

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Programa de Pós-Graduação em Administração

Paulo Augusto de Souza Moreira

**IMPACTO DA GOVERNANÇA DE DADOS NO USO DE *BUSINESS*
INTELLIGENCE:**

Estudo de caso em uma empresa de logística de grande porte

Belo Horizonte

2024

Paulo Augusto de Souza Moreira

**IMPACTO DA GOVERNANÇA DE DADOS NO USO DE *BUSINESS*
INTELLIGENCE:**

Estudo de caso em uma empresa de logística de grande porte

Dissertação apresentada ao Programa de Pós-Graduação em Administração da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para o Mestrado em Administração.

Orientador: Prof. Dr. Rodrigo Baroni de Carvalho

Belo Horizonte

2024

FICHA CATALOGRÁFICA

Elaborada pela Biblioteca da Pontifícia Universidade Católica de Minas Gerais

M838i	Moreira, Paulo Augusto de Souza Impacto da governança de dados no uso de <i>Business Intelligence</i>: estudo de caso em uma empresa de logística de grande porte / Paulo Augusto de Souza Moreira. Belo Horizonte, 2024. 116 f. : il. Orientador: Rodrigo Baroni de Carvalho Dissertação (Mestrado) - Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Administração 1. Sistemas de informação gerencial. 2. Inteligência competitiva (Administração). 3. Gerenciamento de recursos de informação. 4 Sistemas de suporte de decisão. 5. Governança de TI. 6. Business intelligence. 7. Método de estudo de casos. I. Carvalho, Rodrigo Baroni de. II. Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Administração. III. Título. SIB PUC MINAS CDU: 658.011.56
-------	---

Paulo Augusto de Souza Moreira

**IMPACTO DA GOVERNANÇA DE DADOS NO USO DE *BUSINESS*
INTELLIGENCE:**

Estudo de caso em uma empresa de logística de grande porte

Dissertação apresentada ao Programa de Pós-Graduação
em Administração da Pontifícia Universidade Católica de
Minas Gerais, como requisito parcial para o Mestrado em
Administração.

Prof. Dr. Rodrigo Baroni... – PUC Minas (Orientador)

Prof. Dr. Renato Penha – Uninove (Banca Examinadora)

Prof. Dr. Humberto Elias Garcia Lopes – PUC Minas (Banca Examinadora)

Belo Horizonte, 19 de dezembro de 2024

AGRADECIMENTOS

Agradeço primeiramente a minha família, que sempre me apoia em todos os passos que escolho para minha vida pessoal, profissional e acadêmica. À minha mãe Suely, meu pai Flávio e meus irmãos Pedro e Ana Flávia, o apoio de vocês é fundamental para meu crescimento.

Não citei acima pois quero fazer um agradecimento especial, à minha esposa Gabriela. Muito obrigado por sempre acreditar em mim e insistir no meu potencial. Agradeço o apoio dado em toda essa trajetória.

Agradeço ao professor e orientador Rodrigo Baroni pela paciência, excelente orientação e os ensinamentos que possibilitaram a realização dessa pesquisa. Sua capacidade em me guiar durante esses dois anos, as trocas que tivemos e os ensinamentos que foram dados marcaram essa minha trajetória e levarei sempre comigo.

Aos meus colegas de trabalho, agradeço o que estamos construindo em conjunto. Essa pesquisa reflete a transformação que temos feito.

RESUMO

As tecnologias de *Business Intelligence* (BI) potencialmente habilitam melhores tomadas de decisão estratégicas e operacionais. Contudo, principalmente com o advento do modelo *self-service BI* (SSBI), a difusão destas ferramentas pode colocar em risco a segurança e privacidade dos dados na organização, além de poder comprometer a qualidade dos dados que estão sendo utilizados e a assertividade das informações geradas. O objetivo do trabalho é avaliar o impacto da governança de dados na utilização de ferramentas de *self-service BI* nas organizações. Para tal finalidade, foi realizado um estudo de caso único em uma empresa brasileira de grande porte do setor de logística. Uma abordagem quantitativa foi empregada para avaliar os fatores de impacto da governança de dados no *self-service BI* e uma abordagem qualitativa analisou o processo de adoção dos critérios de governança por meio de entrevistas semiestruturadas e análise documental, caracterizando assim métodos mistos. Os resultados mostraram que a aplicação da governança de dados proporciona mais segurança e privacidade, maior qualidade nos dados, melhor difusão do conhecimento organizacional por meio da gestão dos metadados e maior eficiência no desenvolvimento dos relatórios pelas equipes. Nas situações em que a governança de dados se fez presente, um consumo otimizado dos recursos tecnológicos foi constatado, proporcionando uma otimização de custos para a organização e uma abordagem mais sustentável. A pesquisa evidenciou a necessidade de uma prévia capacitação dos desenvolvedores de relatórios de SSBI no modelo de governança da organização antes do início da migração dos relatórios para o modelo governado. No intuito de obter uma maior maturidade dos indicadores que serão construídos e de evitar retrabalhos, uma recomendação gerencial relevante é a preservação da equipe técnica para manter uma continuidade no trabalho. Uma contribuição da pesquisa é um *framework* que associa a governança de dados com o SSBI, evidenciando os requisitos mínimos para o sucesso na implantação da governança.

Palavras-chave: Governança de dados, *Business Intelligence*, *Self-service BI*, Logística, Governança da Informação, *Analytics*.

ABSTRACT

The Business Intelligence (BI) technologies may provide better strategic and operational decisions. However, due to the emergence of self-service BI (SSBI), the dissemination of these tools may put organizational data security and privacy at risk, compromising the data quality and the information assertiveness as well. The research objective is to evaluate the impact of data governance on the use of self-service BI in organizations. A single case study was carried out in a large Brazilian company in the logistics sector. A quantitative approach was used to evaluate the impact factors of data governance on self-service BI and a qualitative approach analyzed the process of data governance adoption through semi-structured interviews and document analysis, characterizing mixed methods. The results showed that the application of data governance provides better security and privacy, greater data quality, a better dissemination of organizational knowledge through metadata management and more efficiency in the development of reports by teams. In the situations where data governance was present, an optimized consumption of technological resources was observed, providing cost optimization for the company and a more sustainable approach. The research highlighted the need for prior training of developers in the company's governance model before you start migrating the reports to the governed model. A relevant managerial recommendation is to preserve a technical team to assure continuity in the work, achieving a higher maturity of the indicators and avoiding rework. A research contribution is a framework that associates data governance with SSBI, pointing out the minimum requirements for successful implementation of governance.

Keywords: Data Governance, Business Intelligence, Self-service BI, Logistics, Information Governance, Analytics.

LISTA DE FIGURAS

FIGURA 1 – Publicações por ano	14
FIGURA 2 – Cenário tecnológico de BI&A	18
FIGURA 3 – Objetivos do SSBI	19
FIGURA 4 – Níveis de <i>self-service</i>	21
FIGURA 5 – Governança Corporativa e dos Principais Ativos	24
FIGURA 6 – Relação entre governança: governança corporativa, governança de TI e governança de dados	26
FIGURA 7 – DAMA-DMBOK2 <i>Data Management Framework</i>	29
FIGURA 8 – Governança e Administração de Dados	31
FIGURA 9 – Organização da governança de dados	35
FIGURA 10 – Conceitos de interação com a TI	39
FIGURA 11 – Modelo Conceitual Preliminar	40
FIGURA 12 – Modelo incorporado	46
FIGURA 13 – Linha do tempo de adoção de <i>Business Intelligence</i>	50
FIGURA 14 – Universo de usuários das ferramentas de BI	51
FIGURA 15 – Modelo arquitetural de dados	52
FIGURA 16 – Processo de implementação de produtos de dados	53
FIGURA 17 – Estrutura de implementação de produtos de dados	54
FIGURA 18 – Estrutura organizacional da área de Tecnologia da Informação	55
FIGURA 19 – Distribuição das respostas das questões referentes à Consumo dos Relatórios	59
FIGURA 20 – Distribuição das respostas das questões referentes à Facilidade de Uso	60
FIGURA 21 – Distribuição das respostas das questões referentes à Velocidade de Implementação	61
FIGURA 22 – Distribuição das respostas das questões referentes à Acesso a Fontes de Dados	62
FIGURA 23 – Distribuição das respostas das questões referentes à Segurança e Privacidade	65
FIGURA 24 – Distribuição das respostas das questões referentes à Qualidade dos dados	67
FIGURA 25 – Distribuição das respostas das questões referentes à Gestão de Metadados	69
FIGURA 26 – Distribuição das respostas das questões referentes à Eficiência no Desenvolvimento	71

FIGURA 27 – Modelo Conceitual Final	74
FIGURA 28 – Consumo de UCs do ambiente de produção	76
FIGURA 29 – Consumo de UCs do ambiente de experimentação	77

LISTA DE TABELAS

TABELA 1 – Frameworks de governança de dados	28
TABELA 2 – Tabela de hipóteses do modelo conceitual	40
TABELA 3 – Itens e variáveis atreladas ao construto Self-Service BI	42
TABELA 4 – Itens e variáveis atreladas ao construto Governança de Dados	43
TABELA 5 – Experiência profissional dos respondentes	58
TABELA 6 – Tempo de empresa dos respondentes	58
TABELA 7 – Tempo de experiência com produtos de BI	58
TABELA 8 – Resultados da avaliação de SSBI	63
TABELA 9 – Resultados da avaliação de Governança de Dados	72
TABELA 10 – Resultados da regressão logística multivariada	73
TABELA 11 – Tabela de hipóteses final	75
TABELA 12 – Performance dos ambientes	77
TABELA 13 – Perfil dos entrevistados técnicos	79
TABELA 14 – Perfil dos entrevistados usuários	82

SUMÁRIO

1. INTRODUÇÃO	11
1.1. Problema de pesquisa	11
1.2. Justificativa.....	13
1.3. Objetivos	15
2. REFERENCIAL TEÓRICO	17
2.1. <i>Business Intelligence e Analytics</i>	17
2.2. Governança de Dados.....	23
2.3. Síntese do Referencial Teórico	38
3. METODOLOGIA	45
3.1. Estratégia e método	45
3.2. Unidade de análise	49
3.3. Coleta de dados	55
4. ANÁLISE DOS RESULTADOS QUANTITATIVOS	57
4.1. Caracterização da Amostra.....	57
4.2. Resultados <i>Self-Service BI</i>	58
4.3. Resultados Governança de Dados	64
4.4. Análise Documental	76
5. ANÁLISE DOS RESULTADOS QUALITATIVOS	79
5.1. Entrevistas Técnicas	79
5.2. Migração de Modelo de Governança	82
6. DISCUSSÃO DOS RESULTADOS.....	88
7. CONCLUSÃO	90
REFERÊNCIAS	92
APÊNDICE A – Formulário para <i>survey</i>	97
APÊNDICE B – Roteiro para entrevista técnica	100
APÊNDICE C – Roteiro para entrevista de avaliação da transição de modelo	102
APÊNDICE D – Fórmulas	104
APÊNDICE E – Código R	106
APÊNDICE F – Dados de análise de ACP	110

1. INTRODUÇÃO

Este capítulo apresenta o problema de pesquisa, na sequência a justificativa do desenvolvimento do trabalho, e em seguida os objetivos e o detalhamento da estrutura do trabalho.

1.1. Problema de pesquisa

As tecnologias de *Business Intelligence* (BI) são definidas como o conjunto de ferramentas de análise que combinam os dados operacionais para apresentação de informações que subsidiam a tomada de decisão nas organizações (Negash, 2004). Elas têm se tornado cada vez mais utilizadas nas organizações devido à necessidade de tomar decisões de forma mais rápida e assertiva, estendendo seu uso das informações estratégicas para a execução de tarefas operacionais, aumentando assim o número de profissionais que necessitam usá-las (Imhoff e White, 2011; Böhringer et al., 2010). Aliada às constantes mudanças dos negócios, a demanda por profissionais alocados nas equipes de Tecnologia da Informação (TI), mais especificamente por especialistas em BI, tem aumentado, gerando um gargalo na produtividade das organizações (Kobielus, 2009).

A partir disso, as tecnologias de *Self-Service Business Inteligente* (SSBI) surgem, permitindo aos usuários da organização acessarem os dados e conduzirem suas próprias análises sem a necessidade de envolvimento dos departamentos de TI (Lennerholt et al., 2018), promovendo um empoderamento dos usuários para a exploração dos dados (Alpar e Schulz, 2015). As tecnologias de SSBI visam remover gargalos devido ao aumento dos dados disponíveis nas organizações e o aumento dos usuários, e desde então passaram a ser consideradas como um movimento no campo do BI (Lennerholt et al., 2023).

O desejo por fácil acesso aos dados de forma independente para usuários de negócio das organizações já existe há muito tempo (Alpar e Schulz, 2015). Com as ferramentas de SSBI, os usuários finais das informações mudam sua ação de simplesmente consumir as informações para tentar diferentes formas de visualização dos dados (Stodder, 2015).

Apesar dos benefícios previstos nas tecnologias de SSBI, sua implantação apresenta vários desafios. Alpar e Schulz (2015) e Lennerholt et al. (2018) argumentam ser necessário ter pesquisas mais detalhadas sobre como as organizações devem implementar o SSBI. Lennerholt et al. (2018) apresentam os principais desafios envolvidos, que podem ser agrupados em duas categorias: acesso e uso dos dados, e independência dos usuários. Os desafios relacionados ao acesso e uso dos dados abrangem a necessidade de tornar as fontes de dados fáceis de acessar, com critérios para que o usuário possa selecionar os dados a serem utilizados com base em sua qualidade, das fontes corretas, e com políticas de governança de dados implementadas para

garantir a integridade e segurança dos dados. Já os desafios relacionados à independência dos usuários tratam sobre tornar as ferramentas de BI fáceis de utilizar, os resultados fáceis de consumir e o treinamento dos usuários nas melhores práticas de uso dos dados para gerar as melhores tomadas de decisão.

Entre as possíveis soluções para esses desafios do uso do SSBI, entende-se que a governança de dados pode ser um caminho, visando a disponibilização assertiva dos dados com qualidade e garantindo uma correta gestão de acesso aos dados. A governança de dados pode ser entendida como “o exercício da autoridade e controle (planejamento, monitoramento e aplicação) sobre o gerenciamento de ativos de dados” (DAMA, 2007, p. 67). A governança de dados guia todas as frentes do gerenciamento do dado como um ativo e visa garantir esse correto gerenciamento, a partir da implantação de políticas e boas práticas (DAMA, 2007).

A implementação de uma governança de dados eficaz pode auxiliar as organizações a habilitar uma melhor tomada de decisão, diminuir o atrito organizacional, proteger as necessidades dos stakeholders, construir padrões, reduzir custos e garantir a transparência dos processos (Al-Ruithe e Benkhelifa, 2017). Além disso, a governança de dados permite que as organizações se adequem e garantam a correta utilização dos dados, com a disponibilização do dado para quem de fato precisa e o controle dos objetivos para os quais os dados estão sendo utilizados, respeitando regulamentações relacionadas ao correto uso dos dados, como no caso brasileiro a Lei Geral de Proteção de Dados (LGPD) (Abraham et al., 2019). De modo geral, a governança de dados é peça fundamental para a organização, garantindo dados de qualidade e com segurança (Ethan et al., 2023).

Dessa forma, a aplicação da governança de dados apresenta grande importância para o cenário de uso de SSBI nas organizações, mas é preciso ter equilíbrio. Segundo Tallon et al. (2004), é importante controlar as restrições aplicadas pela governança de dados, pois essas restrições aplicadas em excesso fazem com que os usuários tentem contornar as políticas, assumindo riscos desnecessários e propiciando maior uso de ferramentas *shadow IT*.

A relação entre a governança de dados e o SSBI é um assunto com alto interesse na indústria, porém atualmente faltam pesquisas acadêmicas sobre esse assunto (Clarke et al., 2016). Clarke et al. (2016) apresentam um *framework* que relaciona o SSBI e a governança dos consumidores de dados, principalmente focado no controle do treinamento das pessoas que farão uso do dado para então terem acesso ao mesmo.

Gröger (2018) apresenta ainda a necessidade do estudo da aplicação da governança para *analytics* e BI, e a necessidade da construção de um *framework* para implantação que apresente alguns aspectos relevantes: balancear a acurácia dos dados e a flexibilidade do uso, integrar o

modelo às plataformas de BI já existentes na organização e gerenciar o processo de tomada de decisão baseada nos dados.

Dessa forma, torna-se necessária a realização de estudos aprofundados sobre o impacto da implantação da governança de dados nas organizações e a utilização de ferramentas de SSBI. Lennerholt et al. (2018) e Stodder (2015) abordam que uma governança de dados eficaz pode fazer com que os usuários entendam o valor dos dados, evitando a criação de estruturas de *shadow IT*, que podem gerar problemas para a organização. *Shadow IT* é um cenário comumente encontrado nas organizações onde hardwares, softwares ou serviços de tecnologia são utilizados no ambiente de trabalho sem a explícita aprovação ou conhecimento dos responsáveis na organização (Haag e Eckhardt, 2017), caracterizando uma estrutura paralela, redundante ou até mesmo conflitante com a estrutura oficialmente homologada na organização. Dessa forma, torna-se necessário o entendimento dos fatores de sucesso para a implantação da governança de dados relacionados ao uso de iniciativas de SSBI (Michalczyk et al., 2020).

Considerando os aspectos e critérios apresentados sobre a governança de dados e os fatores motivadores para a implantação de ferramentas de SSBI nas organizações, formula-se a seguinte pergunta de pesquisa: *Como a governança de dados impacta a utilização de ferramentas de SSBI?*

1.2. Justificativa

As organizações em todo o mundo têm buscado se tornar empresas mais *data-driven*, ou seja, empresas que tomam decisões operacionais e estratégicas baseadas em dados. De acordo com estudo da *Harvard Business Review*, encomendado pela Google, empresas que tem suas ações baseadas em dados se mostraram mais resilientes e preparadas para manter sua performance durante o período de pandemia do COVID-19, sendo que o uso de dados se provou fundamental para a rápida tomada de decisões em momentos críticos (Bak, 2023). Ainda de acordo com este estudo, organizações líderes de mercado estão usando dados e Inteligência Artificial (IA) para potencializar todos os demais objetivos estratégicos, apresentando reflexos na melhoria da eficiência operacional, receita, retenção de clientes, satisfação dos empregados e custos (Bak, 2023).

Arelado ao uso de dados nas organizações, segundo estudo divulgado pelo Gartner (2022), 81% das empresas líderes em seu segmento no mercado mundial utilizam ferramentas de BI para a tomada de decisão. No Brasil, este percentual é de 54% e pequenas e médias empresas adotam as ferramentas de BI tanto quanto grandes empresas. Espera-se que o mercado de ferramentas de BI dobre em 2025, em relação a 2020, alcançando um valor de 13 bilhões de

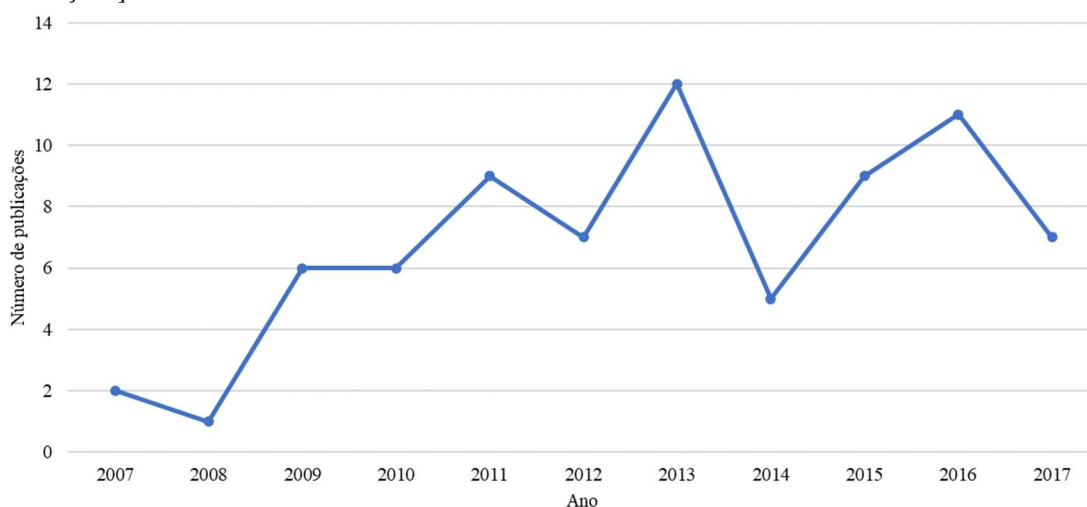
dólares anuais (Gartner, 2022). Além disso, de acordo com o estudo, em 28% das empresas, a tomada de decisão sobre as ferramentas de BI que são adquiridas é feita em conjunto pela equipe de TI e os demais departamentos da organização, o que contribui para a visão crescente de utilização de ferramentas de SSBI (Gartner, 2022).

Por outro lado, de acordo com um estudo feito pela McKinsey, a média das grandes organizações gastam 29% do tempo de seus empregados em tarefas que não geram valor devido à baixa disponibilidade e qualidade de dados (Petzold et al., 2020). Uma pesquisa feita pela Enterprise Strategy Group (2022) com 220 grandes empresas que possuem mais de 1.000 empregados e faturamento anual maior que US\$ 100 MM e que adotaram uma governança de dados mostra que 41% dessas empresas adotaram a governança de dados para a melhoria da qualidade dos dados e 25% para suportar uma melhor tomada de decisão, fatores que favorecem uma utilização mais adequada do SSBI (Leone e Walker, 2022).

Além disso, o estudo sobre governança de dados se faz necessário, pois as organizações ainda têm enfrentado um grande desafio relacionado a esse tema (Gartner, 2023). De acordo com projeção feita pelo Gartner (2023), estima-se que, em 2027, 75% dos produtos de dados serão redundantes e 60% das organizações terão falhado na geração de valor com uso de inteligência artificial.

Avaliando o contexto acadêmico, Nielsen (2017) mostra que a temática de pesquisa focada em governança de dados se mostra recente, apresentando poucas publicações no passado e um crescimento destas publicações nos anos mais recentes (Figura 1).

Figura 1
Publicações por ano



Fonte: Nielsen, O. B. (2017). A comprehensive review of data governance literature. Selected Papers of the IRIS, Issue Nr 8. 3. 120-133.

Por meio de uma busca na central de periódicos da CAPES, utilizando como palavras-chave “*Self-service BI*”, foram encontrados 130 documentos, sendo 105 artigos, 8 capítulos de livros e 17 demais. Quando feita a busca utilizando como palavras-chave “*data governance*”, filtrando apenas documentos em que essas palavras estão contempladas no título, são encontrados 664 documentos, sendo 389 artigos, 179 capítulos de livros e 96 demais. Porém, quando feita uma busca utilizando o critério “*self-service BI AND data governance*”, são encontrados 6 documentos apenas, sendo 2 artigos, 3 capítulos de livros e um livro. Entre os artigos encontrados, Riggins e Klamm (2017) apresentam um caso de ensino para ilustrar a necessidade de implantação de governança de dados, ao passo que Stone e Woodcock (2014) propõem uma abordagem de avaliação de competências e estruturas organizacionais que alie as temáticas de marketing digital e BI.

Portanto, constata-se que existem pesquisas com foco em SSBI e governança de dados separadamente, mas faltam pesquisas que relacionem os assuntos. Assim, o estudo da temática é justificado pelo tamanho e o nível de crescimento do mercado de soluções de BI, principalmente alavancados pela utilização de ferramentas de SSBI, aliado à importância atribuída a governança de dados e ainda a pequena quantidade de pesquisas vinculadas ao tema.

1.3. Objetivos

O objetivo geral desta pesquisa é analisar o impacto da governança de dados no uso de ferramentas de BI em uma empresa de logística de grande porte, sob a ótica de desenvolvedores e consumidores de relatórios produzidos em ferramentas de SSBI.

A pesquisa se propõe a confrontar a avaliação de usuários desenvolvedores e consumidores de SSBI embarcados em um modelo de governança estabelecido na organização, usuários desenvolvedores e consumidores de SSBI não embarcados no modelo de governança e consumidores de relatórios de BI que não fazem uso de ferramentas *self-service*, a fim de comparar os benefícios e restrições percebidas pelos mesmos que afetam a tomada de decisão na organização. Para tal, a pesquisa visa:

Objetivos específicos

- Avaliar a percepção de valor dos diferentes tipos de usuários sobre os critérios de governança de dados alocados ao modelo SSBI;
- Propor parâmetros sobre as definições da governança de dados no cenário de SSBI, a fim de evitar a geração de estruturas *shadow IT*;
- Definir requisitos necessários para o sucesso na implantação da governança de dados em um cenário de SSBI.

Para além desta introdução, a dissertação está organizada da seguinte forma: o capítulo 2 compreende o referencial teórico sobre as temáticas de governança de dados e *business intelligence* e *analytics*; o capítulo 3 detalha a metodologia do estudo de caso; o capítulo 4 contém a análise dos resultados da parte quantitativa da pesquisa e análise documental; o capítulo 5 apresenta a análise das informações obtidas na parte qualitativa da pesquisa; o capítulo 6 contempla a discussão dos resultados com as contribuições teóricas e práticas da pesquisa; e o capítulo 7 apresenta a conclusão da pesquisa.

2. REFERENCIAL TEÓRICO

Nas próximas seções, são abordados os temas de *Business Intelligence e Analytics* e Governança de Dados. Ao final, ambos os assuntos são relacionados para concepção de um modelo conceitual articulando um diálogo entre as duas temáticas.

2.1. *Business Intelligence e Analytics*

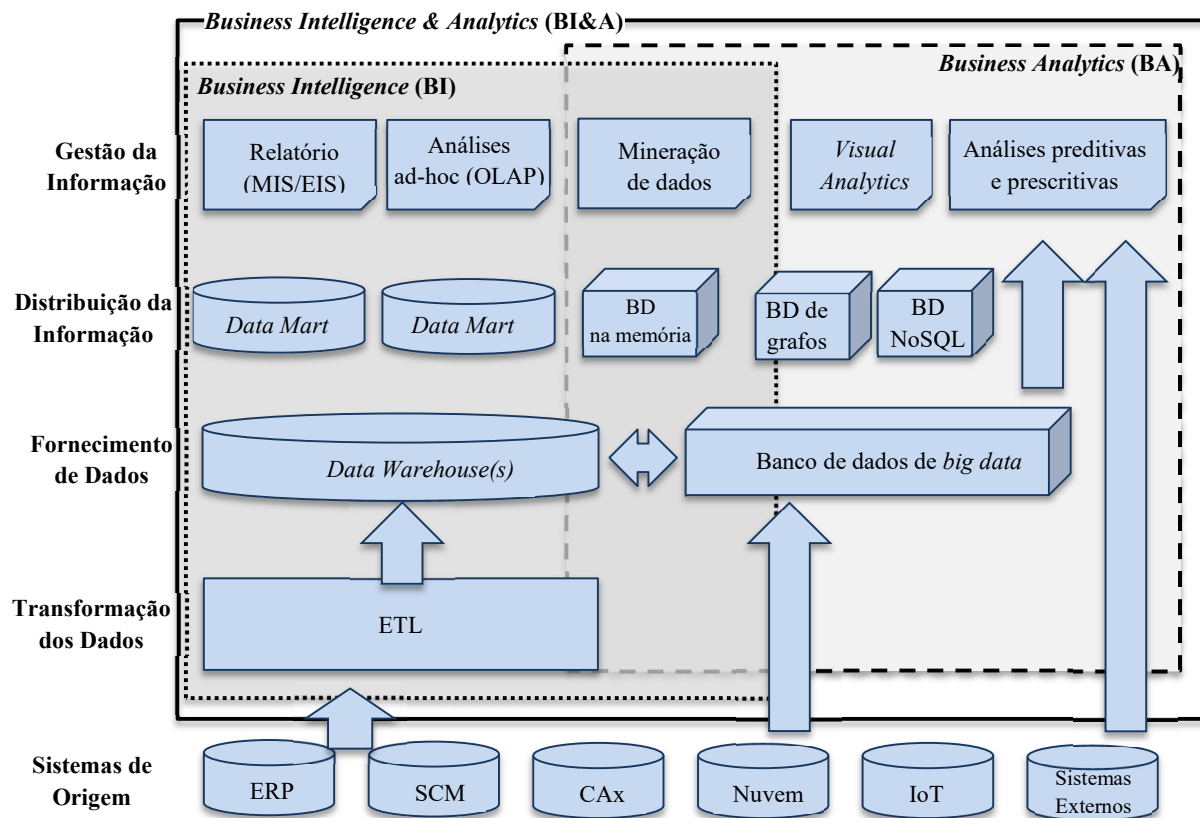
Business Intelligence (BI) pode ser definido como o conjunto de ferramentas que combinam dados operacionais com ferramentas de análise para apresentação de informações para a tomada de decisão (Negash, 2004). O BI compreende o conjunto de ferramentas utilizadas desde o processo de ETL (*Extract, Treat and Load*), que extraem os dados nos sistemas de origem, tratam os dados e carregam em um repositório, passando pelas ferramentas de *Data Warehouse*, que compreende o armazenamento dos dados tratados, ferramentas de *Data Mart*, onde os dados são separados por assuntos de interesse, e por último as ferramentas de gestão da informação, com a construção de relatórios analíticos e análises *ad-hoc* (Ereth e Kemper, 2016).

O conceito de BI se apoia principalmente no *Data Warehouse*, que armazena todos os dados importantes, sejam eles internos ou externos (Lennerholt et al., 2018). O *Data Warehouse* (DW) é uma coleção de dados integrada, não volátil e variável ao longo do tempo, que suporta a tomada de decisão, tendo como principal característica a capacidade de converter, reformatar e resumir os dados de diversas fontes, padronizando o armazenamento e consumo desses dados (Inmon, 2002).

Um conceito posterior ao BI é o *Business Analytics* (BA), que por sua vez trata-se de metodologias e ferramentas que utilizam estatística e análises quantitativas, modelos exploratórios e preditivos para suportar a tomada de decisão (Cao et al., 2015). Assim, BA, diferente do BI, utiliza-se de ferramentas para armazenamento de grandes volumes de dados (*big data*), passando por distribuição dos dados em bancos de dados não relacionais (NoSQL) e de grafos, utilizado quando bancos de dados relacionais não atendem à finalidade, e por fim, ferramentas de análises preditivas e prescritivas de dados. O *framework* apresentado por Ereth e Kemper (2016) exemplifica os conceitos de BI e BA (Figura 2).

Figura 2

Cenário tecnológico de BI&A



Fonte: Ereth, J., Kemper, H.G. (2016). Business Analytics und Business Intelligence. Controlling. 28. 458-464. (DOI: 10.15358/0935-0381-2016-8-9-458).

Nota: ERP (Sistema Integrado de Gestão Empresarial), SCM (Gestão da Cadeia de Suprimentos), CAx (Conjunto de sistemas fabris), IoT (Internet das Coisas).

