mais distribuída e aderente à realidade do negócio.

2.4.3 Plataforma Self-Service de Dados

O terceiro princípio refere-se à criação de uma plataforma de dados self-service, ou seja, um conjunto de ferramentas, padrões e serviços que permita aos times de domínio publicar, consultar e compor dados de forma autônoma. Wider et al. (2023) explicam que essa plataforma deve abstrair a complexidade técnica, oferecendo componentes como fluxos padronizados de dados (pipelines), catálogos, testes automatizados, versionamento e monitoramento, de forma que os domínios não dependam de especialistas para operacionalizar seus dados.

2.4.4 Governança Computacional Federada

A governança computacional federada é o princípio que garante o equilíbrio entre autonomia local dos domínios e padronização organizacional. Segundo Dolhopolov et al. (2024), essa governança combina padrões centralmente definidos com execução descentralizada, apoiada em automações e contratos de dados. Isso inclui controle de acesso, conformidade legal, regras de qualidade e interoperabilidade entre domínios, permitindo escalar o uso dos dados sem comprometer segurança ou conformidade regulatória (compliance). Cabe destacar que os desafios e estratégias práticas relacionados à implementação da governança federada serão discutidos de forma aprofundada nos capítulos de resultados (seções 4.1 e 4.2), com exemplos aplicados e evidências empíricas observadas na literatura recente.

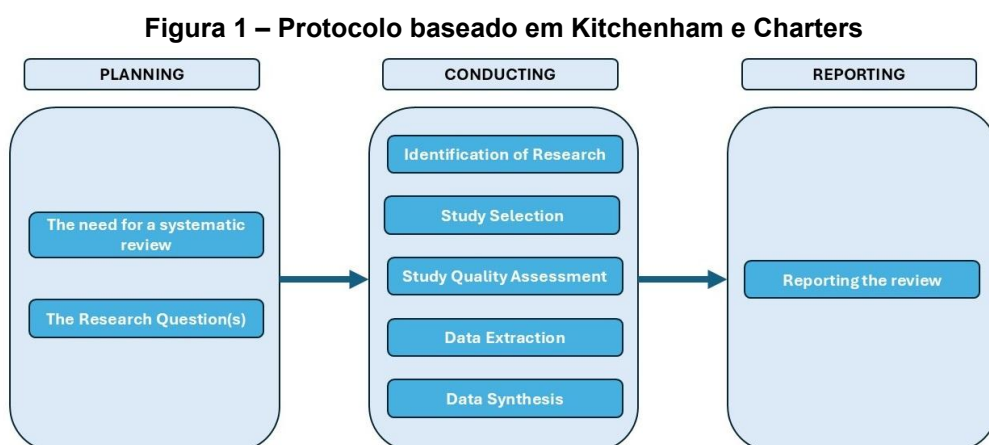
3. METODOLOGIA

Para a realização deste estudo, foi adotado o modelo metodológico proposto por Kitchenham e Charters (2007), que estrutura o processo de mapeamento em três etapas fundamentais: planejamento, execução e relato dos resultados. Cada uma dessas fases é composta por atividades específicas que garantem a condução rigorosa da pesquisa.

Na fase de planejamento, que se divide em dois momentos, ocorre primeiramente a identificação da necessidade do mapeamento, considerando a relevância do tema e a lacuna existente na literatura. Em seguida, são formuladas as questões de pesquisa, que orientarão a busca e análise dos estudos.

A fase de execução engloba um conjunto estruturado de procedimentos para garantir a seleção criteriosa dos estudos. Essa etapa é subdividida em cinco processos: a identificação dos artigos nas bases de dados, a triagem e seleção dos estudos com base nos critérios previamente definidos, a avaliação da qualidade das publicações selecionadas, a extração das informações relevantes e, por fim, a síntese dos achados, permitindo a construção de uma visão consolidada sobre o tema investigado.

A última etapa, referente ao relato dos resultados, compreende a organização e a apresentação dos achados do mapeamento, garantindo que as conclusões sejam expostas de forma clara e objetiva. O detalhamento do fluxo metodológico adotado pode ser visualizado na Figura 1, que esquematiza as etapas seguidas ao longo do estudo.



Fonte: O autor (2025).

A metodologia de Kitchenham e Charters (2007) foi selecionada devido à sua abordagem consolidada no campo da Engenharia de Software e Ciência de Dados, garantindo um processo estruturado para a identificação, seleção e análise de estudos relevantes. Seu rigor metodológico possibilita a coleta e síntese de informações de forma imparcial, assegurando que o mapeamento contemple os estudos mais significativos e confiáveis dentro do escopo definido. Além disso, a aplicação de critérios de qualidade na seleção dos estudos permite uma interpretação mais precisa dos achados, evitando distorções que possam comprometer a validade das conclusões. Esse protocolo visa não apenas mapear os desafios e benefícios da governança descentralizada no modelo Data Mesh, mas também consolidar práticas emergentes identificadas na literatura, fornecendo subsídios para pesquisas futuras e para a implementação desse modelo em diferentes contextos organizacionais.

3.1 Definição das questões de Pesquisa

Este estudo tem como objetivo analisar a governança descentralizada de dados no contexto do Data Mesh, sendo estruturada por três questões centrais: a primeira (RQ1) explora os desafios enfrentados na implementação da governança descentralizada; a segunda (RQ2) analisa as estratégias e práticas emergentes utilizadas para superá-los; e a terceira (RQ3) busca identificar lacunas e oportunidades de pesquisa. As questões e suas respectivas motivações estão detalhadas na Tabela 1. Essa abordagem permite consolidar um panorama abrangente sobre os desafios, benefícios e tendências do Data Mesh, auxiliando tanto pesquisadores quanto profissionais da área a entenderem os impactos dessa nova arquitetura de governança de dados.

Tabela 1 – Questões de Pesquisa (Research Questions ou RQ) e Motivações

Pergunta de Pesquisa	Motivação
RQ1. Quais são os principais desafios enfrentados pelas organizações ao implementar a governança descentralizada no Data Mesh?	Mapear barreiras técnicas, culturais e operacionais que podem dificultar a adoção e implementação eficaz do Data Mesh.
RQ2. Quais estratégias e práticas emergentes têm sido adotadas para superar os desafios da governança descentralizada no Data Mesh?	Analisar soluções aplicadas e tendências que ajudam a mitigar os desafios na adoção do Data Mesh.
RQ3. Quais lacunas e oportunidades de pesquisa ainda existem na literatura sobre governança descentralizada em arquiteturas de Data Mesh?	Identificar aspectos pouco explorados na literatura para orientar pesquisas futuras e aprimorar o modelo de governança descentralizada.

Fonte: O autor (2025).

3.2 Estratégia de Busca

Para identificar estudos relevantes sobre governança descentralizada de dados no modelo Data Mesh, foram definidas palavras-chave estratégicas alinhadas às questões de pesquisa. A seleção dos termos considerou os conceitos mais frequentes na literatura, abrangendo tanto aspectos técnicos quanto organizacionais. A construção da string de busca foi refinada a partir de testes preliminares em bases acadêmicas, utilizando operadores booleanos para ampliar a recuperação de publicações dentro do escopo do mapeamento. A formulação final adotada foi:

("Data Mesh" OR "Decentralized Data Governance" OR "Data Decentralization" OR "Self-Serve Data Infrastructure") AND ("Data Governance" OR "Data Architecture" OR "Data Management" OR "Data Infrastructure" OR "Data Platform" OR "Scalability")

A busca foi realizada de forma automatizada nas principais bases eletrônicas que reúnem periódicos científicos e anais de conferências de alta qualidade. Foram consideradas fontes amplamente reconhecidas, incluindo ACM Digital Library, IEEE Digital Library e Scopus (consulte a Tabela 2).

Tabela 2 – Bibliotecas utilizadas

Pergunta de Pesquisa	Motivação
ACM Digital Library	https://dl.acm.org/
IEEE Digital Library	https://ieeexplore.ieee.org/
Scopus	https://www.scopus.com/

Fonte: O autor (2025).

A seleção das bases ACM Digital Library, IEEE Xplore e Scopus deve-se ao seu reconhecimento internacional e à curadoria rigorosa por pares. Essas bibliotecas oferecem ampla cobertura de publicações acadêmicas relevantes ao tema da pesquisa. Além disso, a escolha também foi influenciada pela familiaridade prévia do autor com essas plataformas, adquirida em trabalhos e estudos anteriores, o que facilitou a construção e aplicação eficiente das estratégias de busca. Embora outras fontes, como SpringerLink, Web of Science e bases de literatura cinzenta (como arXiv ou Zenodo), contenham materiais pertinentes, optou-se por restringir a busca às bases referidas, priorizando publicações com elevado rigor científico e revisadas por pares. A literatura cinzenta foi considerada apenas indiretamente, via comparação com a SGLR de Goedegebuure et al. (2024) na seção de Discussão (Cap. 5).

3.3 Critérios de Inclusão e Exclusão

Esta pesquisa concentra-se exclusivamente em estudos primários recentes, garantindo a inclusão apenas de trabalhos diretamente relacionados à governança descentralizada de dados no modelo Data Mesh. Para assegurar a qualidade e relevância das publicações, foram estabelecidos critérios rigorosos de inclusão e exclusão.

