

ESTRATÉGIA DE DADOS E GOVERNANÇA DE DADOS: EVIDÊNCIAS DE SEUS EFEITOS SOBRE CAPACIDADES ANALÍTICAS DE BIG DATA NO CONTEXTO DE ECOSISTEMAS.

FELIPE FONSECA SALERNO
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

ANTONIO CARLOS GASTAUD MAÇADA
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

Agradecimento à órgão de fomento:
Agradecemos à CAPES e ao CNPq.

ESTRATÉGIA DE DADOS E GOVERNANÇA DE DADOS: EVIDÊNCIAS DE SEUS EFEITOS SOBRE CAPACIDADES ANALÍTICAS DE BIG DATA NO CONTEXTO DE ECOSISTEMAS.

1. INTRODUÇÃO

A intensificação do volume, variedade e velocidade dos dados tem levado organizações de diferentes setores a repensarem como estruturam e mobilizam seus recursos informacionais para fins analíticos e estratégicos (Xu *et al.*, 2024). Em particular, nos ecossistemas de dados, ambientes compostos por múltiplos atores interdependentes que compartilham, processam e utilizam dados além das fronteiras institucionais, esses desafios tornam-se ainda mais complexos, dada a necessidade de coordenação entre organizações com objetivos distintos (Schreieck *et al.*, 2021; Oliveira & Loscio, 2018). A fragmentação de sistemas, a ausência de diretrizes estratégicas claras e a insuficiência de mecanismos de governança articulados tendem a limitar o uso eficaz dos dados, comprometendo sua utilidade para decisões baseadas em evidências (Liao *et al.*, 2023).

Nesse cenário, desponta a necessidade articular a gestão estratégica dos dados e a implementação de práticas de governança nos ecossistemas. A literatura recente tem sinalizado que tanto a estratégia de dados, definida como o conjunto de diretrizes que orienta a priorização e o uso dos dados (Gür *et al.*, 2021); quanto a governança de dados, entendida como o arranjo institucional que regula responsabilidades, padrões e processos (Abraham *et al.*, 2019); influenciam diretamente a capacidade organizacional de analisar grandes volumes de dados e, conseqüentemente, gerar valor através dos dados.

Sob essa perspectiva, o desenvolvimento das capacidades analíticas de big data (CABD), definidas como um conjunto de competências que viabilizam a coleta, integração, análise e interpretação de grandes volumes de dados (Karim *et al.*, 2024; Cui *et al.*, 2022) pode ser comprometido pela ausência de diretrizes estratégicas e de mecanismos de governança que assegurem padrões mínimos para que os dados possam ser dados analisados. Ou seja, não há como analisar dados sem que eles estejam minimamente adequados para uso, ou sem que haja diretriz que indique onde se deseja chegar com os resultados da análise.

No entanto, ainda são escassos os estudos empíricos que analisam, de forma integrada, os efeitos da estratégia e da governança de dados sobre as capacidades analíticas de big data, especialmente no contexto de ecossistemas. Nos ecossistemas de dados, o desenvolvimento dessas capacidades depende não apenas da adoção de tecnologias, mas da existência de um ambiente organizacional no qual a estratégia de dados e os mecanismos de governança estejam adequadamente estruturados e articulados (Weber *et al.*, 2024; Liao *et al.*, 2023). Apesar de estudos anteriores indicarem que as CABD são fundamentais para traduzir dados em valor organizacional (Hassani & Marentes, 2024), ainda persiste uma lacuna quanto à compreensão de seus fatores motivadores. Em especial, permanece pouco explorado como a combinação entre uma direção estratégica para os dados e um arranjo de governança robusto influencia o desenvolvimento dessas capacidades em ecossistemas de dados, onde a interdependência entre atores e a necessidade de alinhamento institucional são relevantes. Estudos que abordam isoladamente esses elementos já foram conduzidos (Fredriksson & Hagberg, 2023; Lis *et al.*, 2023), mas análises que os articulem em um mesmo modelo empírico ainda são incipientes.

Dessa forma, este artigo busca responder à seguinte questão de pesquisa: *qual é o efeito da estratégia de dados e da governança de dados nas capacidades analíticas de big data?* Para isso, esta pesquisa testa empiricamente um modelo teórico com base na Teoria da Orquestração de Recursos (Sirmon *et al.*, 2011), que oferece um referencial para compreender como recursos

relacionados aos dados podem ser estruturados e mobilizados com vistas ao desenvolvimento de capacidades analíticas. Para tanto, são desenvolvidos os construtos de estratégia de dados, governança de dados, e capacidades analíticas de big data, contextualizadas aos ecossistemas de dados, medindo suas relações e efeitos através de equações estruturadas (PLS-SEM) e complementado com modelo de covariância (CB-SEM). Ao integrar esses construtos em um único modelo, a pesquisa contribui para o avanço teórico da literatura sobre gestão e governança de dados, oferecendo subsídios práticos para organizações interessadas em aprimorar o uso analítico dos dados em contextos de ecossistemas.

2. REFERENCIAL TEÓRICO

Este capítulo apresenta a teoria que fundamenta os construtos desta pesquisa visando responder *qual é o efeito da estratégia de dados e da governança de dados nas capacidades analíticas de big data?* Para tanto, é desenvolvida a Teoria da Orquestração de Recursos (TOR), assim como os três principais construtos sob análise: estratégia de dados, governança de dados, e capacidades analíticas de big data.

2.1 Teoria da Orquestração de Recursos

A Teoria da Orquestração de Recursos (TOR) oferece um referencial teórico para compreender como organizações estruturam, agrupam e alavancam seus recursos a fim de gerar valor para suas atividades (Sirmon *et al.*, 2011). Derivada da Visão Baseada em Recursos (*Resource-Based View*) (Barney, 1991), a TOR expande essa abordagem ao incorporar os processos gerenciais responsáveis por intermediar a conversão de recursos em valor, tais como melhor desempenho e inovação, com ênfase em ambientes dinâmicos e interorganizacionais (Chirico *et al.*, 2023).