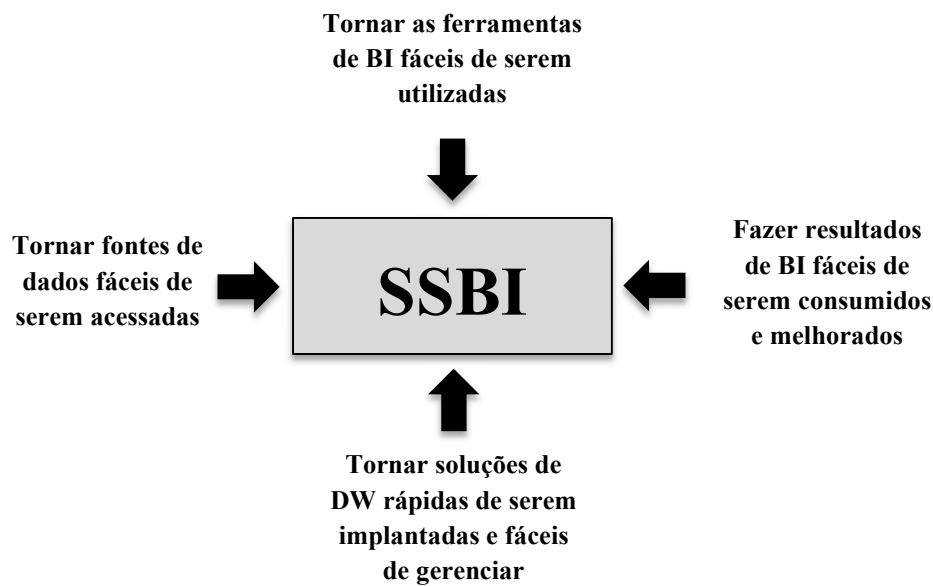
Assim, pode-se entender o conjunto de tecnologias, práticas e metodologias com uso de ferramentas estatísticas e análise de dados para apresentação de informações que suportam análises de negócio, mercado e uma melhor tomada de decisão como o conceito de *Business Intelligence e Analytics* (BI&A) (Chen et al. 2012). Muitas organizações já adotam o modelo de BI&A em suas tomadas de decisão e, por esse motivo, foram desenvolvidas ferramentas no mercado para atenderem às diversas necessidades. As ferramentas de BI&A abrangem análises desde questões estratégicas a tarefas operacionais, aumentando a necessidade de utilização nas organizações (Böhringer et al., 2010). Este desenvolvimento tem criado mais demanda por relatórios de BI e padrões de busca de dados (Alpar e Schulz, 2016).

Com a crescente utilização de ferramentas de BI&A e com a baixa necessidade de conhecimentos técnicos e estatísticos para a construção de análises de BI, surgiram ferramentas que implementaram o conceito de *Self-service BI* (SSBI), principalmente vinculadas à etapa de gestão da informação. A principal ideia do SSBI é o empoderamento de usuários casuais, também referenciados como *Citizen Developers*, para criação dos próprios relatórios em

sistemas de BI de fácil utilização, sem a necessidade de longas tratativas com os departamentos de TI (Michalczyk et al., 2020).

A tecnologia de SSBI se apoia no conceito tecnológico de “faça você mesmo” (do inglês, *Do It Yourself – DIY*), e apresenta 4 objetivos principais a serem atingidos com sua utilização (Imhoff e White, 2011) ilustrados na Figura 3.

Figura 3
Objetivos do SSBI



Fonte: Imhoff, C., White, C. (2011). Self-service Business Intelligence: Empowering Users to Generate Insights. TDWI Best practices report.

Assim, as seguintes hipóteses referentes aos objetivos do uso da tecnologia de SSBI podem ser formuladas:

- H1: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao consumo dos relatórios;
- H2: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à facilidade de uso destas ferramentas;
- H3: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à velocidade de implementação de relatórios;
- H4: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao acesso às fontes de dados.

Michalczyk et al. (2020) trazem ainda três perspectivas que a literatura aborda na discussão de SSBI: visão centrada nos artefatos, visão centrada no usuário e visão centrada na governança.

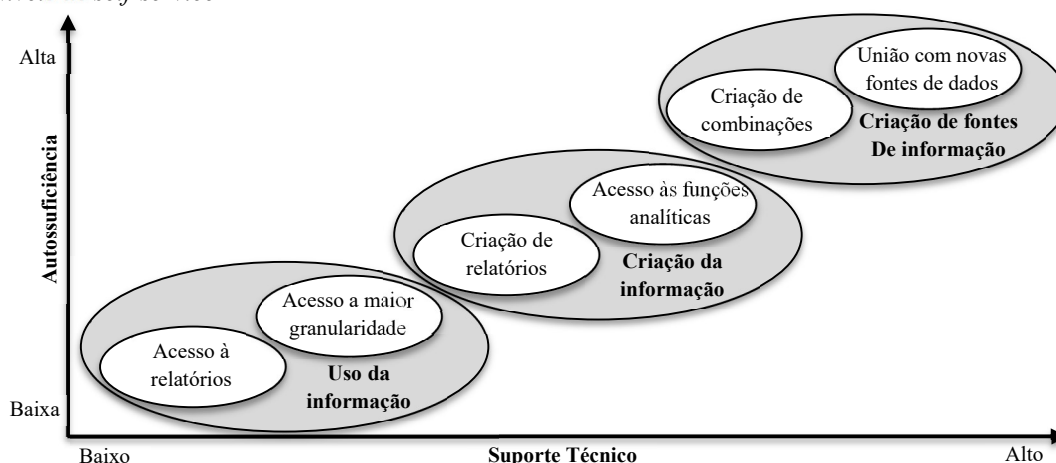
A visão centrada nos artefatos se refere a uma visão com maior foco nas fontes geradoras dos dados. O objetivo é a interação contínua dos usuários com as aplicações para a busca e integração dos dados situacionais, sem a necessidade de atuação dos programadores (Abelló et al., 2013). Portanto, a lógica é apoiar os utilizadores na tomada de decisões bem-informadas, ajudando-os a navegar pelos dados através de soluções técnicas (Michalczyk et al., 2020).

Na visão centrada nos usuários, os autores sugerem uma visão com maior foco na facilitação da usabilidade e construção de ferramentas de SSBI para os usuários, a fim de aumentar a aceitabilidade dos sistemas de SSBI. Imhoff e White (2011) sugerem, para tal fim, quatro objetivos a serem alcançados: facilidade de acesso às fontes de dados, melhor suporte às ferramentas de análise de dados, fácil implantação com uso da computação em nuvem e interface simples e customizável.

Por último, a visão centrada na governança apresenta uma perspectiva com maior enfoque na governança e na segurança dos dados, sem tirar a autonomia e velocidade sugerida pelo SSBI. Para tal, ao invés da utilização de *Data Lakes*, que são repositórios centralizados para armazenamento de dados não tratados, em seu formato original (Nambiar e Mundra, 2022), devem ser utilizados *Data Warehouses*, que são repositórios centralizados que armazenam os dados tratados, incluindo vínculo entre diversas fontes, de forma bem estruturada e prontos para o consumo (Chandra e Gupta, 2018). Essa adoção visa a mitigação de erros no processo de tomada de decisão, garantindo a qualidade dos dados, e que esses estão sendo extraídos da fonte correta. Além disso, a adoção de conceitos de metadados e governança de dados tem papel fundamental para a preparação dos dados para o SSBI (Michalczyk et al., 2020). Assim, pela visão centrada na governança, Imhoff e White (2011) propõem um sistema de classificação para definição da qualidade dos relatórios ou de seus componentes. Esta ação permite a curadoria dos relatórios para toda a organização (Michalczyk et al., 2020).

O conceito de *self-service* pode ser implementado em diferentes etapas, que se dispõem nos níveis de *self-service* para uso dos dados nas organizações: acesso a relatórios preparados ou fontes de dados, acesso direto aos dados, acesso às funções ou a criação de novas fontes de dados (Alpar e Schulz, 2016).

Figura 4
Níveis de self-service



Fonte: Alpar, P., Schulz, M., *Self-service Business Intelligence*. Business Information Systems Engineering. 58. 151-155.

Os níveis ilustrados na Figura 4 são assim definidos por Alpar e Schultz (2016):

- **Acesso à relatórios:** os usuários possuem acesso a informações que já existem (relatórios existentes) ou em locais onde eles apenas irão definir parâmetros e processar as informações. Neste caso, os usuários terão acesso a todos os relatórios que podem ser relevantes a eles;
- **Acesso a maior granularidade:** os usuários possuem acesso a relatórios por meio dos quais podem iniciar suas análises em um nível mais alto da informação, de forma agregada, e podem descer a um menor nível de granularidade da informação, por caminhos pré-definidos nos quais possuem informações mais detalhadas;
- **Criação de relatórios:** os usuários possuem acesso aos dados na maior granularidade de forma desagregada para a criação de novos relatórios. Esse nível possui um risco de agregações e associações incorretas dos dados, visto que os usuários casuais podem ter menor entendimento do relacionamento dos dados;
- **Acesso às funções analíticas:** além do acesso para criação de relatórios, os usuários possuem acesso para utilização de ferramentas de *advanced analytics*, podendo fazer análises de predição, por exemplo;
- **Criação de combinações:** os usuários podem combinar diferentes funcionalidades reutilizando componentes já preparados pelos profissionais de TI, para criação de novos componentes combinados;
- **União com novas fontes de dados:** usuários possuem acesso para utilização de dados de novas fontes de dados, unindo aos dados de fontes já existentes e utilizadas, mesmo que esses dados não tenham sido pré-processados pela TI.

O aumento do nível de *self-service* para utilização dos dados na organização por usuários casuais provê maior autonomia a esses usuários, aumentando a possibilidade de experimentações na organização, porém apresenta alguns riscos. Alpar e Schulz (2016) apresentam um exemplo em que uma organização compilou as informações necessárias para a tomada de decisão provenientes de 26.000 relatórios criados após alguns anos utilizando SSBI em 300 relatórios padrões. O exemplo ilustra que a utilização de funcionalidades analíticas para um alto número de usuários não necessariamente é o cenário desejado para obtenção das melhores informações das melhores formas (Alpar e Schulz, 2016).

Além disso, alguns desafios podem ser encontrados na implantação do SSBI nas organizações (Lenerholt et al., 2018):

- **Acesso e uso dos dados**
 - **Tornar as fontes de dados fácil de acessar e usar:** os usuários precisam ter livre acesso aos dados, para geração de novas análises e modificação de relatórios;
 - **Identificar critérios de seleção dos dados:** os usuários precisam poder selecionar os dados a serem utilizados com base em sua qualidade, definida por critérios específicos, como acurácia, atualização, completude, entre outros;
 - **Usar busca de dados corretas:** os usuários precisam ter acesso a metadados a fim de identificar suas características e usos para evitar erros na busca e junção de dados, o que pode levar a tomadas de decisão incorretas com base nos relatórios;
 - **Controlar da integridade, segurança e distribuição dos dados:** ao passo que mais dados externos se tornam disponíveis, a importância sobre a integridade e segurança dos dados aumenta. O desafio de SSBI é controlar quem pode usar os dados, garantir o acesso adequado e a qualidade dos dados;
 - **Definir políticas para gerenciamento e governança dos dados:** se organizações implementam ferramentas de SSBI sem implementação de uma governança de dados apropriada, usuários podem errar na utilização dos dados, não entender seu valor e criar estruturas de *shadow IT*, o que pode levar a problemas de acesso e uso dos dados corretamente;
 - **Preparar dados para análises visuais:** ferramentas de análise visual de dados são importantes ferramentas pois podem melhorar a produtividade.
- **Usuários independentes**

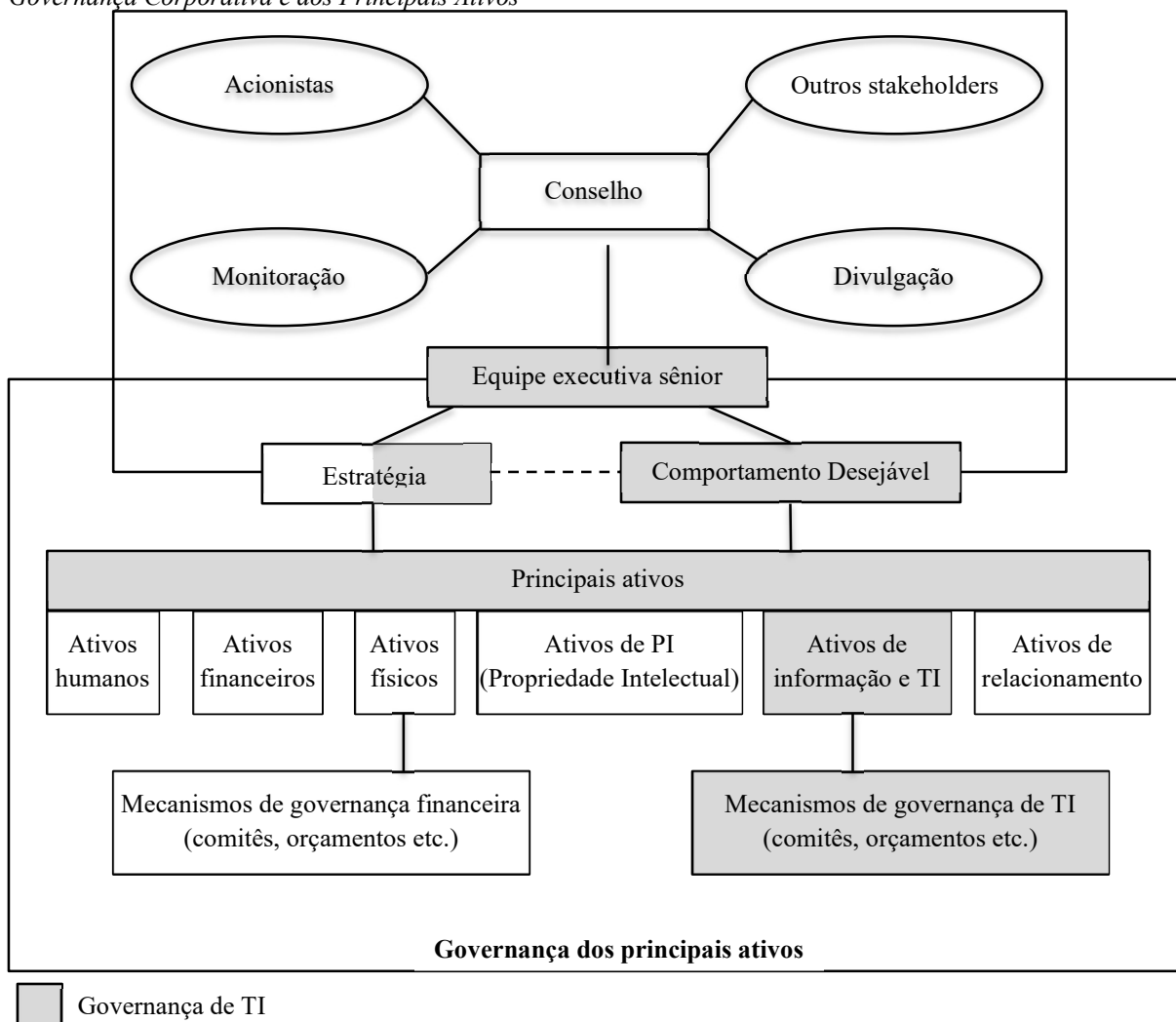
- **Tornar ferramentas de BI fáceis de utilizar:** se as ferramentas de BI não forem fáceis de utilizar, os usuários tendem a “criar” suas próprias ferramentas, o que pode comprometer a correta utilização dos dados;
- **Tornar resultados de BI fáceis de consumir e analisar:** ferramentas de BI precisam ser fáceis de utilizar e devem apresentar os resultados no formato correto e nos dispositivos e interface desejáveis pelos usuários;
- **Dar as ferramentas corretas aos usuários corretos:** diferentes usuários possuem habilidades diferentes e necessidades diferentes, exigindo assim, ferramentas diferentes dependendo das tarefas a serem realizadas;
- **Educar os usuários em como selecionar, interpretar e analisar os dados para a tomada de decisão:** os usuários precisam ser treinados em como utilizar corretamente as ferramentas e validar seus resultados quando da utilização dos dados.

Dessa forma, a importância da governança de dados e da segurança aumentam à medida que são utilizados dados nas organizações. Idealmente, sistemas centralizados e ferramentas de direitos de acessos podem ser utilizados enquanto estruturas de dados temporárias ou permanentes são criadas. Caso contrário, SSBI pode possibilitar falhas na segurança da informação das organizações (Alpar e Schulz, 2016).

2.2. Governança de Dados

Em um contexto mais amplo, governança corporativa é definida como a forma como uma organização define, monitora e alcança seus objetivos estratégicos (Rau, 2004), podendo ser aplicável para diversas camadas e ativos da organização (Figura 5).

Figura 5
Governança Corporativa e dos Principais Ativos



Fonte: Weill, P., Ross, J. (2004). IT Governance: How Top Performers Manage IT Decision Rights for Superior Results.

O espectro da governança abarca desde a governança corporativa, que define uma estrutura para determinação dos objetivos organizacionais, as políticas que embasam esses objetivos e o monitoramento para garantir seu atingimento (Weill e Ross, 2004), até a governança de áreas e temáticas específicas da organização, como a governança de tecnologia da informação (TI) destacada em cinza na Figura 5 e a governança de dados.

A governança de dados começou a ser pesquisada nos anos 1980, tendo ainda atualmente boa parte das pesquisas existentes focadas em algumas áreas específicas. A primeira tentativa de criação de um *framework* para a governança de dados foi publicada em 2007 por Kristin Wende em uma publicação na Conferência Australiana de Sistemas da Informação (ACIS) (Niemi, 2011).

Weber et al. (2009, p. 46) definem que a governança de dados “especifica a estrutura para os direitos de decisão e responsabilização para incentivar o comportamento desejável no

uso de dados, desenvolvendo e implementando políticas, diretrizes e padrões de dados em toda a empresa que sejam consistentes com a missão da organização, estratégia, valores, normas e cultura”. Por sua vez, Panian (2010, p. 939) define governança de dados como “os processos, políticas, padrões, organização e tecnologias necessárias para gerenciar e garantir a disponibilidade, acessibilidade, qualidade, consistência, auditabilidade e segurança dos dados em uma organização”.

Outra definição para a governança de dados é dada por Abraham et al. (2019, p. 8) como “especificar uma estrutura multifuncional para o gerenciamento de dados como um ativo empresarial estratégico. Ao fazê-lo, a governança de dados especifica direitos de decisão e responsabilização para a tomada de decisões de uma organização sobre seus dados. Além disso, a governança de dados formaliza políticas, padrões e procedimentos de dados e monitora a conformidade”.

Ao longo do tempo, algumas instituições foram constituídas para a difusão de conceitos de dados e a tentativa de criação de padrões e boas práticas. Uma destas instituições é o Instituto de Governança de Dados (DGI – *Data Governance Institute*), criado em 2003 nos Estados Unidos por Gwen Thomas, que tem como principal função a difusão do *framework* de Governança de Dados do DGI e boas práticas relacionadas à governança de dados. O DGI (ano 2015) define a governança de dados como “um sistema de direitos de decisão e responsabilização para processos relacionados à informação, executado de acordo com modelos acordados que descrevem quem pode tomar quais ações com quais informações e quando, sob que circunstâncias e usando quais métodos”.

Outra instituição é a Associação de Gestão de Dados (DAMA – *Data Management Association*), associação sem fins lucrativos criada em 1980 em Los Angeles, com a finalidade de avançar conceitos e práticas de gerenciamento de informações e dados. O DAMA (2007, p. 67) define a governança de dados como “o exercício de autoridade, controle e tomada de decisão compartilhada (planejamento, monitoramento e aplicação) sobre o gerenciamento de ativos de dados”.

Dessa forma, mesmo sem uma definição uniforme sobre governança de dados, alguns pontos em comum podem ser vistos nas definições encontradas na literatura, resultando na proposição de uma definição consolidada: a governança de dados se destina à definição de papéis e responsabilidades e a atribuição de autoridade, direitos a tomada de decisão e responsabilização sobre a gestão e uso de dados, por meio do estabelecimento de procedimentos, normas e políticas.

Alguns outros conceitos podem ser confundidos com a governança de dados e, dessa forma, se torna importante a diferenciação entre eles. Primeiramente, é importante a diferenciação entre a governança de dados e a governança de TI. A governança da TI está associada à definição do processo de tomada de decisão e responsabilizações para estimular comportamentos desejáveis no uso da TI (Weill e Ross, 2004). A governança de TI foca, ainda, na definição de processos e mecanismos relacionados aos ativos de TI, bem como a relação de critérios regulatórios envolvidos (Van Grembergen e De Haes, 2010). Assim, pode-se entender que a governança de TI tem foco em governar os ativos de tecnologia e digital como um todo, ao passo que a governança de dados se descola, tendo foco em governar os dados da organização como um ativo.

A governança de dados e a governança de TI podem ser vistas como um subdomínio da governança corporativa, visto que compartilham das políticas e definições que ela define (Figura 6). Porém, a governança de dados se apresenta hoje com certa independência da governança de TI, ao passo que os dados têm se tornado um ativo da organização e não apenas propriedade da tecnologia. Os dados são custodiados e habilitam as ferramentas de TI, mas tem um caráter independente, sendo propriedade da organização de forma geral (Barbieri, 2019).

Figura 6

Relação entre governança: governança corporativa, governança de TI e governança de dados



Fonte: Barbieri, Carlos (2019). Governança de Dados: Prática Conceitos e Novos Caminhos. (1ª ed.). Rio de Janeiro: Atlas, 2019.

Dessa forma, a governança corporativa estabelece a estratégia e o comportamento desejável de forma geral na organização, sendo que estas definições são permeadas por todos seus ativos, desdobradas por meio das governanças específicas. Os dados, entendidos como ativo organizacional, são abarcados de forma complementar pela governança de dados, definindo a estratégia e o comportamento desejável de parâmetros específicos do ativo.

Além disso, há ainda a distinção entre a governança de dados e a governança da informação. Primeiramente, deve-se entender que a informação consiste no resultado gerado pelo processamento de dados, objetivando responder a perguntas como quem, o que, quando, onde e quanto (Ackoff, 1989). Assim, a governança da informação visa a definição de padrões, procedimentos e políticas a fim de proporcionar o uso efetivo dos recursos e capacidades dos ativos da informação (Gianella e Gujer, 2006). O principal objetivo da governança da informação é auxiliar a organização a entender como as informações podem ser úteis para o negócio e qual informação consegue abarcar o contexto da organização e seu ecossistema (Beijer e Kooper, 2010). Dessa forma, a governança da informação terá como foco governar as informações, seu uso e o valor a ser gerado, ao passo que a governança de dados terá como foco governar os dados, desde sua origem até seu uso final, não abrangendo a quais perguntas a informação gerada responde.

Outra importante distinção se dá entre a governança de dados e o gerenciamento de dados. Enquanto a governança de dados se concentra na definição das políticas, procedimentos e normas para regulamentar os dados na organização, o gerenciamento de dados se concentra na execução dessas definições e na utilização de ferramentas que suportem os procedimentos definidos (DAMA, 2017).

Além disso, é importante avaliar a crescente necessidade de controle sobre o armazenamento de dados, identificando quais dados são armazenados, onde são armazenados, como e por quem são utilizados, tendo em vista a implementação de requerimentos regulatórios, no caso brasileiro pela Lei Geral de Proteção de Dados (LGPD). Esse fator é agravado pela maior adoção de relatórios *self-service* pelas organizações, o que amplia a utilização dos dados, mas exige maior controle (Abraham et al., 2019).

Por outro lado, deve-se atentar para as restrições abarcadas pela implantação de uma governança de dados madura para que essas não distanciem as áreas de negócio, visto que, principalmente com o crescimento do modelo *self-service* para produtos tecnológicos, as áreas de negócio precisam estar envolvidas no processo de governança. O excesso de limitações em um modelo de governança de dados faz com que os usuários tentem contornar as políticas, tomando decisões para as quais não estão habilitados tecnicamente e assumindo riscos desnecessariamente. Assim, pode-se assumir uma relação curvilínea entre a governança de dados e o desempenho organizacional, pois a governança em excesso pode comprometer o desempenho da organização (Tallon et al., 2004).

Poucos são os *frameworks* desenvolvidos por artigos científicos, sendo grande parte aqueles publicados por empresas de consultoria ou instituições de governança e gerenciamento de dados. A Tabela 1 apresenta as referências dos *frameworks* encontrados.

Tabela 1

Frameworks de governança de dados

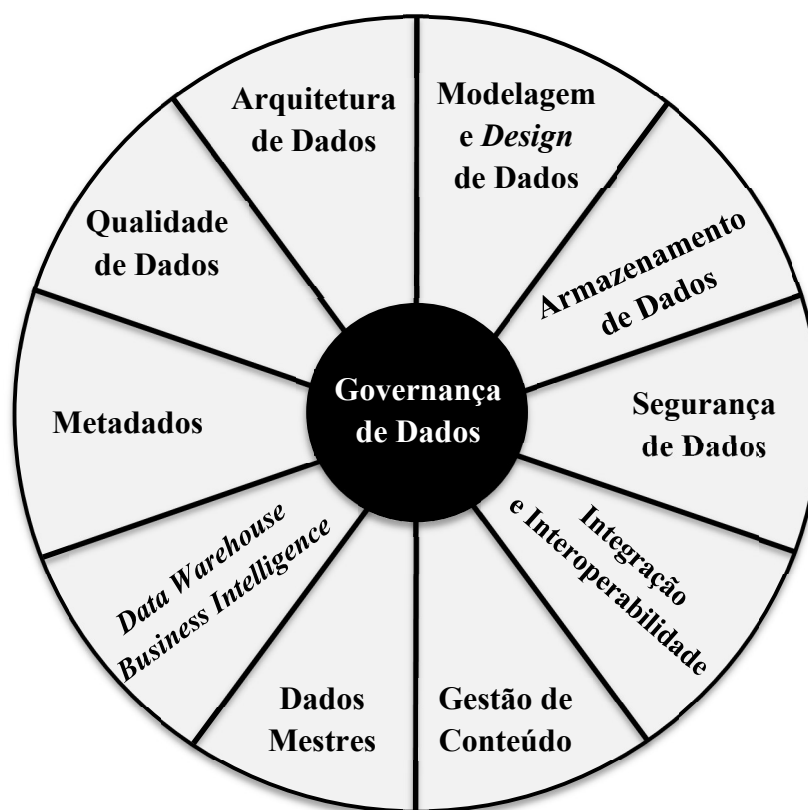
Tipo	Referência
Artigo Acadêmico	Abraham, Brocke & Schneider (2019).
Artigo Acadêmico	Khatri & Brown (2010).
Artigo Acadêmico	Panian (2010).
Consultoria	Baltassis, Gourévitch & Quarta (2019) pelo Boston Consulting Group.
Consultoria	Wells (2019) pelo Eckerson Group
Consultoria	Petzold, Roggendorf, Rowshankish & Sporleder (2020) pela Mckinsey.
Consultoria	Oracle (2011).
Consultoria	PricewaterhouseCoopers (2019).
Consultoria	SAS Institute (2018).
Instituição	DAMA International (2017).
Instituição	The Data Governance Institute (2015).

Fonte: Dados da pesquisa.

Observa-se a utilização do modelo conceitual e das definições apresentadas pelo DAMA no DMBok (*Data Management Body of Knowledge*) como base para os modelos acadêmicos apresentados. Isso se deve à abrangência e relevância do material e, diferente do DGI, os conceitos e modelos de governança de dados apresentados pelo DAMA fazem parte de um conjunto de disciplinas de boas práticas que suportam a adoção de um modelo de governança de dados completo. Além disso, o modelo do DAMA tem natureza generalista, diferentemente dos modelos apresentados pelas consultorias, que se adequam a práticas específicas atendidas por elas. Dessa forma, o presente trabalho tem como base o *framework* apresentado no DMBok do DAMA (2017), complementando, quando necessário, os conceitos com as informações apresentadas pelos demais materiais.

São apresentadas no DMBok boas práticas para o gerenciamento de dados em uma organização, divididas em 11 disciplinas. A primeira disciplina apresentada é a governança de dados (Figura 7).

Figura 7
DAMA-DMBOK2 Data Management Framework



Fonte: DAMA International. (2017). DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition). Technics Publications, LLC, Denville, NJ, USA.

A governança de dados visa, principalmente, a garantia da gestão dos dados de acordo com políticas e boas práticas. Para tal, a depender de escopos particulares e necessidades organizacionais, a maior parte dos programas de governança de dados incluem (DAMA, 2017):

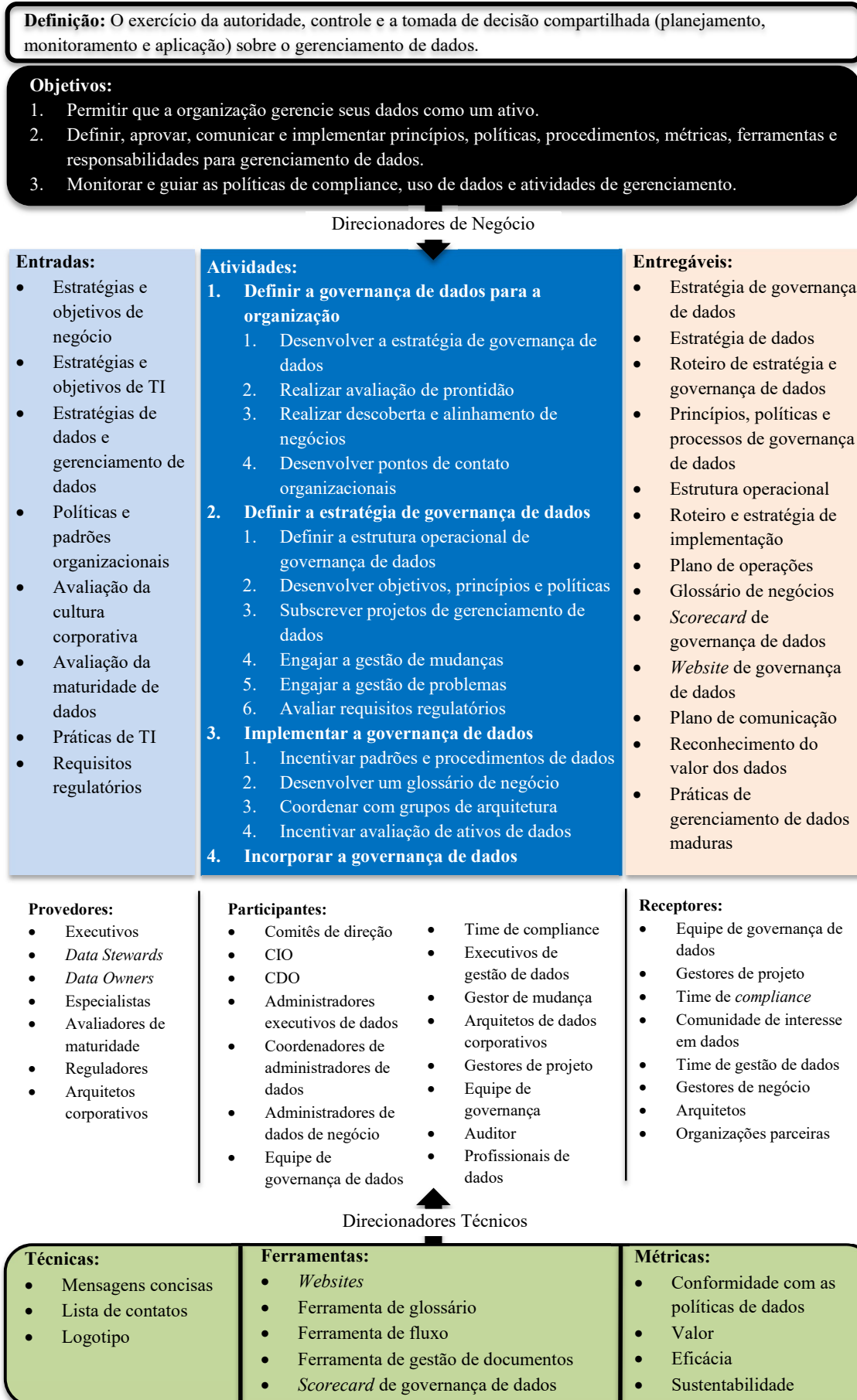
- **Estratégia:** definir, comunicar e executar a estratégia de dados e a estratégia de governança de dados;
- **Política:** definir e garantir políticas relacionadas ao gerenciamento de dados e metadados, acesso, uso, segurança e qualidade;
- **Padrões e qualidade:** definir e garantir padrões de arquitetura de dados e qualidade de dados;
- **Supervisão:** fornecer observação prática, auditoria e correção em áreas-chave de qualidade, política e gerenciamento de dados;
- **Compliance:** garantir que a organização esteja adequada a requisitos regulatórios relacionados aos dados;
- **Gerenciamento de problemas:** identificar, definir, escalar e resolver problemas relacionados à segurança de dados, acesso a dados, qualidade de dados, conformidade

regulatória, propriedade de dados, políticas, padrões, terminologia e procedimentos de governança de dados;

- **Projetos de gerenciamento de dados:** patrocinar esforços para melhoria das práticas de gerenciamento de dados;
- **Valoração dos ativos de dados:** definir padrões e processos para definir consistentemente o valor comercial dos ativos de dados.