Foram considerados apenas artigos completos, publicados em periódicos científicos ou anais de conferências reconhecidos, que abordassem governança descentralizada de dados ou Data Mesh de forma significativa. Estudos que apresentassem frameworks, estudos de caso ou aplicações práticas foram priorizados, pois fornecem uma perspectiva mais concreta sobre a implementação e os desafios desse modelo.

Por outro lado, foram excluídos trabalhos duplicados, literatura cinzenta e artigos indisponíveis na íntegra ou de acesso pago. Além disso, estudos que não tratassem diretamente do tema ou que tivessem sido publicados antes de 2020 foram descartados, garantindo que o mapeamento contemplasse apenas pesquisas alinhadas com os avanços mais recentes.

O recorte temporal recente foi essencial para refletir a evolução do conceito de Data Mesh e governança descentralizada, considerando práticas emergentes, desafios contemporâneos e novas abordagens tecnológicas. Dessa forma, evitou-se a inclusão de estudos que não dialogam com a maturidade atual do tema. Os critérios completos de inclusão e exclusão estão detalhados nas Tabelas 3 e 4.

Tabela 3 – Critérios de Inclusão

#	Critérios de Inclusão
C1	Artigos completos, acessíveis em formato integral
C2	Publicações do tipo journal article (artigos em revistas científicas) ou conference paper/proceedings (trabalhos apresentados em conferências)
C3	Estudos que abordem governança descentralizada de dados ou Data Mesh de forma significativa
C4	Estudos que apresentem discussões, frameworks, estudos de caso ou aplicações práticas do modelo de Data Mesh
C5	Estudos publicados de 2020 até o presente

Fonte: O autor (2025).

Tabela 4 – Critérios de Exclusão

#	Critérios de Exclusão
C1	Estudos fora do período definido (publicados antes de 2020)
C2	Artigos não disponíveis na íntegra ou de acesso restrito/pago
C3	Dissertações, teses não publicadas, relatórios técnicos.
C4	Trabalhos duplicados
C5	Artigos que não estejam diretamente relacionados ao tema de governança descentralizada de dados no modelo de Data Mesh

Fonte: O autor (2025).

3.4 Critérios de Qualidade

Para garantir que os estudos selecionados no mapeamento sobre governança descentralizada de dados no modelo Data Mesh sejam metodologicamente rigorosos e cientificamente relevantes, será aplicada uma avaliação de qualidade baseada em critérios estruturados. Segundo Kitchenham e Charters (2007), a aplicação de critérios de qualidade em mapeamentos reduz vieses, aumenta a validade interna e externa e assegura a confiabilidade dos achados. Dessa forma, os resultados sintetizados poderão ser melhor aplicados em diferentes contextos de pesquisa e prática.

Os critérios de qualidade foram definidos para avaliar a fundamentação teórica, a metodologia empregada, a apresentação dos resultados e a contribuição científica dos estudos analisados. Cada critério será pontuado com Sim (S=1), Não (N=0) ou Parcialmente (P=0.5), permitindo uma análise comparativa entre as publicações selecionadas. A Tabela 5 apresenta os critérios adotados nesta avaliação. A aplicação desses critérios possibilita a seleção de estudos com alto rigor metodológico, garantindo que o levantamento realizado sintetize informações confiáveis sobre os desafios, benefícios e práticas emergentes na governança descentralizada de dados com Data Mesh.

Tabela 5 – Critérios de Avaliação de Qualidade

#	Pergunta	Resposta
Q1	O estudo apresenta uma justificativa clara para sua realização?	S=1, N=0, P=0.5
Q2	Os objetivos da pesquisa estão claramente definidos?	S=1, N=0, P=0.5
Q3	O contexto do estudo é descrito de forma detalhada?	S=1, N=0, P=0.5
Q4	Os resultados são apresentados de forma clara e bem estruturada?	S=1, N=0, P=0.5
Q5	A pesquisa inclui uma avaliação da aplicabilidade dos achados?	S=1, N=0, P=0.5
Q6	O estudo reconhece suas próprias limitações metodológicas e discute possíveis ameaças à validade dos resultados?	S=1, N=0, P=0.5

Fonte: O autor (2025).

3.5 Seleção dos Estudos

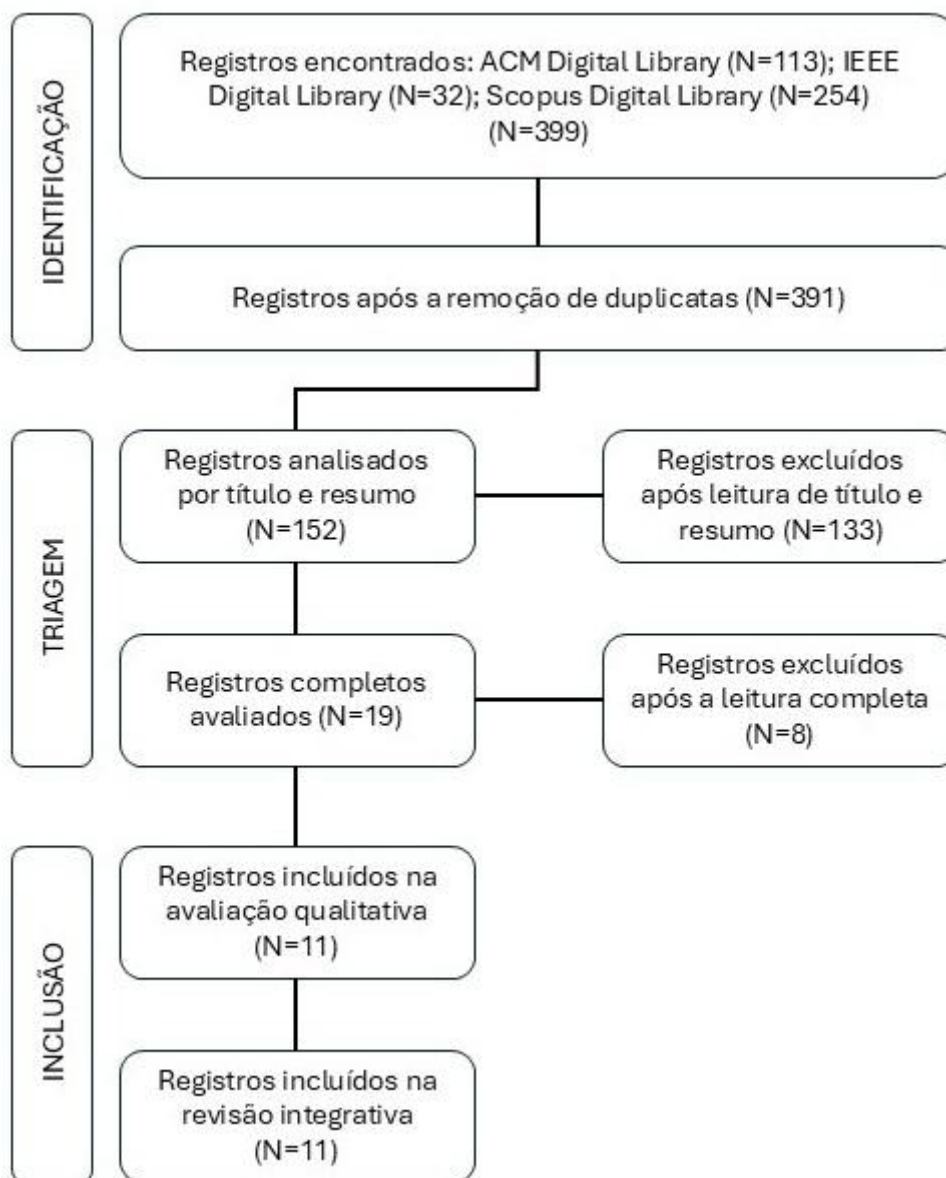
A etapa de triagem dos artigos foi conduzida com base nos critérios de exclusão previamente estabelecidos, garantindo a seleção de estudos alinhados aos objetivos da pesquisa. Inicialmente, 399 artigos foram identificados nas bases de dados ACM Digital Library, IEEE Xplore e Scopus, porém grande parte foi descartada por não atender aos critérios definidos.

Na primeira filtragem, 66 artigos foram excluídos por estarem fora do intervalo de análise (2020 - 2025), pois não refletiam os avanços mais recentes sobre governança descentralizada de dados no modelo Data Mesh. Além disso, 34 registros não eram artigos científicos, sendo classificados como resumos, apresentações, tutoriais ou materiais sem revisão por pares. Outros 6 artigos estavam em idiomas diferentes do inglês, inviabilizando sua inclusão. Após essa etapa, 293 artigos permaneceram para análise.

Na verificação de disponibilidade, constatou-se que apenas 152 artigos possuíam acesso ao texto completo, sendo os demais descartados devido a restrições de acesso. Esses 152 artigos foram submetidos à análise de título e resumo, na qual se avaliou sua relevância para o tema da pesquisa. Durante essa triagem, foram eliminados os estudos que não abordavam diretamente governança descentralizada de dados no modelo Data Mesh, resultando na seleção de 19 artigos para leitura completa.