A teoria define três processos principais: estruturação (*structuring*), que envolve a identificação, seleção e organização dos recursos; agrupamento (*bundling*), que se refere à combinação desses recursos para formar capacidades integradas; e alavancagem (*leveraging*), que consiste na mobilização das capacidades para alcançar objetivos específicos e gerar valor (Sirmon *et al.*, 2011). Esses processos são considerados interativos, contínuos e sensíveis ao contexto, e ajudam a explicar que apenas a posse de recursos não é suficiente para geração de valor, e sim sua orquestração (Chirico *et al.*, 2023).

No campo dos dados, a TOR tem sido utilizada para investigar como diferentes recursos informacionais e institucionais, quando bem coordenados, contribuem para a geração de valor a partir dos dados (Karim *et al.*, 2024; Cui *et al.*, 2022). No presente estudo, a TOR fundamenta a hipótese de que tanto a estratégia de dados quanto a governança de dados contribuem para os processos de estruturação e agrupamento ao estabelecerem direções, papéis e regras que organizam o uso dos dados de forma alinhada aos objetivos envolvendo dados. A partir do desenvolvimento de estratégia e governança para os dados, criam-se as condições para que as organizações estruturem e combinem recursos informacionais, tecnológicos e humanos, facilitando a utilização de dados em processos analíticos (alavancagem).

A aplicação da Teoria da Orquestração de Recursos permite, assim, interpretar as capacidades analíticas de big data como um dos resultados de processos organizacionais por meio dos quais os dados são organizados, integrados e mobilizados de forma alinhada aos objetivos definidos pela estratégia e mecanismos de governança. Em vez de serem concebidas como atributos técnicos ou pela disponibilidade de ferramentas analíticas, as CABD são compreendidas, à luz da TOR, como capacidades construídas por meio de mecanismos que viabilizam a coordenação eficaz dos recursos relacionados aos dados (Xu e Pero, 2023; Zhang *et al.*, 2022). Ou seja, inclui a definição de diretrizes sobre os dados (estratégia) e o estabelecimento de regras formais e estruturas de responsabilização sobre seu uso (governança).

Esses mecanismos atuam de forma complementar, promovendo um ecossistema no qual os dados não apenas circulam, mas são tratados como insumos estratégicos, acessíveis, confiáveis e com qualidade suficiente para sua utilização analítica (Cui *et al.*, 2022). A TOR, portanto, fornece um referencial teórico para compreender como práticas e estruturas organizacionais são orquestradas de forma a suportar o desenvolvimento de capacidades analíticas, responsáveis por transformar dados em valor.

2.1 Estratégia de Dados

A estratégia de dados refere-se à formulação de uma orientação organizacional que estabelece como os dados serão utilizados para apoiar os objetivos organizacionais e gerar valor através dos dados (Marr, 2021). Ela orienta a integração dos dados ao planejamento e à operação das organizações, estabelecendo prioridades, definindo metas e alinhando os ativos de dados à lógica institucional. A estratégia de dados, assim, posiciona os dados como elementos centrais do funcionamento organizacional.

Em ecossistemas de dados, nos quais múltiplas instituições interagem com diferentes capacidades e interesses, a existência de uma estratégia para os dados permite alinhar as ações dos atores em torno de objetivos comuns. Isso reduz a dispersão de esforços e contribui para a coordenação de práticas orientadas por dados, facilitando a integração entre diferentes fontes e usos da informação (Schreieck *et al.*, 2021). A estratégia, nesse sentido, atua pela convergência entre os participantes do ecossistema em prol dos objetivos envolvendo dados.

A literatura aponta que estratégias de dados bem formuladas estabelecem, entre outros elementos, as prioridades em relação ao uso de dados, os critérios para investimento relacionados aos dados, e os indicadores para avaliação do progresso dos objetivos envolvendo dados (Gür *et al.*, 2021). Essas estratégias favorecem a consolidação de uma cultura orientada por dados e criam condições para o uso estruturado da informação de forma contínua, alinhada com os objetivos da organização e do ecossistema (Salerno & Maçada, 2025a). Isso inclui planos estratégicos voltados ao uso de dados, a incorporação de dados no processo decisório, e a definição de metas orientadas por evidências. Tais elementos indicam que a mobilização dos dados não ocorre de forma isolada, mas está integrada à lógica de funcionamento das organizações que compõem o ecossistema.

Pelas lentes da Teoria da Orquestração de Recursos (Sirmon *et al.*, 2011), a estratégia de dados cumpre uma função central no processo de estruturação ao delinear os caminhos pelos quais os recursos informacionais serão organizados e articulados. Essa estruturação suporta os processos de agrupamento e de alavancagem, sendo necessária para garantir coerência entre os recursos disponíveis, as capacidades a serem desenvolvidas e os objetivos institucionais a serem alcançados (Chirico *et al.*, 2023).

Assim, a estratégia de dados constitui um componente estruturante das iniciativas analíticas, criando as condições para que os dados possam ser mobilizados de maneira orientada e consistente, tanto no nível organizacional quanto no nível interorganizacional.

2.2 Governança de Dados

Governança de dados refere-se aos princípios, estruturas formais e mecanismos institucionais que orientam o uso ético e confiável dos dados pelas organizações (Abraham *et al.*, 2019; Otto, 2011). Ela abrange aspectos como a definição de papéis e responsabilidades, a criação de instâncias deliberativas, o estabelecimento de políticas institucionais e a supervisão dos processos relacionados aos dados, visando garantir que eles sejam utilizados, armazenados e compartilhados a partir de critérios transparentes de qualidade, acesso, segurança, responsabilidade e conformidade (Lis *et al.*, 2023).

Em ecossistemas de dados, nos quais múltiplos atores compartilham dados para diversos usos, a governança torna-se ainda mais relevante, uma vez que a ausência de regras claras para os dados pode limitar o uso analítico da informação. A literatura argumenta que mecanismos de governança bem definidos promovem interoperabilidade, alinhamento institucional e maior previsibilidade no uso dos dados, fatores que favorecem o desenvolvimento de capacidades analíticas em contextos colaborativos (Lin *et al.*, 2023; Schreieck *et al.*, 2021). Estudos como Mikalef & Krogstie (2018) modelaram a governança de dados como um construto multidimensional, composto por mecanismos estruturais, procedurais e relacionais.