Dessa forma, o DAMA (2017) define um *framework* para adoção de uma governança de dados madura, definindo as atividades necessárias para execução dessa governança (Figura 8).

Figura 8
Governança e Administração de Dados



Fonte: DAMA International. (2017). DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition). Technics Publications, LLC, Denville, NJ, USA.

Os impulsionadores de negócio para a adoção da governança de dados, e consequentemente os objetivos para avaliação do sucesso em sua adoção são, de acordo com DAMA (2017):

- **Gerenciamento geral dos riscos:** supervisão dos riscos que os dados representam para as finanças ou a reputação, incluindo resposta a questões legais e regulatórias;
- **Segurança dos dados:** proteção dos dados através de controles para a disponibilidade, usabilidade, integridade, consistência, auditabilidade e segurança dos dados;
- **Privacidade:** controle da privacidade de dados pessoais através de políticas e monitoramento de compliance;
- **Conformidade regulatória:** a capacidade de responder de forma eficiente e consistente a requisitos regulatórios;
- **Melhoria da qualidade de dados:** a capacidade de contribuir para a melhoria da performance de negócio tornando os dados mais confiáveis;
- **Gerenciamento de metadados:** estabelecimento de um glossário de negócios para definir e localizar dados em uma organização;
- **Eficiência no desenvolvimento de projetos:** melhorias no ciclo de vida do desenvolvimento de softwares para abordar problemas e oportunidades no gerenciamento de dados em toda a organização, incluindo o gerenciamento de dívida técnica específica de dados por meio da governança do ciclo de vida dos dados;
- **Gestão de fornecedores:** controle de contratos que tratam de dados, como armazenamento em nuvem, compra de dados externos, venda de dados como um produto e terceirização de operações de dados.

Diante do exposto, formulam-se outras quatro hipóteses da pesquisa, que serão detalhadas mais à frente:

- H5: A aplicação da governança de dados gera uma melhor percepção relativa à segurança e privacidade dos dados na utilização do SSBI;
- H6: A aplicação da governança de dados gera uma melhor percepção relativa à qualidade dos dados na utilização do SSBI;
- H7: A aplicação da governança de dados gera uma melhor percepção relativa à gestão de metadados na utilização do SSBI;
- H8: A aplicação da governança de dados gera uma melhor percepção relativa à eficiência no desenvolvimento na utilização do SSBI.

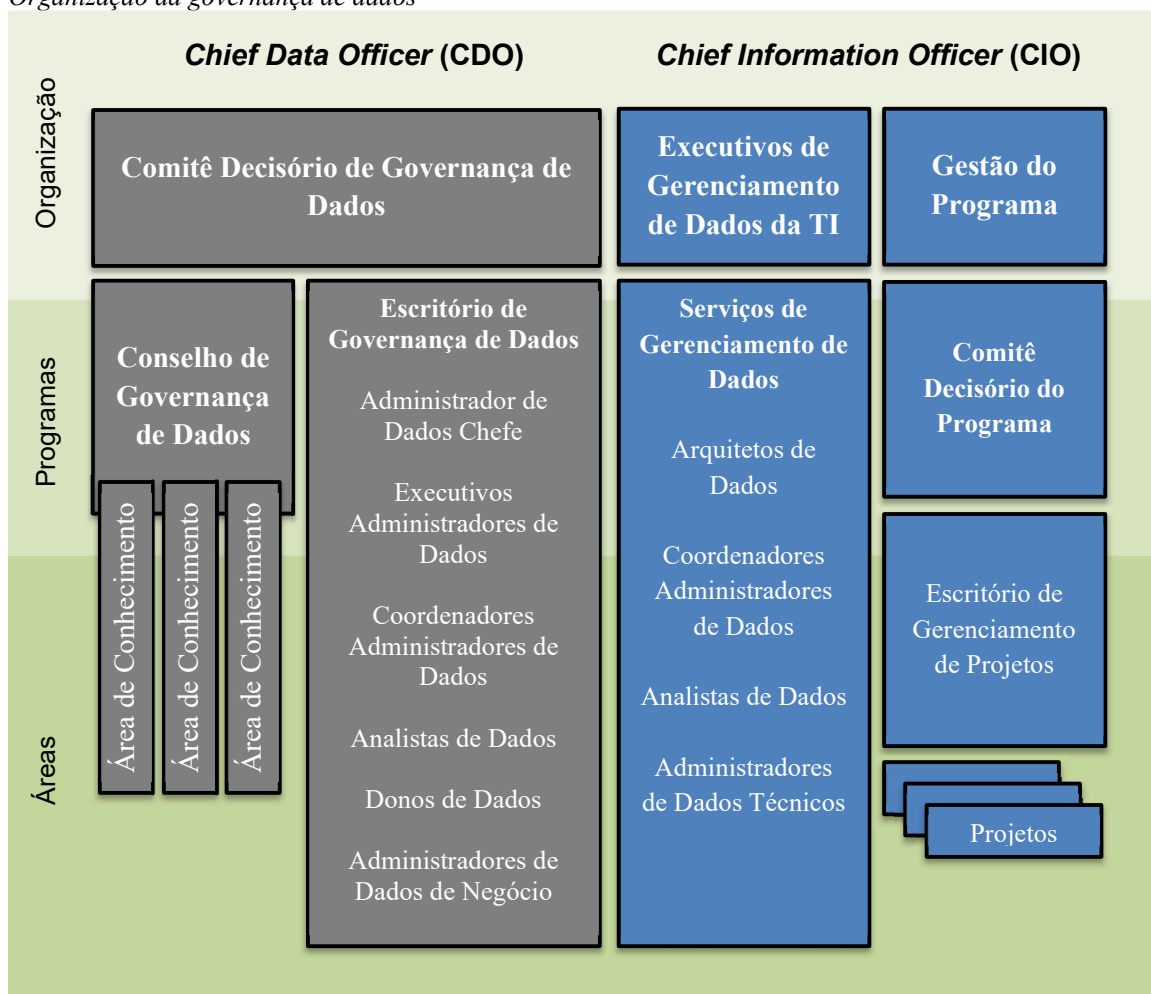
As atividades propostas no *framework* para a implantação de uma governança de dados em uma organização são:

- **Definir a governança de dados para a organização**
 - **Desenvolver a estratégia de governança de dados:** a estratégia de governança de dados define o âmbito e a abordagem dos esforços de governança, e deve ser definida de forma abrangente e articulada em relação à estratégia empresarial global, bem como à gestão de dados e às estratégias de TI;
 - **Realizar a avaliação de prontidão:** avaliação que descreve o estado atual das capacidades, maturidade e efetividade da gestão de dados da organização. A avaliação também é importante para medir a eficácia do programa de implementação da governança de dados;
 - **Realizar a descoberta e alinhamento de negócios:** identificação e avaliação da eficácia das políticas e diretrizes existentes – quais riscos abordam, quais comportamentos encorajam e quão bem foram implementadas. Também pode-se identificar oportunidades para a governança de dados melhorar a utilidade dos dados;
 - **Desenvolver pontos de contato organizacionais:** avaliação dos pontos de contato organizacionais para o trabalho de governança de dados.
- **Definir a estratégia de governança de dados**
 - **Definir a estrutura operacional de governança de dados:** determina a responsabilização pelas atividades de administração dos dados bem como quem é o proprietário dos dados. O modelo operacional também deve definir a interação entre a governança organizacional e as pessoas responsáveis pelos projetos ou iniciativas de gestão de dados, o modelo de gestão de mudanças e o modelo de gestão de problemas por meio da governança;
 - **Desenvolver objetivos, princípios e políticas:** definição dos objetivos, princípios e políticas, derivadas da estratégia da governança de dados, que irão guiar a organização para o futuro desejado;
 - **Subscrever projetos de gerenciamento de dados:** as organizações que desejam obter mais valor dos seus dados devem priorizar o desenvolvimento ou a melhoria das capacidades de gerenciamento de dados;
 - **Engajar a gestão de mudanças:** para muitas organizações, a formalidade e a disciplina inerentes da governança de dados difere das práticas existentes. Adotar a governança de dados então, exige que as pessoas mudem seus comportamentos e interações. Um programa formal para gestão de mudança na organização é crítico para

impulsionar as mudanças comportamentais requeridas para sustentar a governança de dados;

- **Engajar a gestão de problemas:** processo para identificar, quantificar, priorizar e resolver problemas relacionados à governança de dados. Deve definir as regras de escalonamento dos problemas para os administradores de dados e o conselho de governança de dados;
- **Avaliar requisitos regulatórios:** toda organização é afetada por regulamentações governamentais, incluindo regulamentações que ditam como os dados devem ser geridos. Parte da função da governança de dados é de monitorar e garantir a conformidade com essas regulamentações;
- **Implementar a governança de dados**
 - **Incentivar padrões e procedimentos de dados:** estabelecimento de padrões e procedimentos para aqueles conceitos que são padronizáveis, como a arquitetura, a modelagem, o armazenamento, a segurança, a integração e a qualidade de dados, documentações, ferramentas de *data warehousing* e *Business Intelligence*, metadados, entre outros;
 - **Desenvolver um glossário de negócio:** o glossário de negócio se faz necessário, pois as organizações, normalmente, desenvolvem seu próprio vocabulário, e assim é necessário otimizar a comunicação. A criação e manutenção do glossário de negócios, normalmente, é responsabilidade dos Administradores de Dados;
 - **Coordenar com grupos de arquitetura:** o conselho de governança de dados deve auxiliar e aprovar artefatos de arquitetura, interagindo constantemente com o comitê de arquitetura corporativo;
 - **Incentivar a avaliação de ativos de dados:** a governança de dados deve canalizar os esforços e padronizar a forma de valorar os dados, como um ativo.
- **Incorporar a governança de dados:** um dos objetivos da governança de dados é incorporar uma série de processos relacionados à gestão de dados como um ativo.

As organizações que adotam a governança de dados podem definir ainda diferentes níveis estruturais para endereçar as tratativas relacionadas à governança. Dessa forma, faz-se necessária a definição de papéis e responsabilidades dentro de um modelo de governança de dados, conforme Figura 9 (DAMA, 2007).

Figura 9*Organização da governança de dados*

Fonte: DAMA International. (2017). DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition). Technics Publications, LLC, Denville, NJ, USA.

- **Comitê Decisório de Governança de Dados:** mais alta autoridade para governança de dados na organização, responsável por definir e supervisionar as atividades de governança de dados;
- **Conselho de Governança de Dados:** gerencia as iniciativas de governança de dados, como o desenvolvimento de políticas e métricas, tratativa de problemas e escalação;
- **Escritório de Governança de Dados:** foco contínuo nas definições e padrões de gerenciamento de dados em nível corporativo;
- **Administrador de Dados Chefe:** preside o escritório de governança de dados;
- **Executivo Administrador de Dados:** gerentes sêniores que servem ao conselho de governança de dados;
- **Coordenador Administrador de Dados:** lidera times de administradores de dados de negócio e técnicos em discussões entre equipes;

- **Dono de dados:** administradores de dados com autoridade para aprovação sobre decisões a respeito dos dados sobre seu domínio;
- **Administradores de Dados de Negócio:** profissionais das áreas de negócio com amplo conhecimento de determinados assuntos, responsável por um subconjunto de dados, trabalhando com outros profissionais para definir e controlar os dados;
- **Administradores de Dados Técnicos:** profissionais de TI responsáveis por uma das áreas de conhecimento, como especialistas de integração de dados, administradores de banco de dados, especialistas de *Business Intelligence*, analistas de qualidade de dados e administradores de metadados.

O DAMA (2007) citado por Caballero e Piattini (2023) ainda define uma mensuração em 6 níveis para o nível de maturidade na governança de dados para a mensuração das organizações:

- **Ausência de governança (nível 0):** não existem práticas de governança de dados ou um processo formal para gestão dos dados na organização;
- **Inicial (nível 1):** as principais disciplina dos dados são gerenciadas usando um conjunto limitado de ferramentas e com pouca ou nenhuma governança, sendo dependente de poucos especialistas;
- **Repetível (nível 2):** a organização apresenta maior supervisão sobre o gerenciamento dos dados, e possui ferramentas robustas e centralizadas e a definição de papéis e responsabilidades para suporte aos dados;
- **Definido (nível 3):** a organização possui processos de gerenciamento de dados escalados e institucionalizados como habilitadores organizacionais;
- **Gerenciado (nível 4):** possui aumentos mensuráveis na qualidade dos dados e nas capacidades ao longo da organização, com um modelo de governança centralizado e planejado;
- **Otimizado (nível 5):** a governança de dados está implementada e difundida, contando com automações de processos e gestão de mudança na organização. A organização foca na melhoria contínua, ao invés da implementação de conceitos de governança. Além disso, métricas são utilizadas para gerenciar e medir a qualidade dos dados e de processos.

Pode-se então definir 4 fatores objetivos da implementação de uma governança de dados:

- **Segurança e privacidade:** os dados precisam estar disponíveis e corretos para os usuários autorizados, e apenas aos usuários autorizados (DAMA, 2007). Além disso, é necessário atender a requisitos regulatórios de privacidade e segurança de dados. Pode ser avaliado por meio de 5 conceitos (DAMA, 2007):
 - *Stakeholders*: organizações precisam reconhecer a necessidade por privacidade e confidencialidade de seus *stakeholders* (partes interessadas), como clientes, fornecedores, acionistas, entre outros, e garantir esta privacidade;
 - Regulamentações: garantir as restrições de acesso e a proteção das informações exigidas por regulamentações governamentais, como é o caso da LGPD no Brasil;
 - Propriedade de negócios: organizações possuem dados próprios de negócios que auxiliam na obtenção de vantagens competitivas. Estes dados precisam ser protegidos evitando o vazamento, o que pode comprometer as vantagens da organização;
 - Acesso legitimado: garantir a concessão dos acessos aos dados necessários a quem de fato precisa do acesso àqueles dados para a realização do trabalho;
 - Obrigações contratuais: garantir a segurança de dados prevista em contratos e acordos de acordos de não divulgação (NDA – *Non-Disclosure Agreements*).
- **Qualidade dos dados:** pode ser avaliado por meio de 4 conceitos (Khatri e Brown, 2010):
 - Acurácia: se refere à conformidade dos valores dos dados armazenados em relação ao valor real, respeitando o objetivo de uso;
 - Pontualidade: indica a atualização do dado quando requisitado;
 - Completude: sugere que os valores necessários estão armazenados e que estão na granularidade necessária;
 - Credibilidade: indica a confiabilidade das fontes de dados, bem como seus conteúdos.
- **Gerenciamento de metadados:** metadado é definido como os “dados sobre os dados”, ou seja, descreve as características e descrições do dado (Khatri e Brown, 2010). A governança de dados deve prover as melhores formas de capturar, armazenar, gerenciar e disponibilizar os metadados (SAS Institute, 2018). Os metadados são avaliados por 4 conceitos (DAMA, 2007):
 - Conhecimento organizacional: documentar e gerenciar o conhecimento organizacional relacionado às terminologias de negócio;
 - Integração entre fontes de dados: integrar metadados de diversas fontes para entendimento das similaridades e diferenças dos dados na organização;

- Qualidade dos metadados: garantir a qualidade, consistência e segurança dos metadados;
- Acesso aos metadados: prover acesso aos metadados para os usuários da organização;
- Troca de dados: possibilitar a troca de dados entre sistemas, devido ao entendimento dos metadados;
- **Eficiência no desenvolvimento:** auxilia no desenvolvimento de aplicações e novos produtos de dados, acelerando o acesso aos dados e garantindo sua correta disponibilidade. Pode ser avaliado por 2 fatores (Khatri e Brown, 2010):
 - Acesso aos dados: estabelece a forma de identificação das necessidades de dados da empresa e aborda salvaguardas para garantir a confidencialidade, integridade e disponibilidade de dados. O acesso aos dados se baseia na capacidade dos beneficiários dos dados de atribuir um valor a diferentes categorias de dados;
 - Ciclo de vida dos dados: com base em como os dados são utilizados e por quanto tempo os dados precisam ser armazenados, define como a organização pode desenvolver abordagens para mapear os padrões de uso para otimizar o armazenamento de dados, reduzindo custos. Entender como os dados se comportam ao longo do seu ciclo de vida é um ponto central para definição de uma governança de dados.

Por fim, para atingir esses objetivos, um programa de governança de dados deve ser (DAMA, 2017):

- **Sustentável:** a governança de dados não é um projeto com um fim definido, mas sim um processo contínuo que requer o comprometimento da organização. A governança de dados necessita de mudanças em como os dados são geridos e usados;
- **Incorporado:** a governança de dados não é um processo complementar. As atividades de governança de dados precisam ser incorporadas nos métodos de desenvolvimento de software, na utilização de dados para análise, na gestão de dados mestres e na gestão de riscos;
- **Mensurado:** uma governança de dados bem implementada tem impacto financeiro positivo na organização, mas demonstrar esse impacto necessita de entender o ponto inicial e planejar melhorias mensuráveis.

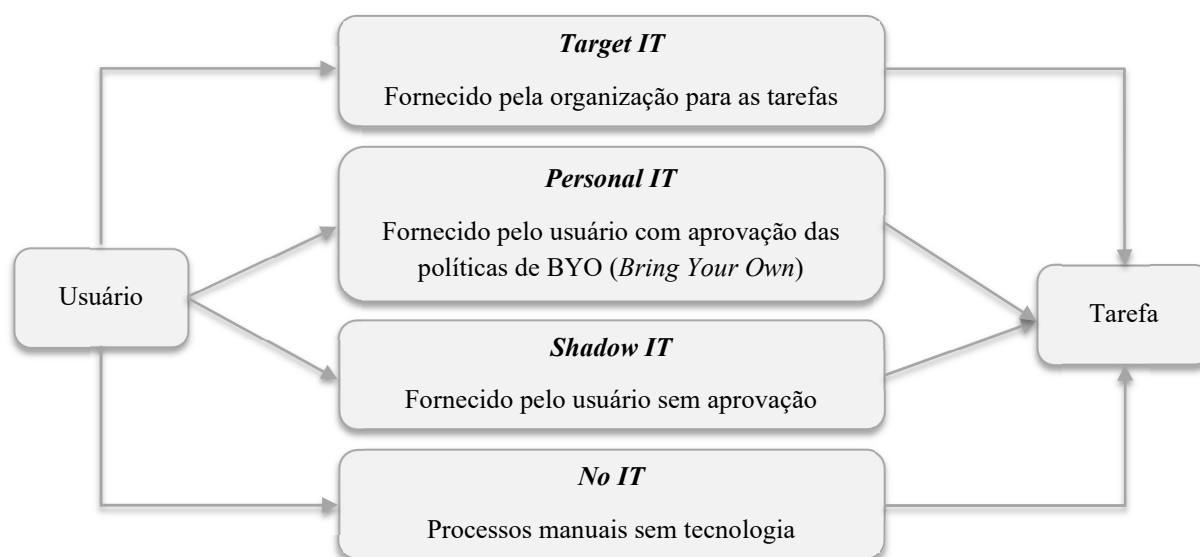
2.3. Síntese do Referencial Teórico

Com base nos conceitos apresentados, é estabelecida a importância no estabelecimento de políticas de governança de dados para organizações que implementam SSBI, a fim de definir

regras de acesso e padrões de qualidade para os dados. Sem esse estabelecimento de políticas, cria-se um cenário propício ao desenvolvimento de estruturas de *shadow IT*, podendo levar a problemas de segurança e privacidade dos dados (Lennerholt et al., 2018). A distinção das estruturas *shadow IT* para demais estruturas de TI pode ser observada na Figura 10, conforme apresentado por Haag e Eckhardt (2017). Além disso, o uso de SSBI sem uma governança de dados adequada pode levar ao desperdício de esforço trabalhando com dados incorretos. Assim, os programas de implantação e disseminação das tecnologias de SSBI também agem como um catalisador para a adoção de modelos de governança de dados nas organizações (Meyers, 2014).

Figura 10

Conceitos de interação com a TI



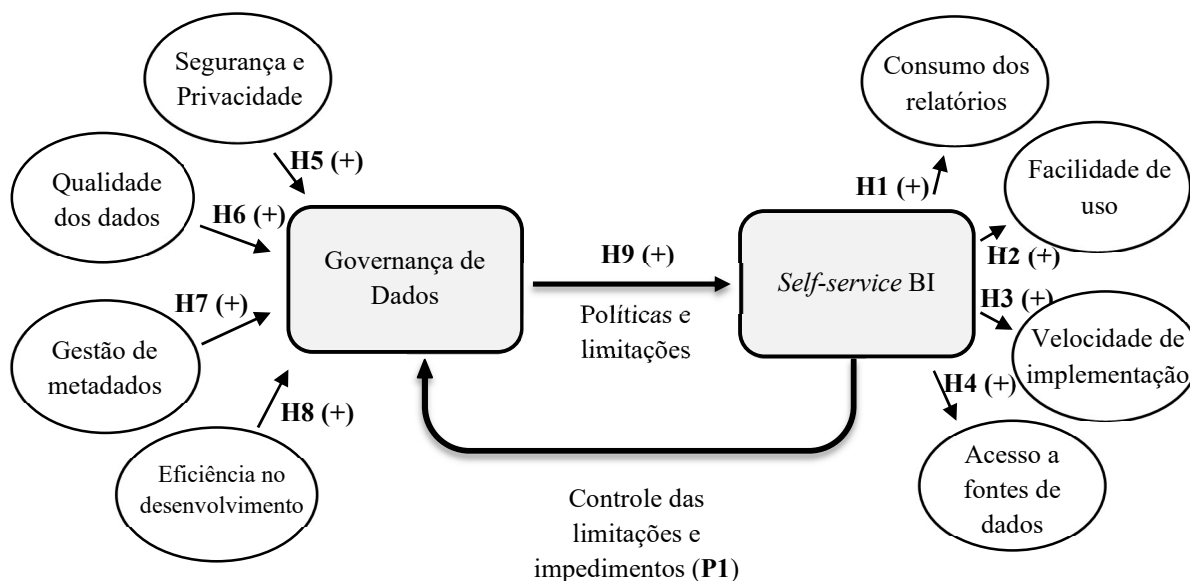
Fonte: Haag, S., Eckhardt, A. (2017). Shadow IT. *Business & Information Systems Engineering*, 59(6), 469-473.

Dessa forma, entende-se que a adoção do SSBI proporciona maior autonomia para as áreas de negócio, mas precisa ser embasado em uma governança de dados eficaz, que irá definir os limites até os quais o *self-service* pode ir de modo a garantir a confiança das análises e a segurança dos dados, mas não limitando em excesso de forma a fazer as áreas assumirem riscos desnecessariamente. Assim, pode-se formular a última hipótese da pesquisa.

- H9: A aplicação da governança de dados impacta positivamente a utilização do SSBI, gerando uma melhor percepção sobre a qualidade das ferramentas.

A presente pesquisa opta pela utilização de um *framework* baseado na visão centrada na governança para SSBI (Figura 11). As hipóteses estão sumarizadas na Tabela 2.

Figura 11
Modelo Conceitual Preliminar



Fonte: Dados da pesquisa.

Tabela 2
Tabela de hipóteses do modelo conceitual

Hipótese	Descrição
H1	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao consumo dos relatórios.
H2	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à facilidade de uso destas ferramentas.
H3	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à velocidade de implementação de relatórios.
H4	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao acesso às fontes de dados.
H5	A aplicação da governança de dados gera uma melhor percepção relativa à segurança e privacidade dos dados na utilização do SSBI.
H6	A aplicação da governança de dados gera uma melhor percepção relativa à qualidade dos dados na utilização do SSBI.
H7	A aplicação da governança de dados gera uma melhor percepção relativa à gestão de metadados na utilização do SSBI.
H8	A aplicação da governança de dados gera uma melhor percepção relativa à eficiência no desenvolvimento na utilização do SSBI.
H9	A aplicação da governança de dados impacta positivamente a utilização do SSBI, gerando uma melhor percepção sobre a qualidade das ferramentas.

Fonte: Dados da pesquisa.

Como consolidação dos conceitos, é apresentado o modelo conceitual (Figura 11), sendo que a governança de dados é implementada com objetivo de atuar em 4 variáveis avaliadas sobre os itens apresentados na Tabela 4. Para atingimento dos objetivos, a governança de dados implementa políticas e limitações que atuam, entre outros aspectos, no modelo de SSBI das organizações, de modo a agir sobre o escopo de atuação dos usuários com regras e

definição de comportamentos desejáveis, gerando uma percepção de benefícios para os usuários (H9). O SSBI visa atingir 4 objetivos, conforme apresentado na Tabela 3, definidos com base no trabalho seminal de Imhoff e White (2011), o qual iniciou as discussões sobre o conceito de SSBI. Junto com seus usuários, o modelo de SSBI reage às políticas e limitações impostas pela governança de dados (seta indicativa da proposição P1), fazendo com que a organização precise balancear estas limitações a fim de evitar a geração de estruturas de *shadow IT*. Formulam-se as seguintes hipóteses e a proposição como reação à aplicação da governança:

P1: Os usuários do modelo de *Self-service BI* reagem ao controle das limitações e impedimentos, impactando a Governança de Dados.

A diferença de nomenclatura entre hipóteses e proposição se deve porque as hipóteses foram investigadas quantitativamente ao passo que a proposição foi averiguada qualitativamente, sendo isto detalhado no próximo capítulo sobre a metodologia de pesquisa.

As variáveis de impacto do SSBI podem ser explicadas com base nos itens apresentados na Tabela 3.

Tabela 3*Itens e variáveis atreladas ao construto Self-Service BI*

Construto	Variáveis	Itens	Descrição Itens	Referências
<i>Self-service BI</i>	Consumo dos relatórios	Facilidade no consumo	Usuários podem acessar os recursos por uma interface que já estão familiarizados.	Alpar e Schulz (2016)
		Clareza das definições	Clareza nas definições e regras de negócio e facilidade de acesso à essas definições.	Imhoff e White (2011)
		Retroalimentação de conhecimento	Aumento do conhecimento por meio de interações com os resultados.	Imhoff e White (2011)
	Facilidade de uso	Facilidade de desenvolvimento	Facilidade no desenvolvimento de novos relatórios possibilitando o desenvolvimento por usuários casuais.	Lennerholt et al. (2018)
		Curva de aprendizado	Facilidade de criação de relatórios por novos usuários.	Imhoff e White (2011)
	Velocidade de implementação	Velocidade de implementação	Menor tempo de desenvolvimento que outras ferramentas de <i>Business Intelligence</i> .	Alpar e Schulz (2016)
		Custo de implementação	Menor custo de implementação de novos desenvolvimentos.	Imhoff e White (2011)
	Acesso a fontes de dados	Disponibilidade de acesso	Autonomia no acesso aos dados e criação relatórios.	Lennerholt et al. (2021)

As variáveis de impacto da governança de dados podem ser explicadas com base nos itens apresentados na Tabela 4.

Tabela 4*Itens e variáveis atreladas ao construto Governança de Dados*

Construto	Variáveis	Itens	Descrição Itens	Referências
Governança de Dados	Segurança e Privacidade	<i>Stakeholders</i>	Privacidade e confidencialidade dos dados de <i>stakeholders</i> .	DAMA (2017)
		Regulamentações	Cumprimento de regulamentações governamentais associadas à segurança e privacidade de dados.	DAMA (2017)
		Propriedade de negócios	Segurança de dados de negócio que asseguram vantagens competitivas.	DAMA (2017)
		Acesso legitimado	Concessão de acessos aos dados a quem de fato precisa.	DAMA (2017)
		Obrigações contratuais	Cumprimento de contratos e acordos de não divulgação de dados.	DAMA (2017)
	Qualidade dos Dados	Acurácia	Conformidade dos valores dos dados em relação ao valor real.	Khatri e Brown (2010)
		Pontualidade	Atualização do dado quando requisitado.	Khatri e Brown (2010)
		Compleitude	Disponibilidade dos valores necessários na granularidade necessária.	Khatri e Brown (2010)
		Credibilidade	Confiabilidade das fontes de dados e de seus conteúdos.	Khatri e Brown (2010)
	Gestão de metadados	Conhecimento organizacional	Documentação e disseminação do conhecimento organizacional.	DAMA (2017)
		Integração entre fontes de dados	Integração do conceito de dados entre sistemas.	DAMA (2017)
		Qualidade dos metadados	Garantia da qualidade, consistência e segurança dos metadados.	DAMA (2017)
		Acesso aos metadados	Concessão de acesso aos metadados na organização.	DAMA (2017)
		Troca de dados	Troca de dados entre sistemas.	DAMA (2017)
	Eficiência no Desenvolvimento	Acesso aos dados	Disponibilidade dos dados pela necessidade da organização.	Khatri e Brown (2010)

Construto	Variáveis	Itens	Descrição Itens	Referências
		Ciclo de vida dos dados	Otimização do armazenamento dos dados, e custos, entendendo o ciclo de vida dos dados.	Khatri e Brown (2010)

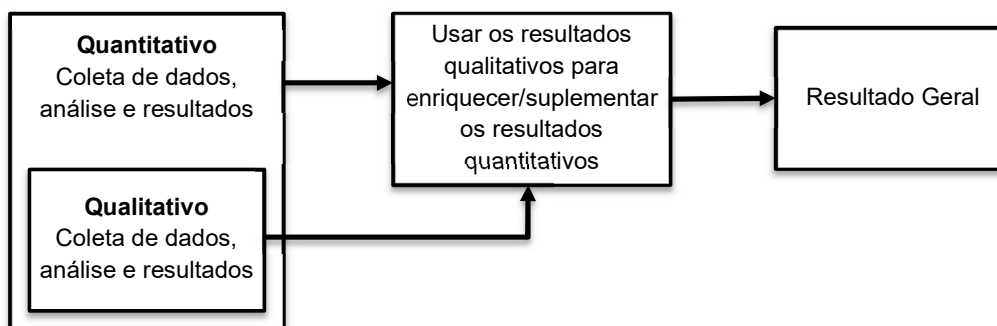
3. METODOLOGIA

3.1. Estratégia e método

No que tange à natureza do trabalho desenvolvido, a pesquisa emprega métodos mistos e está dividida em dois momentos: um momento quantitativo e um momento qualitativo. A pesquisa quantitativa pode ser entendida como uma pesquisa que emprega métodos e declarações empíricos, sendo as declarações empíricas afirmações descritivas sobre o que “é” ao invés do que “deveria ser”. Para tal, as declarações empíricas normalmente são expressas em termos numéricos. As avaliações empíricas são definidas como uma forma de determinar o quanto um programa específico cumpre empiricamente um determinado padrão (Cohen, 1980). Por seu turno, a pesquisa qualitativa pode ser entendida como o emprego de entrevistas, notas, gravações, entre outros meios, aplicando práticas interpretativas, a fim de interpretar um fenômeno com base no significado que as pessoas dão a ele (Denzin e Lincoln, 2008).

A utilização de um método misto de pesquisa agrega as vantagens tanto dos métodos quantitativos quanto dos métodos qualitativos, possibilitando avaliar diferentes tipos de dados, como números e tendências ao mesmo tempo que contextos e discursos, proporcionando um entendimento mais completo do fenômeno pesquisado (Clark et al., 2010). Para tal, a pesquisa utiliza o modelo incorporado (Figura 12) em que um tipo de dado, para a presente pesquisa o dado quantitativo, desempenha o papel principal dentro da pesquisa, enquanto o outro tipo de dado desempenha um papel secundário, suportando as análises e conclusões geradas e assim, a pesquisa visa responder duas partes diferentes de uma pergunta (Clark et al., 2010). Isso se justifica devido ao fato de a parte qualitativa da pesquisa endereçar variáveis complementares à análise principal, referente à parte quantitativa da pesquisa. A pesquisa qualitativa subsequente pode auxiliar na interpretação dos resultados quantitativos, e por isso, os resultados da análise dos dados da pesquisa quantitativa foram apresentados aos entrevistados no início das entrevistas.

Figura 12
Modelo incorporado



Fonte: Clark, V. L. P., Creswell, J. W., Green, D. O., Shope, R. J. (2010). Mixing Quantitative and Qualitative Approaches: An Introduction to Emergent Mixed Methods Research. In: Handbook of Emergent Methods, 363-387.