Na última fase, os 19 artigos selecionados foram avaliados detalhadamente de acordo com os critérios de qualidade previamente estabelecidos, garantindo que apenas estudos com metodologia sólida, alinhamento com pesquisas recentes e relevância prática fossem considerados no mapeamento. Após essa avaliação, 11 artigos foram escolhidos para compor a análise final, assegurando que apenas os trabalhos mais robustos e alinhados ao escopo da pesquisa fossem incluídos. A Figura 2 ilustra o processo de seleção dos artigos, enquanto a Tabela 6 apresenta detalhes dos estudos escolhidos, incluindo ordem, ano de publicação, autores e título.

Figura 2 – Fluxo do processo de seleção dos artigos.



Fonte: elaborado de acordo com o modelo de Moher et al. (2009)

A lista completa dos 152 artigos identificados com acesso ao texto completo, incluindo o registro dos critérios aplicados na etapa de triagem e avaliação com base na leitura de título e resumo, encontra-se disponível publicamente no repositório [GitHub](#). O repositório inclui os dados brutos da análise, contribuindo para a transparência, reprodutibilidade e possibilidade de replicação futura deste mapeamento por outros pesquisadores.

Tabela 6 – Síntese dos artigos/estudos analisados

Nº	Ano	Base	Autores	Título
1	2024	Scopus	Abdulaziz Almaslukh, Abdulmajeed Alameer, Hamad Alsaleh, Fahad Alkadyan, Nasser Allheeib, Abdulaziz Alhadlag, Yazeed Alabdulkarim	Data Mesh Meets Blockchain
2	2024	Scopus	Ivo Blohm, Felix Wortmann, Christine Legner, Felix Köbler	Data products data mesh and data fabric: New paradigm(s) for data and analytics?
3	2024	Scopus	Anton Dolhopolov, Arnaud Castelltort, and Anne Laurent	Implementing Federated Governance in Data Mesh Architecture
4	2024	Scopus	Michalis Pingos, Panayiotis Christodoulou, and Andreas S. Andreou	Security and Ownership in User-Defined Data Meshes
5	2024	Scopus	Roberto González-Velázquez, Izaskun Fernández, Rubén Ferreira, Carlos Carballo, Álvaro García, Bruno Santamaría, and Diego González	Smart Factory Hub - Towards a Data Mesh in Smart Manufacturing
6	2024	IEEE Digital Library	Jan Bode, Niklas Kühl, Dominik Kreuzberger, and Carsten Holtmann	Toward Avoiding the Data Mess: Industry Insights From Data Mesh Implementations
7	2024	ACM	Kitipong Viseshart and Worapol Pongpech	Marine Transport Business: A Datamesh-Based Data Asset Strategy
8	2023	Scopus	Otmane Azeroual & Radka Nacheva	Data Mesh for Managing Complex Big Data Landscapes and Enhancing Decision Making in Organizations

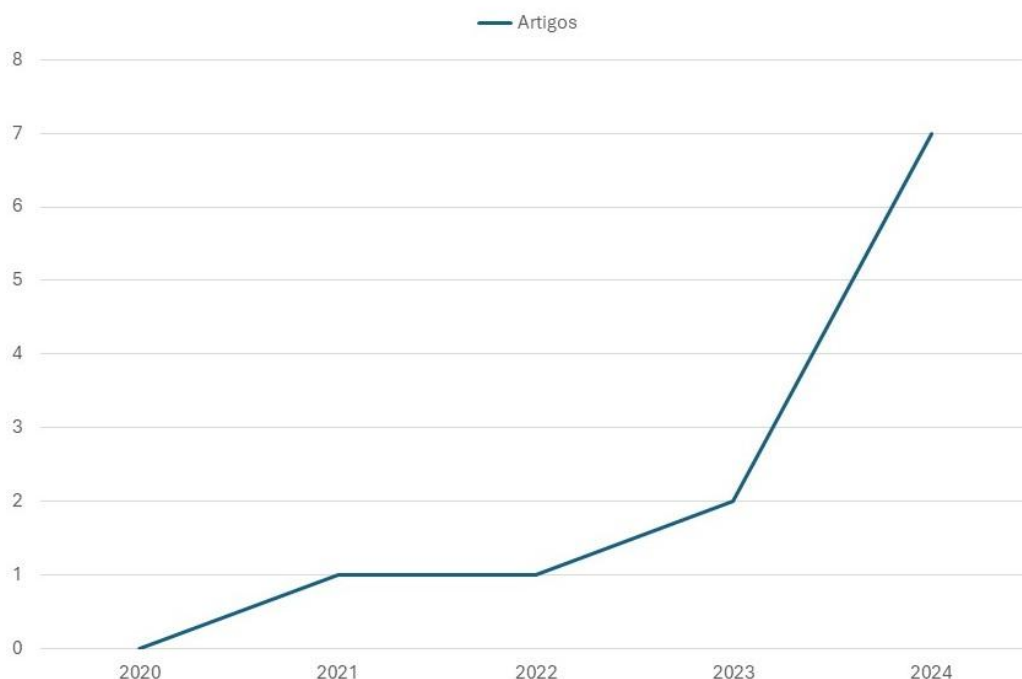
9	2023	IEEE Digital Library	Arif Wider, Sumedha Verma, and Atif Akhtar	Decentralized Data Governance as Part of a Data Mesh Platform: Concepts and Approaches
10	2022	Scopus	Kathrine Vestues, Geir Kjetil Hanssen, Marius Mikalsen, Thor Aleksander Buan, and Kieran Conboy	Agile Data Management in NAV: A Case Study
11	2021	Scopus	Inês Araújo Machado, Carlos Costa, Maribel Yasmina Santos	Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures

Fonte: O autor (2025).

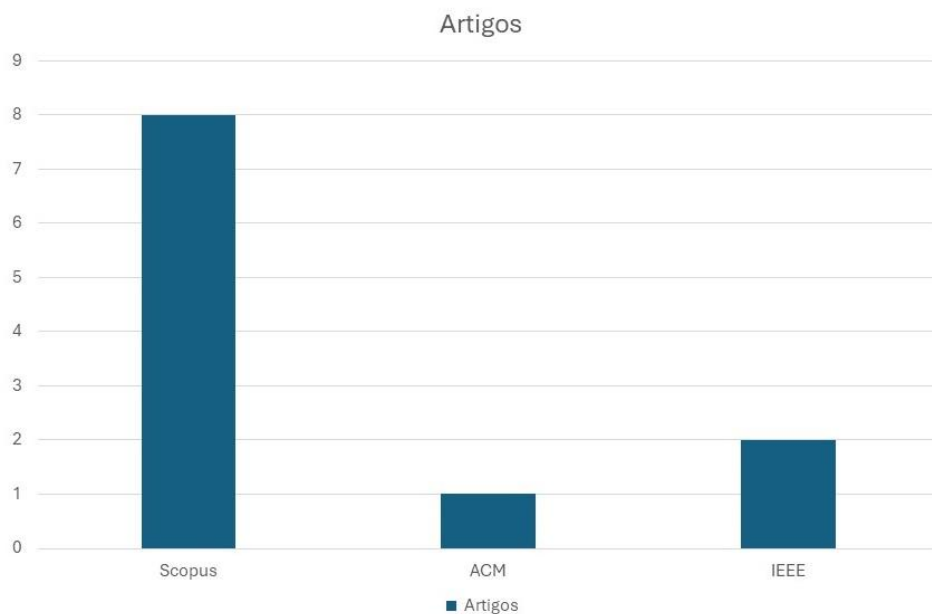
Apesar de todas as etapas da triagem terem sido conduzidas por um único revisor, foram adotadas estratégias para mitigar possíveis vieses individuais. A definição prévia de critérios objetivos de exclusão e inclusão (Seções 3.3 e 3.4), aliada à aplicação estruturada de uma matriz de avaliação da qualidade (Tabela 5), visou assegurar consistência na seleção dos estudos. Em pesquisas futuras, recomenda-se a participação de múltiplos revisores ou aplicação de validação cruzada para ampliar a robustez metodológica.

Após a seleção dos artigos, foram elaboradas representações gráficas para fornecer uma visão mais detalhada sobre a produção acadêmica e as fontes bibliográficas utilizadas nesse estudo. O Gráfico 1 apresenta a distribuição temporal das publicações, evidenciando um crescimento ao longo dos anos. Esse aumento indica um interesse crescente no tema, além de refletir a evolução das tecnologias e práticas que impulsionam as pesquisas sobre Data Mesh e governança descentralizada de dados.

Complementando essa análise, o Gráfico 2 ilustra a distribuição das bases de dados utilizadas nas referências dos artigos selecionados. Esse levantamento permite identificar quais fontes são mais recorrentes entre os pesquisadores, destacando as plataformas mais influentes na construção do conhecimento sobre o tema.

Gráfico 1 – Gráfico da quantidade de artigos publicados por ano

Fonte: O autor (2025).

Gráfico 2 – Gráfico da distribuição dos artigos nas diferentes bibliotecas digitais utilizadas na pesquisa

Fonte: O autor (2025).