Os mecanismos estruturais dizem respeito à formalização de papéis, comitês e estruturas organizacionais voltadas à supervisão do uso dos dados. Os mecanismos procedurais incluem as políticas, normas e padrões que orientam as práticas envolvendo dados. Já os mecanismos relacionais envolvem os canais de comunicação e treinamentos sobre dados e atores envolvidos (Abraham *et al.*, 2019). A articulação desses três mecanismos de governança de dados promove um ambiente em que a análise de dados pode ocorrer de forma mais segura, coordenada e efetiva. Nesta linha, Bernardo *et al.* (2024) indicam que estruturas de governança bem estabelecidas contribuem para a criação de um contexto favorável à análise de dados, ao garantir confiabilidade, rastreabilidade e legitimidade no uso das informações. Isso é particularmente relevante para o desenvolvimento das capacidades analíticas de big data, que dependem da disponibilidade de dados consistentes, da clareza sobre responsabilidades e da existência de políticas que viabilizem seu uso integrado.

2.4 Capacidades Analíticas de Big Data

As capacidades analíticas de big data (*big data analytics capabilities*) (CABD) representam a capacidade de uma organização de utilizar grandes volumes de dados, de diferentes fontes, para extrair insights relevantes, embasar a tomada de decisões e impulsionar o sucesso do negócio (Sumrit, 2025). Essas capacidades não se limitam ao domínio técnico, mas envolvem também aspectos humanos, institucionais e organizacionais, como a qualificação das equipes, a estruturação de processos analíticos, a integração de fontes de dados e a capacidade de transformar resultados analíticos em ações (Huynh *et al.*, 2023).

A literatura destaca que as CABD são capacidades construídas a partir da combinação de recursos tecnológicos, informacionais e gerenciais, que se tornam valiosas à medida que são aplicadas de forma coordenada e com propósito definido (Hassani & Marentes, 2024). Sua efetividade depende da existência de uma cultura orientada a dados que assegure tanto a disponibilidade e a qualidade dos dados quanto a capacidade de utilizá-los oportunamente (Korayim *et al.*, 2024). Dessa forma, as CABD representam um estágio superior de maturidade organizacional no uso de dados, indo além da posse de ferramentas analíticas ou de infraestrutura computacional (Langer, 2025).

No contexto de ecossistemas de dados, o desenvolvimento de capacidades analíticas envolve desafios adicionais, como a necessidade de interoperabilidade entre sistemas, o alinhamento entre diferentes atores institucionais e o estabelecimento de regras compartilhadas sobre o uso da informação (Liao *et al.*, 2023). Nesses ambientes, as CABD emergem de arranjos colaborativos que viabilizam o acesso a dados de diferentes atores, o uso conjunto de infraestrutura e a articulação de conhecimentos especializados. Por essa razão, o fortalecimento de capacidades analíticas exige, além de investimento técnico, o amadurecimento institucional de elementos como estratégia de dados e governança (Huynh *et al.*, 2023).

A aplicação da Teoria da Orquestração de Recursos (Sirmon *et al.*, 2011) permite compreender as CABD como capacidades que resultam principalmente dos processos de estruturação e agrupamento propostos pela teoria. A estratégia de dados orienta o direcionamento acerca do

uso dos dados, definindo como e para quais finalidades os dados serão utilizados. A governança, por sua vez, estabelece os parâmetros institucionais que garantem a integridade, a segurança e a legitimidade do uso dos dados. Em conjunto, elas criam as condições necessárias para que os recursos informacionais possam ser alavancados de forma eficaz através das capacidades analíticas (Weber *et al.*, 2024; Cui *et al.*, 2022).

3. DESENVOLVIMENTO DAS HIPÓTESES

Este capítulo apresenta o desenvolvimento das duas hipóteses que serão testadas nesta pesquisa. Inicialmente, argumenta-se que a estratégia de dados atua como um mecanismo que direciona o uso dos dados de forma planejada, coerente e orientada aos objetivos institucionais (Marr, 2021; Gür *et al.*, 2021). No contexto da Teoria da Orquestração de Recursos, esse construto se relaciona diretamente ao processo de estruturação, uma vez que orienta como os dados serão utilizados no ecossistema (Salerno e Maçada, 2025; Sirmon *et al.*, 2011). Em ambientes compostos por múltiplos atores, como os ecossistemas de dados, a presença de uma estratégia bem definida contribui para alinhar esforços e orientar a aplicação dos dados em práticas analíticas (Sarwar *et al.*, 2025; Schreieck *et al.*, 2021).

A literatura destaca que uma estratégia de dados formulada de forma clara e implementada com consistência favorece a criação de um ambiente propício ao desenvolvimento de capacidades analíticas ao definir prioridades e objetivos em relação aos dados (Gür *et al.*, 2021). Esse direcionamento estratégico viabiliza o agrupamento de recursos técnicos, humanos e informacionais, apoiando a criação de capacidades analíticas de big data (Karim *et al.*, 2024; Huynh *et al.*, 2023). Assim, considera-se que, ao orientar para o uso dos dados, a estratégia de dados exerce influência direta sobre o desenvolvimento das CABD ao criar as condições institucionais e operacionais para que os dados possam ser usados de forma alinhada e integrada aos objetivos organizacionais. Assim, a primeira hipótese é:

H1: A estratégia de dados tem efeito positivo nas capacidades analíticas de big data.

Em relação à segunda hipótese, tem-se que a governança de dados compreende o conjunto de estruturas formais, normas e processos que orientam o uso responsável, seguro e efetivo dos dados dentro e entre organizações (Abraham *et al.*, 2019; Otto, 2011). Em ambientes interorganizacionais, como os ecossistemas de dados, a presença de mecanismos de governança claros e articulados promove a interoperabilidade e a confiança entre os participantes (Lis *et al.*, 2023; Schreieck *et al.*, 2021). Pela perspectiva da Teoria da Orquestração de Recursos, a governança contribui para os processos de estruturação e agrupamento ao definir regras, papéis e padrões que permitem a coordenação e a integração dos recursos voltados à análise de dados (Xie *et al.*, 2023).