A opção da utilização de pesquisa quantitativa se deve à possibilidade de abranger um maior número de pessoas na coleta das declarações, a fim de avaliar as premissas definidas no modelo conceitual (Figura 11). Para tal, utiliza-se uma pesquisa por meio de *survey*, que corresponde a um método de busca de informação por meio de autorrelato descritivo de comportamentos e outras características de uma população (Rosenfeld, Edwards e Thomas, 1995). Por fim, faz-se necessária a aplicação de análise multivariada devido a sua capacidade de incorporar múltiplas variáveis para entender um complexo fenômeno organizacional, como o fenômeno teorizado (Swanson e Holton, 2005).

A pesquisa visa avaliar as hipóteses estabelecidas:

- H1: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao consumo dos relatórios;
- H2: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à facilidade de uso destas ferramentas;
- H3: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à velocidade de implementação de relatórios;
- H4: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao acesso às fontes de dados;
- H5: A aplicação da governança de dados gera uma melhor percepção relativa à segurança e privacidade dos dados na utilização do SSBI;
- H6: A aplicação da governança de dados gera uma melhor percepção relativa à qualidade dos dados na utilização do SSBI;

- H7: A aplicação da governança de dados gera uma melhor percepção relativa à gestão de metadados na utilização do SSBI;
- H8: A aplicação da governança de dados gera uma melhor percepção relativa à eficiência no desenvolvimento na utilização do SSBI;
- H9: A aplicação da governança de dados impacta positivamente a utilização do SSBI, gerando uma melhor percepção sobre a qualidade das ferramentas.

Para análise, primeiramente foi aplicado o método ACP (Análise dos Componentes Principais) de modo a reorientar os dados para se ter poucos componentes que expliquem o maior número possível de informações disponíveis, assim reduzindo as dimensões avaliadas para tornar a visualização dos dados mais direta e a análise mais administrável (Lattin et al., 2011). A aplicação da ACP foi feita sobre o resultado da aplicação do *survey*, compondo os itens e variáveis apresentados nas Tabelas 3 e 4.

Com os componentes definidos, para cada hipótese estudada, foram aplicados testes de correlação entre os componentes gerados pela ACP, variáveis X, e duas variáveis dicotômicas Y1 e Y2, sendo que a variável Y1 foi avaliada para as hipóteses H1 a H4 e a variável Y2 foi avaliada para as hipóteses H5 a H9, sendo:

- Y1: Os produtos de dados do respondente foram feitos com o uso de ferramentas de SSBI;
- Y2: Os produtos de dados do respondente seguem o modelo de governança de dados da organização.

Para a avaliação das hipóteses H1 a H8, foi calculado Coeficiente de Correlação Ponto Bisserial, sendo esse derivado do Coeficiente de Correlação de Pearson, avaliando a correlação entre uma variável contínua X e uma variável dicotômica Y (Ferguson, 1981). Além disso, para a utilização do coeficiente de correlação ponto bisserial, a variável Y não necessita ser contínua e nem normalmente distribuída (Bunchaft e Kellner, 1999). As fórmulas podem ser verificadas no Apêndice D. A análise considerou um nível de significância de 0,05 por gerar um maior equilíbrio entre erros do tipo I (rejeição de uma hipótese nula verdadeira) e do tipo II (não rejeição de uma hipótese nula falsa), permitindo um melhor equilíbrio entre o nível de exigência do teste e o tamanho da amostra, para uma pequena amostra (Cohen, 2017).

Para a avaliação da hipótese H9, foi avaliada a correlação entre os componentes gerados pela ACP das variáveis relacionadas à governança de dados, sendo essas segurança e privacidade, qualidade dos dados, gestão de metadados e eficiência no desenvolvimento. Visto que não se pode assumir uma correlação teórica direta entre os componentes gerados, foi feita

análise de regressão logística multivariada. A análise considerou apenas as variáveis que apresentaram a existência de uma correlação na análise univariada, ou seja, valor-p menor que 0,05 na avaliação do coeficiente de correlação ponto bisserial.

Como apresentado por Hosmer e Lemeshow (2000), a regressão logística multivariada objetiva avaliar a correlação entre uma variável dependente e múltiplas variáveis independentes, onde a variável dependente é uma variável dicotômica. As fórmulas da regressão também podem ser verificadas no Apêndice D.

Devido à natureza probabilística da regressão logística multivariada, o R^2 tradicional não é aplicável, e por isso, calculou-se o pseudo R^2 como coeficiente de determinação para avaliar o ajuste do modelo, conforme apresentado por Tjur (2009).

A utilização de pesquisa qualitativa de característica longitudinal se deve ao interesse em avaliar o processo de transição de produtos de dados entre o modelo de *shadow IT* e o modelo governado. Além disso, a natureza qualitativa se mostra adequada devido ao fato de avaliar o fenômeno em um contexto real, onde as fronteiras entre o fenômeno e o contexto não estão claramente definidas (Yin, 2015). A característica longitudinal da pesquisa visa a descrição e análise de mudanças observadas ao longo do tempo (Aaboen, Dubois e Lind, 2012).

Além disso, opta-se pela utilização de pesquisa qualitativa para avaliação de critérios de segurança envolvidos no modelo de governança de dados devido à necessidade de avaliação junto ao corpo técnico de TI, restringindo-se a um pequeno número de pessoas, além da avaliação de documentos, o que é possibilitado pela pesquisa qualitativa por permitir a utilização de múltiplas evidências e a consideração de condições contextuais (Yin, 2016).

Para a análise dos dados qualitativos obtidos, a pesquisa utiliza o método de análise de conteúdo. Conforme definido por Bardin (2016), esse método visa avaliar os significados atribuídos pelos sujeitos participantes da pesquisa, considerando seu contexto, ao objeto de estudo. Esse método se faz adequado por considerar a subjetividade das manifestações do sujeito, de modo que o pesquisador se afaste de suas crenças e convicções para interpretação dos dados, contextualizando-os ao referencial teórico (Bardin, 2016).

Ainda segundo Bardin (2016), esse método pode ser dividido em: pré-análise, quando o pesquisador irá sistematizar as ideias preliminares, a fim de avaliar a pertinência dos materiais a serem utilizados; exploração do material, transformando os dados brutos em unidades significativas para análise, podendo ser palavras, frases, entre outros; e tratamento dos resultados e interpretação, quando o pesquisador dá sentido às manifestações do sujeito e às relaciona ao objeto de estudo.

Quanto ao método, a presente pesquisa utiliza-se do estudo de caso, o qual visa se concentrar na compreensão da dinâmica presente em ambientes únicos (Eisenhardt, 1989). A escolha do método foi feita devido ao caráter descritivo da pesquisa, objetivando uma descrição completa do fenômeno dentro de seu contexto (Yin, 2015). Além disso, pode-se entender o presente estudo como um estudo de caso único holístico, visto que visa avaliar a natureza global da organização (Yin, 2015).

3.2. Unidade de análise

A empresa objeto de estudo desse caso é uma organização brasileira de capital fechado, de grande porte do setor logístico, que atua com logística multimodal, contemplando a operação de transporte ferroviário e operação de portos e terminais. A organização é fruto de um *spin-off* realizado há 13 anos e atualmente possui 7 mil empregados atuando em 250 municípios das regiões Sudeste, Centro-Oeste, Norte e Nordeste do Brasil, sendo sua matriz localizada na cidade de Belo Horizonte/MG.

A empresa utiliza como ERP, desde 2017, o sistema SAP ECC (*Enterprise Core Component*), responsável pela gestão financeira e processos transacionais administrativos. A empresa também possui sistemas transacionais desenvolvidos de forma customizada por meio da contratação de fábricas de software, com dados armazenados em sistemas gerenciadores de banco de dados Oracle e Microsoft Azure. Em janeiro de 2024, os dados armazenados pelos sistemas transacionais totalizavam um armazenamento de 1,02 *petabytes*.

Com base na pesquisa documental, constatou-se que a utilização de produtos de *Business Intelligence* se iniciou em 2017, com a implantação da solução de Microsoft Power BI, sendo que à época as áreas de negócio puderam desenvolver relatórios com atualizações manuais, inicialmente adotando um modelo *shadow IT*. Em 2018, foi feita a implantação de um *data warehouse* para melhor ingestão e tratamento de dados para a construção de uma solução específica de BI apresentando relatórios financeiros e contábeis. Esse produto é responsável por todas as apurações financeiras e contábeis da organização.

Em 2019, iniciou-se a implementação de um produto de BI sem a utilização do modelo *self-service*, denominado Centro de Inteligência. O Centro de Inteligência é um sistema desenvolvido por fábricas de software para cálculo e exibição dos indicadores referentes à operação da empresa, e já iniciou seu desenvolvimento consumindo dados do *data warehouse* da empresa. Mesmo com um desenvolvimento tecnicamente adequado das soluções, os produtos de dados desenvolvidos ainda enfrentavam um problema relacionado à gestão do conhecimento sobre os desenvolvimentos feitos, ficando dependente do conhecimento

adquirido pelos profissionais envolvidos no desenvolvimento. Por este motivo, em 2019, foi feita a implantação da ferramenta da Oracle denominada *SQL Developer Data Modeler*, visando o registro dos metadados para assegurar a correta gestão de conhecimento sobre os desenvolvimentos feitos.

Até 2019, os desenvolvimentos estavam sendo feitos utilizando ferramentas adequadas e robustas para utilização, porém sem um modelo de governança de dados que garantisse a construção de *templates* e padrões para auxiliar nos desenvolvimentos. Face tais necessidades, em 2020, iniciou-se a implantação do modelo de governança de dados da organização. Ainda em 2020, após o custo com o processamento e armazenamento de dados chegar a mais de R\$ 100 mil por mês, foi implementado um processo de gestão de ambientes e custos de modo a otimizar os desenvolvimentos e produtos e fazer uma otimização dos custos destinados. Por fim, os relatórios de SSBI ainda estavam sendo construídos em estruturas *shadow IT*, e em 2022, iniciou-se a migração dos relatórios para um modelo governado, além do desenvolvimento de novos produtos que já iniciam no modelo governado. A Figura 13 sumariza os principais eventos em uma linha do tempo.

Figura 13

Linha do tempo de adoção de Business Intelligence

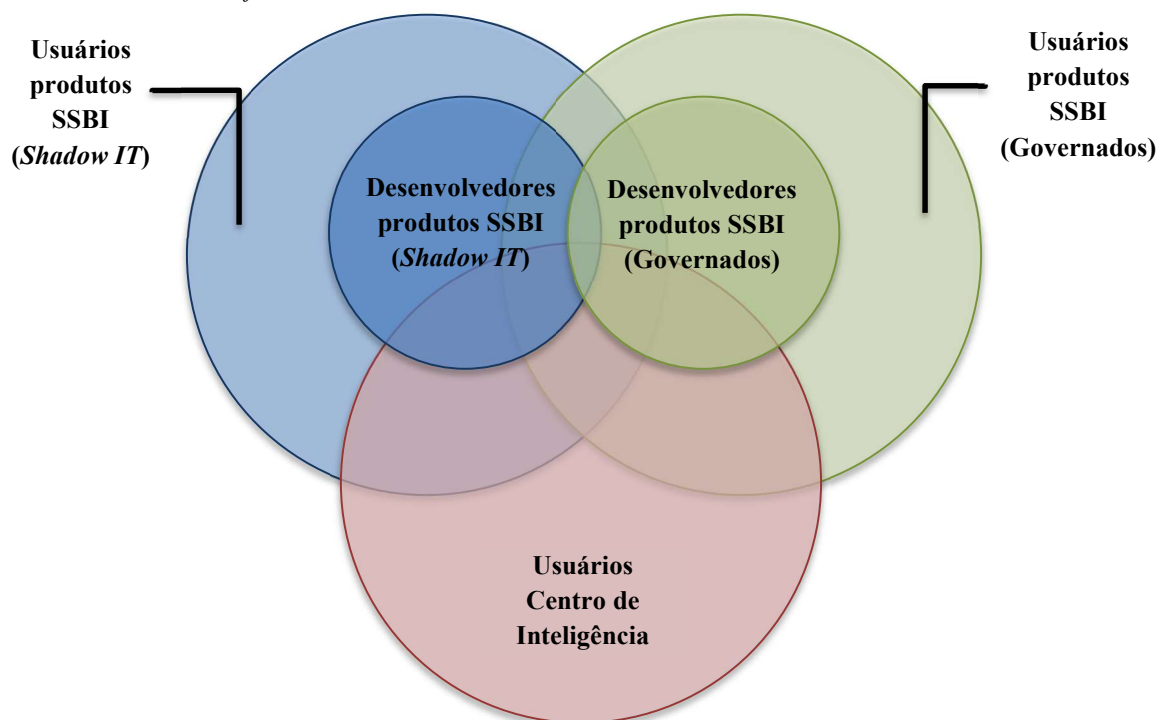


Fonte: O Autor, com base na análise documental.

Atualmente, a empresa conta com produtos de *Business Intelligence* tanto desenvolvidos por equipes técnicas de TI quanto desenvolvidos no modelo de SSBI. Com base na avaliação das informações disponíveis em abril de 2024 no Portal do Administrador do Microsoft Fabric e no Portal do Microsoft Azure, ferramentas disponibilizadas pela Microsoft para avaliação e monitoramento do ambiente, a empresa possui atualmente 8.275 relatórios produzidos por tecnologia de SSBI, tanto em estruturas *shadow IT*, quanto seguindo o modelo de governança de dados estabelecido. Desses, 705 relatórios foram utilizados nos últimos 30 dias a contar da data consultada. O *data warehouse* da empresa possui 174 fatos, 453 dimensões e 117 *reports* totalizando 1,12 *terabytes* armazenados.

Figura 14

Universo de usuários das ferramentas de BI



Fonte: Dados da pesquisa.

Ainda de acordo com as informações de abril de 2024, a empresa possui 1.139 usuários ativos que utilizam ferramentas de BI. Esses usuários foram categorizados pela natureza dos produtos que eles consomem (Figura 14), sendo que para aqueles usuários que consomem produtos de mais uma natureza foi feita a separação atrelando-os ao universo de produtos que eles mais consomem. Por conseguinte, foi obtido o seguinte cenário:

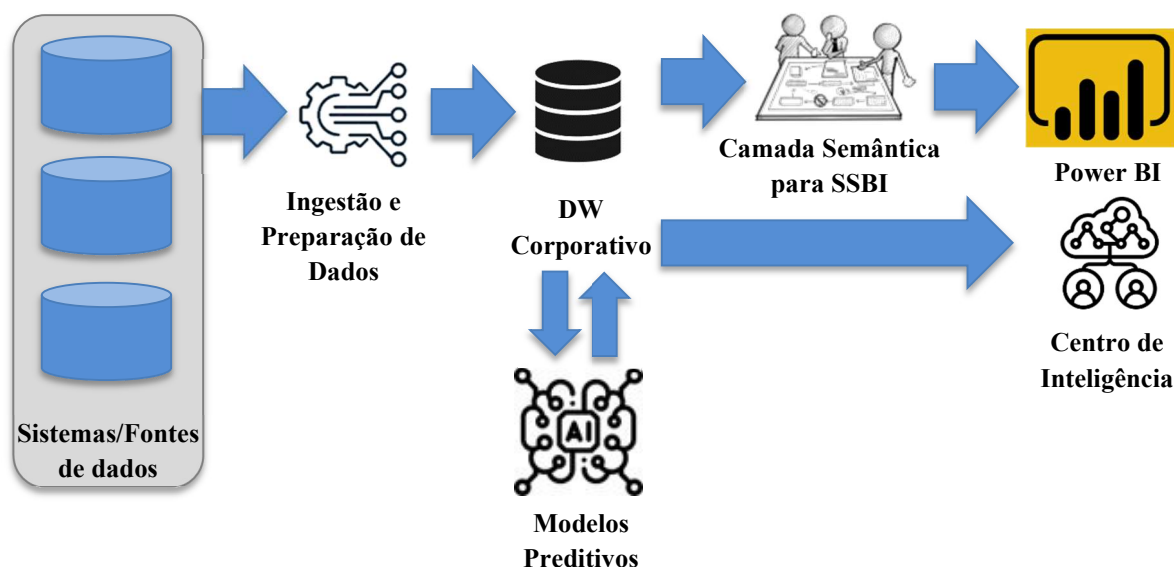
- Usuários de produtos SSBI (*Shadow IT*): 597
- Desenvolvedores de produtos SSBI (*Shadow IT*): 98
- Usuários de produtos SSBI (Governados): 163
- Desenvolvedores de produtos SSBI (Governados): 26
- Usuários do Centro de Inteligência: 255

Devido à empresa ter um histórico ainda recente de uso de BI e SSBI, de modelo de governança e uma distribuição de usuários e desenvolvedores de produtos governados e não governados, ela possui alta relevância de estudo para esta pesquisa, pois possibilita contrapor a visão de usuários que usufruem da governança de dados dentro do uso de SSBI e usuários que utilizam o modelo *shadow IT*. Isso possibilita a extrapolação dos resultados e análises para empresas do mesmo setor e de diferentes setores que possuam um cenário semelhante de usuários.

O modelo arquitetural de desenvolvimento de produtos de dados da empresa se apresenta conforme Figura 15. Para produtos desenvolvidos em estruturas *shadow IT*, as áreas utilizam a ferramenta Microsoft Power BI com alimentação manual das bases de dados ou utilização de armazenamento via plataforma de Microsoft Sharepoint.

Figura 15

Modelo arquitetural de dados



Fonte: O Autor, com base na análise documental.

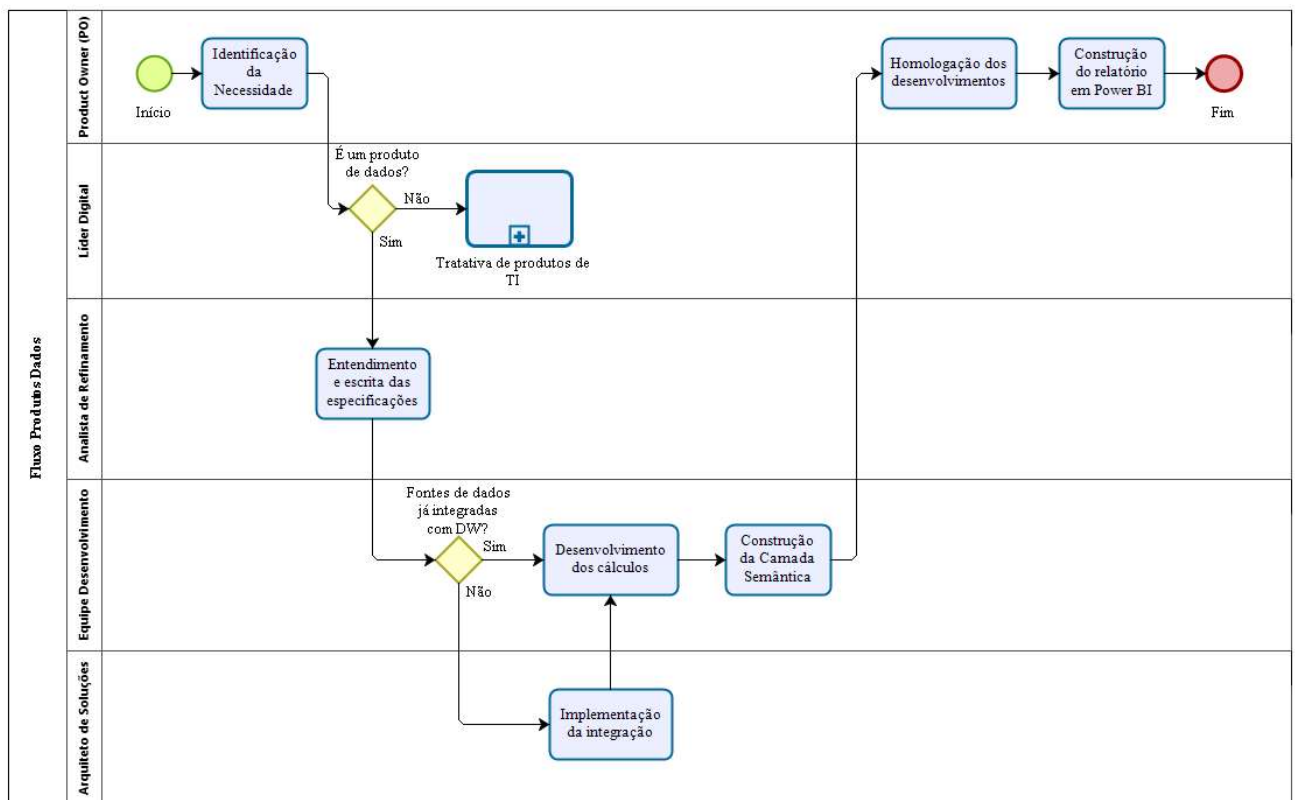
Para a construção de novos relatórios e produtos de dados, as áreas de negócio possuem licenças de MS Power BI, as quais podem fazer uso para construção de relatórios em estruturas *shadow IT* com carga de dados de forma manual. As áreas de negócio podem ainda definir uma iniciativa digital para a construção de relatórios no modelo governado ou migração de relatórios existentes. Para essa construção ou migração de relatórios, a empresa dispõe do fluxo apresentado na Figura 16, por meio do qual as equipes de negócio que identificam as necessidades apresentam uma demanda ao time de Transformação Digital, que alocará um analista para entendimento da natureza da demanda. Após entendimento inicial, é alocado um analista de refinamento que entenderá todo o processo ao qual o relatório atenderá e, assim, avaliará os conceitos definidos, entendendo se são conceitos já implementados em outros relatórios ou processos da empresa, a fim de convergir conceitos, e construindo especificações que traduzam os relatórios desejados para um formato que a equipe técnica consiga entender.

Após essa especificação, é feita a alocação de um engenheiro de dados que fará o entendimento dos dados e cálculos necessários a fim de avaliar se são dados e cálculos já existentes dentro do *data warehouse* da empresa, para fins de convergência e reaproveitamento de desenvolvimentos já feitos. Após esse entendimento, é concluída a especificação técnica dos

relatórios a serem desenvolvidos. Nesse momento, o time técnico de desenvolvimento, composto por arquitetos de solução, é mobilizado caso sejam necessárias a definição e implementação de novas integrações de sistemas com o *data warehouse*, sendo esta a etapa de ingestão de dados, engenheiros de dados, para desenvolvimento das etapas de preparação dos dados e desenvolvimento dos cálculos dentro do *data warehouse* e das camadas semânticas, e cientistas de dados, caso seja necessária a aplicação de modelos estatísticos e preditivos.

Figura 16

Processo de implementação de produtos de dados



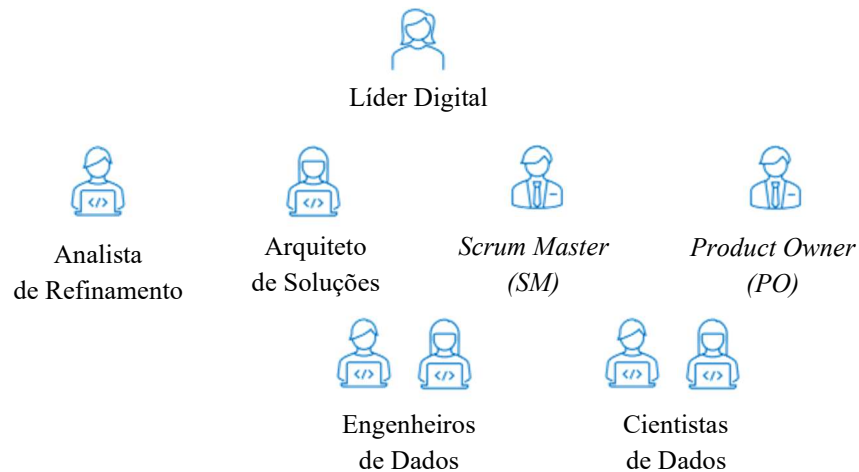
Fonte: O Autor, com base na análise documental.

Durante toda a implementação, seguindo preceitos da abordagem ágil, uma pessoa da área de negócio é responsável por ser o *Product Owner* (PO) do produto a ser desenvolvido, sendo responsável por definir todas as regras de negócio que serão implementadas, priorizar os desenvolvimentos e homologar o que for desenvolvido. Além disso, ao final da entrega da camada semântica, desenvolver o relatório em MS Power BI é responsabilidade do PO, a fim de consumir os dados e cálculos implementados para criar os elementos visuais que serão apresentados. Outros papéis necessários durante a implementação ágil dos produtos de dados são o *Scrum Master* (SM), responsável por garantir a condução do projeto de implantação pela metodologia *Scrum*, e o Líder Digital, responsável por gerenciar o time de implementação do

produto. Dessa forma, para as implementações de produtos de dados, é constituída a estrutura apresentada na Figura 17.

Figura 17

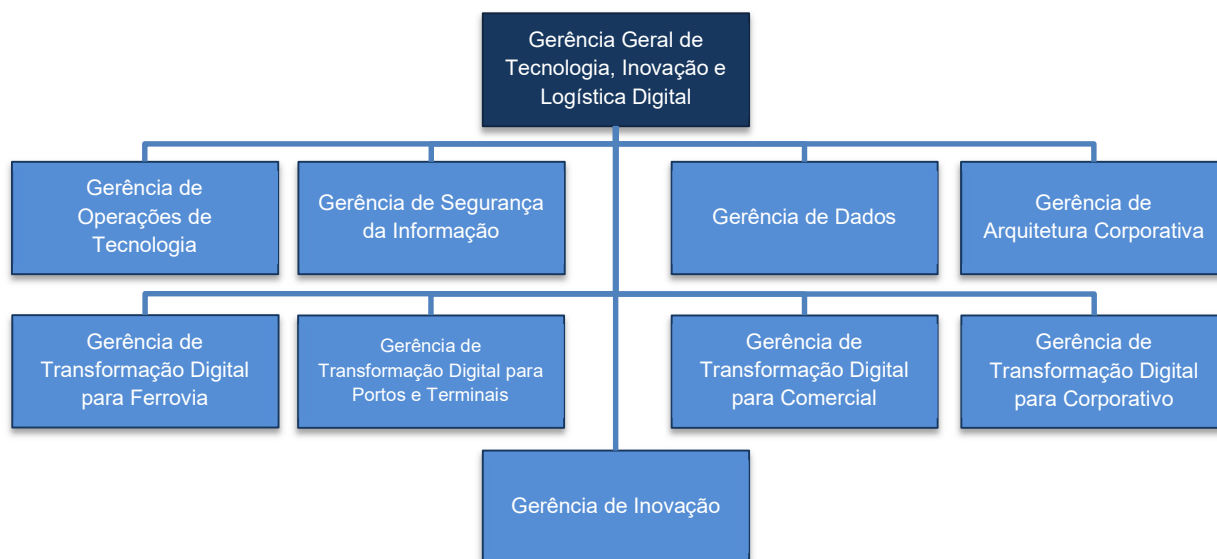
Estrutura de implementação de produtos de dados



Fonte: Dados da pesquisa.

Dessa forma, pelos conceitos definidos por Michalczyk et al. (2020), o modelo da organização pode ser entendido como um modelo de visão centrada na governança, com o *self-service* no nível 3 – Criação de relatórios, de acordo com o modelo definido por Alpar e Schultz (2016).

Toda a governança e implantação de produtos de dados, bem como a gestão de sistemas de tecnologia é feita pela área de TI da organização, que se estrutura como apresentado na Figura 18.

Figura 18*Estrutura organizacional da área de Tecnologia da Informação*

Fonte: Dados da pesquisa.

3.3. Coleta de dados

Foi feita uma pesquisa no modelo de *survey* com uma escala tipo Likert com 11 itens, apresentado no Apêndice A, utilizando a ferramenta Microsoft Forms, com os diferentes conjuntos de usuários de produtos de BI, conforme apresentado na Figura 14, a fim de avaliar a percepção de valor observada no cenário de adoção do modelo de governança de dados.

Após a análise dos dados quantitativos, foram feitas entrevistas com 6 profissionais da estrutura de governança de dados e segurança da informação, utilizando-se o roteiro apresentado no Apêndice B, a fim de avaliar a percepção de valor sob o aspecto de segurança e privacidade dos dados. Foi feita ainda entrevistas com 7 desenvolvedores de produtos de SSBI concebidos em estruturas *shadow IT* e migrados para o modelo governado, conforme roteiro apresentado no Apêndice C, a fim de avaliar o processo de migração dos produtos entre as estruturas. No início destas entrevistas, foi apresentado aos entrevistados um sumário dos resultados da *survey* para capturar a interpretação dos respondentes sobre os resultados obtidos na fase quantitativa.

Por último, foi feita uma análise documental de controles, a fim de apresentar indicadores numéricos de performance, qualidade e segurança dos dados quando da utilização dos modelos de governança. Essa análise foi feita com base na avaliação da performance dos servidores, análise de *logs* e volumetrias obtidos nas ferramentas do Portal do Administrador do Microsoft Fabric, que apresenta os indicadores referentes ao Microsoft Power BI, e do Portal do Microsoft Azure, que apresenta os indicadores referentes ao banco de dados e ao *Data*

Warehouse da empresa. Além disso, foram utilizados os documentos de governança de dados da empresa, que contemplam os MVA's (*Minimum Viable Architecture*) dos produtos de dados governados, documentos que indicam o modelo de arquitetura aplicado, fontes de dados, entre outras informações técnicas, e o catálogo de dados da organização.

4. ANÁLISE DOS RESULTADOS QUANTITATIVOS

Nesse capítulo são apresentados os resultados da parte quantitativa da pesquisa, com base no *survey* aplicado. O capítulo está dividido em quatro seções: caracterização da amostra; análise dos resultados das variáveis de *self-service BI*, conforme apresentado no modelo conceitual (Figura 11), onde serão avaliadas as hipóteses de 1 a 4; análise dos resultados das variáveis de governança de dados, onde serão avaliadas as hipóteses de 5 a 9; análise documental, com base nos *logs* e volumetria obtidos nas ferramentas do Portal do Administrador do Microsoft Fabric e do Portal do Microsoft Azure.

4.1. Caracterização da Amostra

No período de maio a junho de 2024, a pesquisa aplicada foi respondida por 368 pessoas. Os respondentes estão distribuídos da seguinte forma, com base na categorização dos usuários apresentada na Figura 14:

- Usuários de produtos SSBI (*Shadow IT*): 113
- Desenvolvedores de produtos SSBI (*Shadow IT*): 39
- Usuários de produtos SSBI (Governados): 96
- Desenvolvedores de produtos SSBI (Governados): 20
- Usuários do Centro de Inteligência: 100

Para reduzir o impacto dos estratos dos profissionais publicados, o *survey* foi aplicado apenas aos profissionais em cargos de analistas, supervisores e gerentes, não sendo aplicado aos altos executivos da organização, limitando o público aos profissionais de nível operacional e tático.

Os respondentes apresentaram tempo médio de experiência profissional de 16 anos, tempo médio de empresa de 7 anos e tempo médio de experiência com produtos de BI de 4 anos, caracterizando um perfil bastante experiente da amostra. Visto o histórico apresentado de produtos de BI e governança de dados na empresa, conforme a distribuição dos respondentes apresentada nas Tabelas 5, 6 e 7, grande parte dos entrevistados presenciaram a implementação dos produtos de BI e a evolução da governança de dados ocorrida na empresa.

Tabela 5*Experiência profissional dos respondentes*

Tempo de experiência profissional	Quantidade
Até 5 anos	36
De 5 a 10 anos	75
De 10 a 20 anos	163
De 20 a 30 anos	71
Acima de 30 anos	23

Fonte: Dados da pesquisa.

Tabela 6*Tempo de empresa dos respondentes*

Tempo de empresa	Quantidade
Até 5 anos	188
De 5 a 10 anos	104
De 10 a 20 anos	64
Acima de 20 anos	12

Fonte: Dados da pesquisa.

Tabela 7*Tempo de experiência com produtos de BI*

Tempo de experiência em BI	Quantidade
Até 1 ano	68
De 1 a 5 anos	248
De 5 a 10 anos	46
Acima de 10 anos	6

Fonte: Dados da pesquisa.