Após a seleção dos 11 artigos, foi aplicada a avaliação de qualidade, garantindo que os estudos incluídos no mapeamento atendessem aos critérios metodológicos estabelecidos. O objetivo dessa etapa foi verificar se os estudos apresentavam rigor metodológico, clareza na apresentação dos resultados e

relevância para a pesquisa sobre governança descentralizada de dados no modelo Data Mesh. A análise indicou que todos os 11 artigos selecionados atenderam aos critérios exigidos, alcançando pontuação superior a 80% na avaliação. Esse resultado confirma que os estudos possuem fundamentação teórica sólida, metodologia bem definida e aplicabilidade prática, assegurando sua inclusão na análise final do mapeamento. A Tabela 7 apresenta os resultados da avaliação de qualidade, detalhando a pontuação de cada artigo com base nos critérios aplicados.

Tabela 7 – Distribuição dos artigos/estudos em relação aos critérios de avaliação de qualidade

Artigo	CQ1	CQ2	CQ3	CQ4	CQ5	CQ6	Total
1	1	1	1	1	0.5	1	91.7%
2	1	1	1	1	0.5	0.5	83%
3	1	1	1	1	0.5	0.5	83%
4	1	1	1	1	0.5	0.5	83%
5	1	1	1	1	0.5	0.5	83%
6	1	1	1	1	0.5	1	91.7%
7	1	1	1	1	0.5	0.5	83%
8	1	1	0.5	1	1	0.5	83%
9	1	1	1	1	0.5	0.5	83%
10	1	1	1	1	0.5	0.5	83%
11	1	1	1	1	0.5	0.5	83%

Fonte: O autor (2025).

Com a conclusão das etapas de seleção, avaliação e refinamento dos estudos, o mapeamento está preparado para fornecer uma análise aprofundada sobre os desafios, benefícios e práticas emergentes na governança descentralizada de dados no modelo Data Mesh. Embora técnicas qualitativas como análise temática ou codificação aberta possam enriquecer a categorização dos achados, este trabalho optou por uma abordagem descritiva estruturada, diretamente guiada pelas questões de pesquisa (RQ1–RQ3). Essa escolha se deve à natureza exploratória do estudo e ao número reduzido de artigos selecionados (n=11), o que favoreceu uma síntese objetiva e comparativa. Em investigações futuras, será considerada a

aplicação de métodos qualitativos sistemáticos, com suporte de ferramentas como NVivo, para aprofundar a análise e a relação semântica entre os conceitos emergentes. O próximo capítulo apresenta os resultados obtidos, consolidando os achados e discutindo suas implicações no contexto acadêmico e organizacional.

4. RESULTADOS

Este capítulo apresenta os principais achados do estudo sobre a abordagem descentralizada de gestão de dados proposta pelo Data Mesh. A análise conduzida tem como objetivo responder às questões de pesquisa, explorando os principais desafios enfrentados pelas organizações na sua implementação, estratégias adotadas para superá-los e, por fim, as lacunas identificadas na literatura. A abordagem adotada permite consolidar um panorama abrangente sobre o estado da arte do tema, destacando tanto avanços quanto questões ainda não solucionadas.

4.1 Quais são os principais desafios enfrentados pelas organizações ao implementar a governança descentralizada no Data Mesh?

A governança descentralizada proposta pelo Data Mesh representa uma mudança significativa na forma como os dados são gerenciados nas organizações, atribuindo às áreas de negócio maior autonomia e responsabilidade sobre os dados que produzem. Embora traga ganhos como escalabilidade organizacional e alinhamento com os objetivos de negócio, essa abordagem também impõe uma série de desafios técnicos, organizacionais, regulatórios e culturais (Blohm et al., 2024; Wider et al., 2023; González-Velázquez et al., 2024).

4.1.1 Desafios técnicos de interoperabilidade e qualidade dos dados

A interoperabilidade, nesse contexto, refere-se à capacidade de diferentes sistemas e equipes trocarem dados entre si de forma eficiente, clara e padronizada. Já a qualidade dos dados envolve aspectos como precisão, completude, consistência e atualidade das informações, elementos frequentemente citados como pilares essenciais da governança de dados (Dolhopolov et al., 2024; Blohm et al., 2024; Wider et al., 2023).

No modelo descentralizado, cada domínio possui autonomia para gerenciar os dados sob sua responsabilidade. Isso significa que diferentes equipes podem utilizar abordagens distintas para nomear, organizar e armazenar seus dados. Por exemplo, uma equipe pode utilizar o termo “cliente ativo” para se referir a um

consumidor com compra recente, enquanto outra pode usar o mesmo termo com um critério diferente, como o simples fato de estar cadastrado no sistema. Essa falta de padronização dificulta a integração das informações e pode gerar inconsistências nos relatórios e análises organizacionais (Wider et al., 2023; Blohm et al., 2024).

Na prática, esse problema pode ser observado quando duas áreas da empresa registram dados similares, mas com estruturas ou critérios diferentes, impossibilitando o cruzamento das informações de forma direta. Essa situação compromete não apenas a confiabilidade dos dados, mas também o tempo e esforço necessários para preparar essas informações para uso estratégico. Esse tipo de desafio foi identificado, por exemplo, no setor industrial analisado por Wider et al. (2023), onde a ausência de padrão dificultava o reaproveitamento de dados de produção entre linhas de montagem diferentes, gerando retrabalho e duplicação de esforços. Da mesma forma, no estudo de Viseschart & Pongpech (2023), que aborda o caso da empresa PayPal, observou-se que a falta de contratos de dados com critérios claros de interoperabilidade dificultava a integração entre os dados dos domínios de crédito e risco, comprometendo a efetividade das análises internas e aumentando a necessidade de retrabalho analítico.

Esses exemplos evidenciam que os desafios técnicos de interoperabilidade e qualidade dos dados são recorrentes em contextos reais e impactam diretamente a capacidade das organizações de extrair valor dos seus dados em ambientes descentralizados. Portanto, o alinhamento entre os domínios em termos de estrutura e interpretação das informações é um aspecto crítico que precisa ser enfrentado desde os estágios iniciais da adoção do modelo Data Mesh.

4.1.2 Desafios organizacionais e ausência de estruturas de governança

A descentralização da governança de dados proposta pelo modelo Data Mesh desafia diretamente a lógica tradicional de estruturas organizacionais centralizadas. Um dos principais obstáculos enfrentados nesse processo está relacionado à ausência de mecanismos institucionais sólidos para sustentar a nova distribuição de responsabilidades. Em muitos casos, as organizações não possuem uma estrutura de governança madura o suficiente para lidar com a descentralização, o que dificulta a implementação de práticas coerentes e consistentes de gestão de dados (Wider et al., 2023; Blohm et al., 2024; Bode et al., 2023). Essa ausência se manifesta de

diferentes formas. Frequentemente, não há uma definição clara sobre quais são os papéis e responsabilidades que devem ser assumidos pelos domínios, o que gera insegurança e sobreposição de atribuições entre as equipes. A indefinição quanto à figura do responsável pelo dado compromete a responsabilização e dificulta o controle sobre a qualidade, integridade e segurança das informações (Dolhopolov et al., 2024; Bode et al., 2023).

Além disso, a falta de diretrizes organizacionais formais para orientar os domínios pode levar à adoção de práticas divergentes, criando ambientes de governança fragmentados e desarticulados. Sem uma estrutura mínima de governança que estabeleça princípios, normas e critérios comuns, os domínios tendem a seguir direções distintas, o que compromete a coerência da organização como um todo (Wider et al., 2023; González-Velázquez et al., 2024).

Além desses obstáculos estruturais, é importante reconhecer que a adoção de uma governança descentralizada também implica impactos no funcionamento cotidiano das equipes. As mudanças institucionais exigidas pelo modelo não se limitam a questões formais, mas têm desdobramentos relevantes sobre a forma como os profissionais se relacionam com os dados. Esses aspectos comportamentais, que envolvem a transformação da cultura organizacional e o engajamento dos times, serão detalhados na subseção 4.1.4.

4.1.3 Segurança da informação e conformidade regulatória

A adoção da governança descentralizada no modelo Data Mesh também impõe desafios importantes no que diz respeito à segurança da informação e ao cumprimento de regulamentações legais sobre proteção de dados. À medida que a responsabilidade pelo gerenciamento dos dados é distribuída entre vários domínios, torna-se mais difícil manter um controle centralizado sobre questões sensíveis como acessos, consentimentos e auditorias. Essa fragmentação da responsabilidade pode abrir brechas no sistema de proteção de dados da organização e gerar riscos operacionais e jurídicos (Pingos et al., 2024; Dolhopolov et al., 2024).

Esses riscos são especialmente críticos diante de legislações como a Lei Geral de Proteção de Dados (LGPD), no Brasil, e o Regulamento Geral de Proteção de Dados (GDPR), na União Europeia. Ambas as normas estabelecem diretrizes rigorosas sobre o tratamento de dados pessoais, exigindo que as empresas adotem

medidas claras para garantir a privacidade, a rastreabilidade e o uso legítimo das informações de seus clientes, usuários e colaboradores. Entre os requisitos previstos por essas legislações estão, por exemplo, o registro dos consentimentos dos titulares dos dados, a limitação do acesso a dados sensíveis, a definição de prazos de retenção e a possibilidade de auditoria completa das operações realizadas com esses dados (Bode et al., 2024; Pingos et al., 2024).