Mecanismos de governança bem estabelecidos criam um contexto favorável para a alavancagem dos recursos envolvendo dados ao garantir a confiabilidade, a rastreabilidade e a conformidade no seu uso (Bernardo *et al.*, 2024). Isso é especialmente relevante para o desenvolvimento das capacidades analíticas de big data uma vez que demandam dados acessíveis, seguros e de qualidade (Hassani & Marentes, 2024; Liao *et al.*, 2023). A governança de dados atua, portanto, como um facilitador que sustenta as condições necessárias para a construção e a aplicação eficaz das CABD. Logo, a segunda hipótese é

H2: A governança de dados tem efeito positivo nas capacidades analíticas de big data.

Definidas as duas hipóteses de pesquisa, a Figura 1 ilustra o modelo a ser testado para responder qual é o efeito da estratégia de dados e da governança de dados nas capacidades analíticas de big data?. Neste estudo, estratégia de dados e governança de dados são investigadas como

variáveis explicativas das capacidades analíticas de big data. Parte-se do pressuposto que ecossistemas de dados com maior maturidade em estratégia e governança de dados criam condições mais robustas para o desenvolvimento de capacidades analíticas como, por exemplo, ao promover regras claras, legitimidade na tomada de decisão e confiança que sustentam o uso sistemático e estruturado dos dados. Conseqüentemente, as CABD são concebidas como variável dependente, representando o grau que as organizações inseridas em ecossistemas de dados estão preparadas para coletar, processar, interpretar e comunicar informações a partir de grandes volumes de dados.

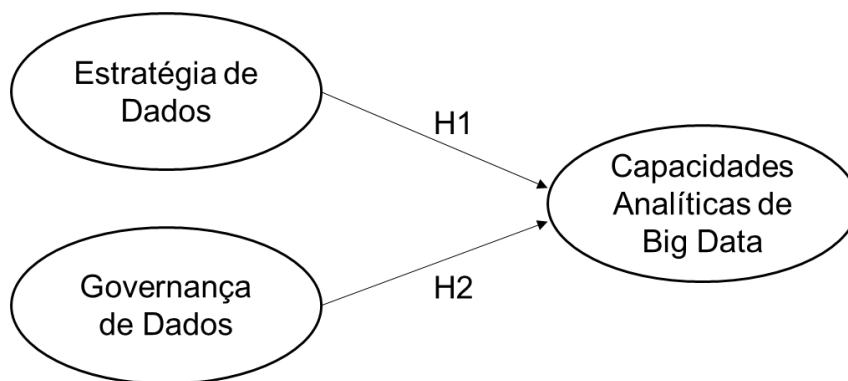


Figura 1: Modelo teórico. Elaborado pelos autores.

4. METODOLOGIA.

Este capítulo aborda os procedimentos metodológicos adotados nesta pesquisa, sendo operacionalizada a partir do desenvolvimento, validação e coleta de dados a partir de um questionário, e utilização de equações estruturadas (PLS-SEM) e análise de covariância (CB-SEM), conforme detalhado nas próximas seções.

4.1 Desenvolvimento e validação do questionário

Com o objetivo de mensurar os três construtos deste estudo (governança de dados, estratégia de dados e capacidades analíticas de big data), foi desenvolvido um questionário contendo 22 itens. A governança de dados foi concebida como um construto de ordem superior, composto pelos três mecanismos indicados pela literatura (Abraham *et al.*, 2019): estruturais, procedurais, e relacionais, alinhado com estudos anteriores como Mikalef e Krogstie (2018). Os construtos de estratégia de dados e CABD foram especificados como reflexivos e de primeira ordem, com cinco e quatro indicadores, respectivamente.

A estrutura inicial do instrumento foi submetida a um processo de validação qualitativa, com a aplicação da técnica de *card sorting* para avaliar o agrupamento dos itens em relação aos construtos teóricos (Spencer, 2009). Participaram dessa etapa dez especialistas, e os resultados revelaram um nível de concordância acima de 80%, o que sinalizou consistência entre a estrutura teórica e as percepções dos especialistas. Alguns itens foram reformulados com foco em aprimoramento semântico, mas sem alteração conceitual.

Para assegurar a validade de conteúdo do instrumento, foi conduzida uma segunda rodada de avaliação, na qual oito especialistas analisaram cada item quanto à clareza, relevância e adequação ao construto proposto, sendo calculado o Índice de Validade de Conteúdo (IVC) (Alexandre e Coluci, 2011). A maior parte dos itens apresentou IVC superior a 0,80, sendo que apenas um indicador exigiu revisão textual por ter obtido índice inferior ao valor de referência.

4.2 Amostra e coleta de dados

A obtenção dos dados ocorreu ao longo de um período de cinco meses, entre 15 de outubro de 2024 e 20 de março de 2025. A aplicação do questionário foi realizada através de canais digitais, com destaque para a divulgação em plataformas sociais e redes profissionais, especialmente o LinkedIn, visando alcançar profissionais com atuação direta em áreas ligadas à gestão, análise e governança de dados.

Com o objetivo de garantir o dimensionamento adequado da amostra, foi utilizado o software G*Power (Faul *et al.*, 2009) para a estimativa do tamanho mínimo necessário, com base em parâmetros apropriados à modelagem por equações estruturais via PLS-SEM. Considerando um tamanho de efeito pequeno ($f^2 = 0,10$) (Cohen, 1992), nível de significância de 5% ($\alpha = 0,05$) e poder estatístico de 90% ($1 - \beta = 0,90$), o cálculo apontou a necessidade de pelo menos 130 participantes.

Ao final do período de coleta, foram inicialmente registradas 364 respostas. Foram removidos os casos que apresentaram padrão de resposta uniforme (*straightlining*) ou que não forneceram informações mínimas sobre o ecossistema de dados no qual atuam (Meade e Craig, 2012). Após essa triagem, 38 registros foram excluídos, resultando em uma amostra final composta por 326 respondentes. As principais características sociodemográficas e profissionais da amostra estão descritas no Quadro I.