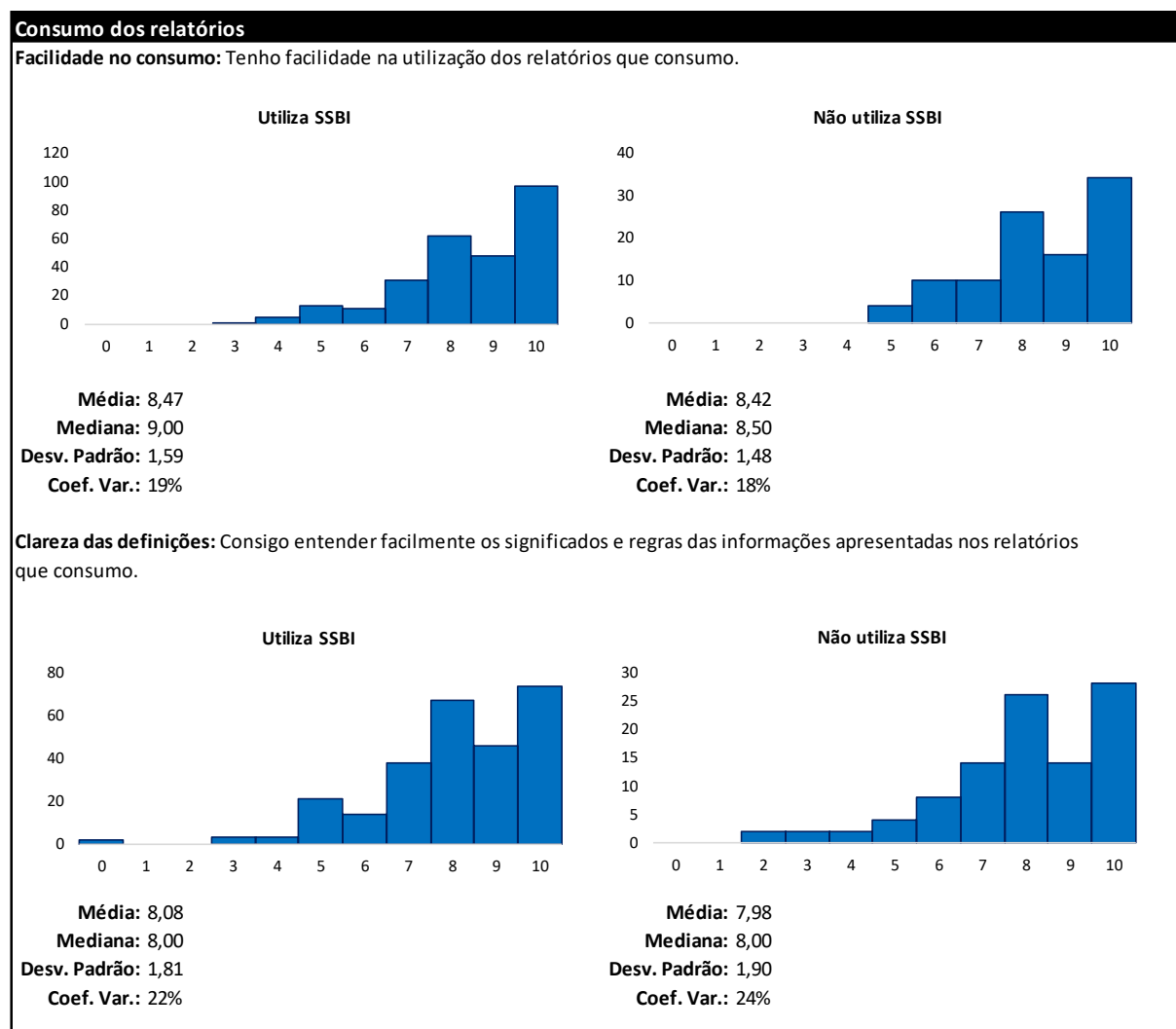
4.2. Resultados *Self-Service BI*

O uso de ferramentas de SSBI foi avaliado junto aos usuários com 7 questões aplicadas no *survey*, endereçando as 4 variáveis apresentadas na Tabela 3. A distribuição das avaliações referentes às variáveis Consumo dos Relatórios e Facilidade de uso pode ser vista nas Figuras 19 e 20, respectivamente.

Para análise descritiva das respostas, apesar de comporem dados ordinais, devido ao número de categorias ser grande (acima de 10) e da homogeneidade das categorias, estas poderão ser interpretadas como intervalares, permitindo o uso de técnicas de estatística quantitativa (Norman, 2010). Portanto, foram calculados, além da mediana, a média, desvio padrão e o consequente coeficiente de variação.

Figura 19

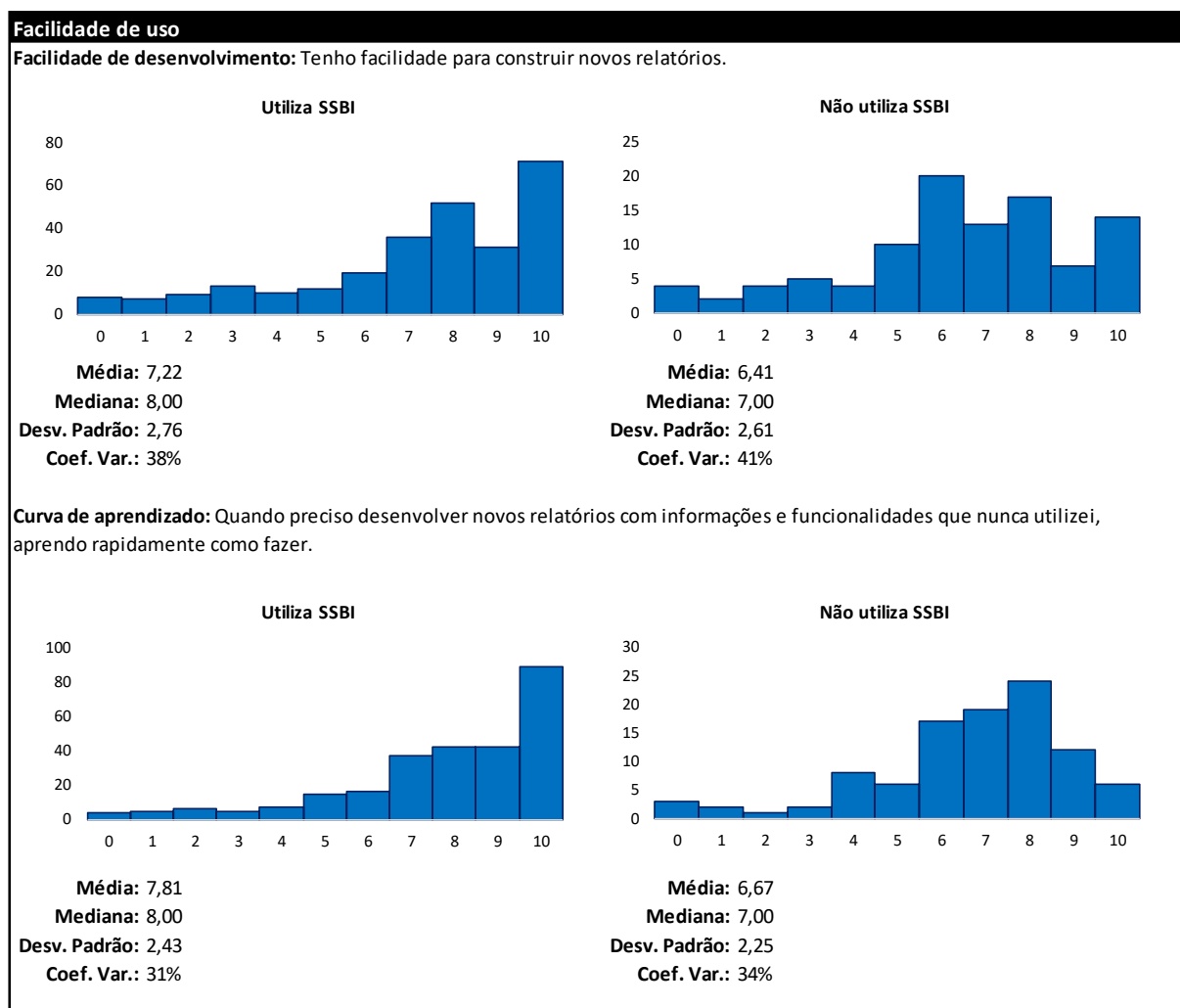
Distribuição das respostas das questões referentes à Consumo dos Relatórios



Fonte: Dados da pesquisa.

Figura 20

Distribuição das respostas das questões referentes à Facilidade de Uso

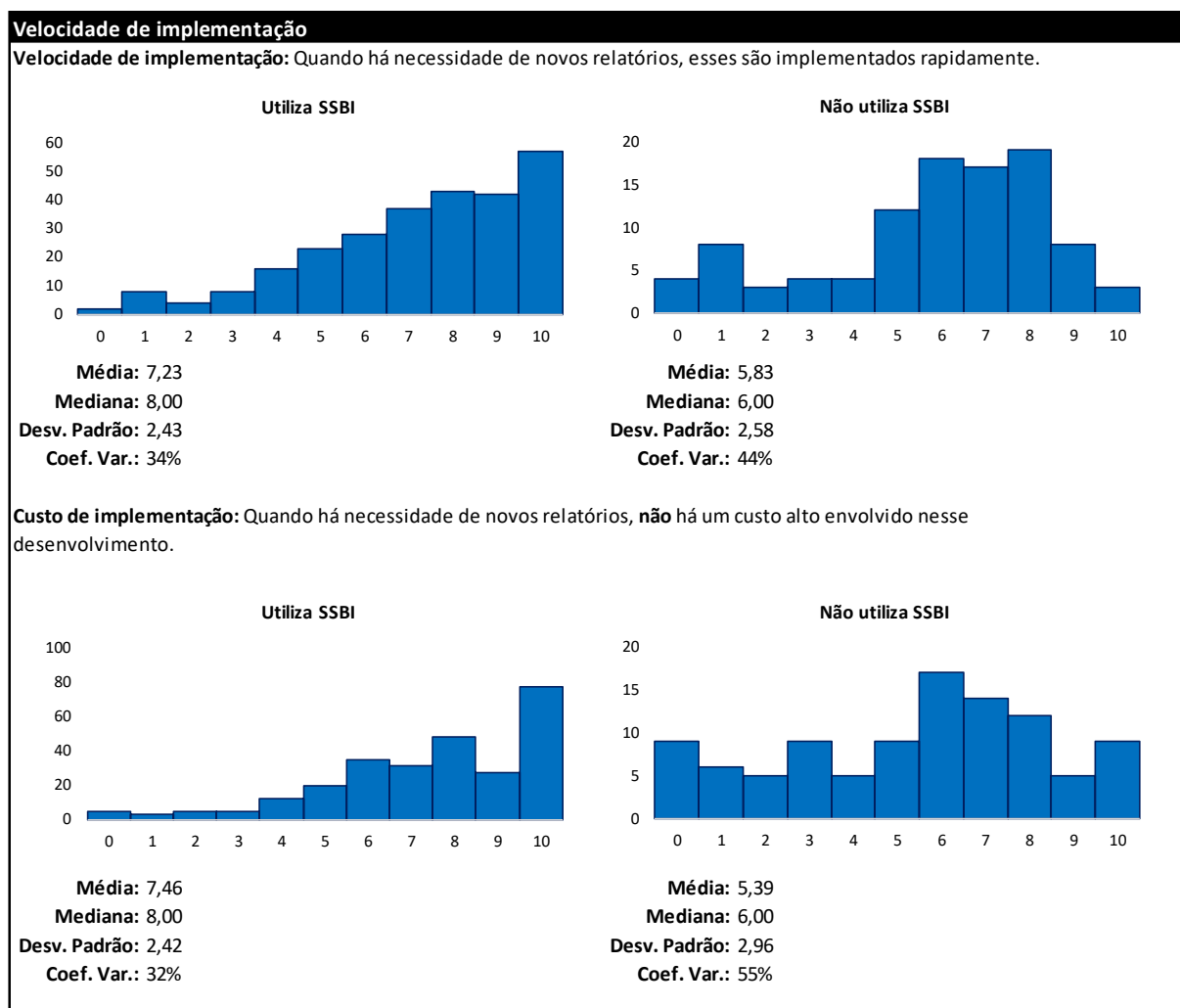


Fonte: Dados da pesquisa.

Observa-se uma grande diferença de média nas respostas da variável “facilidade de uso”. Esta diferença condiz com o próprio objetivo das ferramentas de SSBI, que visam ser ferramentas de fácil utilização, não requerendo um grande conhecimento técnico dos desenvolvedores, de modo que as próprias equipes de negócio possam desenvolver seus próprios relatórios. A seguir, a Figura 20 exibe a distribuição das respostas sobre a velocidade e o custo da implementação do SSBI.

Figura 21

Distribuição das respostas das questões referentes à Velocidade de Implementação

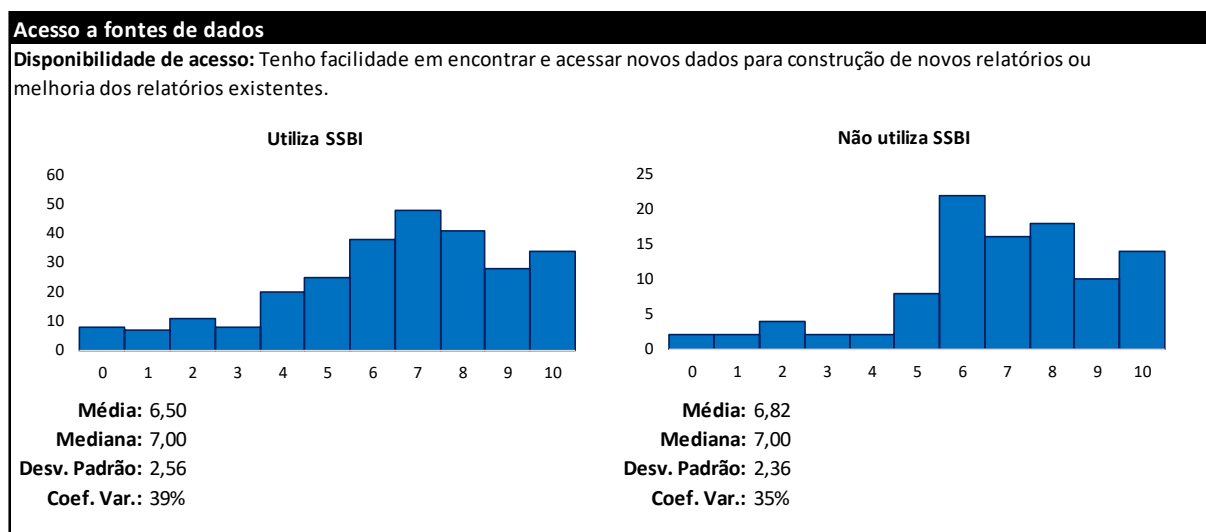


Fonte: Dados da pesquisa.

As maiores diferenças entre as médias das respostas de quem utiliza SSBI para quem não utiliza estão nos itens da variável “velocidade de implementação”, o que demonstra, como esperado, que os usuários dos produtos de BI não *self-service* tem a percepção da construção e manutenção dos relatórios ser demorada e cara. Esse fator se deve à autonomia que as ferramentas de SSBI proveem aos usuários de negócio, que são os mesmos detentores das regras a serem aplicadas, de modo que eles mesmos podem fazer os ajustes e as criações quando precisam. Este desenvolvimento não gera um custo novo nem para a área de TI nem para a organização, visto que utiliza uma mão de obra já existente.

Figura 22

Distribuição das respostas das questões referentes à Acesso a Fontes de Dados



Fonte: Dados da pesquisa.

Além disso, observa-se, de forma geral, uma menor média nas avaliações da variável “Acesso a fontes de dados” (Figura 21). Pode-se considerar que essa menor avaliação se deve ao estágio de *self-service* em que a empresa se encontra, alocada no nível 3 – Criação de relatórios, de acordo com o modelo definido por Alpar e Schultz (2016) – Figura 4. Já era esperado um resultado menor nessa variável devido à restrição de acesso apresentada no modelo de governança atual da empresa quando avaliadas as fontes de dados. À medida que a empresa avança no nível de *self-service*, alcançando o nível 4 - Acesso às funções analíticas, pode-se esperar uma maior avaliação na variável de “Acesso a fonte de dados”.

Para todas as variáveis foi aplicada Análise dos Componentes Principais (ACP) para redução dos itens avaliados em uma única variável, sendo depois calculado Coeficiente de Correlação Ponto Bisserial, com base na metodologia proposta, correlacionando com a variável Y1. Os produtos de dados do respondente foram feitos com o uso de ferramentas de SSBI, sendo 1 quando o respondente utiliza ferramentas de SSBI, e 0 quando não utiliza. Para os cálculos foi utilizado o *software* R, gerando os resultados que podem ser observados na Tabela 8.

Tabela 8
Resultados da avaliação de SSBI

Variável	Quantidade Componentes	Representatividade Componente Principal	Valor-p	Correlação
Consumo dos relatórios	2	85%	0,6289	-0,2528
Facilidade de uso	2	85%	5,31e-09	0,2983
Velocidade de implementação	2	81%	8,65e-18	0,4276
Acesso a fonte de dados	1	100%	0,2779	-0,0567

Fonte: Dados da pesquisa.

A aplicação do teste permitiu a rejeição da hipótese nula para as variáveis “Facilidade de uso” e “Velocidade de implementação”, possibilitando assumir uma correlação. Por outro lado, o teste não permitiu a rejeição da hipótese nula para as variáveis “Consumo dos relatórios” e “Acesso a fonte de dados”.

Portanto, pode-se assumir, pela não rejeição, que as hipóteses 2 e 3 da pesquisa são pertinentes e há a rejeição das hipóteses 1 e 4. Isso demonstra que a utilização de ferramentas de SSBI possibilita uma maior facilidade no desenvolvimento de novos relatórios, um aprendizado mais rápido para novos desenvolvimentos, implementação de novos relatórios de formas mais rápidas e um menor custo para os desenvolvimentos.

Porém, o consumo dos relatórios não apresentou correlação com o uso das ferramentas de SSBI, o que demonstra que relatórios de BI desenvolvidos sem estas ferramentas, apesar de apresentarem maior custo e demora no desenvolvimento, ainda podem ser fáceis de serem utilizados e os indicadores apresentados podem estar claramente definidos aos usuários. Assim, observa-se uma igual facilidade e clareza nas definições para consumo das informações tanto em ferramentas de SSBI quanto em ferramentas de BI desenvolvidas pelas equipes de TI. Tanto para usuários de produtos de SSBI quanto para relatórios fora do *self-service*, a média de avaliações da facilidade no consumo foi alta, mostrando que não é um fator que tem influência da aplicação das ferramentas de SSBI, mas é uma característica positiva para a organização, pois demonstra a qualidade geral dos produtos desenvolvidos.

Além disso, o uso de ferramentas de SSBI não apresentou correlação com maior acesso às fontes de dados. Esse item pode ser observado no caso estudado devido ao estágio de *self-service* em que a empresa se encontra, alocada no nível 3 (Criação de Relatório), com base no modelo definido por Alpar e Schultz (2016). Dessa forma, os *Citizen Developers* possuem menor autonomia para acessar os dados de novas fontes de dados diretamente, estando restritos

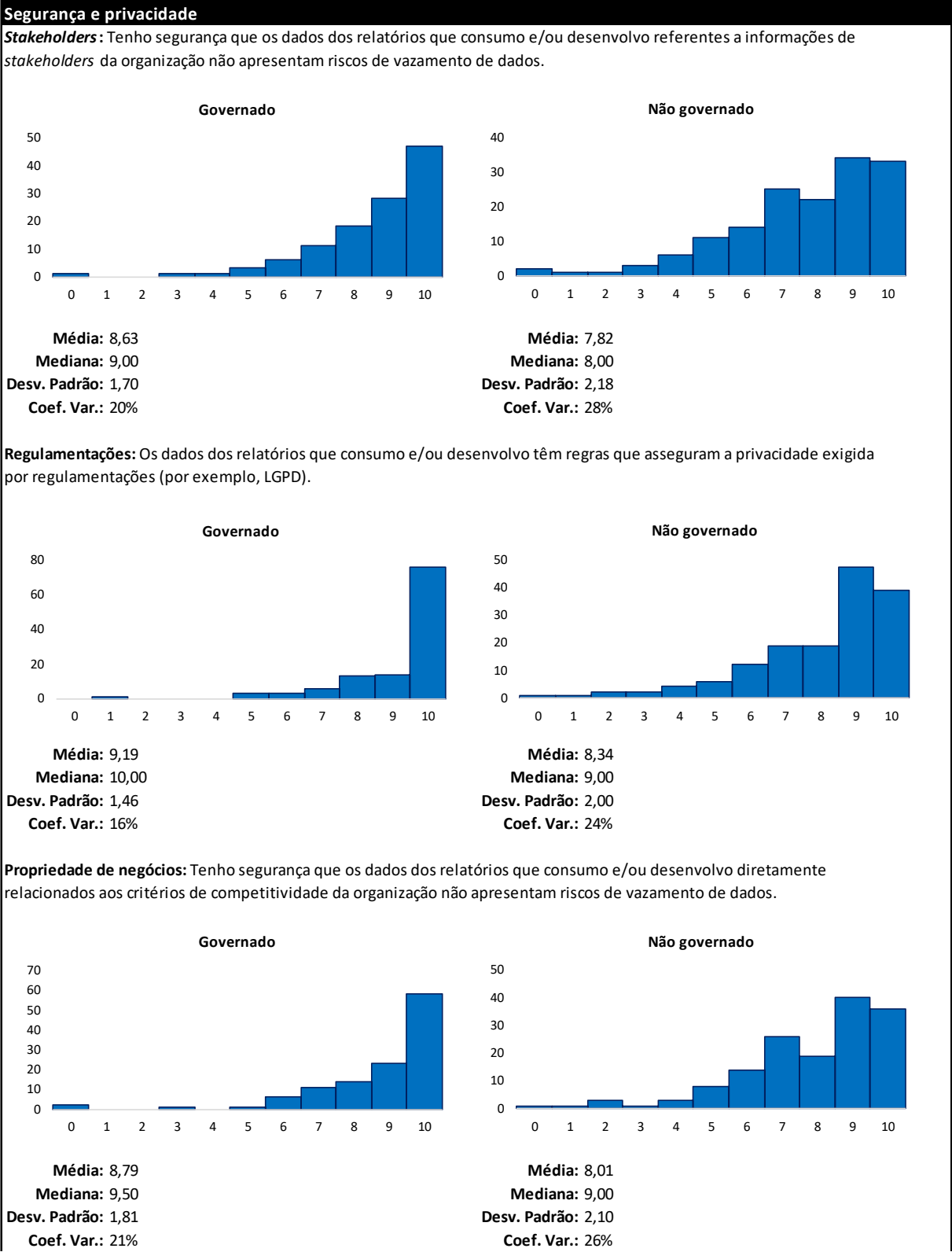
aos dados fornecidos pelas ferramentas analíticas. À medida que a empresa avança no nível de *self-service*, alcançando o nível 4 (Acesso às funções analíticas), pode-se esperar uma melhor avaliação nesta variável. Devido às restrições aplicadas pela governança de dados, esta falta de percepção na autonomia para acesso a fontes de dados pelos desenvolvedores e usuários pode levar a um aumento no desenvolvimento de soluções *shadow IT*, por apresentarem uma alternativa mais flexível de acesso mais rápido aos dados, mesmo que com menor qualidade.

Para as outras variáveis foram verificadas correlações positivas, demonstrando que a utilização de ferramentas de SSBI possibilita uma maior facilidade no desenvolvimento de novos relatórios, um aprendizado mais rápido para novos desenvolvimentos, implementação de novos relatórios de forma mais rápida e um menor custo para os desenvolvimentos.

4.3. Resultados Governança de Dados

A aplicação da governança de dados, por sua vez, foi avaliada junto aos usuários com 16 questões aplicadas no *survey*, endereçando as quatro variáveis apresentadas na Tabela 4. Considerando a proposta de avaliação da aplicação da governança de dados para o uso de ferramentas de SSBI, para essa análise foram isoladas apenas as respostas dos usuários e desenvolvedores de relatórios que utilizam ferramentas de SSBI. A distribuição das avaliações sobre Segurança e Privacidade pode ser vista na Figura 22.

Figura 23
Distribuição das respostas das questões referentes à Segurança e Privacidade



Governado

Nota	Frequência
0	2
1	0
2	0
3	2
4	0
5	2
6	6
7	11
8	13
9	23
10	58

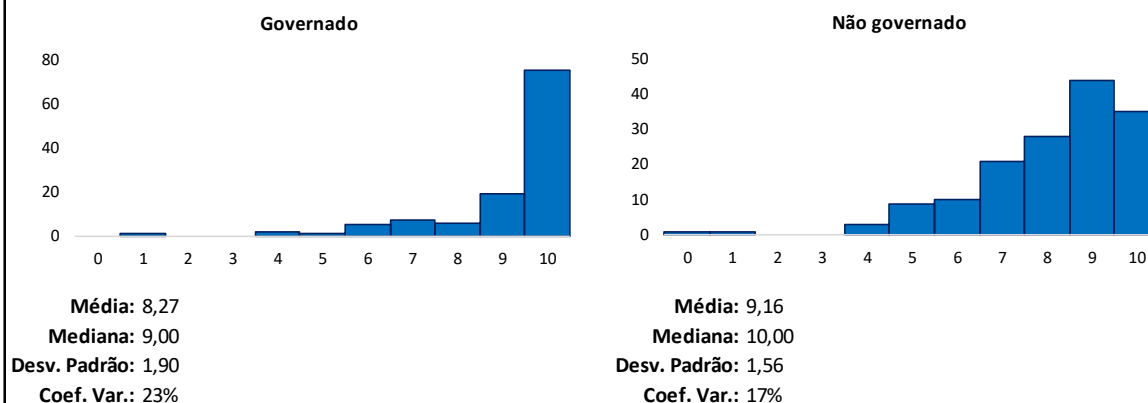
Média: 8,79
Mediana: 9,50
Desv. Padrão: 1,81
Coef. Var.: 21%

Não governado

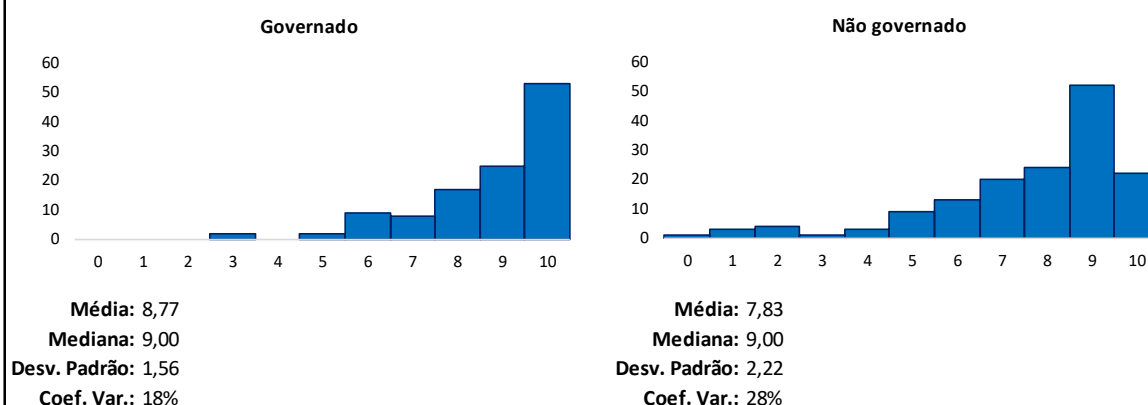
Nota	Frequência
0	1
1	0
2	3
3	1
4	3
5	8
6	14
7	26
8	19
9	40
10	36

Média: 8,01
Mediana: 9,00
Desv. Padrão: 2,10
Coef. Var.: 26%

Acesso legitimado: Todos os relatórios que acesso apresentam informações que são importantes para minhas atividades do dia a dia.



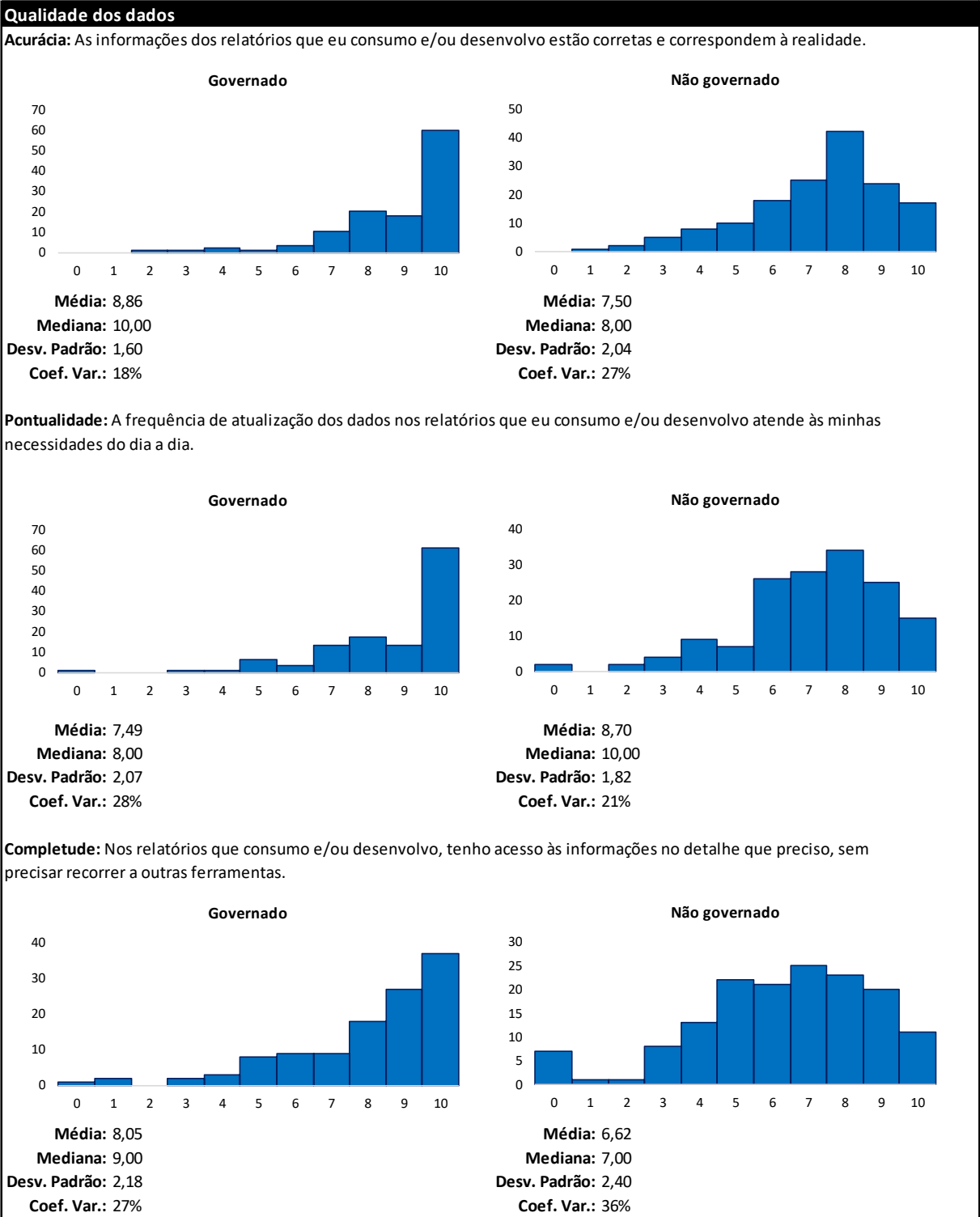
Obrigações contratuais: Tenho segurança que os dados contratualmente sigilosos dos relatórios que consumo e/ou desenvolvo não apresentam riscos de vazamento de dados.