Em um ambiente descentralizado, o cumprimento dessas exigências se torna mais complexo, justamente porque diferentes domínios passam a ser responsáveis por partes específicas dos dados. Por exemplo, se cada domínio decide individualmente como armazenar ou excluir registros de clientes, sem seguir um padrão organizacional, a empresa pode enfrentar dificuldades para comprovar, durante uma fiscalização, que está em conformidade com a legislação. Essa lacuna é particularmente preocupante em setores regulados, como o financeiro, o de saúde ou o setor público, onde a rastreabilidade e a segurança das informações são exigidas com ainda mais rigor (Pingos et al., 2024; Bode et al., 2023). Na prática, esse desafio foi evidenciado em alguns estudos revisados, como os de Pingos et al. (2024), que apontam que a ausência de mecanismos comuns de controle de acesso e auditoria pode permitir que dados confidenciais sejam compartilhados sem autorização adequada, expondo a organização a riscos de vazamento de dados e sanções legais. Tais necessidades reforçam a importância de mecanismos automatizados de rastreabilidade, controle de acesso e monitoramento, como será detalhado na seção seguinte.

Dessa forma, garantir a segurança da informação e a conformidade regulatória não é apenas uma questão técnica, mas também organizacional e jurídica. Mesmo que os domínios tenham autonomia, é necessário assegurar que existam diretrizes mínimas comuns, capazes de manter o sistema de governança coeso e seguro. A literatura destaca que esse é um dos pontos mais sensíveis da descentralização e que, se não for tratado adequadamente desde o início da implementação do Data Mesh, pode colocar em risco toda a estrutura de dados da organização (Dolhopolov et al., 2024).

4.1.4 Mudança cultural e capacitação dos times

Além dos desafios técnicos e regulatórios, a implementação da governança descentralizada no modelo Data Mesh requer uma mudança cultural profunda nas organizações. Essa transformação não se limita à adoção de novas ferramentas ou estruturas tecnológicas, mas envolve principalmente uma redefinição do papel dos profissionais em relação aos dados (Wider et al., 2023; Bode et al., 2023).

Tradicionalmente, os times de negócio são vistos apenas como consumidores de dados, dependendo de uma equipe centralizada de tecnologia ou ciência de dados para obter as informações de que precisam. No Data Mesh, essa lógica se inverte e os times de negócio passam a ser também produtores ativos dos seus próprios dados, sendo responsáveis por sua qualidade, documentação e disponibilização aos demais domínios da organização (Wider et al., 2023; Dolhopolov et al., 2024).

Essa nova abordagem exige que os profissionais desenvolvam habilidades técnicas básicas, como compreender o significado e a estrutura dos dados que manipulam, registrar informações relevantes sobre os dados (metadados) e colaborar com outras equipes para garantir que esses dados possam ser utilizados com segurança e clareza (Wider et al., 2023; Dolhopolov et al., 2024). A ausência de iniciativas voltadas à construção dessa cultura pode gerar diversos problemas, como resistência dos colaboradores, baixa adesão às práticas de governança, duplicidade de esforços entre domínios e subutilização dos dados disponíveis. Por isso, a literatura enfatiza que o sucesso da descentralização depende diretamente do engajamento humano e da preparação organizacional para a mudança (Wider et al., 2023; Blohm et al., 2024).

4.2 Quais estratégias e práticas emergentes têm sido adotadas para superar os desafios da governança descentralizada no Data Mesh?

As dificuldades enfrentadas pelas organizações na implementação do Data Mesh têm motivado o desenvolvimento de estratégias estruturadas para apoiar a governança descentralizada. A literatura recente aponta quatro grandes pilares de resposta: padronização, automação, capacitação e transformação organizacional.

Essas dimensões compõem os principais eixos de resposta identificados na literatura especializada, detalhados a seguir.

4.2.1 Padronização da governança e dos metadados

A padronização da governança de dados é uma estratégia essencial para garantir que, mesmo com a autonomia dos domínios, os dados sigam critérios comuns de organização, significado e acesso em toda a empresa. Isso significa que cada equipe pode cuidar de seus próprios dados, mas deve seguir regras compartilhadas que permitam que esses dados sejam compreendidos e utilizados por outras áreas da organização (Dolhopolov et al., 2024; Blohm et al., 2024; Wider et al., 2023).

Um dos instrumentos mais eficazes para viabilizar essa padronização é o uso de catálogos corporativos de dados, que funcionam como bibliotecas organizadas onde os dados disponíveis na empresa são registrados, descritos e classificados (Dolhopolov et al., 2024; Wider et al., 2023; González-Velázquez et al., 2024). Esses catálogos ajudam as equipes a saber o que cada dado significa, de onde ele veio, quem é o responsável por ele e como ele pode ser utilizado. Assim, mesmo que um dado tenha sido produzido por outro domínio, ele pode ser compreendido e aproveitado por qualquer equipe (Bode et al., 2023). Ferramentas tecnológicas como o DataHub, o Apache Atlas e o OpenMetadata são exemplos de plataformas utilizadas para construir esses catálogos. Elas permitem registrar informações técnicas sobre os dados como o tipo de informação, frequência de atualização, responsáveis, regras de acesso e o data lineage, ou linhagem de dados, que representa o histórico de origem, transformação e uso de um dado ao longo de seu ciclo de vida (Wider et al., 2023; González-Velázquez et al., 2024).

Além disso, outro recurso importante são as ontologias e taxonomias semânticas. As ontologias estruturam o conhecimento da organização, indicando como os conceitos se relacionam entre si. Por exemplo, podem esclarecer que “cliente ativo” é um cliente com transações nos últimos 90 dias, enquanto “cliente inativo” é aquele sem movimentações no mesmo período. As taxonomias, por sua vez, agrupam e classificam esses conceitos em categorias organizadas, facilitando a busca, a interpretação e a integração dos dados entre diferentes áreas (Pingos et al., 2024; Dolhopolov et al., 2024).

A padronização também inclui a definição de contratos de dados, que estabelecem formalmente os compromissos de um domínio produtor ao entregar dados para outro domínio consumidor. Esses contratos especificam, por exemplo, o nível mínimo de qualidade dos dados, os prazos de atualização e os responsáveis em caso de falhas, criando assim uma relação de confiança e previsibilidade entre as equipes (Blohm et al., 2024; Bode et al., 2023).

Um exemplo prático de aplicação bem-sucedida dessa estratégia é o caso do Smart Factory Hub, projeto que estruturou seus dados industriais com base em catálogos semânticos e domínios técnicos, organizando a informação de forma padronizada e permitindo o reaproveitamento dos dados por diferentes áreas de produção. Essa iniciativa demonstrou que a padronização de metadados e vocabulário organizacional é fundamental para garantir a interoperabilidade e reduzir o retrabalho em ambientes descentralizados (González-Velázquez et al., 2024).

4.2.2 Automação da qualidade, rastreabilidade e segurança

A automação tem se consolidado como uma das estratégias mais importantes para apoiar a governança descentralizada de dados. Em um ambiente com múltiplos domínios atuando de forma autônoma, o uso de ferramentas automatizadas permite acompanhar, controlar e garantir a qualidade, a segurança e a rastreabilidade dos dados sem depender exclusivamente de ações manuais, que seriam inviáveis em larga escala (Dolhopolov et al., 2024; Wider et al., 2023).

Uma das práticas mais recorrentes é o uso de scorecards automatizados de qualidade, que funcionam como painéis digitais para acompanhar o desempenho e a conformidade dos dados (Dolhopolov et al., 2024; Wider et al., 2023). Esses scorecards monitoram, por exemplo, se os campos obrigatórios estão preenchidos corretamente, se os dados estão sendo atualizados na frequência esperada, se seguem os padrões estabelecidos nos contratos de dados e se apresentam inconsistências frequentes. Em outras palavras, são ferramentas que funcionam como uma espécie de “checklist inteligente”, garantindo que os dados estejam dentro dos critérios exigidos pela governança organizacional (Dolhopolov et al., 2024; Blohm et al., 2024). Essas ferramentas também ajudam a acompanhar a linhagem de dados (data lineage), que torna possível identificar, por exemplo, quais sistemas alimentaram determinada informação, quais domínios a modificaram e em

quais análises ela foi utilizada. Essa rastreabilidade é essencial tanto para garantir a qualidade dos dados quanto para atender requisitos regulatórios e de auditoria (Wider et al., 2023; Bode et al., 2023).

Outro recurso tecnológico que tem sido explorado por algumas organizações é o uso de contratos inteligentes (smart contracts) baseados em tecnologia blockchain (Pingos et al., 2024; Almaslukh et al., 2024). Esses contratos funcionam como regras programadas que são executadas automaticamente quando determinadas condições são atendidas. Por exemplo, um domínio consumidor pode acessar um conjunto de dados somente se tiver as permissões necessárias e se os dados estiverem com a documentação e qualidade adequadas. Como os contratos são registrados em blockchain, todas as transações são seguras, auditáveis e imutáveis, ou seja, não podem ser alteradas sem deixar rastro (Pingos et al., 2024; Almaslukh et al., 2024).