Quadro I: Dados da amostra.

| Tipo de organização | Percentual | Número de empregados | Percentual | Área de atuação | Percentual |
|---------------------|------------|----------------------|------------|-----------------------------------|------------|
| Privada | 57,6% | Até 100 | 10,3% | Tecnologia da Informação | 37,6% |
| Pública | 41,2% | Entre 101 e 500 | 16,8% | Finanças/Contabilidade/Auditoria | 24,7% |
| Outra | 1,2% | Mais de 501 | 70,6% | Gestão Estratégica | 11,5% |
| | | Não informado | 2,4% | Logística/Administração/Operações | 9,1% |
| | | | | Outras | 17,1% |

Fonte: Elaborado pelos autores

4.3 Modelagem de equações estruturais por mínimos quadrados parciais (PLS-SEM)

A modelagem das equações estruturais neste estudo foi conduzida com base no procedimento *disjoint two-stage*, indicado para situações em que há construtos de ordem superior (*Higher-Order Constructs* - HOC) (Sarstedt *et al.*, 2019), como é o caso da governança de dados, modelada como um HOC formativo-reflexivo composto pelos mecanismos estrutural, processual e relacional. Esse procedimento permite estimar separadamente os escores das dimensões reflexivas de primeira ordem e, em seguida, utilizar esses escores como indicadores formativos do HOC no modelo estrutural final (Hair *et al.*, 2019; Sarstedt *et al.*, 2019).

Na primeira etapa, foi especificado um modelo provisório em que as dimensões de primeira ordem da governança de dados foram incluídas diretamente no lugar do HOC, com objetivo de gerar os escores latentes para utilização na segunda etapa. Foram avaliados os critérios de confiabilidade composta, alfa de Cronbach, validade convergente (por meio do AVE - *Average Variance Extracted*) e validade discriminante, com base no critério HTMT (*Heterotrait-Monotrait Ratio*) (Hair *et al.*, 2019; Sarstedt *et al.*, 2019).

Na segunda etapa, esses escores foram incorporados ao modelo final como indicadores formativos do HOC governança de dados, juntamente com os demais construtos do modelo

(estratégia de dados e capacidades analíticas de big data). A validação do HOC considerou os critérios aplicáveis a construtos formativos: verificação da colinearidade entre os indicadores (VIF), e a significância e relevância dos pesos externos (*outer weights*) (Hair *et al.*, 2019; Sarstedt *et al.*, 2019).

Com a estrutura do modelo completa, foram então conduzidas as análises das relações estruturais entre os construtos, incluindo a estimação dos coeficientes de caminho (β), do poder explicativo (R^2), o cálculo dos tamanhos de efeito (f^2) e a avaliação da relevância preditiva ($Q^2_{predict}$), conforme proposto por Shmueli *et al.* (2019). Todas as estimativas e análises foram realizadas com o auxílio do software SmartPLS 4, com 10000 repetições de bootstrapping e 5% de significância estatística (Hair *et al.*, 2019; Sarstedt *et al.*, 2019).

4.4 Modelagem complementar com base na covariância (CB-SEM)

Visando complementar as análises por PLS-SEM, foi conduzido o CB-SEM no software SmartPLS 4 (Hair *et al.*, 2025). A utilização do CB-SEM permite verificar se os resultados obtidos pelo PLS-SEM se mantêm consistentes quando analisados por um método alternativo de estimação, baseado na covariância, fortalecendo a robustez do modelo proposto. Para tanto, o construto governança de dados foi tratado como reflexivo de segunda ordem, uma vez que o CB-SEM não permite a modelagem de construtos formativos. As estimativas foram obtidas com 1000 repetições de bootstrapping e 5% de significância. Foram observadas as cargas fatoriais, o alfa de Cronbach, a confiabilidade composta, a AVE e o HTMT, além dos índices de ajuste χ^2 , RMSEA, SRMR, CFI e TLI (Hair *et al.*, 2025).

5. RESULTADOS

5.1 Estimação dos escores latentes das dimensões da governança de dados

A primeira etapa do procedimento *disjoint two-stage* teve como objetivo estimar os escores latentes dos três mecanismos que compõem a governança de dados: estruturais, procedurais e relacionais. Essas dimensões foram modeladas como construtos reflexivos de primeira ordem, permitindo que fossem avaliadas individualmente quanto à qualidade do modelo antes da composição do construto de ordem superior (Mikalef e Krogstie, 2018).

Para cada mecanismo, foram verificados os critérios definidos na metodologia, e todas as três dimensões apresentaram resultados satisfatórios, com valores de confiabilidade superiores a 0,70, AVE acima de 0,50, e HTMT abaixo de 0,85, permitindo a extração dos escores latentes (Hair *et al.*, 2019; Sarstedt *et al.*, 2019).

5.2 PLS-SEM: Validação do HOC e dos construtos reflexivos

Na segunda etapa, os escores latentes dos mecanismos estruturais, procedurais e relacionais foram utilizados como indicadores formativos do construto de segunda ordem governança de dados, o qual foi incorporado ao modelo estrutural final. Esse modelo incluiu também os construtos estratégia de dados e capacidades analíticas de big data (CABD), conforme proposto no modelo teórico (Figura 1).

A validação do HOC foi realizada dentro do modelo estrutural final, por meio da análise dos pesos externos (*outer weights*) dos indicadores formativos, bem como da verificação da colinearidade entre os indicadores (VIF). Embora o peso externo do mecanismo estrutural não tenha sido estatisticamente significativo, sua carga (*loading*) absoluta foi elevada (0,830), indicando uma associação substancial com o construto governança de dados. Conforme sugerem Hair *et al.* (2022), em modelos formativos, indicadores com pesos não significativos,

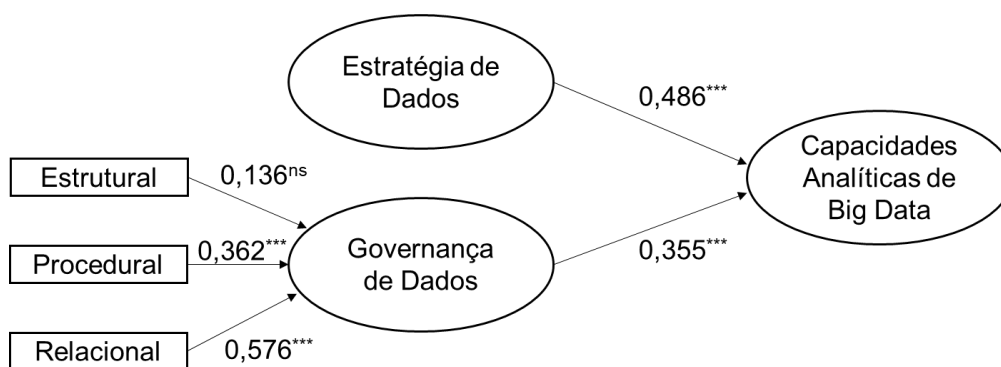
mas com cargas superiores a 0,50, podem ser mantidos quando há justificativa teórica. No presente estudo, o mecanismo estrutural foi mantido no modelo final por refletir aspectos formais e institucionais reconhecidos na literatura como componentes fundamentais da governança de dados (Mikalef & Krogstie, 2018). Por fim, os VIFs ficaram abaixo de 5, indicando ausência de multicolinearidade e validade do modelo formativo (Hair *et al.*, 2022).