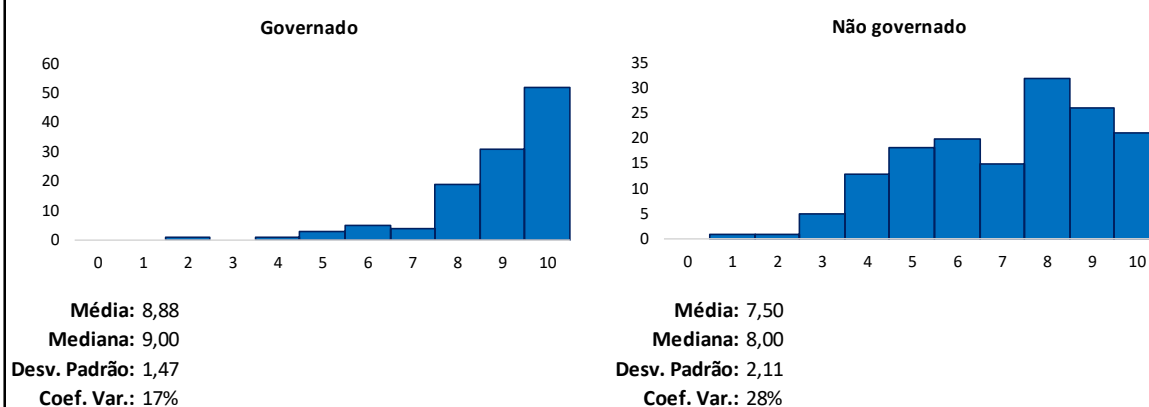
Fonte: Dados da pesquisa.

As respostas dos usuários nos itens relativos à variável “segurança e privacidade” apresentam baixa diferença entre os governados (8,40) e os não governados (7,78). Esse fator pode ser explicado pela natureza técnica dessa variável. Ainda assim, a diferença observada se deve ao processo de aplicação da governança, pois as equipes de negócio são treinadas e orientadas sobre os processos e os motivos desses processos de governança, mostrando os critérios de segurança que são aplicados e gerando a percepção de melhora.

Figura 24
Distribuição das respostas das questões referentes à Qualidade dos dados



Credibilidade: As informações apresentadas nos relatórios que eu consumo e/ou desenvolvo buscam os dados nos sistemas corretos.



Fonte: Dados da pesquisa.

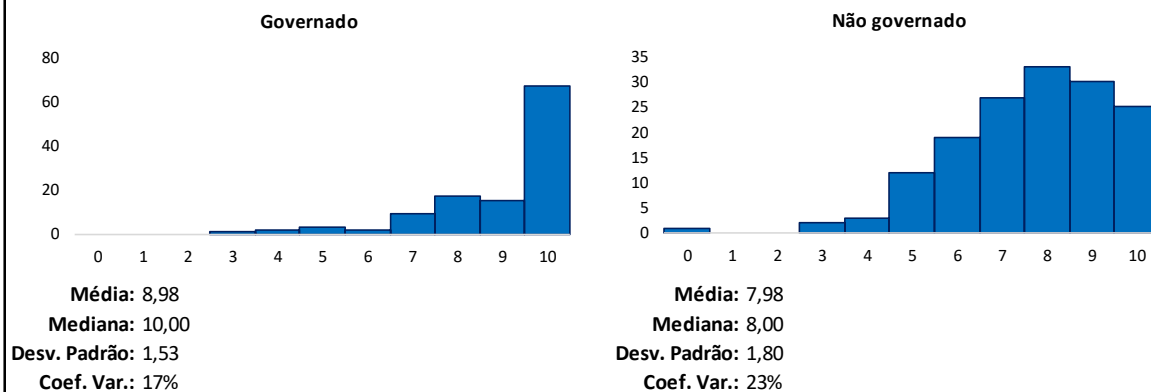
Outra avaliação que pode ser destacada é relativa ao item “completude” da variável “qualidade dos dados” (Figura 23) que apresentou uma baixa média na avaliação para os usuários não governados (6,45), e para os usuários governados, esse item teve resultado satisfatório (8,03). Esta baixa avaliação e grande diferença podem ser explicadas pelo fato de que os usuários de produtos de dados não governados obtêm seus dados diretamente de funcionalidades construídas nos sistemas transacionais. Por mais que várias informações são inseridas nos sistemas durante a operação da empresa, as funcionalidades construídas podem ter regras aplicadas e trarão granularidades da informação (níveis de detalhe) previamente definidas. Por outro lado, para a construção de relatórios em produtos de dados governados, o *Data Warehouse* da organização se integra ao banco de dados dos sistemas transacionais, capturando as réplicas das tabelas lá presentes. Assim, são replicados para o ambiente todos os dados em todas as granularidades, possibilitando obter as informações na completude desejada e facilidades para *zoom-in* e *zoom-out*. Dessa forma, para maior completude das informações, os usuários de produtos de dados não governados ainda podem depender de desenvolvimentos a serem aplicados nos sistemas que devem ser feitos pelas equipes técnicas.

Figura 25

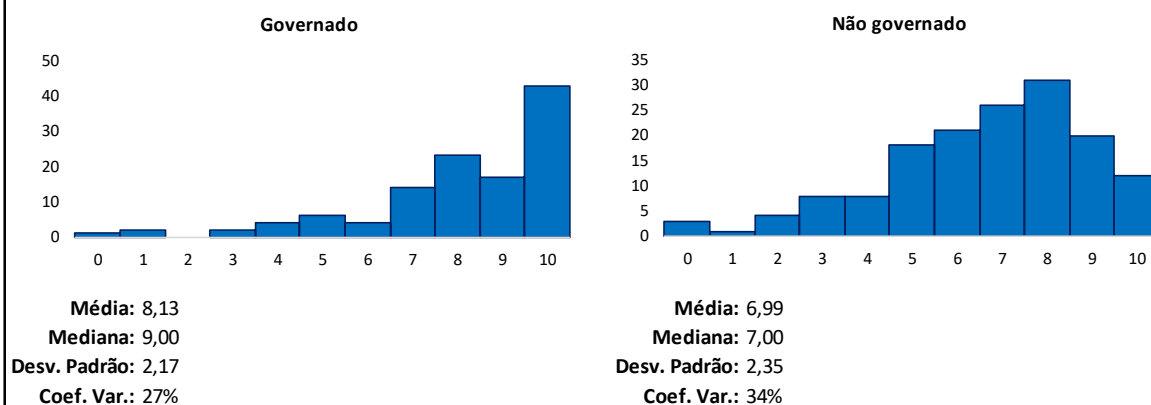
Distribuição das respostas das questões referentes à Gestão de Metadados

Gestão de metadados

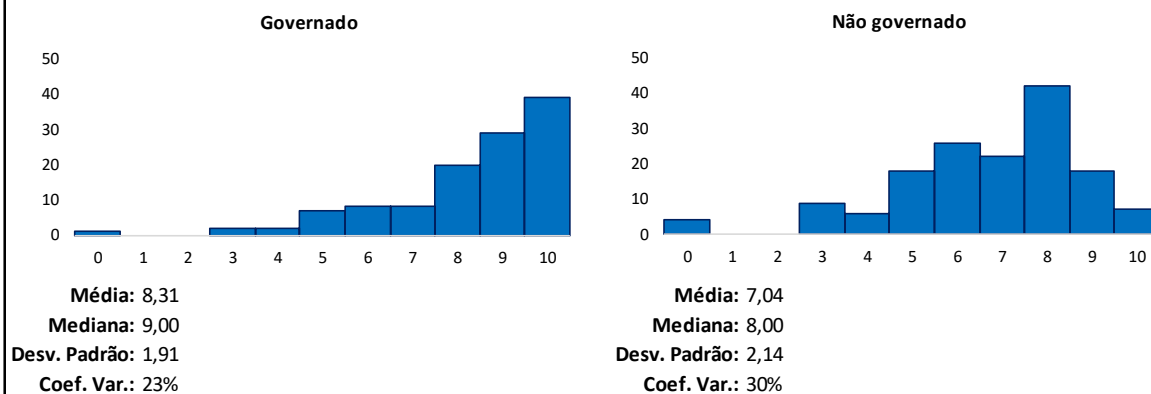
Conhecimento organizacional: Entendo exatamente o que representam as informações presentes nos relatórios que consumo e/ou desenvolvo.



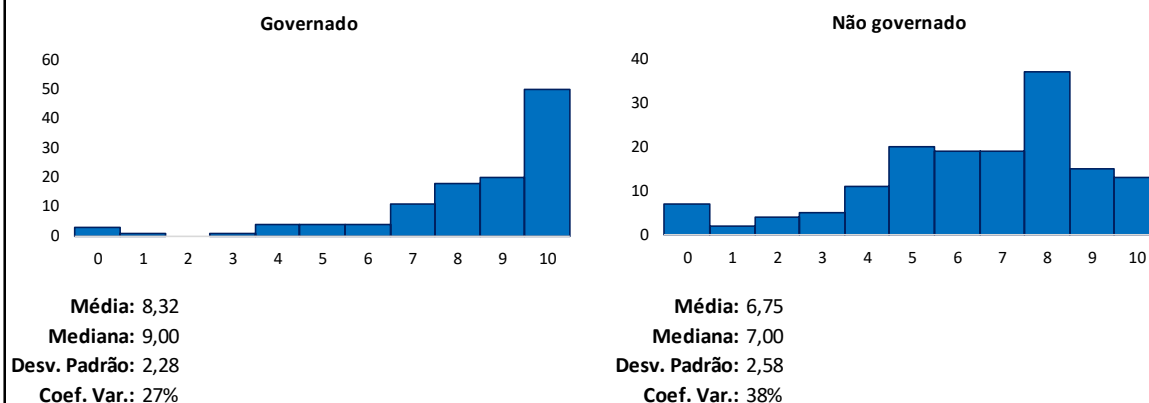
Integração entre fontes de dados: Os relatórios que consumo e/ou desenvolvo convergem nos conceitos similares (por exemplo, o cliente tem a mesma definição em todos os relatórios).



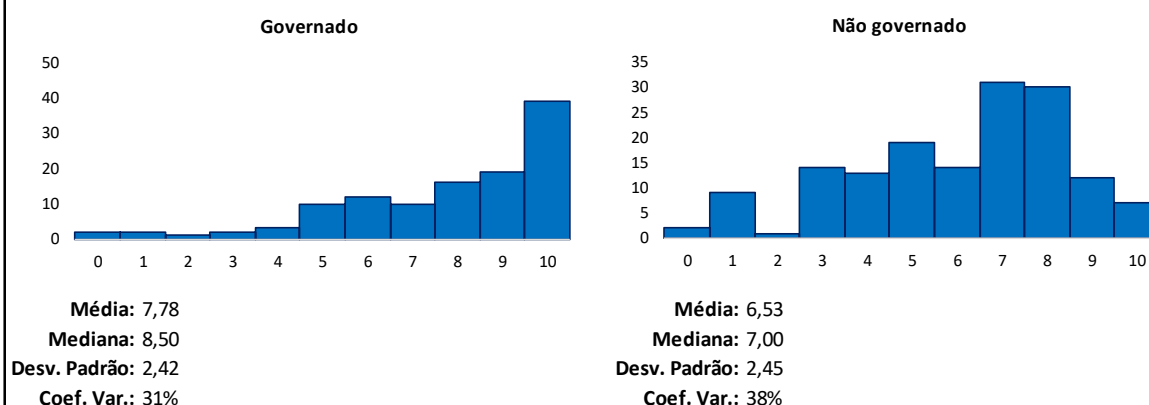
Qualidade dos metadados: O conteúdo dos relatórios que consumo e/ou desenvolvo apresentam claramente as definições necessárias para entender o que significam.



Acesso aos metadados: Tenho acesso às definições e significados dos dados referentes aos relatórios que consumo e/ou desenvolvo.



Troca de dados: Consigo aproveitar dados de relatórios já existentes na organização, não precisando criar uma nova forma de obtenção e tratamento dos dados.



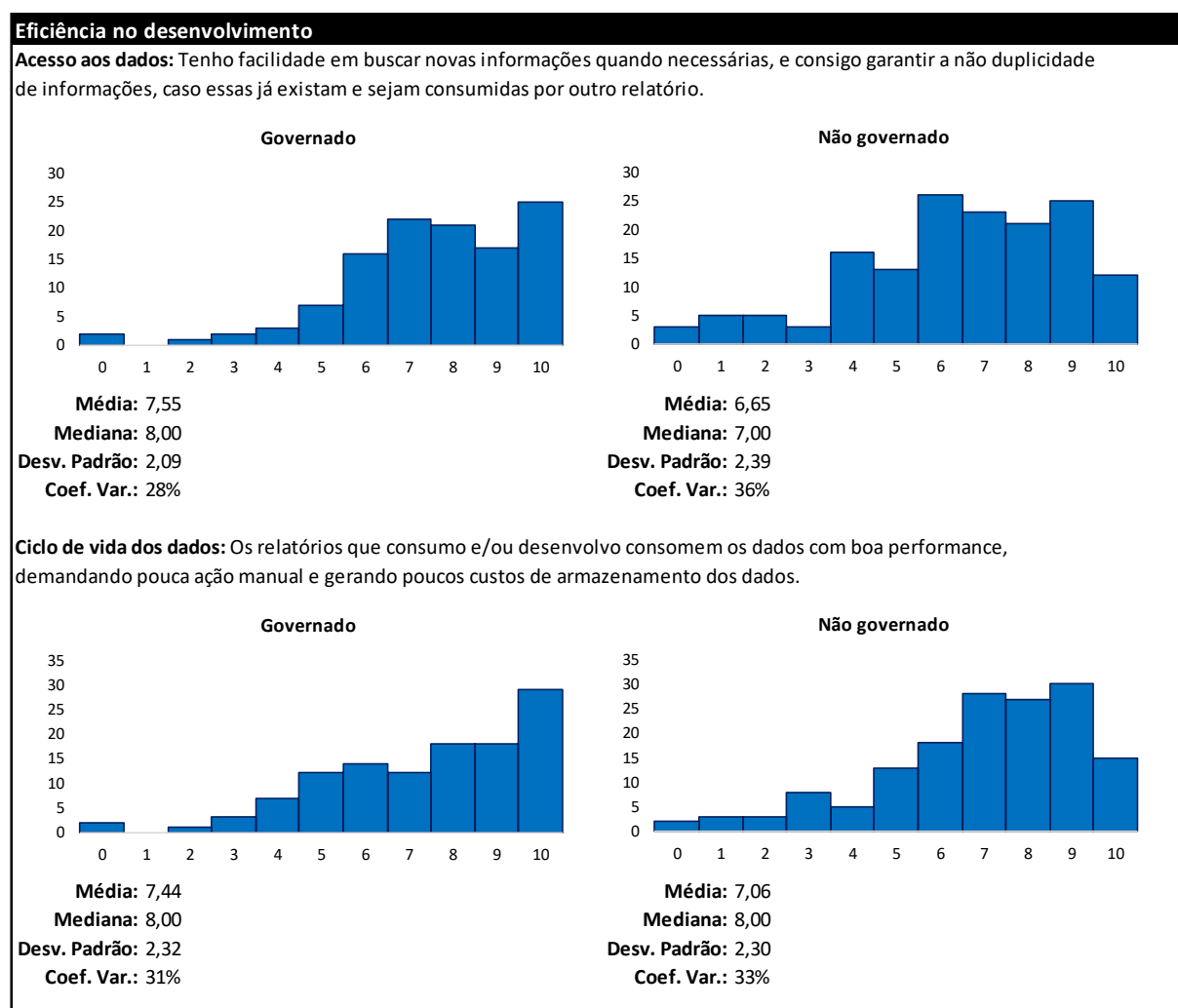
Fonte: Dados da pesquisa.

Foram observadas ainda grandes diferenças de médias nos itens relativos à variável “gestão de metadados” (Figura 24). Essas diferenças indicam uma baixa disseminação do conhecimento organizacional para os conceitos explorados individualmente em produtos não governados, podendo dificultar o entendimento dos usuários dos relatórios sobre o significado dos dados apresentados. Para os produtos governados, esse fator é melhorado, pois todos os conceitos, conhecimentos e metadados gerados são registrados em sistemas de fácil acesso. Além disso, destaca-se a menor avaliação da variável “troca de dados”, que obteve média de 5,90 para usuários de produtos não governados e 8,28 para usuários de produtos governados. Essa diferença se deve principalmente ao não reaproveitamento de dados para construção de relatórios de áreas de negócio distintas, quando da construção de produtos não governados, evidenciando potencial situação de retrabalho. Tal fato é eliminado para produtos governados, pois o processo de governança garante a verificação e reutilização dos dados para produtos diferentes que possuam os mesmos conceitos, além de uma metodologia de gerenciamento de

dados mestres aplicada. Por exemplo, caso a equipe de Contas a Receber da organização queira construir um relatório de prazo médio de pagamento por cliente e a equipe Comercial já tenha um relatório de pedidos por cliente, a primeira, caso produza relatórios no modelo não governado, precisará extrair os dados dos clientes do sistema de cadastro de clientes e inserir em seu relatório, extração esta que já é feita pela equipe do Comercial, gerando assim retrabalho. Caso sejam feitos relatórios no modelo governado, o processo garante que será utilizada a mesma extração de dados dos clientes, apenas relacionando a visões diferentes a depender da necessidade de cada área. Tal situação traz indícios de que o custo da governança de dados pode ser abatido com a economia gerada pela redução do retrabalho.

Figura 26

Distribuição das respostas das questões referentes à Eficiência no Desenvolvimento



Fonte: Dados da pesquisa.

Observa-se uma menor média nas avaliações da variável “Eficiência no desenvolvimento” (Figura 25). Este item possui uma relação conceitual com a variável de baixa avaliação para a parte de SSBI da pesquisa, “Acesso a fonte de dados”. Essa avaliação baixa pode estar relacionada ao histórico ainda recente de uso de ferramentas de BI e da governança de dados na empresa, apresentando baixa maturidade na obtenção de dados quando necessário pelos usuários e pelos desenvolvedores de relatórios na empresa.

Assim como na análise dos resultados das variáveis de SSBI, para as variáveis de governança de dados foi aplicada Análise dos Componentes Principais (ACP) para redução dos itens avaliados em uma única variável, sendo depois calculado Coeficiente de Correlação Ponto Bisserial, com base na metodologia proposta, correlacionando com a variável Y2. Os produtos de dados do respondente seguem o modelo de governança de dados da organização, sendo 1 quando o respondente utiliza produtos que seguem o modelo de governança, e 0 quando não seguem. Os cálculos também foram feitos no *software* R e os resultados podem ser observados na Tabela 9.

Tabela 9
Resultados da avaliação de Governança de Dados

Variável	Quantidade Componentes	Representatividade Componente Principal	Valor-p	Correlação
Segurança e Privacidade	5	71%	0,0002	0,2258
Qualidade dos Dados	4	70%	1,31e-12	0,4155
Gestão de Metadados	5	69%	4,10e-15	0,4552
Eficiência no Desenvolvimento	2	84%	0,0226	0,1393

Fonte: Dados da pesquisa.

Com o nível de significância de 0,05, a aplicação do teste permitiu a rejeição da hipótese nula para todas as variáveis. Portanto, as 4 hipóteses apresentadas referentes à aplicação da governança de dados para o uso e desenvolvimento de ferramentas de SSBI se mostram pertinentes, visto que não se rejeita a hipótese nula. Isso demonstra que a aplicação da governança de dados gera uma melhor percepção dos usuários no uso de relatórios de BI, tanto relacionado à segurança e qualidade dos dados, quanto relativo ao acesso aos metadados e aos próprios dados em si.

Para as variáveis “Segurança e Privacidade” e “Eficiência no Desenvolvimento”, foi encontrada uma correlação fraca. Para a primeira, a correlação fraca pode indicar uma menor

percepção dos usuários frente à variável segurança e privacidade, visto o aspecto técnico dessa variável. Esse fator poderá ser observado na segunda parte da pesquisa, com as entrevistas com as equipes técnicas da empresa.

A variável “eficiência no desenvolvimento” apresentou uma correlação fraca, demonstrando que a governança não apresenta grande relevância nesta variável. Conceitualmente, pode haver uma relação entre esta eficiência de desenvolvimento e a baixa percepção relativa ao acesso as fontes de dados observada para as ferramentas de SSBI. Esse fator pode estar relacionado ao histórico ainda recente de uso de ferramentas de BI e da governança de dados na empresa, apresentando baixa maturidade na obtenção de dados quando necessário pelos usuários e pelos desenvolvedores de relatórios na empresa.

Visto valor-p abaixo de 0,05 observado nas análises univariadas referentes às variáveis de governança de dados, todas as variáveis podem ser utilizadas para aplicação do modelo de regressão logística multivariada, a fim de avaliar o impacto geral da governança de dados no uso de ferramentas de SSBI. Assim, foi feita a aplicação do modelo para análise da correlação dos componentes obtidos para cada uma das variáveis com a variável Y2. O resultado obtido pode ser observado na Tabela 10, conforme aplicado no software R.

Tabela 10
Resultados da regressão logística multivariada

Variável	Valor-p
Geral	5,91e-08
Segurança e Privacidade	0,1996
Qualidade dos Dados	0,0048
Gestão de Metadados	6,10e-09
Eficiência no Desenvolvimento	8,42e-06

Fonte: Dados da pesquisa.

Conforme resultado obtido, pode-se assumir uma correlação entre a governança de dados e uma melhor percepção dos usuários de ferramentas de SSBI, pela não rejeição da hipótese nula. Essa correlação, conforme pseudo R^2 de Tjur pode ser considerada positiva e moderada, indicando que a aplicação da governança de dados impacta positivamente o uso de ferramentas de SSBI, com base na Hipótese 9 (H9: A aplicação da governança de dados impacta

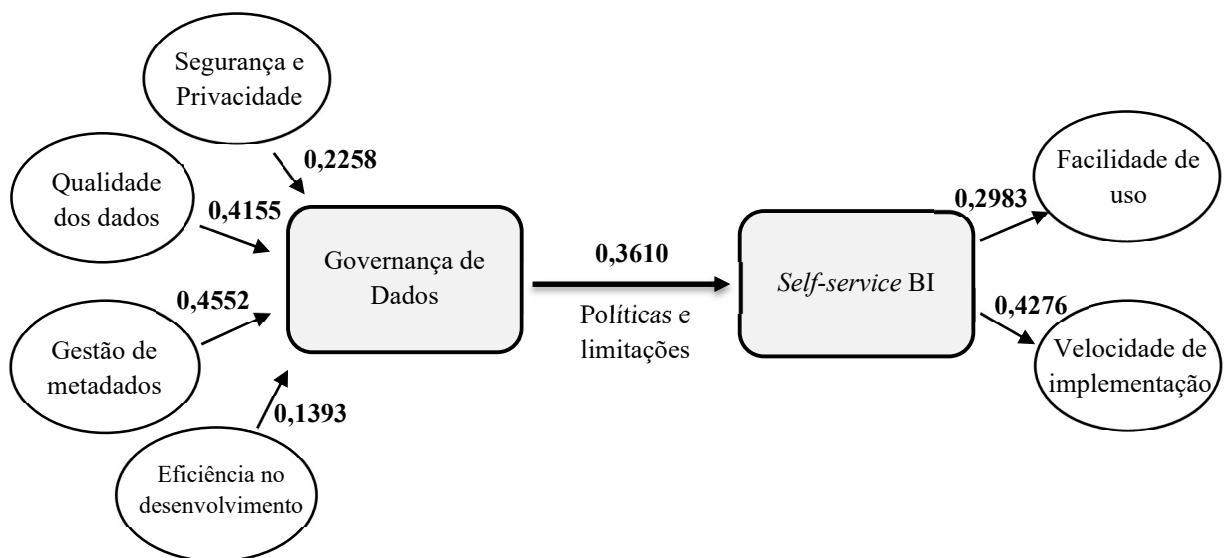
positivamente a utilização do SSBI, gerando uma melhor percepção sobre a qualidade das ferramentas).

Além disso, conforme cálculo de valor-p individual para cada variável dentro do modelo de regressão, pode ser observado um alto valor-p para a variável segurança e privacidade, o que significa que a correlação positiva da aplicação da governança de dados com o uso das ferramentas de SSBI não pode ser explicada por essa variável. Essa avaliação pôde ser observada na análise univariada quando a segurança e privacidade foi isolada, visto que, apesar de os dados permitirem assumir a correlação entre a variável e o uso de SSBI, esta é considerada fraca.

Pode-se construir o modelo final de relação da governança de dados com o uso de ferramentas de SSBI como apresentado na Figura 26.

Figura 27

Modelo Conceitual Final



Fonte: Dados da pesquisa.

Nota: Todos os valores apresentados foram significativos ao nível de 0,05.

Tabela 11
Tabela de hipóteses final

Hipótese	Descrição	Resultado
H1	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao consumo dos relatórios.	Rejeitada
H2	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à facilidade de uso destas ferramentas.	Não rejeitada
H3	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à velocidade de implementação de relatórios.	Não rejeitada
H4	A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao acesso às fontes de dados.	Rejeitada
H5	A aplicação da governança de dados gera uma melhor percepção relativa à segurança e privacidade dos dados na utilização do SSBI.	Não rejeitada
H6	A aplicação da governança de dados gera uma melhor percepção relativa à qualidade dos dados na utilização do SSBI	Não rejeitada
H7	A aplicação da governança de dados gera uma melhor percepção relativa à gestão de metadados na utilização do SSBI	Não rejeitada
H8	A aplicação da governança de dados gera uma melhor percepção relativa à eficiência no desenvolvimento na utilização do SSBI.	Não rejeitada
H9	A aplicação da governança de dados impacta positivamente a utilização do SSBI, gerando uma melhor percepção sobre a qualidade das ferramentas.	Não rejeitada

Fonte: Dados da pesquisa.

Portanto, é demonstrada a percepção positiva dos usuários com o uso de ferramentas de SSBI, quando confrontadas com ferramentas tradicionais de BI para uso de equipes de TI, bem como uma percepção positiva dos benefícios gerados pela aplicação da governança de dados no uso das ferramentas de SSBI, e quais variáveis evidenciam esses benefícios.

Durante aplicação do *survey*, foi perguntado em uma questão aberta como os respondentes enxergam a governança de dados na organização. Quando se avalia isoladamente os desenvolvedores de produtos de SSBI governados, observa-se uma percepção positiva do ponto de vista da confiança na qualidade dos dados utilizados. Porém, é percebida uma baixa flexibilidade e autonomia dos produtos, o que conversa com o modelo de *self-service* adotado pela empresa, que apresenta visão centrada na governança. No caso estudado, mostra-se importante evoluir em manter uma segurança e qualidade dos dados, dando maior autonomia para mudanças que competem aos *Citizen Developers*.

Por outro lado, os desenvolvedores de produtos de SSBI não governados possuem pouca visibilidade da governança aplicada na organização, mostrando uma necessidade de melhor comunicação do modelo de governança. Os respondentes também foram questionados das dificuldades enfrentadas para a gestão de indicadores e dados. Os desenvolvedores de produtos de SSBI não governados apontaram principalmente a dificuldade de uso devido à necessidade de alimentação manual dos relatórios, baixa automação, falta de centralização dos dados e muita utilização de dados duplicados, referindo-se à mesma informação, mas sendo buscados de

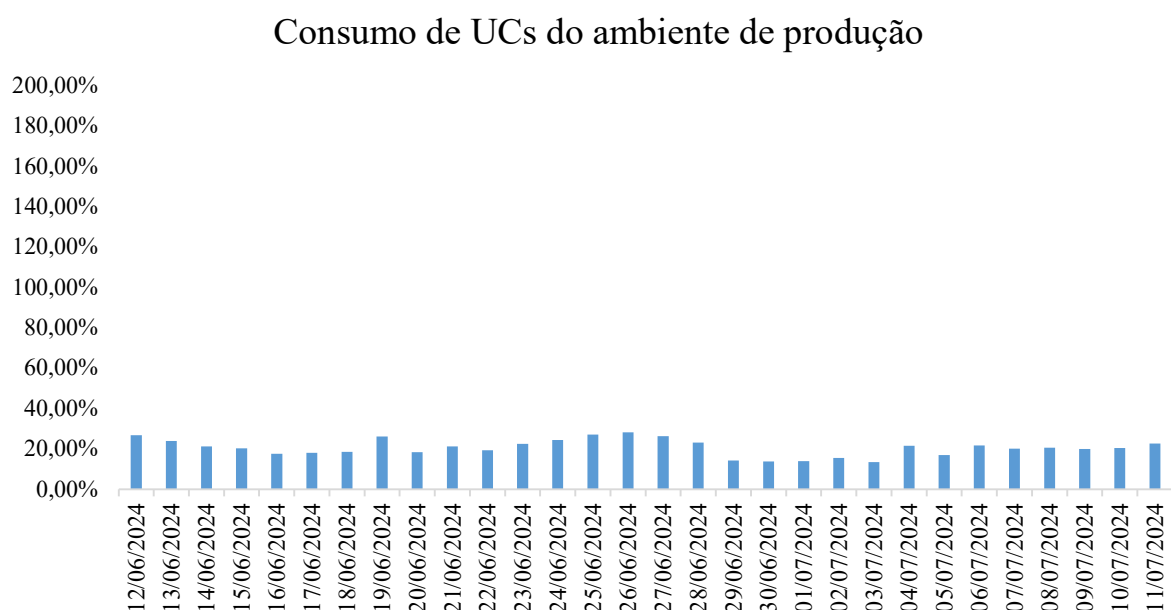
fontes diferentes e com resultados divergentes. Trata-se claramente de um cenário de baixa produtividade e retrabalho que pode comprometer a qualidade e o *timing* da tomada de decisão. Em suma, os fatores apontados são fatores endereçados pela governança de dados e que apresentam maior percepção de valor quando aplicados, pela qualidade dos dados e a gestão dos metadados.

4.4. Análise Documental

Por meio de análise feita no Portal do Administrador do Microsoft Fabric, o ambiente do MS Power BI da organização é composto de dois ambientes contratados para processamento dos produtos de SSBI desenvolvidos: um ambiente de experimentação, onde são publicados os produtos desenvolvidos fora do modelo de governança de dados; e um ambiente de produção, onde são publicados os produtos governados. Ambos os ambientes possuem a mesma configuração, sendo um ambiente do tipo P1, conforme definições da Microsoft, que possui 8 núcleos virtuais e 64 UCs (Unidades de Capacidade, do inglês *Capacity Unit*), que conforme documentação apresentada, totaliza 25 *gigabytes* de memória RAM. Nas Figuras 27 e 28, pode ser observado o consumo de UCs dos ambientes de produção e experimentação, respectivamente, ao longo de 1 mês e na tabela 12 podem ser observados os números de performance dos ambientes.

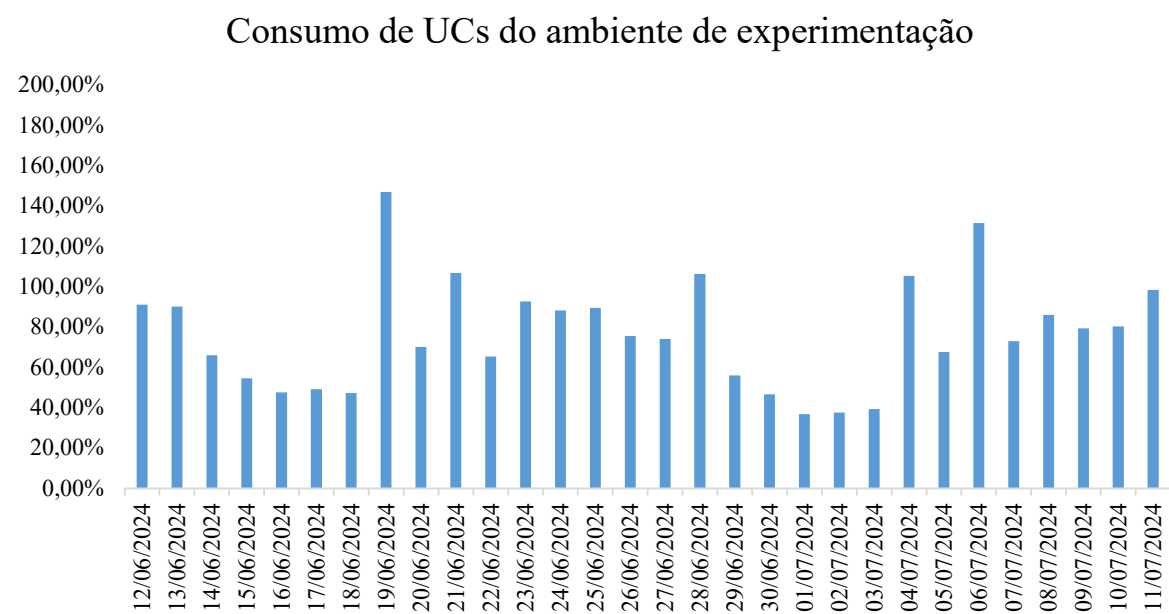
Figura 28

Consumo de UCs do ambiente de produção



Fonte: Dados da pesquisa.