Esse tipo de solução tecnológica fortalece não apenas a governança, mas também a segurança da informação, uma vez que impede acessos não autorizados, registra todas as interações com os dados e assegura a conformidade com políticas organizacionais. O estudo de Pingos et al. (2024) destaca que esses mecanismos automáticos são particularmente úteis em organizações de grande porte, onde o controle manual seria ineficiente e sujeito a falhas. Um exemplo prático dessa estratégia pode ser observado no caso do PayPal, descrito por Viseschart & Pongpech (2023), em que contratos inteligentes e mecanismos de validação automatizada foram adotados para garantir que os dados dos domínios de crédito e risco estivessem devidamente documentados, protegidos e com acesso controlado. Essa abordagem reduziu o risco de inconsistências e reforçou a confiabilidade do sistema de dados da organização.

Como mostram esses exemplos, a automação tem se mostrado uma aliada fundamental para tornar viável a descentralização da governança de dados, oferecendo agilidade, confiabilidade e segurança em um ambiente cada vez mais complexo e dinâmico.

4.2.3 Plataformas self-service e escalabilidade operacional

Outra estratégia importante identificada na literatura para viabilizar a governança descentralizada de dados no modelo Data Mesh é a adoção de plataformas self-service, ou seja, plataformas de autosserviço. O objetivo dessas plataformas é oferecer às equipes dos diferentes domínios os recursos necessários para que elas próprias consigam acessar, transformar, analisar e publicar dados sem depender de um time centralizado de tecnologia (Wider et al., 2023; González-Velázquez et al., 2024). Essas plataformas são estruturadas com base no conceito de data as a platform, que significa tratar a infraestrutura de dados como um serviço oferecido internamente à organização. Ou seja, cada domínio passa a ter à sua disposição ferramentas e recursos padronizados para desenvolver seus próprios produtos de dados, com segurança e autonomia. Essa abordagem permite aumentar a escalabilidade operacional, pois reduz os gargalos causados pela dependência de equipes técnicas centrais e amplia a capacidade das organizações de lidar com grandes volumes de dados em diferentes áreas simultaneamente (Dolhopolov et al., 2024; Bode et al., 2023).

Na prática, essas plataformas costumam disponibilizar interfaces amigáveis, que permitem que os profissionais manipulem os dados sem precisar ter conhecimento aprofundado em programação (Wider et al., 2023). Além disso, incluem funcionalidades como pipelines reutilizáveis (que são fluxos automatizados para extração, transformação e carregamento de dados), controle de acesso baseado em papéis (RBAC), dashboards personalizáveis, templates de contratos de dados e ferramentas para documentação colaborativa (Wider et al., 2023; González-Velázquez et al., 2024). O RBAC (Role-Based Access Control) é um mecanismo que garante que cada pessoa tenha acesso apenas aos dados necessários para o desempenho da sua função, evitando acessos indevidos e reforçando a segurança da informação. Já os pipelines reutilizáveis facilitam o desenvolvimento de novos produtos de dados, permitindo que os times se concentrem no conteúdo, e não na infraestrutura técnica necessária para manipular os dados (Wider et al., 2023; González-Velázquez et al., 2024).

Um exemplo concreto da aplicação dessa estratégia pode ser observado no projeto Smart Factory Hub, que estruturou suas operações com base em uma plataforma self-service para integração e reutilização de dados industriais. Essa

plataforma possibilitou que os dados oriundos de diferentes domínios técnicos, como sensores de máquinas, registros de produção e relatórios de qualidade, fossem organizados de forma modular, com semântica comum e acesso controlado, promovendo maior fluidez operacional e interoperabilidade entre áreas distintas (González-Velázquez et al., 2024). Esse tipo de estrutura foi também observado por Vestues et al. (2022), em um estudo de caso sobre gestão ágil de dados, no qual a adoção de plataformas de autosserviço possibilitou maior autonomia dos domínios e maior velocidade na entrega de produtos de dados, reduzindo a dependência de equipes centralizadas e promovendo maior eficiência operacional em ambientes dinâmicos.

Dessa forma, as plataformas self-service têm se mostrado fundamentais para viabilizar, na prática, a descentralização preconizada pelo Data Mesh, oferecendo às equipes autonomia com responsabilidade, e promovendo um ambiente organizacional mais ágil, colaborativo e orientado a dados.

4.2.4 Estruturas organizacionais híbridas e apoio interdomínio

Embora o modelo Data Mesh proponha uma descentralização da governança de dados, a prática tem mostrado que nem todas as organizações conseguem aplicar esse modelo de forma totalmente autônoma em todos os domínios desde o início. Isso ocorre, principalmente, porque os níveis de maturidade e preparo das equipes variam significativamente dentro da mesma organização. Diante disso, uma estratégia frequentemente utilizada é a adoção de estruturas organizacionais híbridas, que combinam autonomia local com mecanismos de suporte institucional (Bode et al., 2023; Wider et al., 2023). Nesse modelo híbrido, cada domínio continua sendo responsável pelos dados que produz e utiliza, mas passa a contar com o apoio de núcleos coordenadores, que atuam como instâncias de orientação técnica e organizacional.

Uma das formas mais comuns de operacionalizar esse apoio são os comitês interdomínio, também chamados de federated governance boards. Esses comitês reúnem representantes de diferentes domínios da organização para definir padrões comuns, supervisionar a aplicação das diretrizes de governança e promover a troca de experiências entre as equipes. Além disso, funcionam como espaços para resolver conflitos, alinhar expectativas e reforçar a colaboração entre áreas que, de

outra forma, poderiam operar de forma isolada (Wider et al., 2023). Por exemplo, esses comitês podem ser responsáveis por revisar glossários corporativos, acompanhar a execução dos contratos de dados, supervisionar métricas de qualidade e avaliar riscos relacionados ao compartilhamento de dados sensíveis.

Um exemplo real da aplicação desse modelo híbrido pode ser observado na empresa Equinor, do setor de energia. Segundo González-Velázquez et al. (2024), a organização estruturou seus dados por meio de domínios especializados e criou instâncias de coordenação interdomínio que utilizam ferramentas como o LinkedIn DataHub para garantir rastreabilidade, interoperabilidade e alinhamento estratégico entre equipes. Essa estrutura permitiu que os domínios tivessem autonomia para desenvolver seus produtos de dados, mas dentro de um arcabouço comum que assegura consistência e governança organizacional.

Dessa forma, as estruturas híbridas possibilitam que a descentralização ocorra de maneira gradual e segura, adaptando-se ao nível de maturidade dos diferentes domínios sem comprometer a integridade dos dados nem o alinhamento institucional.

4.2.5 Capacitação e mudança cultural

A capacitação contínua dos profissionais e a mudança cultural nas organizações são dimensões indispensáveis para consolidar a governança descentralizada de dados no modelo Data Mesh. A transformação proposta por esse modelo não se limita à adoção de ferramentas tecnológicas ou ao redesenho de estruturas organizacionais. Trata-se, sobretudo, de uma redefinição da relação dos profissionais com os dados, exigindo novas posturas, responsabilidades e competências (Pingos et al., 2024; Wider et al., 2023).

No Data Mesh, há uma demanda para os domínios por uma capacitação técnica básica, como familiaridade com princípios de qualidade, compreensão de metadados e noções de uso de plataformas self-service (Dolhopolov et al., 2024; Bode et al., 2023). Entretanto, mais do que habilidades técnicas, o modelo exige uma mudança de mentalidade organizacional. É necessário promover a cultura da corresponsabilidade entre os domínios, em que os dados são reconhecidos como ativos estratégicos e compartilháveis, e não como informações restritas a áreas específicas (Blohm et al., 2024; Wider et al., 2023).

Nesse sentido, diversas organizações têm adotado estratégias voltadas ao engajamento dos colaboradores e à disseminação da cultura de dados. Um dos instrumentos mais citados na literatura são as data guilds, comunidades internas formadas por profissionais de diferentes áreas que se reúnem para compartilhar boas práticas, discutir padrões de governança e promover aprendizado colaborativo (Bode et al., 2023; González-Velázquez et al., 2024). Essas guilds funcionam como espaços de aprendizagem informal e são eficazes na incorporação dos princípios do Data Mesh no cotidiano organizacional.

Empresas como Netflix, Zalando e Saxo Bank são frequentemente citadas na literatura por investirem fortemente nessas iniciativas, obtendo resultados como maior engajamento dos profissionais, melhora na qualidade dos dados e maior adesão aos princípios da governança federada. Tais evidências reforçam que a transformação cultural e a capacitação organizacional são pilares centrais para o sucesso da descentralização proposta pelo Data Mesh.

4.3 Quais lacunas e oportunidades de pesquisa ainda existem na literatura sobre governança descentralizada em arquiteturas de Data Mesh?

Apesar dos avanços teóricos e práticos observados nos últimos anos sobre o modelo Data Mesh, a literatura ainda apresenta diversas lacunas que precisam ser aprofundadas por estudos futuros. O mapeamento conduzido neste trabalho revelou que, embora muitos dos princípios e estratégias do Data Mesh estejam bem definidos, ainda há áreas críticas que carecem de maior investigação, validação empírica e aplicação prática em diferentes contextos organizacionais (Machado et al., 2022; Dolhopolov et al., 2024; Wider et al., 2023).