Concluída a avaliação do HOC governança de dados, a avaliação dos dois construtos reflexivos do modelo de mensuração (estratégia de dados e CABD) seguiu os mesmos critérios da etapa anterior: confiabilidade composta, alfa de Cronbach, AVE e HTMT. Ambos os construtos apresentaram valores consistentes com os padrões de qualidade recomendados na literatura (Hair *et al.*, 2022; 2019). Foram também avaliados os índices de VIF para os construtos preditores, com valores inferiores a 3, afastando a possibilidade de multicolinearidade entre governança e estratégia de dados.

5.3 PLS-SEM: Avaliação do modelo estrutural

O modelo estrutural final teve como foco a análise dos efeitos da governança de dados e da estratégia de dados sobre o construto CABD. Ambos os efeitos foram estatisticamente significativos, com estatísticas t superiores a 5 e p-valores abaixo de 0,001. Isso indica que, com 1% de significância, tanto a governança de dados quanto a estratégia de dados exercem efeito positivo sobre as capacidades analíticas de big data.

O coeficiente de determinação (R^2) para o construto CABD foi de 0,623, o que significa que 62,3% da variância em CABD é explicada pelos dois preditores. De acordo com Hair *et al.* (2019), esse valor pode ser interpretado como expressivo no contexto das ciências sociais. Os tamanhos de efeito (f^2) foram calculados para medir a contribuição individual de cada construto sobre CABD, sendo $f^2 = 0,168$ para governança de dados, e $f^2 = 0,325$ para estratégia de dados. Esses valores indicam efeitos de médio porte conforme os padrões propostos por Cohen (1992). A Figura 2 ilustra o modelo final.



***: p-valor<0,001, ns: não significativo.

Figura 2: Resultado do Modelo PLS-SEM. Elaborado pelos autores

Em relação à qualidade preditiva (Tabela 1), os valores de Q^2 predict variaram entre 0,371 (CABD4) e 0,580 (CABD3), indicando que o modelo apresenta capacidade preditiva para todos os indicadores da variável latente CABD. De acordo com Hair *et al.* (2022), valores acima de zero evidenciam relevância preditiva, sendo que valores acima de 0,25 indicam uma previsão de qualidade média.

Ao comparar os erros de predição (RMSE e MAE) do modelo PLS-SEM com os benchmarks, observa-se que, para CABD1, CABD2 e CABD4, o modelo PLS-SEM apresenta menor erro do que o modelo de regressão linear, tanto em termos de RMSE quanto de MAE. Para CABD3, o desempenho do modelo PLS-SEM é semelhante ao do modelo LM, com uma diferença mínima entre os erros. Esses resultados indicam que o modelo PLS-SEM apresenta desempenho preditivo superior ou equivalente ao modelo linear, evidenciando sua capacidade de generalização preditiva e reforçando a robustez do modelo proposto.

Tabela 1: Resultados preditivos

| | PLS-SEM RMSE | PLS-SEM MAE | Q ² predict | LM RMSE | LM MAE |
|-------|-----------------|----------------|------------------------|------------|-----------|
| CABD1 | 1.1633 | 0.8899 | 0.4861 | 1.1832 | 0.8928 |
| CABD2 | 1.2148 | 0.9426 | 0.4482 | 1.2311 | 0.9568 |
| CABD3 | 1.0462 | 0.7770 | 0.5799 | 1.0400 | 0.7594 |
| CABD4 | 1.2102 | 0.9199 | 0.3711 | 1.2082 | 0.9313 |

Fonte: Elaborado pelos autores.

Logo, verifica-se que o modelo apresentou adequada qualidade de mensuração, relevância preditiva e significância estatística nas relações estruturais estimadas, confirmando as duas hipóteses de pesquisa.

5.4 CB-SEM: Avaliação do modelo de mensuração

Visando complementar os resultados do PLS-SEM, foi conduzido o CB-SEM (Hair *et al.*, 2025). Os resultados indicaram que todas as cargas fatoriais apresentaram valores superiores a 0,70. A confiabilidade interna foi verificada por meio do Alpha de Cronbach (α) e da confiabilidade composta (CR), ambos superiores a 0,70 para os três construtos. A validade convergente foi confirmada com valores de AVE superiores a 0,50, enquanto a validade discriminante foi confirmada pelo HTMT, apresentando valores abaixo de 0,85. Logo, o modelo de mensuração atende aos critérios definidos por Hair *et al.* (2025).