Figura 29
Consumo de UCs do ambiente de experimentação



Fonte: Dados da pesquisa.

Tabela 12
Performance dos ambientes

Ambientes	Produção	Experimentação
Quantidade de <i>Datasets</i>	79	253
Tamanho Total (Gb)	27,7	146,5
Tamanho médio (Gb)	0,35	0,58
Consumo médio de UCs	20,5%	76,6%
Tempo médio de atualização (min)	26,4	46,2

Fonte: Dados da pesquisa.

Analisando os números, é possível observar que no ambiente de produção o tamanho médio dos *datasets* referentes aos dados dos produtos de SSBI é menor que no ambiente de experimentação, indicando menor volumetria de dados armazenada diretamente no ambiente do MS Power BI. Esse fator se deve à presença do componente do DW Corporativo antes da composição da camada semântica para os relatórios, conforme demonstrado na Figura 15. Com esse componente, os dados brutos dos sistemas estão armazenados no DW, além dos principais cálculos e filtros. Assim, a camada semântica do Power BI recebe apenas os dados filtrados e na granularidade necessária para o relatório, otimizando o armazenamento dos dados. Esta otimização do armazenamento está diretamente relacionada com a performance do ambiente, refletindo no menor tempo médio de atualização dos relatórios e no menor consumo do ambiente.

O menor percentual de consumo do ambiente, conforme demonstrado, está diretamente relacionado ao custo despendido pela organização. Enquanto o ambiente de produção tem um

consumo médio de 20,5% de UCs do ambiente, possibilitando maior capacidade sobressalente para absorção de novos dados e relatórios, o ambiente de experimentação possui consumo médio de 76,6%, apresentando em determinados momentos picos de consumo de até 150% do ambiente, demonstrando não haver capacidade sobressalente para absorver as oscilações naturais de consumo nem capacidade para absorver novos dados e relatórios.

Visto os fatores observados na pesquisa quantitativa para relatórios não governados de baixa troca de dados e integração entre as fontes de dados, e baixa disseminação do conhecimento organizacional, o alto quantitativo apresentado de *datasets* e de volume de informações armazenada no ambiente de experimentação indica a probabilidade de uma quantidade grande de relatórios duplicados, de baixa qualidade e desnecessários. Com os picos de consumo indicados do ambiente, é de se esperar uma necessidade de aumento ou criação de novos ambientes para suportar o volume de processamentos, aumentando os custos da organização. Além disso, considerando que esses relatórios possuem atualizações manuais dos dados, há um gasto de horas dos empregados excessiva com essas atualizações, o que pode ser automatizado com a aplicação da governança, para geração de relatórios que por vezes podem não ser mesmo consumidos, ou que podem estar gerando informações incoerentes, conflitantes, duplicadas ou de baixa qualidade. Em suma, há indícios de que o custo da governança de dados pode gerar economias oriundas não só da melhor produtividade, mas também da redução dos custos de processamento e, por conseguinte, de energia elétrica em uma perspectiva de sustentabilidade.

Além disso, a diferença na quantidade de *datasets* pode demonstrar uma menor otimização das informações, de modo que o significativo maior número de *datasets* do ambiente de experimentação infere na possibilidade de duplicidade de informações, demandando maior consumo do ambiente sem necessidade, e ao final, um maior custo para a organização.

5. ANÁLISE DOS RESULTADOS QUALITATIVOS

Nesse capítulo são apresentados os resultados da parte qualitativa da pesquisa, com base nas entrevistas conduzidas. O capítulo está dividido em duas seções: análise dos resultados das entrevistas feitas com as equipes técnicas de governança de dados e segurança da informação a respeito dos impactos da governança para a segurança e privacidade dos dados; análise dos resultados das entrevistas feitas com os desenvolvedores que migraram seus produtos de SSBI do modelo não governado para o modelo governado.

5.1. Entrevistas Técnicas

As entrevistas foram feitas com seis pessoas das equipes de governança de dados e segurança da informação (Tabela 13), de forma *online*, entre os dias 14 e 31 de outubro de 2024, a respeito dos impactos da governança de dados aplicada a SSBI do ponto de vista de segurança e privacidade dos dados. O roteiro semiestruturado das entrevistas se encontra no Apêndice B.

Tabela 13
Perfil dos entrevistados técnicos

#	Área da organização	Cargo	Tempo de empresa (anos)	Tempo de atuação profissional (anos)	Duração da entrevista (min)
1	Dados	Especialista de Dados	8	19	49
2	Dados	Especialista de Dados	3	21	55
3	Dados	Administrador de Dados	3	32	32
4	Dados	Arquiteto de Dados	5	32	45
5	Segurança da Informação	Especialista de Segurança	4	23	39
6	Segurança da Informação	Arquiteto de Segurança	2	20	36
-	-	Média dos Tempos	4	25	43

Fonte: Dados da pesquisa.

Ao serem apresentados ao resultado da pesquisa quantitativa, os entrevistados entenderam ser esperado o resultado superior em todas as variáveis quando avaliado o público que possui BIs no modelo governado. No entanto, os entrevistados 1 a 4, que fazem parte da

equipe de Dados, ficaram surpresos ao verem uma menor diferença de avaliação do ponto de vista da variável Eficiência no Desenvolvimento. Eles acreditavam que essa variável seria a de maior diferença quando comparado o modelo governado do modelo não governado, tendo em vista a disponibilização dos dados ser feita por uma equipe técnica, não sendo necessário que o desenvolvedor de negócio busque os dados que precisa manualmente.

Quando questionados sobre a gestão de acesso aos dados no modelo de governança de dados da organização (questão 4), os entrevistados pertencentes à equipe de Governança de Dados afirmaram que o modelo de governança prevê um nível de segurança no acesso aos dados, iniciando com os dados oriundos do sistema transacional. O arquiteto de dados (entrevistado 4) explicou que o acesso ao dado do sistema transacional é feito por um usuário de serviço, ao invés de um usuário nominal, que realiza a integração entre o sistema e o banco de dados analítico. A senha desse usuário é inserida no sistema analítico por um profissional de segurança devidamente capacitado e é mascarada. Quando questionada com a mesma pergunta, uma evolução apontada pela equipe de Segurança da Informação para esse acesso, é por meio da utilização de uma ferramenta de PAM (*Privileged Access Management*), de modo que nenhuma pessoa na organização tenha a senha para não possibilitar um acesso direto ao banco de dados transacional.

Após essa integração, o administrador de dados (entrevistado 3) informou que os engenheiros de dados apenas possuem acesso ao ambiente analítico após liberação feita pelo DBA (*DataBase Administrator*), desde que devidamente autorizada por um gerente de tecnologia da organização, e todos esses acessos são documentados. Esse acesso é revisado periodicamente para garantir que nenhum usuário tenha um acesso que não necessita mais ter, conforme depoimentos a seguir.

E_3: “...A cada dois meses, é feito um trabalho de validação da continuidade dos acessos, questionando as lideranças se o profissional precisa continuar tendo aquele acesso.”

E_2: “...No momento da liberação dos acessos, é feita uma avaliação de sensibilidade do dado. Dados classificados como sensíveis têm um nível a mais de autorização necessária. Por exemplo, dados de funcionários, como salário, são colocados em uma estrutura apartada e a liberação de acesso precisa ser validada pela diretora de RH.”

A questão 2 abordou o impacto da governança de dados na aplicação dos critérios de segurança provenientes de regulamentações. Para esse cenário de atendimento à LGPD (Lei Geral de Proteção de Dados), os entrevistados explicaram que todos os dados de pessoa física possuem uma marcação no ambiente e toda modelagem e utilização desses dados passa por aprovação do DPO (*Data Protection Officer*) da organização. Assim, os entrevistados informaram que, a qualquer momento, podem fazer uma consulta para saber quais dados de pessoa física estão sendo utilizados e para qual finalidade, conforme ilustrado pelo trecho a seguir.

E_1: “...Conseguimos fazer uma consulta rápida e saber todos os dados de pessoa física que temos hoje no ambiente analítico, como e para qual finalidade está sendo utilizado. Temos também o registro das aprovações do DPO para a utilização desses dados.”

Quando questionados do impacto do modelo de governança na segurança dos dados estratégicos da organização (questão 3), os entrevistados 5 e 6 pertencentes à equipe de Segurança da informação afirmaram que o armazenamento dos dados no modelo governado possui maior nível de segurança quando comparado com o modelo não governado. Segundo o especialista em segurança (entrevistado 5), no modelo não governado, os dados ficam armazenados no computador do usuário desenvolvedor, expondo a alguns riscos, como o risco de intrusão devido às redes as quais o usuário se conecta e um risco de vazamento de senha dos usuários. Já no modelo governado, os dados ficam armazenados em um banco de dados com critérios de segurança elevados, onde os dados do ambiente de produção só podem ser acessados por usuários de serviço, garantindo a segurança das senhas que possuem acesso. Além disso, o entrevistado 6 informou que estes ambientes estão hospedados em uma rede interna à organização, não estando expostos à Internet.

Além disso, quando a equipe de Governança de Dados foi questionada pela mesma pergunta, informaram que todas as execuções que são feitas no ambiente analítico apresentam registro de *log* no ambiente, de modo que caso haja algum vazamento de dados, é possível avaliar quem executou alguma ação no ambiente como uma consulta ou uma extração de cada dado, conforme evidenciado nos seguintes relatos.

E_5: “...Muitas vezes, os usuários usam o mesmo usuário e senha da empresa para se cadastrarem em diversos sites da Internet. Isso gera um risco de exposição do

usuário e senha, que pode ser utilizado para tentativas de intrusão nos sistemas da empresa e o consequente vazamento de dados.”

E_4: “...Conseguimos ter acesso a tudo que é feito nos ambientes analíticos. Nenhuma consulta é feita sem que essa ação seja registrada para podermos rastrear posteriormente, caso necessário.”

5.2. Migração de Modelo de Governança

As entrevistas foram feitas com 7 pessoas distribuídas em diferentes funções e níveis na organização, que são responsáveis pelo desenvolvimento de relatórios em SSBI e migraram seus relatórios para um modelo governado ao longo dos últimos 2 anos (Tabela 14). Foram feitas entrevistas no formato *online*, entre os dias 01 e 10 de outubro de 2024

Tabela 14

Perfil dos entrevistados usuários

#	Área da organização	Cargo	Experiência profissional (anos)	Experiência com BI (anos)	Início da migração	Duração da entrevista (min)
1	Suprimentos	Analista de Suprimentos	13	5	2022	58
2	Suprimentos	Gerente de Suprimentos	23	3	2022	45
3	Financeiro	Analista Financeiro	29	3	2022	62
4	Segurança Operacional	Analista de Segurança	8	4	2024	55
5	Segurança Operacional	Analista de Segurança	30	3	2024	58
6	Recursos Humanos	Especialista de Gente	21	8	2023	48
7	Recursos Humanos	Analista de Gente	7	3	2023	52
-	-	Média dos Tempos	19	4	-	54

Fonte: Dados da pesquisa.

Primeiramente, quando apresentados aos resultados quantitativos da pesquisa, todos os entrevistados concordaram com os resultados apresentados. Os entrevistados que iniciaram a migração em 2022 afirmaram que apresentaram algumas dificuldades no processo de migração,

porém, após concluída a migração, os resultados foram bons e estão bem refletidos nas avaliações quantitativas.

Quando perguntados a respeito do nível de conhecimento das políticas e diretrizes de dados da organização (questão 2), todos os entrevistados mostraram o mesmo nível de conhecimento. Porém, as primeiras migrações, que iniciaram no ano de 2022, foram feitas sem um treinamento prévio do modelo governança para os desenvolvedores, enquanto as migrações iniciadas em 2023 e 2024 foram precedidas de uma capacitação. Assim, para os três entrevistados que iniciaram a migração em 2022, seria importante ter esse conhecimento prévio, para que a especificação das necessidades seja mais bem feita e evite retrabalhos futuros.

E_1: “...acho que esse conhecimento eu precisava ter lá no começo, para melhorar o que eu iria pedir. Nós tivemos que refazer algumas modelagens, pois eu havia especificado uma necessidade de ter todos os dados, para que eu pudesse ter mais autonomia no BI, mas os relatórios ficaram pouco performáticos¹. Depois eu entendi que eu poderia dividir minha necessidade dos dados, para ter dados mais sintéticos alimentando os relatórios.”

Além disso, esse conhecimento prévio do modelo de governança pode demonstrar para os desenvolvedores a importância de se ter as regras de negócio claramente definidas antes de serem feitos os desenvolvimentos. Pode-se entender isso pelas respostas à questão 5, a respeito da percepção dos usuários dos relatórios que produzem após a mudança. Nesta questão, os mesmos três entrevistados que iniciaram a migração em 2022 responderam que existe um incômodo dos usuários por uma demora nas mudanças que porventura precisam ser feitas em regras dos indicadores. A restrição da autonomia, que os desenvolvedores têm após a migração, faz com que mudanças futuras de regras sejam mais demoradas para serem executadas. Por outro lado, os entrevistados que iniciaram a migração em 2023 e 2024 não abordaram essa preocupação. Quando questionados sobre este aspecto, eles informaram que não têm esse problema, pois não possuem mudanças constantes nas regras.

Dessa forma, mostra-se importante que o desenvolvedor meça os indicadores manualmente junto aos futuros usuários dos relatórios, de modo a amadurecerem as regras a serem aplicadas, para depois fazerem a implantação de relatórios em uma arquitetura de dados robusta e governada, para que as mudanças de regras não sejam frequentes.

E_2: “...A morosidade do processo é enxergada pelos usuários por eles quererem constantemente mudar a regra. E não é mais tão simples, eles precisam esperar um tempo para mudar.”

E_6: “...Nós já gerávamos os indicadores há muitos anos e apenas migramos o cálculo, então já estava muito maduro. Então, não temos mudança nenhuma nos cálculos, não tem morosidade de mudanças.”

Por outro lado, ao responderem à pergunta sobre as principais mudanças que a migração trouxe para seu dia a dia e para os produtos (questão 4), todos os desenvolvedores entrevistados e os usuários dos relatórios que eles produzem observaram uma grande evolução na velocidade e frequência de atualização dos dados, gerando mais possibilidades de análises. O entrevistado 1 afirmou que o modelo governado permite agora que ele tenha atualizações diárias ou até mesmo várias vezes ao longo do dia, pois, como a atualização dos relatórios é feita por integrações entre as ferramentas, e não mais uma atualização manual do desenvolvedor, estas podem ser processadas automaticamente durante à noite. Isso faz com que os dados possam refletir melhor o dinamismo do negócio.

Além disso, os entrevistados informaram que o tempo que era gasto para atualizar esse relatório agora é utilizado pela equipe para analisar os resultados e gerar ações para melhoria, otimizando o trabalho dos profissionais e gerando possibilidades de custos evitados pelas áreas, possibilitando até mesmo uma otimização do quadro de funcionários para execução das atividades necessárias, conforme comprovado pelos relatos a seguir.

E_1: “...Hoje, BI não é mais problema aqui na área. Na verdade, surgiu outro problema: como os BIs atualizam tão rápido e com tanta frequência, as pessoas veem que o resultado mudou e começam a perguntar o porquê mudou. Estavam acostumados com uma frequência de atualização muito baixa, e aí quando atualizava, às vezes algumas mudanças eram imperceptíveis. Agora, houve uma transação nova, ele atualiza, e aí a pessoa começa a ver os números dela mudarem muito rapidamente.”

E_3: “...Em um relatório que tínhamos, gastávamos um dia inteiro de uma pessoa para extrair os dados, fazer os cálculos e montar o relatório. Por isso, atualizávamos ele apenas no fechamento do mês. Hoje, com o modelo governado, ele é totalmente automatizado e podemos ter ele atualizado diariamente.”

Quando questionados sobre como o uso do *Power BI* mudou após a migração para o modelo governado (questão 1), quatro dos entrevistados, que possuem uma maior quantidade de relatórios que consomem uma grande quantidade de dados, apresentaram como benefício o fato de os relatórios consumirem menos recursos das máquinas, tornando o desenvolvimento mais fácil e rápido. Como a grande massa de dados é processada e armazenada dentro do ambiente analítico, e a ferramenta de BI se conecta por modelo *direct query* aos dados, ela se torna mais leve para gerar novos relatórios ou modificar relatórios já existentes.

Além disso, alguns dos entrevistados trouxeram como benefício o fato de os desenvolvedores agora se preocuparem com a criação ou mudança de visuais dos relatórios, e não mais em diariamente verificar fazer os cálculos e verificarem se foi calculado corretamente. Isso gera um ganho de produtividade para os desenvolvedores, segundo os depoimentos coletados.

E_1: “...Tenho uma performance muito melhor. Eu abri um BI que está no modelo governado para construir um visual novo, e é muito mais rápido, muito mais leve. Tenho muito mais facilidade de construir.”

E_7: “...Além disso, tiramos o erro da nossa mão, pois compor (o indicador) manualmente, possibilita o erro humano de cálculo e composição do relatório. Agora, não tem mais erro.”

A questão 7 se referiu a como tem sido o suporte e envolvimento da TI e seus fornecedores para o desenvolvimento dos relatórios. Os desenvolvedores que iniciaram a migração em 2022 e 2023 e passaram por um período de mudança da equipe técnica relataram a importância da manutenção de uma equipe com bom conhecimento da organização. Assim, para este caso de estudo em que equipe técnica é composta de profissionais de um fornecedor, e não próprios da organização, uma mudança de fornecedor pode gerar um impacto na produtividade do processo de migração.

Os entrevistados afirmaram que o profissional precisa construir o conhecimento do negócio e dos sistemas que existem na empresa, para que consiga entender a necessidade do desenvolvedor do relatório, buscar as fontes dos dados, e fazer bons modelos de dados, que atendam as necessidades finais e que sejam eficientes em termos de desempenho de máquina. Assim, segundo eles, quando ocorre uma mudança no fornecedor devido a um *turnover* da

equipe técnica, ou para outros fornecedores, o conhecimento construído se perde e a produtividade cai durante a nova obtenção de conhecimento.

E_4: “...A mudança do arquiteto gerou uma grande dor, pois foi necessário construir uma nova modelagem e o novo arquiteto não sabia nem da existência de alguns sistemas muito relevantes da empresa. A falta de conhecimento do todo da empresa fez com que tivéssemos que fazer e refazer algumas vezes a modelagem.”

Por fim, com relação às dificuldades enfrentadas atualmente para a gestão e desenvolvimento de indicadores (questão 8), alguns entrevistados responderam que é importante, para além da governança dos dados, ser feita uma governança dos indicadores mais importantes da organização, principalmente no que diz respeito àquelas informações que são direcionadas à alta direção da organização. Os entrevistados informaram que uma dificuldade que ainda enfrentam é relativa às informações que, mesmo tendo a mesma natureza, são vistas de formas diferentes por diferentes áreas, demandando um grande esforço para traduzir essa diferença para a alta direção, conforme detalhado no próximo depoimento.

E_2: “...Quando avaliamos o estoque de materiais de manutenção, nós, do time de Suprimentos, avaliamos apenas o estoque que está nos armazéns. Porém, o Financeiro avalia o estoque contábil, que leva em consideração materiais que estão em uso pelo fornecedor, em trânsito, por exemplo. Isso faz com que o número final de valor de estoque seja diferente entre as áreas. Precisamos ter um normativo que defina como será calculado o estoque que será observado pela alta direção, e que tenha uma forma lógica de cascatear esse número para as diferentes visões que os times precisam ver e agir.”

Dessa forma, com base nas respostas das entrevistas, alguns requisitos mínimos podem ser definidos para se ter sucesso na migração de produtos de SSBI de um modelo não governado para um modelo com aplicação da governança de dados:

- Capacitação prévia dos desenvolvedores de relatórios de SSBI quanto ao modelo de governança de dados da organização e sua aplicação para BI;
- Regras de negócio claramente definidas antes do início dos desenvolvimentos;
- Indicadores maduros e que já sejam medidos antes do início dos desenvolvimentos dos relatórios conectados à arquitetura governada de dados;

- Manutenção dos profissionais da equipe técnica responsável pela definição da arquitetura, modelagem e desenvolvimento dos dados no ambiente analítico.

6. DISCUSSÃO DOS RESULTADOS

No aspecto acadêmico, esta pesquisa pode contribuir preenchendo uma lacuna teórica de avaliação da correlação entre a governança de dados e o uso das ferramentas de BI, visto a presença de estudos isolados sobre governança de dados nas empresas e sobre a evolução do uso de ferramentas de BI e SSBI, mas com uma ausência de estudo correlacionando estas duas variáveis. Do ponto de vista de SSBI, este estudo diverge parcialmente do modelo apresentado por Imhoff e White (2011) por não poder confirmar que o SSBI provê maior facilidade no consumo dos resultados e torna as fontes de dados fáceis de serem acessadas.

Por outro lado, os resultados encontrados corroboram o risco da disseminação do SSBI sem uma governança de dados, conforme apresentado por Alpar e Schulz (2016). No caso estudado, os relatórios não governados geram uma carga de trabalho para atualização e conferência dos dados para os usuários, além de possibilitar a divergência de conceitos, a alta carga de trabalho para obtenção de dados e geração de relatórios que possuem a mesma finalidade e a apresentação de informações incorretas. Além disso, a pesquisa confirma o estudo de Lennerholt et al. (2018), ao reforçar a importância da correta capacitação dos usuários de ferramentas de SSBI, tanto no uso dos sistemas analíticos quanto na aplicação da governança, reforçando esse como um desafio na implantação das ferramentas de SSBI.

Para a governança de dados, o estudo confirma os objetivos apresentados pelo DAMA (2017) e Khatri et al. (2010), evidenciando os benefícios específicos da governança quando relacionada com o SSBI e reforçando também a percepção pelos usuários e desenvolvedores do valor dos dados, conforme apresentado por Lennerholt et al. (2018) e Stodder (2015).

Além disso, a pesquisa complementa os requisitos gerais que um programa de governança de dados deve ter para atingir seus objetivos na organização, apresentados pelo DAMA (2017), com um olhar focado para o sucesso da governança do ponto de vista de BI e SSBI. Inicialmente, a pesquisa quantitativa apresentou uma divergência do apresentado pelo DAMA (2017), ao trazer uma baixa correlação direta entre a governança de dados e a segurança e privacidade dos dados, observada pelos usuários de negócio. No entanto, ao se entrevistar a equipe técnica (pesquisa qualitativa), pôde ser observado que essa correlação existe e é alta, porém é uma temática técnica, que é pouco observada pelas áreas de negócio e altamente valorizada pela equipe de segurança.

Ao avaliar, por meio das entrevistas, o incômodo apresentado pelos usuários dos relatórios produzidos no modelo governado quanto à demora para as mudanças de regras e alterações de dados nos indicadores, principalmente quando aplicada a governança sobre indicadores com baixa maturidade, a pesquisa concorda com a afirmação apresentada por

Tallon et al. (2004), de que uma governança que aplica alta restrição pode fazer com que os usuários busquem contornar o modelo e aplicar ferramentas de *Shadow IT*.

Quando avaliado o aspecto gerencial, a pesquisa pode contribuir com as empresas que buscam investir na estruturação de uma governança de dados para melhora na utilização das ferramentas de BI. Esta pesquisa possibilita a essas empresas avaliarem os benefícios que podem ser esperados ao fazer a estruturação desse modelo de governança, para justificar o investimento a ser feito, além de demonstrar as dificuldades que podem ser encontradas, bem como os benefícios esperados dentro da evolução no modelo de *self-service* que a empresa pode aplicar. O caso estudado demonstrou que a governança de dados poupa o consumo de energia do processamento, pois torna a demanda computacional mais eficiente, possibilitando o processamento de um maior conjunto de dados de forma mais eficiente. Esta eficiência possibilita ainda uma redução de custo para a organização, à medida que reduz a necessidade de escalonamento da capacidade de processamento e armazenamento de dados contratada pela organização.

Os requisitos mínimos apresentados para sucesso na migração de relatórios de SSBI para o modelo governado também permitem antever possíveis ineficiências que a organização poderia ter ao conduzir este trabalho, ajudando em uma melhor seleção dos escopos e das pessoas a serem envolvidas para a migração.

7. CONCLUSÃO

A proposta desta pesquisa foi de analisar os impactos que a governança de dados tem no uso das ferramentas de SSBI nas organizações, do ponto de vista dos usuários e desenvolvedores de produtos de dados. Por meio de um modelo conceitual proposto, pode-se concluir que a governança de dados exerce um impacto positivo nos produtos de dados, de modo que este benefício pode ser explicado por quatro aspectos: a segurança e privacidade dos dados, a qualidade dos dados, a disponibilização e gestão dos metadados e a eficiência no desenvolvimento dos produtos de dados.

Um dos aspectos com menor percepção de valor pelos usuários e desenvolvedores foi a segurança e a privacidade, mas, como complemento à pesquisa quantitativa aplicada, foram feitas entrevistas técnicas junto aos times de governança de dados e segurança da informação, que reforçou este impacto da governança, apresentando os critérios e benefícios de segurança dos dados. As ferramentas aplicadas junto ao modelo de governança se mostram efetivas para a garantia da segurança, tanto protegendo o acesso aos dados, quanto garantindo uma proteção contra agentes externos à organização.

Outro aspecto de menor percepção de valor foi a eficiência no desenvolvimento, podendo ser justificada pelo fato da governança de dados da organização ter sido implementada há poucos anos, ainda necessitando um maior amadurecimento. Assim, apresenta-se uma limitação da pesquisa, que por ter sido feita em apenas uma organização com uma governança de dados ainda em estágio inicial, não consegue endereçar o potencial máximo que a governança pode impactar. A organização pesquisada apresentou, em avaliação feita em 2023, nível inicial na maturidade da governança de dados (1,50) e nível repetível quando avaliada a dimensão *Data Warehouse* e *Business Intelligence* (2,07). Sugere-se, como estudo futuro, uma avaliação do modelo conceitual aqui apresentado, em uma organização com níveis mais altos de maturidade em governança de dados.

Além disso, a pesquisa conseguiu avaliar os benefícios que o SSBI gera na organização, quando avaliado o ponto de vista do mesmo público. Para este, os benefícios foram apresentados em dois aspectos: a facilidade do uso das ferramentas de desenvolvimento de relatórios e a velocidade de implementação de novos relatórios. Dois aspectos pesquisados não apresentaram uma correlação positiva com o uso das ferramentas de SSBI, foram eles: um melhor e mais fácil consumo dos relatórios pelos usuários finais e um maior acesso às fontes de dados para geração de novos relatórios. O primeiro aspecto mostra que, tanto para as ferramentas de SSBI quanto para as ferramentas tradicionais de BI, o consumo dos relatórios pelos usuários finais é o mesmo, estando mais relacionado à capacidade das pessoas que criam

o relatório e menos às ferramentas utilizadas. Dessa forma, pode-se entender que a capacitação dos desenvolvedores gera relatórios mais fáceis de serem consumidos.

Já o segundo aspecto apresenta potencial de correlação com o uso do SSBI, sendo sua limitação mais relacionada ao nível de *self-service* da organização, onde quanto mais alto esse nível, mais se espera ter um acesso às fontes de dados. Dessa forma, o fato de a pesquisa ter sido aplicada a apenas uma organização em nível 3 de *self-service* mostra mais uma limitação. Assim, sugere-se uma pesquisa dos benefícios do SSBI em organizações com mais alto nível de *self-service* aplicado aos dados.

A pesquisa cumpriu seus objetivos específicos. Primeiramente, a pesquisa contrapôs a percepção de valor dos diferentes tipos de usuários, sendo usuários de produtos de BI tradicionais, usuários de SSBI governado e SSBI *shadow IT*. A pesquisa auxiliou na compreensão dos critérios de migração de estruturas *shadow IT* para estruturas governadas, identificando o que é necessário para evitar a geração ou aumento do primeiro tipo. Também, a pesquisa definiu quatro requisitos mínimos necessários para se ter sucesso na implantação da governança de dados, do ponto no cenário de SSBI.

A proposição apresentada no modelo conceitual pôde também ser confirmada, avaliando que o usuário de SSBI reage ao controle das limitações e impedimentos provenientes da governança de dados, a impactando. Essa reação foi observada quando vista a necessidade de maiores autonomias e maior acesso aos dados apresentados pelos usuários durante as entrevistas. Essa necessidade mostra um potencial de geração de *shadow IT* quando o nível de restrição for muito alto, de modo que o retorno do usuário à equipe de governança precisa ser constante e bem avaliado.

Por fim, conclui-se que a governança de dados apresenta um impacto positivo para a organização, gerando eficiência para as equipes, desonerando os profissionais do trabalho operacional de geração de informações para que possam se dedicar às análises, tomadas de decisão e ação para melhoria dos resultados, e uma eficiência nos custos, possibilitando um processamento maior de dados com um menor custo. É necessário apenas se atentar ao processo de implantação de um modelo de governança de dados na organização, para que se tenha uma correta capacitação de todas as pessoas que são envolvidas, um apoio e controle da equipe de tecnologia quanto à maturidade dos produtos de dados que serão migrados para modelos governados, a fim de evitar retrabalho, e um investimento na manutenção de uma equipe técnica perene para cuidar dos dados da organização.

REFERÊNCIAS

- Aaboen, L., Dubois, A., Lind, F. (2012). Capturing processes in longitudinal multiple case studies. *Industrial Marketing Management*, 41(2), 235-246.
- Abelló, A., Darmont, J., Etcheverry, L., Golfarelli, M., Mazón, J. N., Naumann, F., Pedersen, T., Rizzi, S. B., Trujillo, J., Vassiliadis, P., Vossen, G. (2013). Fusion Cubes: Towards Self-service Business Intelligence, *International Journal of Data Warehousing and Mining*, (9:2), pp. 66–88. (DOI: 10.4018/jdwm.2013040104).
- Abraham, R., Brocke, J. V., Schneider, J. (2019). Data Governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424-438.
- Ackoff, R. L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16(1), 3-9.
- Alpar, P., Schulz, M. (2015). Self-service Business Intelligence. *Business Information Systems Engineering*. 58. 151-155.
- Al-Ruithe, M., Benkhelifa, E., Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous Computing*, 23, 839-859.
- Al-Ruithe, M., Benkhelifa, E. (2017). Analysis and Classification of Barriers and Critical Success Factors for Implementing a Cloud Data Governance Strategy. *The 8th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2017)*. 223-232.
- Bak, J. (2023). *Big on data: Study shows why data-driven companies are more profitable than their peers*. Disponível em: <https://cloud.google.com/blog/transform/data-leaders-more-profitable-innovative-hbr-data>
- Baltassis, E., Gourévitch, A., Quarta, L. (2019). *Good Data Starts with Great Governance*. Disponível em: <https://www.bcg.com/publications/2019/good-data-starts-with-great-governance>
- Barbieri, C. (2019). *Governança de Dados: Prática Conceitos e Novos Caminhos*. (1ª ed.). Atlas, Rio de Janeiro.
- Bardin, L. (2016). *Análise de Conteúdo*. São Paulo, SP, Brasil.
- Beijer, P., & Kooper, M. N. (2010). Information Governance: beyond risk and compliance. In *Anais, European Conference on Management Leadership and Governance*, 6, 10-34.
- Böhringer, M., Gluchowski, P., Kurze, C., Schieder, C. (2010). A business intelligence perspective on the future internet. In: *AMCIS 2010 Proceedings*. 1. 267.
- Bunchaft, G., Kellner, S. R. O. (1999). *Estatística sem mistérios*. v.2. Vozes, Petrópolis, RJ, Brasil.
- Caballero, I., Piattini, M. (2023). *Data Governance: From the Fundamentals to Real Cases*. Springer, Cham, CH.