Uma das lacunas mais relevantes diz respeito à escalabilidade do Data Mesh em grandes organizações, especialmente aquelas que operam com centenas ou milhares de domínios (Machado et al., 2022; Bode et al., 2023). Embora a descentralização reduza a sobrecarga sobre equipes centrais, pouco se sabe sobre os reais impactos desse modelo em estruturas organizacionais altamente complexas. São escassas, por exemplo, as pesquisas que mensuram com precisão custos operacionais, desafios de coordenação entre domínios em larga escala ou níveis de esforço necessários para garantir a interoperabilidade entre áreas com graus distintos de maturidade digital (Machado et al., 2022; Bode et al., 2023). Essas

lacunas, embora distintas, estão fortemente interligadas: a escalabilidade do modelo está diretamente relacionada à maturidade organizacional, enquanto a conformidade regulatória depende de estruturas robustas de rastreabilidade e coordenação interdomínio.

Além disso, a literatura ainda carece de frameworks normativos que relacionem o grau de maturidade organizacional à prontidão para a adoção do Data Mesh. Isso significa que faltam instrumentos que permitam diagnosticar, de forma estruturada, se uma organização está realmente preparada para descentralizar sua governança de dados ou se ainda precisa desenvolver competências técnicas, culturais e processuais antes de iniciar essa transição (Blohm et al., 2024; Bode et al., 2023).

Outro tema emergente que ainda carece de aprofundamento é a integração entre o Data Mesh e tecnologias de Inteligência Artificial (IA). A literatura reconhece o potencial da IA para apoiar a governança descentralizada, por exemplo, por meio da categorização automática de dados, detecção de anomalias ou sugestão inteligente de contratos semânticos. No entanto, são raros os estudos que exploram de forma prática essa integração, seus benefícios, limitações e implicações para os domínios organizacionais (Pingos et al., 2024; Dolhopolov et al., 2024).

A conformidade regulatória em ambientes descentralizados também permanece como um desafio subexplorado na literatura (Bode et al., 2023; Pingos et al., 2024). Embora estudos apontem que a fragmentação da responsabilidade entre domínios possa dificultar o cumprimento de normas como a Lei Geral de Proteção de Dados (LGPD), no Brasil, e o General Data Protection Regulation (GDPR), na União Europeia, ainda são necessários estudos mais aprofundados que avaliem como as organizações podem conciliar autonomia local com mecanismos corporativos de auditoria e rastreabilidade (Pingos et al., 2024; Bode et al., 2024).

Por fim, observa-se que muitas das práticas citadas na literatura ainda estão em fase experimental ou restritas a estudos de caso pontuais, carecendo de validações mais robustas em diferentes setores, portes de empresas e culturas organizacionais. Assim, há espaço para o desenvolvimento de estudos de caso longitudinais, pesquisas comparativas entre modelos de governança e modelagens teóricas que sistematizem boas práticas já consolidadas na indústria (Dolhopolov et al., 2024; González-Velázquez et al., 2024).

Essas lacunas demonstram que, embora o Data Mesh represente um avanço significativo no campo da governança de dados, ainda há um caminho importante a ser percorrido para consolidá-lo como modelo dominante nas organizações. Os resultados obtidos neste mapeamento evidenciam que a descentralização promove inovação e flexibilidade, mas sua efetiva implementação depende de esforços contínuos de padronização, capacitação, automação e amadurecimento organizacional.

5. DISCUSSÃO

A crescente adoção do modelo Data Mesh pelas organizações tem gerado diferentes interpretações e abordagens sobre sua aplicação prática. Nesse sentido, além do mapeamento realizado neste trabalho, destaca-se também a contribuição da Systematic Gray Literature Review (SGLR) de Goedegebuure et al. (2024), que analisou 114 fontes da chamada literatura cinzenta, ou seja, documentos técnicos, relatórios corporativos e experiências compartilhadas por profissionais da indústria. A comparação entre essas duas abordagens permite uma análise mais abrangente e crítica sobre os fundamentos, desafios e estratégias associados à governança descentralizada de dados.

Ambos os trabalhos reconhecem a relevância dos quatro princípios centrais propostos por Dehghani (2019), no entanto, as duas abordagens se diferenciam quanto ao foco de aprofundamento. A literatura acadêmica, conforme analisado no mapeamento, enfatiza principalmente os desafios técnicos e regulatórios, tais como interoperabilidade de dados, conformidade com leis como a LGPD e o GDPR, e segurança da informação (Blohm et al., 2024; Dolhopolov et al., 2024; Pingos et al., 2024). Já a SGLR amplia a discussão ao destacar questões práticas e organizacionais, como a resistência cultural à descentralização, a importância de capacitação contínua, e a necessidade de definir novos papéis de governança dentro das equipes (González-Velázquez et al., 2024; Bode et al., 2023).

Um dos principais avanços trazidos pela literatura cinzenta está na proposição de arquiteturas de referência baseadas em Service-Oriented Architecture (SOA) (Goedegebuure et al., 2024). Essa abordagem, ainda recente na literatura acadêmica, sugere que os princípios do Data Mesh podem ser operacionalizados por meio de estruturas inspiradas em serviços organizados de forma modular e reutilizável, como se fossem pequenas partes independentes de um sistema que se conectam para formar um ecossistema de dados eficiente. Essa analogia permite mapear mais claramente quais papéis são necessários (como produtores, curadores e consumidores de dados) e como cada domínio pode oferecer seus dados como um "serviço" dentro da organização (Marine Transport Business, 2023; Wider et al., 2023).

Outro ponto de diferenciação importante está na abordagem da governança dos dados. Enquanto os estudos acadêmicos tratam a governança federada com

base em padronizações organizacionais e boas práticas documentadas, a literatura cinzenta evidencia o avanço no uso de soluções automatizadas, como contratos inteligentes e políticas como código (policy-as-code). Neste último caso, trata-se de uma prática em que regras de segurança, qualidade e acesso aos dados são escritas como códigos executáveis por sistemas computacionais, permitindo que essas regras sejam aplicadas de forma automática e padronizada (Pingos et al., 2024).

Contudo, apesar da agilidade que essas soluções oferecem, o impacto real dessas ferramentas em contextos mais regulados, como setor público, saúde e energia, ainda precisa ser melhor investigado. Nesses setores, a auditabilidade e a prestação de contas (accountability) exigem mecanismos formais e robustos de validação. A literatura acadêmica ainda carece de estudos que demonstrem como essas práticas automatizadas se comportam em ambientes com exigências legais mais complexas (Pingos et al., 2024).

Além dos aspectos técnicos, a SGLR também aprofunda a discussão sobre a dimensão organizacional da implementação do Data Mesh. Estratégias como capacitação permanente dos colaboradores, formação de equipes interdisciplinares, definição de incentivos por domínio e mapeamento dos processos internos foram destacadas como caminhos relevantes para enfrentar os obstáculos da descentralização. Embora esses elementos também estejam presentes nos estudos acadêmicos, muitas vezes são tratados de forma menos sistematizada, o que reforça a importância de unir teoria e prática (Blohm et al., 2024; Wider et al., 2023; González-Velázquez et al., 2024).

A análise comparativa entre o mapeamento e a SGLR evidencia que, embora ambas as abordagens reforcem os fundamentos essenciais do Data Mesh, a prática organizacional tem avançado com mecanismos complementares, como o uso de tecnologias automatizadas e estratégias de engajamento humano, cuja base teórica ainda precisa ser melhor desenvolvida pela academia. A integração entre as visões acadêmicas e as contribuições práticas da indústria representa uma oportunidade valiosa para amadurecer o campo de estudo, consolidar diretrizes aplicáveis e ampliar a adoção do Data Mesh de forma contextualizada e sustentável (Goedegebuure et al., 2024; González-Velázquez et al., 2024).

6. AMEAÇAS À VALIDADE

A análise das ameaças à validade representa uma etapa essencial para assegurar a consistência metodológica e a confiabilidade dos resultados apresentados neste estudo, conforme ressaltado por Ampatzoglou et al. (2019) como uma prática indispensável para garantir a credibilidade de estudos secundários. Diversos cuidados foram adotados no planejamento e execução da pesquisa com o objetivo de minimizar vieses e garantir maior rigor científico.

A primeira consideração recai sobre a validade da seleção dos estudos. A formulação e validação das strings de busca foram realizadas por meio de testes preliminares, de modo a assegurar a abrangência dos resultados e evitar a inclusão de publicações irrelevantes. Para isso, foram estabelecidos critérios de inclusão e exclusão objetivos, considerando apenas estudos publicados a partir de 2020, com acesso ao texto completo e submetidos à revisão por pares. A padronização da extração de dados foi considerada uma etapa crítica para garantir a consistência dos resultados (Kitchenham & Charters, 2007; Ampatzoglou et al., 2019). Para isso, foi desenvolvido um formulário estruturado que orientou a coleta de informações-chave, como metodologia utilizada, desafios identificados, benefícios destacados e propostas de solução. Além disso, foi realizada uma avaliação de qualidade dos estudos incluídos, restringindo a análise a artigos com pontuação superior a 80%, o que garantiu a robustez do conjunto amostral.

No que se refere à validade externa, destaca-se a heterogeneidade dos contextos organizacionais abordados nos estudos, o que pode limitar a generalização dos resultados. Para mitigar esse efeito, foram incluídas publicações que contemplassem múltiplas perspectivas e setores de atuação, ampliando a representatividade do mapeamento. Ademais, todas as etapas metodológicas foram documentadas de forma detalhada, permitindo sua replicabilidade por outros pesquisadores.