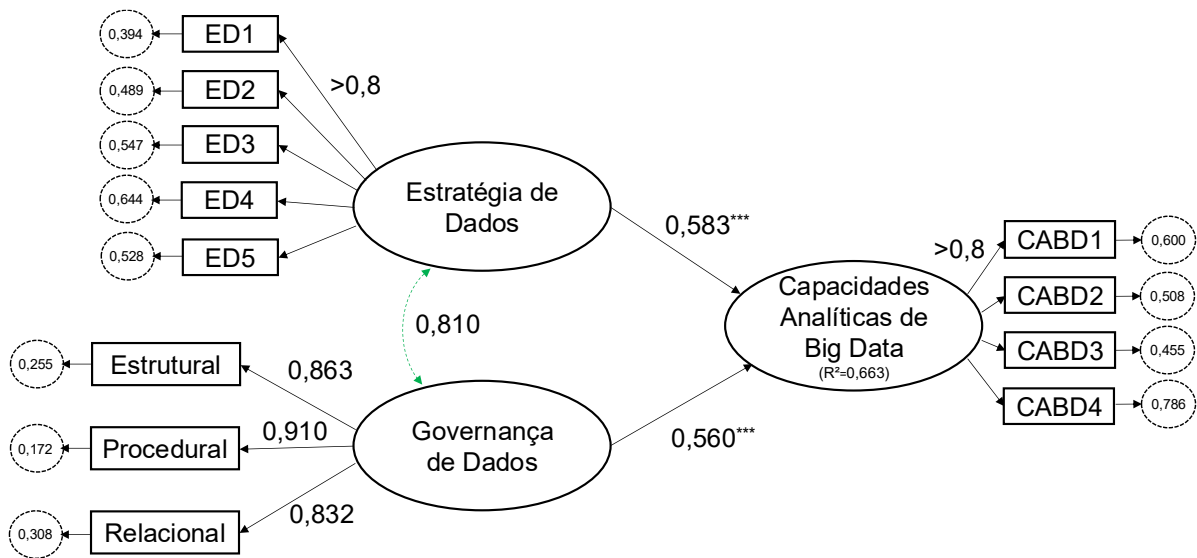
5.5 CB-SEM: Avaliação do modelo estrutural

O modelo estrutural apresentou coeficiente de determinação (R^2) de 0,663 para o construto capacidades analíticas de big data, indicando boa capacidade explicativa. Os dois coeficientes de caminho (β) entre os três construtos foram estatisticamente significativos ($p < 0,001$), com correlação de 0,810 entre estratégia e governança de dados (Figura 3), confirmando as hipóteses propostas. A avaliação do ajuste global do modelo seguiu recomendações de Hair *et al.* (2025) e é apresentada na Tabela 2. O qui-quadrado ($\chi^2 = 3717,34$; g.l. = 51) apresentou razão $\chi^2/g.l. = 2,684$, valor considerado aceitável. O RMSEA apresentou intervalo de confiança de 90% entre 0,057 e 0,087, dentro dos limites recomendados. O SRMR foi de 0,030, indicando baixo nível de resíduos padronizados. Os índices incrementais TLI (0,970) e CFI (0,976) ficaram acima de 0,95, sugerindo ajuste excelente. Esses resultados indicam que o modelo apresenta bom ajuste aos dados, reforçando a robustez dos achados obtidos por meio do PLS-SEM.

Tabela 2: Ajuste global do CB-SEM

| Métrica | Resultado |
|---------------------------|----------------|
| Qui-quadrado (χ^2) | 3717,34 |
| Graus de Liberdade (g.l.) | 51 |
| $\chi^2/g.l.$ | 2,684 |
| RMSEA IC 90% | [0,057; 0,087] |
| SRMR | 0,030 |
| TLI | 0,970 |
| CFI | 0,976 |

Fonte: Elaborado pelos autores.



***: p-valor<0,001

Figura 3: Resultado do Modelo CB-SEM. Elaborado pelos autores.

6. DISCUSSÃO

Os resultados obtidos por meio da modelagem de equações estruturais confirmam as duas hipóteses formuladas neste estudo, e os resultados do CB-SEM confirmaram as relações estruturais com índices de ajuste dentro dos padrões recomendados pela literatura (Hair *et al.*, 2025), o que contribui para aumentar a confiança nos achados empíricos uma vez que as evidências foram replicadas com dois métodos distintos de estimação. As análises indicaram que tanto a estratégia de dados quanto a governança de dados exercem efeitos positivos e estatisticamente significativos sobre as capacidades analíticas de big data (CABD), com coeficiente de determinação ($R^2 = 0,623$) que evidencia um alto poder explicativo (Hair *et al.*, 2019). Esses achados reforçam a importância da estratégia e da governança de dados para favorecer o desenvolvimento de capacidades analíticas em ecossistemas de dados.

Em relação à hipótese H1, os resultados sustentam que a estratégia de dados exerce influência moderada sobre as CABD, com tamanho de efeito ($f^2 = 0,325$) (Cohen, 1992). Este resultado é coerente com a literatura que define a estratégia de dados como um direcionador para o uso dos dados, articulando prioridades, metas e diretrizes que orientam sua mobilização (Marr, 2021; Gür *et al.*, 2021). Pelo viés da TOR, a estratégia atua no processo de estruturação ao estabelecer os caminhos pelos quais os recursos informacionais são organizados e mobilizados em função

de objetivos traçados. Em ecossistemas de dados, marcados por diversidade e interdependência, essa função estruturante da estratégia contribui para alinhar atores distintos em torno de objetivos comuns (Schreieck *et al.*, 2021), viabilizando o desenvolvimento de capacidades analíticas a partir da organização e do uso sistemático dos dados.

No que tange à hipótese H2, os dados também confirmam que a governança de dados tem efeito positivo e significativo sobre as CABD, com tamanho de efeito moderado ($f^2 = 0,168$). A governança, conforme definida Abraham *et al.* (2019), consiste em um conjunto de mecanismos que orientam o uso ético, seguro e coordenado dos dados, por meio da definição de papéis, estruturas formais e políticas. No modelo teórico proposto, a governança de dados foi concebida como um construto de ordem superior composto por mecanismos estruturais, procedurais e relacionais, conforme indicado por Mikalef e Krogstie (2018). Pela lente da TOR, a governança de dados atua nos processos de estruturação e agrupamento ao definir diretrizes que auxiliam a estruturar o fluxo de dados no ecossistema, como também regras para sua combinação entre diferentes fontes (Bernardo *et al.*, 2024). Isso é particularmente relevante em ecossistemas de dados, nos quais a interoperabilidade e o alinhamento entre instituições são fatores condicionantes para o desenvolvimento de capacidades analíticas (Liao *et al.*, 2023).

Ainda sobre H2, mesmo que o peso externo da dimensão estrutural não tenha sido estatisticamente significativo (p -valor $> 0,05$), sua carga elevada (0,830) justifica sua permanência no modelo dado seu respaldo teórico. Uma possível explicação para a ausência de efeito significativo é decorrente da própria natureza do mecanismo. Estruturas formais, comitês e papéis institucionalizados tendem a exercer influência indireta e de longo prazo, servindo como pré-requisitos para a estabilização dos demais mecanismos de governança, especialmente os de natureza procedural e relacional. Esse resultado sugere que, embora relevante, o mecanismo estrutural pode não ser suficiente para impulsionar diretamente a formação de capacidades analíticas. Este achado está em consonância com a lógica proposta pela TOR de que diferentes mecanismos contribuem de forma complementar e não necessariamente simétrica para a construção de capacidades (Sirmon *et al.*, 2011).