- Cao et al., 2015 – Cao, G., Duan, Y., Li, G. (2015). Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis. *IEEE Transactions on Engineering Management*, 62(3), 384-395. <https://doi.org/10.1109/TEM.2015.2441875>
- Chandra, P., Gupta, M. K. (2018). Comprehensive survey on data warehousing research. *International Journal of Information Technology*, 10, 217-224.
- Chen, H., Chiang, R. H. L., and Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS quarterly*, (36:4), pp. 1165–1188. (DOI: 10.2307/41703503).
- Clark, V. L. P., Creswell, J. W., Green, D. O., Shope, R. J. (2010). Mixing Quantitative and Qualitative Approaches: An Introduction to Emergent Mixed Methods Research. In: *Handbook of Emergent Methods*, 363-387.
- Clarke, P., Tyrrell, G., Nagle, T. (2016). Governing self-service analytics. *Journal of Decision Systems*, 25(1), 145-159.
- Cohen, L., Manion, L. (1980). *Research Methods in Education*. Groom Helm Ltda, Education Research Methodology, London.
- Cohen, J. (2017). *Statistical power analysis for the behavioral sciences* (3rd ed.). Routledge.
- DAMA International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd Edition). Technics Publications, LLC, Denville, NJ, USA.
- Denzin, N. K., Lincoln, Y. S. (2008). *The Landscape of Qualitative Research* (3rd edition). Sage Publications, Inc.
- Eisenhardt, K. M. (1989). Building Theory from Case Study Research. *Academy of Management Review*, 14(4), 532-550.
- Ereth, J., Kemper, H.G. (2016). Business Analytics und Business Intelligence. *Controlling*. 28. 458-464. (DOI: 10.15358/0935-0381-2016-8-9-458).
- Ethan, A., Osaka, M. (2023). *Driving Business Intelligence with Big Data and Data Governance*. Disponível em: https://www.researchgate.net/publication/375039481_Driving_Business_Intelligence_with_Big_Data_and_Data_Governance
- Ferguson, G. A. (1981). *Statistical analysis in psychology and education*. McGraw-Hill Kogakusha, Tokyo.
- Gartner (2023). *Predicts 2024: Data and Analytics Governance Requires a Reset*. Disponível em: <https://www.gartner.com/en/documents/5049831>
- Gartner (2022). *Software Market Insights: Business Intelligence (BI) and Data Analytics*. Disponível em: <https://emt.gartnerweb.com/ngw/globalassets/en/digital-markets/documents/software-market-guide-bi-data-analytics.pdf>
- Gianella S., & Gujer W. (2006) Improving the Information Governance of Public Utilities. Through an Organizational Knowledge Base. In J. Matthew, K. J., L. Ma, A. Tan, & D. Anderson (Eds.), *Engineering Asset Management*, 406-41. London: Springer.

- Gröger, C. (2018). Building an Industry 4.0 Analytics Platform: Practical Challenges, Approaches and Future Research. *Datenbank-Spektrum*, 18(1), 5-14.
- Haag, S., Eckhardt, A. (2017). Shadow IT. *Business & Information Systems Engineering*, 59(6), 469-473.
- Hosmer, D.W., Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). Wiley, US.
- Imhoff, C., White, C. (2011). *Self-service Business Intelligence: Empowering Users to Generate Insights. TDWI Best practices report*. Disponível em: <https://tdwi.org/Research/2011/07/2011YiR-Best-Practices-Report-Q3-Self-Service-Business-Intelligence.aspx?m=1>
- Inmon, W. H. (2002). *Building the Data Warehouse* (3rd edition). John Wiley & Sons, Inc.
- Khatrri, V., Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53, 148-152.
- Kobielus, J. (2009). Mighty mashups: do-it-yourself business intelligence for the new economy. *Information and Knowledge Management Professionals*. Disponível em: <https://www.forrester.com/report/Mighty-Mashups-DoItYourself-Business-Intelligence-For-The-New-Economy/RES47806>
- Lattin, J. M., Carroll, J. D., Green, P. E. (2011). *Análise de Dados Multivariados*. Cengage Learning, SP, Brasil.
- Lennerholt, C., Laere, J. V., Söderström, E. (2018). Implementations challenges of *self-service* business intelligence: a literature review. *51st Hawaii International Conference on System Sciences*. 5055-5063.
- Lennerholt, C., Laere, J. V., Söderström, E. (2021). User-Related Challenges of Self-Service Business Intelligence. *Information Systems Management*, 38(4), 309-323.
- Lennerholt, C., Laere, J. V., Söderström, E. (2023). Success factors for managing the SSBI challenges of the AQUIRE framework. *Journal of Decision Systems*. 32 (2). 491-512.
- Leone, M., Walker, K. (2022). 2022 State of Data Governance and Empowerment. *ESG Research Insight Paper*. Disponível em: <https://www.erwin.com/docs/2022-esg-state-of-data-governance-and-empowerment-report-analyst-reports-30424.pdf>
- Meyers, C. (2014). How data management and Governance can enable successful *self-service* BI. *TDWI's Business Intelligence Journal*, 19, 4. Disponível em: <https://tdwi.org/research/2014/12/business-intelligence-journal-vol-19-no-4.aspx>
- Michalczyk, S., Nadj, M., Azarfar, D., Maedche, A., Gröger, C. (2020). A State-of-The-Art Overview and Future Research Avenues of *Self-service* Business Intelligence and Analytics. *Twenty-Eighth European Conference on Information Systems (ECIS2020)*. 1-18.
- Nambiar, A., Mundra, D. (2022). An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management. *Big Data and Cognitive Computing*. 6. 132.

- Negash, S. (2004). Business Intelligence. *Communications of the Association for Information Systems*. 13. 177-195.
- Niemi, E. (2011). *Designing a Data Governance Framework*. Aalto University School of Economics, p. 14. Disponível em: https://www.mn.uio.no/ifi/english/research/news-and-events/events/conferences-and-seminars/iris2013/groups/iris36_submission_4.pdf
- Nielsen, O. B. (2017). A comprehensive review of data governance literature. *Selected Papers of the IRIS, Issue Nr 8*. 3. 120-133.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in health sciences education: theory and practice*. 15. 625-32.
- Oracle. (2011). *Enterprise Information Management: Best Practices in Data Governance*. Disponível em: <https://www.oracle.com/assets/oea-best-practices-data-gov-1357848.pdf>
- Panian, Z. (2010). Some practical experiences in data Governance. *World Academy of Science, Engineering and Technology Management*, 62, 939-946.
- Petzold, B., Roggendorf, M., Rowshankish, K., Sporleder, C. (2020). *Designing data governance that delivers value*. Disponível em: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/designing-data-governance-that-delivers-value>
- PricewaterhouseCoopers. (2019). *Global and industry frameworks for data governance*. Disponível em: <https://www.pwc.in/consulting/technology/data-and-analytics/govern-your-data/insights/global-and-industry-frameworks-for-data-governance.html>
- Rau, K. G. (2004). Effective Governance of IT: Design objectives, roles, and relationships. *Information Systems Management*, 21, 4, 35-42
- Rosenfeld, P., Edwards, J. E., Thomas, M. D. (1995). Surveys. In Nicholson, N. *The Blackwell encyclopedic dictionary of organizational behavior* (pp. 548-549). Wiley-Blackwell, Cambridge, MA.
- SAS Institute. (2018). *The SAS Data Governance Framework: A Blueprint for Success*. Disponível em: <https://www.sas.com/content/dam/SAS/documents/marketing-whitepapers-ebooks/sas-whitepapers/en/sas-data-governance-framework-107325.pdf>
- Satriani, A., Umar, U., Fahlia, F. (2022). The effect of self organized learning environments (SOLE). Approach on student’s leaning motivation and achievement in history lessons. *Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, 10(3), 776-785.
- Stodder, D. (2015). Visual analytics for making smarter decisions faster – applying self-service business intelligence technologies to data-driven objectives. *TDWI Best Practices Report*. Disponível em: https://www.oracle.com/webfolder/s/delivery_production/docs/FY16h1/doc3/TDWI-BPReport-VAnalytics-July2105.pdf
- Swanson, R. A., Holson, E. F. (2005). *Research in Organizations: Foundation and Methods in Inquiry*. Berrett-Koehler Publishers.

- Tallon, P. P., Ramirez, R. V., Short, J. E. (2014). The information artifact in IT governance: toward a theory of information governance. *Journal of Management Information Systems*, 30(3), pp. 141-177.
- The Data Governance Institute (2015). *Data Governance Framework*. Disponível em: <https://datagovernance.com/the-dgi-data-governance-framework/dgi-data-governance-framework-components/>
- The Data Governance Institute (2015). *Definitions of Data Governance*. Disponível em: <https://datagovernance.com/defining-data-governance/>
- Tjur, T. (2009). Coefficients of Determination in Logistic Regression Models – A New Proposal: The Coefficient of Discrimination. *The American Statistician*, 63. 366-372.
- Van Grembergen, W., De Haes, S. (2010). A research journey into enterprise governance of IT, business/IT alignment and value creation. *The International Journal of IT/Business Alignment and Governance (IJITBAG)*, 1(1), 1–13.
- Weber, K., Otto, B., Osterle, H. (2009). One size does not fit all – a contingency approach to data Governance. *ACM Journal of Data and Information Quality*, 1, Artigo 4.
- Weill, P., Ross, J. (2004). *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Harvard Business Review Press.
- Wells, D. (2019). *The Path to Modern Data Governance*. Disponível em: <https://www.eckerson.com/articles/modern-data-governance-problems>
- Yin, R. K. (2015). *Estudo de caso: planejamento e métodos* (5 ed.). Bookman, Porto Alegre.
- Yin, R. K. (2016). *Pesquisa Qualitativa do Início ao Fim*. Penso Editora, Porto Alegre.

Bloco III – Gestão de Metadados										
Conhecimento organizacional: Entendo exatamente o que representam as informações presentes nos relatórios que consumo e/ou desenvolvo.										
Integração entre fontes de dados: Os relatórios que consumo e/ou desenvolvo convergem nos conceitos similares (por exemplo, o cliente tem a mesma definição em todos os relatórios).										
Qualidade dos metadados: O conteúdo dos relatórios que consumo e/ou desenvolvo apresentam claramente as definições necessárias para entender o que significam.										
Acesso aos metadados: Tenho acesso às definições e significados dos dados referentes aos relatórios que consumo e/ou desenvolvo.										
Troca de dados: Consigo aproveitar dados de relatórios já existentes na organização, não precisando criar uma nova forma de obtenção e tratamento dos dados.										
Bloco IV – Eficiência no Desenvolvimento										
Acesso aos dados: Tenho facilidade em buscar novas informações quando necessárias, e consigo garantir a não duplicidade de informações, caso essas já existam e sejam consumidas por outro relatório.										
Ciclo de vida dos dados: Os relatórios que consumo e/ou desenvolvo consomem os dados com boa performance, demandando pouca ação manual e gerando poucos custos de armazenamento dos dados.										

ETAPA 3 – GERAL

Quais dificuldades você enfrenta atualmente para a gestão de indicadores e dados?

Como você enxerga a governança de dados atualmente na organização?

Função na Empresa:

Tempo de Empresa:

Tempo Total de Experiência Profissional (considerar outras empresas também):

Tempo de Experiência com BI ou Self-Service BI (considerar outras empresas também):

APÊNDICE B – Roteiro para entrevista técnica

PESQUISA: IMPACTO DA GOVERNANÇA DE DADOS NO USO DE *BUSINESS INTELLIGENCE*: Estudo de caso em uma empresa de logística de grande porte

ROTEIRO DE ENTREVISTA

ETAPA 1 – Diretrizes e Instruções

- A esta entrevista será gravada com o objetivo de garantir a segurança das informações que serão prestadas e para facilitar a transcrição e compilação dos dados obtidos;
- Sua identidade será preservada, de modo que utilizaremos a referência “entrevistado 1...2...X” para referenciar os dados coletados e todas as informações obtidas nesta entrevista serão utilizadas exclusivamente para fins acadêmicos;
- A entrevista é composta de 9 questões;
- Caso não se sinta confortável ou não tenha conhecimento do assunto para responder qualquer pergunta, pode solicitar que passaremos para a próxima questão.

ETAPA 2 – Objetivo do trabalho

As tecnologias de Self-Service Business Inteligente (SSBI) surgiram com o objetivo de empoderar os usuários de negócio para produzirem suas próprias análises e indicadores. Porém, essa maior liberdade pode gerar alguns efeitos colaterais, comprometendo a qualidade dos dados e a gestão de acesso a eles. A governança de dados pode ser um caminho para mitigar esses efeitos.

O objetivo geral desta pesquisa é analisar o impacto da governança de dados no uso de ferramentas de BI, sob a ótica de desenvolvedores e consumidores de relatórios produzidos em ferramentas de SSBI.

ETAPA 3 – Apresentação dos resultados do *survey* aplicado

ETAPA 4 – Roteiro da entrevista

- 1) Como você percebe que a governança de dados da empresa garante a privacidade e confidencialidade dos dados dos stakeholders?
- 2) Como você percebe que a governança de dados da empresa impacta o cumprimento de regulamentações relacionadas com a segurança e a privacidade dos dados?
- 3) Como você vê que a governança de dados da empresa impacta a segurança de dados estratégicos da organização?
- 4) Como você percebe a gestão de acesso aos dados na governança de dados da organização?

5) Como você percebe que a governança de dados da empresa impacta contratos e acordos de não divulgação de dados na empresa?

6) Função na Empresa:

7) Tempo de Empresa:

8) Tempo Total de Experiência Profissional (considerar outras empresas também):

9) Tempo de Experiência com BI ou Self-Service BI (considerar outras empresas também):

APÊNDICE C – Roteiro para entrevista de avaliação da transição de modelo

PESQUISA: IMPACTO DA GOVERNANÇA DE DADOS NO USO DE *BUSINESS INTELLIGENCE*: Estudo de caso em uma empresa de logística de grande porte

ROTEIRO DE ENTREVISTA

ETAPA 1 – Diretrizes e Instruções

- A esta entrevista será gravada com o objetivo de garantir a segurança das informações que serão prestadas e para facilitar a transcrição e compilação dos dados obtidos;
- Sua identidade será preservada, de modo que utilizaremos a referência “entrevistado 1...2...X” para referenciar os dados coletados e todas as informações obtidas nesta entrevista serão utilizadas exclusivamente para fins acadêmicos;
- A entrevista é composta de 12 questões;
- Caso não se sinta confortável ou não tenha conhecimento do assunto para responder qualquer pergunta, pode solicitar que passaremos para a próxima questão.

ETAPA 2 – Objetivo do trabalho

As tecnologias de Self-Service Business Inteligente (SSBI) surgiram com o objetivo de empoderar os usuários de negócio para produzirem suas próprias análises e indicadores. Porém, essa maior liberdade pode gerar alguns efeitos colaterais, comprometendo a qualidade dos dados e a gestão de acesso a eles. A governança de dados pode ser um caminho para mitigar esses efeitos.

O objetivo geral desta pesquisa é analisar o impacto da governança de dados no uso de ferramentas de BI, sob a ótica de desenvolvedores e consumidores de relatórios produzidos em ferramentas de SSBI.

ETAPA 3 – Apresentação dos resultados do *survey* aplicado

ETAPA 4 – Roteiro da entrevista

- 1) A implantação da governança de dados junto ao uso de ferramentas de SSBI prevê como benefícios: **Segurança e Privacidade, Qualidade dos Dados, Gestão dos Metadados e Eficiência no Desenvolvimento**. Como você percebe que cada uma dessas dimensões mudou no seu uso do Power BI após a migração para o modelo governado?
- 2) O quanto você acredita que conhece das políticas e diretrizes de dados da organização?
- 3) Como você vê que o modelo de governança atual se encaixa nas suas necessidades do dia a dia relacionadas a gestão dos indicadores e dos relatórios de BI que você utiliza ou desenvolve?

- 4) Quais as principais mudanças positivas que a migração para um modelo governado de uso das ferramentas de SSBI trouxe para seu dia a dia e os produtos que você entrega? Quais as principais mudanças negativas?
- 5) Como você enxerga a percepção dos usuários dos relatórios que você produz referente a migração para o modelo governado?
- 6) Como você vê que o modelo de governança de dados da VLI pode melhorar, sob o aspecto de uso das ferramentas de SSBI?
- 7) Como tem sido o suporte e envolvimento da equipe de TI e seus fornecedores para o desenvolvimento dos relatórios que você desenvolve?
- 8) Quais dificuldades você enfrenta atualmente para a gestão e o desenvolvimento de indicadores?
- 9) Função na Empresa:
- 10) Tempo de Empresa:
- 11) Tempo Total de Experiência Profissional (considerar outras empresas também):
- 12) Tempo de Experiência com BI ou Self-Service BI (considerar outras empresas também):

APÊNDICE D – Fórmulas

Coeficiente de Correlação Ponto Bisserial

De acordo com Guilford (1950), o coeficiente de correlação bisserial e o erro padrão são calculados como:

$$p_{pb} = \frac{(\bar{X}_p - \bar{X})}{S_x} \sqrt{\frac{p}{q}} \quad \text{ou}$$

$$p_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_x} \sqrt{pq}$$

onde: p_{pb} é o Coeficiente de Correlação Ponto Bisserial;

$\bar{X}_p$ é a média dos valores de X para o grupo superior (grupo aos quais os valores de Y1 ou Y2 são “sim”);

$\bar{X}$ é a média total de X da amostra;

$\bar{X}_q$ é a média dos valores de X para o grupo inferior (grupo aos quais os valores de Y1 ou Y2 são “não”);

S_x é o desvio padrão total de X da amostra;

p é a proporção de casos do grupo superior (grupo aos quais os valores de Y1 ou Y2 são “sim”, ou seja, assumem valor 1);

q é a proporção de casos do grupo inferior (grupo aos quais os valores de Y1 ou Y2 são “não”, ou seja, assumem valor 0).

$$\sigma_{pb} = \sqrt{\frac{1 - p_{pb}^2}{n - 2}}$$

onde: σ_{pb} é o erro padrão;

p_b^2 é o quadrado do Coeficiente de Correlação Ponto Bisserial;

n é o número de observações da amostra.

Para testar a hipótese nula de que $p_b = 0$ o estimador do erro padrão será:

$$\sigma_{pb} = \frac{\frac{\sqrt{pb}}{y}}{\sqrt{n}}$$

Regressão Logística Multivariada

O modelo de regressão logística multivariada é expresso pela equação de definição do *logit*:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

onde: Y é o *logit*;

$b_0 \dots b_n$ são os coeficientes definidos para a equação;

$X_0 \dots X_n$ são as variáveis dependentes.

Onde o *logit* é definido por:

$$Y = \left(\frac{P(evento)}{1 - P(evento)} \right)$$

Onde P(evento) é a probabilidade de ocorrência do evento.

APÊNDICE E – Código R

#Instalação dos pacotes necessários

```
install.packages("corr")
install.packages("ggcorrplot")
install.packages("FactoMineR")
install.packages("factoextra")
install.packages("ltm")
```

#Leitura dos pacotes

```
library(readxl)
library('corr')
library('ggcorrplot')
library('FactoMineR')
library(factoextra)
library("ltm")
```

#Leitura das bases

```
Base_Pesquisa <- read_excel("C:/Users/morei/Pesquisa/Base Pesquisa.xlsx", sheet =
"Respostas Tratadas")
Usuarios_SSBI <- read_excel("C:/Users/morei/Downloads/Base Pesquisa.xlsx", sheet = "Base
Ordenada SSBI")
Usuarios_Governado <- read_excel("C:/Users/morei/Downloads/Base Pesquisa.xlsx", sheet =
"Base Ordenada Governado")
```

#Conversão das variáveis X

```
Usuarios_SSBI_Num <- as.numeric(as.character(unlist(Usuarios_SSBI)))
Usuarios_Governado_Num <- as.numeric(as.character(unlist(Usuarios_Governado)))
```

#Hipótese 1: correlação de SSBI para consumo

```
Base_Resp_SSBI_Consumo <- data.frame(a=Base_Pesquisa[,1], b=Base_Pesquisa[,2])

corr_matrix <- cor(Base_Resp_SSBI_Consumo)
corr_matrix

data.pca <- princomp(Base_Resp_SSBI_Consumo)
summary(data.pca)
data.pca$loadings

Resultado_Matriz <- Base_Resp_SSBI_Consumo * loadings(data.pca)[,1]
Resultado_Geral <- Resultado_Matriz[,1] + Resultado_Matriz[,2]

correlacao <- cor.test(Usuarios_SSBI_Num, Resultado_Geral)
correlacao$p.value
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_SSBI_Num, use = c("all.obs"), level = 2)
```

#Hipótese 2: correlação de SSBI para facilidade de uso

```
Base_Resp_SSBI_Uso <- data.frame(a=Base_Pesquisa[,3], b=Base_Pesquisa[,4])
```

```
corr_matrix <- cor(Base_Resp_SSBI_Uso)
corr_matrix
```

```
data.pca <- princomp(Base_Resp_SSBI_Uso)
summary(data.pca)
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_SSBI_Uso * loadings(data.pca)[,1]
Resultado_Geral <- Resultado_Matriz[,1] + Resultado_Matriz[,2]
```

```
correlacao <- cor.test(Usuarios_SSBI_Num, Resultado_Geral)
correlacao$p.value
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_SSBI_Num, use = c("all.obs"), level = 2)
```

#Hipótese 3: correlação de SSBI para velocidade

```
Base_Resp_SSBI_Velo <- data.frame(a=Base_Pesquisa[,5], b=Base_Pesquisa[,6])
```

```
corr_matrix <- cor(Base_Resp_SSBI_Velo)
corr_matrix
```

```
data.pca <- princomp(Base_Resp_SSBI_Velo)
summary(data.pca)
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_SSBI_Velo * loadings(data.pca)[,1]
Resultado_Geral <- Resultado_Matriz[,1] + Resultado_Matriz[,2]
```

```
correlacao <- cor.test(Usuarios_SSBI_Num, Resultado_Geral)
correlacao$p.value
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_SSBI_Num, use = c("all.obs"), level = 2)
```

#Hipótese 4: correlação de SSBI para Acesso

```
Base_Resp_SSBI_Acesso <- data.frame(a=Base_Pesquisa[,7])
```

```
Usuarios_Acesso_Num <- as.numeric(as.character(unlist(Base_Resp_SSBI_Acesso)))
```

```
correlacao <- cor.test(Usuarios_SSBI_Num, Usuarios_Acesso_Num)
correlacao$p.value
```

```
correlacao$estimate
```

```
biserial.cor(Usuarios_Acesso_Num, Usuarios_SSBI_Num, use = c("all.obs"), level = 2)
```

#Hipótese 5: correlação de governança com segurança

```
Base_Resp_Gov_Seguranca <- data.frame(a=Base_Pesquisa[,8], b=Base_Pesquisa[,9],  
c=Base_Pesquisa[,10], d=Base_Pesquisa[,11], e=Base_Pesquisa[,12])
```

```
corr_matrix <- cor(Base_Resp_Gov_Seguranca)  
corr_matrix
```

```
data.pca <- princomp(Base_Resp_Gov_Seguranca)  
summary(data.pca)  
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_Gov_Seguranca * loadings(data.pca)[,1]  
Resultado_Geral_Seguranca <- Resultado_Matriz[,1] + Resultado_Matriz[,2] +  
Resultado_Matriz[,3] + Resultado_Matriz[,4] + Resultado_Matriz[,5]
```

```
correlacao <- cor.test(Usuarios_Governado_Num, Resultado_Geral_Seguranca)  
correlacao$p.value  
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_Governado_Num, use = c("all.obs"), level = 2)
```

#Hipótese 6: correlação de governança com qualidade

```
Base_Resp_Gov_Qualidade <- data.frame(a=Base_Pesquisa[,13], b=Base_Pesquisa[,14],  
c=Base_Pesquisa[,15], d=Base_Pesquisa[,16])
```

```
corr_matrix <- cor(Base_Resp_Gov_Qualidade)  
corr_matrix
```

```
data.pca <- princomp(Base_Resp_Gov_Qualidade)  
summary(data.pca)  
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_Gov_Qualidade * loadings(data.pca)[,1]  
Resultado_Geral_Qualidade <- Resultado_Matriz[,1] + Resultado_Matriz[,2] +  
Resultado_Matriz[,3] + Resultado_Matriz[,4]
```

```
correlacao <- cor.test(Usuarios_Governado_Num, Resultado_Geral_Qualidade)  
correlacao$p.value  
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_Governado_Num, use = c("all.obs"), level = 2)
```

#Hipótese 7: correlação de governança com metadados

```
Base_Resp_Gov_Metadados <- data.frame(a=Base_Pesquisa[,17], b=Base_Pesquisa[,18],
c=Base_Pesquisa[,19], d=Base_Pesquisa[,20], e=Base_Pesquisa[,21])
```

```
corr_matrix <- cor(Base_Resp_Gov_Metadados)
corr_matrix
```

```
data.pca <- princomp(Base_Resp_Gov_Metadados)
summary(data.pca)
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_Gov_Metadados* loadings(data.pca)[,1]
Resultado_Geral_Metadados <- Resultado_Matriz[,1] + Resultado_Matriz[,2] +
Resultado_Matriz[,3] + Resultado_Matriz[,4] + Resultado_Matriz[,5]
```

```
correlacao <- cor.test(Usuarios_Governado_Num, Resultado_Geral_Metadados)
correlacao$p.value
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_Governado_Num, use = c("all.obs"), level = 2)
```

#Hipótese 8: correlação de governança com eficiencia

```
Base_Resp_Gov_Eficiencia <- data.frame(a=Base_Pesquisa[,22], b=Base_Pesquisa[,23])
```

```
corr_matrix <- cor(Base_Resp_Gov_Eficiencia)
corr_matrix
```

```
data.pca <- princomp(Base_Resp_Gov_Eficiencia)
summary(data.pca)
data.pca$loadings
```

```
Resultado_Matriz <- Base_Resp_Gov_Eficiencia* loadings(data.pca)[,1]
Resultado_Geral_Eficiencia <- Resultado_Matriz[,1] + Resultado_Matriz[,2]
```

```
correlacao <- cor.test(Usuarios_Governado_Num, Resultado_Geral_Eficiencia)
correlacao$p.value
correlacao$estimate
```

```
biserial.cor(Resultado_Geral, Usuarios_Governado_Num, use = c("all.obs"), level = 2)
```

#Hipótese 9: correlação geral da governança com SSBI

```
m1=glm(Usuarios_Governado_Num~Resultado_Geral_Seguranca+Resultado_Geral_Qualida
de+Resultado_Geral_Metadados+Resultado_Geral_Eficiencia, family =
binomial(link="logit"))
```

```
summary(m1)
```

APÊNDICE F – Dados de análise de ACP

H1: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação ao consumo dos relatórios.

Variáveis:

Y1: Facilidade no consumo

Y2: Clareza das definições

Matriz de Covariância

	Y1	Y2
Y1	1	0,696
Y2	0,696	1

Autovetores dos Componentes

	Comp. 1	Comp. 2
Y1	0,639	0,769
Y2	0,769	-0,639

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2
σ	2,22	0,93
λ	4,92	0,86
Proporção de Variância	0,85	0,15

H2: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à facilidade de uso destas ferramentas.

Variáveis:

Y1: Facilidade de desenvolvimento

Y2: Curva de aprendizado

Matriz de Covariância

	Y1	Y2
Y1	1	0,679
Y2	0,679	1

Autovetores dos Componentes

	Comp. 1	Comp. 2
Y1	0,790	0,613
Y2	0,613	-0,790

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2
σ	3,23	1,38
λ	10,44	1,91
Proporção de Variância	0,85	0,15

H3: A utilização de ferramentas de SSBI melhora a experiência dos usuários com relação à velocidade de implementação de relatórios.

Variáveis:

Y1: Velocidade de implementação

Y2: Custo de implementação

Matriz de Covariância

	Y1	Y2
Y1	1	0,543
Y2	0,543	1

Autovetores dos Componentes

	Comp. 1	Comp. 2
Y1	0,715	0,699
Y2	0,699	-0,715

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2
σ	3,01	1,64
λ	9,07	2,68
Proporção de Variância	0,77	0,23

H5: • A aplicação da governança de dados gera uma melhor percepção relativa à segurança e privacidade dos dados na utilização do SSBI.

Variáveis:

Y1: *Stakeholders*

Y2: Regulamentações

Y3: Propriedade de negócios

Y4: Acesso legitimado

Y5: Obrigações contratuais

Matriz de Covariância

	Y1	Y2	Y3	Y4	Y5
Y1	1	0,639	0,780	0,561	0,699
Y2	0,639	1	0,659	0,340	0,662
Y3	0,780	0,659	1	0,511	0,733
Y4	0,561	0,340	0,511	1	0,497
Y5	0,699	0,662	0,733	0,497	1

Autovetores dos Componentes

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
Y1	0,526	0,102	0,526	0,367	0,552
Y2	0,415	-0,549	-0,540	0,476	0
Y3	0,502	0	0,412	-0,186	-0,735
Y4	0,314	0,819	-0,434	0,153	-0,137
Y5	0,451	-0,113	-0,271	-0,762	0,359

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
σ	3,57	1,45	1,10	1,04	0,72
λ	12,72	2,11	1,22	1,07	0,52
Proporção de Variância	0,71	0,11	0,07	0,06	0,05

H6: A aplicação da governança de dados gera uma melhor percepção relativa à qualidade dos dados na utilização do SSBI.

Variáveis:

Y1: Acurácia

Y2: Pontualidade

Y3: Completude

Y4: Credibilidade

Matriz de Covariância

	Y1	Y2	Y3	Y4
Y1	1	0,636	0,559	0,695
Y2	0,636	1	0,567	0,593
Y3	0,559	0,567	1	0,581
Y4	0,695	0,593	0,581	1

Autovetores dos Componentes

	Comp. 1	Comp. 2	Comp. 3	Comp. 4
Y1	0,442	0,329	0,231	0,802
Y2	0,504	0,352	-0,762	-0,203
Y3	0,558	-0,829	0	0
Y4	0,489	0,285	0,604	-0,561

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2	Comp. 3	Comp. 4
σ	3,47	1,51	1,33	1,04
λ	12,05	2,28	1,76	1,09
Proporção de Variância	0,70	0,13	0,10	0,07

H7: A aplicação da governança de dados gera uma melhor percepção relativa à gestão de metadados na utilização do SSBI.

Variáveis:

Y1: Conhecimento organizacional

Y2: Integração entre fontes de dados

Y3: Qualidade dos metadados

Y4: Acesso aos metadados

Y5: Troca de dados

Matriz de Covariância

	Y1	Y2	Y3	Y4	Y5
Y1	1	0,640	0,616	0,516	0,430
Y2	0,640	1	0,750	0,604	0,597
Y3	0,616	0,750	1	0,670	0,598
Y4	0,516	0,604	0,670	1	0,586
Y5	0,430	0,597	0,597	0,586	1

Autovetores dos Componentes

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
Y1	0,319	0,456	0,216	0,802	0
Y2	0,476	0,318	0,394	-0,487	0,529
Y3	0,421	0,216	0	-0,291	-0,828
Y4	0,505	0	-0,843	0	0,183
Y5	0,489	-0,803	0,287	0,184	0

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
σ	4,20	1,73	1,59	1,21	1,05
λ	17,65	2,99	2,51	1,47	1,10
Proporção de Variância	0,69	0,11	0,10	0,06	0,04

H8: A aplicação da governança de dados gera uma melhor percepção relativa à eficiência no desenvolvimento na utilização do SSBI.

Variáveis:

Y1: Acesso aos dados

Y2: Ciclo de vida dos dados

Matriz de Covariância

	Y1	Y2
Y1	1	0,678
Y2	0,678	1

Autovetores dos Componentes

	Comp. 1	Comp. 2
Y1	0,730	0,683
Y2	0,683	-0,730

Autovalores e Proporção de Variância

	Comp. 1	Comp. 2
σ	3,02	1,32
λ	9,14	1,75
Proporção de Variância	0,84	0,16