Considera-se, portanto, que os cuidados metodológicos adotados contribuíram para garantir a confiabilidade e a transparência dos resultados, fortalecendo a validade dos achados e sua relevância para pesquisas futuras sobre governança descentralizada de dados.

7. CONCLUSÃO E TRABALHOS FUTUROS

Este estudo teve como objetivo identificar os principais fundamentos, desafios e estratégias associados à governança descentralizada de dados no contexto do modelo Data Mesh. Os resultados indicam que, embora esse modelo represente uma alternativa promissora às arquiteturas centralizadas, sua adoção exige transformações estruturais que abrangem desde práticas operacionais e institucionais até mudanças comportamentais na cultura organizacional.

Os resultados evidenciam que o Data Mesh possui potencial para transformar a governança de dados organizacional, promovendo escalabilidade, autonomia dos domínios e maior fluidez operacional. No entanto, sua consolidação depende da capacidade das organizações em alinhar práticas técnicas, institucionais e culturais de forma integrada e contínua. Por outro lado, também foram identificados desafios relevantes, como a necessidade de padronização da interoperabilidade, garantia da qualidade dos dados, conformidade regulatória e superação da resistência organizacional. Estratégias como a adoção de governança federada, capacitação das equipes, uso de catálogos de dados e automação por meio de contratos inteligentes têm sido apontadas como caminhos viáveis para mitigar esses desafios.

Além de mapear lacunas conceituais ainda pouco exploradas na literatura, este trabalho oferece contribuições práticas ao identificar padrões de adoção e estratégias aplicáveis a diferentes realidades organizacionais. Casos como o da Equinor e do Smart Factory Hub evidenciam que a descentralização pode ser operacionalizada com estruturas híbridas, mesmo em setores industriais tradicionais, fornecendo aprendizados valiosos para empresas em estágio inicial de transição.

A análise comparativa com a Systematic Gray Literature Review evidencia que a prática organizacional vem avançando em soluções que ainda carecem de respaldo acadêmico mais robusto. Aspectos como o alinhamento com arquiteturas SOA, a automação da governança e a redefinição dos papéis organizacionais despontam como temas centrais para estudos futuros. Tais elementos devem ser explorados com maior profundidade, especialmente por meio de estudos empíricos e investigações interdisciplinares.

Nesse contexto, destaca-se como linha de investigação promissora a interseção entre Data Mesh e Inteligência Artificial. A associação entre esses dois campos pode viabilizar avanços significativos na governança descentralizada,

sobretudo com o uso de IA para classificação automatizada de metadados, validação de contratos de dados e detecção de anomalias. Pretende-se explorar essa interseção em estudos futuros, com foco na aplicação de LLMs (Large Language Models) para apoiar a governança federada e na criação de assistentes inteligentes para curadoria de dados em plataformas self-service. Essa linha representa uma continuidade natural desta pesquisa e poderá ser aprofundada em uma futura investigação em nível de mestrado.

Estudos de caso e experimentações controladas em ambientes reais podem ser caminhos valiosos para aprofundar a aplicabilidade do modelo (Dolhopolov et al., 2024) e avaliar sua eficácia em diferentes graus de maturidade organizacional. Tais abordagens permitiriam capturar a evolução progressiva da adoção do Data Mesh, seus impactos nos processos organizacionais e as reais barreiras enfrentadas em distintos contextos setoriais.

Acredita-se que, com o avanço das práticas de engenharia de dados, políticas automatizadas e capacitação organizacional, o Data Mesh tende a evoluir de uma tendência emergente para consolidar-se como um modelo dominante de governança em ecossistemas digitais complexos. Dessa forma, o presente trabalho contribui não apenas para o entendimento conceitual e aplicado da abordagem, mas também para a consolidação de um campo de estudo promissor, cuja relevância tende a crescer à medida que a complexidade dos dados e das organizações se intensifica.

REFERÊNCIAS

- [1] ALMASLUKH, Abdulaziz et al. Data Mesh Meets Blockchain. *International Journal of Computational Intelligence Systems*, v. 17, 2024, 27 p. DOI: <https://doi.org/10.1007/s44196-024-00404-z>.
- [2] AMPATZOGLOU, Apostolos et al. Identifying, categorizing and mitigating threats to validity in software engineering secondary studies. *Information and Software Technology*, v. 106, p. 201–230, 2019. DOI: <https://doi.org/10.1016/j.infsof.2018.10.006>.
- [3] AZEROUAL, Otmane; NACHEVA, Radka. Data Mesh for Managing Complex Big Data Landscapes and Enhancing Decision Making in Organizations. In: *15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2023) - Volume 3: KMIS*. p. 202–212. DOI: <https://doi.org/10.5220/0012195700003598>.
- [4] BLOHM, Ivo et al. Data products, data mesh, and data fabric: New paradigms for data and analytics? *Business and Information Systems Engineering*, v. 66, n. 5, p. 643–652, 2024. DOI: <https://doi.org/10.1007/s12599-024-00876-5>.
- [5] BODE, Jan et al. Toward Avoiding the Data Mess: Industry Insights From Data Mesh Implementations. *IEEE Access*, v. 12, p. 95402–95416, 2024. DOI: <https://doi.org/10.1109/ACCESS.2024.3417291>.
- [6] DEHGHANI, Zhamak. How to move beyond a monolithic data lake to a distributed data mesh. Martin Fowler, 2019. Disponível em: <https://martinfowler.com/articles/data-monolith-to-mesh.html>. Acesso em: 18 mar. 2025.
- [7] DOLHOPOLOV, Anton; CASTELLTORT, Arnaud; LAURENT, Anne. Implementing Federated Governance in Data Mesh Architecture. *Future Internet*, v. 16, n. 4, p. 115, 2024. DOI: <https://doi.org/10.3390/fi16040115>.
- [8] GOEDEGEBUURE, Abel et al. Data Mesh: A Systematic Gray Literature Review. New York: ACM, 2024. Disponível em: <https://arxiv.org/abs/2304.01062>. Acesso em: 10 mar. 2025.
- [9] GONZÁLEZ-VELÁZQUEZ, Roberto et al. Smart Factory Hub – Towards a Data Mesh in Smart Manufacturing. *Procedia Computer Science*, v. 232, p. 2709–2719, 2024. DOI: <https://doi.org/10.1016/j.procs.2024.02.088>.
- [10] KITCHENHAM, Barbara; CHARTERS, Stuart. Guidelines for performing Systematic Literature Reviews in Software Engineering. Keele University; University of Durham, 2007. (EBSE Technical Report EBSE-2007-01). Disponível em: https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf. Acesso em: 10 mar. 2025.
- [11] LOUKIALA, Antti et al. Migrating from a Centralized Data Warehouse to a Decentralized Data Platform Architecture. [Evento não especificado], 2023. p. 1–12.

- [12] MACHADO, Inês Araújo; COSTA, Carlos; SANTOS, Maribel Yasmina. Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures. *Procedia Computer Science*, v. 196, p. 263–271, 2022. DOI: <https://doi.org/10.1016/j.procs.2021.12.013>.
- [13] MOHER, David et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, v. 6, n. 7, p. e1000097, 2009. DOI: <https://doi.org/10.1371/journal.pmed.1000097>.
- [14] NARGESIAN, Fatemeh et al. Data Lake Management: Challenges and Opportunities. *Proceedings of the VLDB Endowment*, v. 12, n. 12, p. 1986–1989, 2019. DOI: <https://doi.org/10.14778/3352063.3352116>.
- [15] PINGOS, Michalis; CHRISTODOULOU, Panayiotis; ANDREOU, Andreas S. Security and Ownership in User-Defined Data Meshes. *Algorithms*, v. 17, n. 4, p. 169, 2024. DOI: <https://doi.org/10.3390/a17040169>.
- [16] SAWADOGO, Pegdwendé; DARMONT, Jérôme. On Data Lake Architectures and Metadata Management. *arXiv preprint arXiv:2107.11152*, 2021. Disponível em: <https://arxiv.org/abs/2107.11152>. Acesso em: 10 mar. 2025.
- [17] VESTUES, Kathrine et al. Agile Data Management in NAV: A Case Study. In: STRAY, V. et al. (Org.). *XP 2022: Agile Processes in Software Engineering and Extreme Programming. Lecture Notes in Business Information Processing*, v. 445, p. 220–235, 2022. DOI: https://doi.org/10.1007/978-3-031-08169-9_14.
- [18] VISESCHART, Kitipong; PONGPECH, Worapol. Marine Transport Business: A Data Mesh-Based Data Asset Strategy. In: *2023 8th International Conference on Information Systems Engineering (ICISE 2023)*, Bangkok, Thailand. ACM, 2023. p. 1–9. DOI: <https://doi.org/10.1145/3641032.3641060>.
- [19] WIDER, Arif; VERMA, Sumedha; AKHTAR, Atif. Decentralized Data Governance as Part of a Data Mesh Platform: Concepts and Approaches. *IEEE*, 2023. Disponível em: <https://ieeexplore.ieee.org/document/10248262>. Acesso em: 10 mar. 2025.