Aprofundando as contribuições teóricas, os resultados reforçam que as CABD resultam de processos gerenciais e institucionais que viabilizam a coordenação e o uso eficaz dos dados nos ecossistemas (Karim *et al.*, 2024, Hassani & Marentes, 2024). A estratégia de dados fornece orientação sobre o que fazer com os dados, enquanto a governança estabelece as condições para que esse uso ocorra de maneira legítima, segura e coordenada. Juntas, essas duas dimensões criam o ambiente organizacional necessário para que os recursos informacionais sejam estruturados, agrupados e alavancados para promover a geração de valor através dos dados (Salerno e Maçada, 2025). Assim, este estudo avança ao integrar, em um mesmo modelo, dois construtos frequentemente tratados de forma isolada na literatura: a estratégia de dados e a governança de dados, demonstrando que eles operam de forma complementar na formação das CABD, contribuindo para o amadurecimento conceitual da TOR.

Do ponto de vista prático, os achados oferecem diretrizes para organizações que atuam em ecossistemas e estão interessadas em aprimorar suas capacidades analíticas. Em vez de concentrarem esforços na adoção de tecnologias ou ferramentas analíticas, as organizações devem investir em estruturas que sustentem o uso dos dados de forma estratégica e coordenada. Isso envolve a formulação de estratégia para o uso de dados, com metas e prioridades bem definidas, assim como o fortalecimento de mecanismos de governança que assegurem responsabilidade, qualidade e legitimidade no acesso e uso dos dados (Gür *et al.*, 2021; Otto, 2011). Tais ações são especialmente relevantes em ecossistemas de dados, nos quais a construção de capacidades analíticas depende não apenas da ação de uma organização isolada, mas da articulação entre múltiplos atores com diferentes interesses (Liao *et al.*, 2023; Schreieck *et al.*, 2021).

7. CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo investigar o efeito da estratégia de dados e da governança de dados nas capacidades analíticas de big data (CABD), contextualizado para organizações que atuam em ecossistemas de dados. A partir da aplicação da Teoria da Orquestração de Recursos (Sirmon *et al.*, 2011), propôs-se um modelo teórico no qual a estratégia e a governança de dados são concebidas como mecanismos que viabilizam os processos de estruturação e agrupamento dos recursos informacionais para o desenvolvimento de capacidades analíticas de big data. Para tanto, foram definidas e confirmadas duas hipóteses de pesquisa utilizando equações estruturais (PLS-SEM) e CB-SEM, cujos resultados confirmaram que os dois construtos exercem efeitos significativos sobre as CABD.

A análise demonstrou que a estratégia de dados apresenta maior magnitude de efeito sobre as capacidades analíticas, reforçando sua função como orientadora do uso dos dados em consonância com os objetivos traçados (Marr, 2021; Gür *et al.*, 2021). A governança de dados, por sua vez, também revelou contribuição moderada, o que é consistente com sua atuação como mecanismo que estabelece regras, papéis e processos que regulam o uso dos dados (Abraham *et al.*, 2019; Otto, 2011). Os mecanismos estruturais da governança não apresentaram peso estatisticamente significativo (p -valor $> 0,05$), porém sua carga elevada sugere que este mecanismo possui uma influência indireta, servindo de base para os mecanismos procedurais e relacionais. A TOR respalda esse achado ao argumentar que a construção de capacidades decorre da ação complementar e não simétrica dos diferentes mecanismos organizacionais (Sirmon *et al.*, 2011).

Do ponto de vista teórico, o estudo contribuiu ao integrar, em um mesmo modelo, dois conceitos frequentemente abordados de maneira separada na literatura: a estratégia de dados e a governança de dados. Ao demonstrar que ambos são relevantes para o desenvolvimento das CABD, o estudo avança na compreensão das capacidades analíticas em contextos interorganizacionais como os ecossistemas de dados. Nesse sentido, contribui para a consolidação da TOR como referencial teórico para analisar os processos de estruturação, agrupamento e alavancagem envolvendo dados (Cui *et al.*, 2022; Sirmon *et al.*, 2011).

A contribuição prática é que o desenvolvimento de capacidades analíticas de big data requer estratégias bem definidas e mecanismos de governança. Isso é especialmente importante em ambientes nos quais múltiplos atores compartilham dados e recursos, sendo necessária uma coordenação estruturada para garantir o uso seguro e orientado dos dados (Liao *et al.*, 2023; Schrieck *et al.*, 2021). Organizações que desejam avançar em sua maturidade analítica devem, portanto, priorizar ações dentro do ecossistema de dados que promovam o alinhamento estratégico envolvendo dados e a institucionalização de práticas de governança de dados.

Entre as limitações desta pesquisa, a amostra foi obtida por meios digitais, o que pode limitar a generalização dos resultados. Além disso, o modelo focou apenas nos efeitos diretos da estratégia e da governança de dados sobre as capacidades analíticas, sem considerar possíveis mediações ou outros fatores contextuais que possam influenciar essas relações. Recomenda-se que pesquisas futuras explorem o papel de variáveis mediadoras e moderadoras que possam aprofundar a compreensão dos mecanismos que influenciam as CABD. Sugere-se também a aplicação do modelo em contextos setoriais distintos.

Agradecimentos: Agradecemos à CAPES e ao CNPq.

REFERÊNCIAS

- Abraham, R., Scheneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Alexandre, N. M. C., & Coluci, M. Z. O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência & Saúde Coletiva*, 16(7). <https://doi.org/10.1590/S1413-81232011000800006>
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Bernardo, B. M. V., Mamede, H. S., Barroso, J. M. P., & dos Santos, J. M. P. (2024). Data governance & quality management—Innovation and breakthroughs across different fields. *Journal of Innovation & Knowledge*, 9(4), 1–35. <https://doi.org/10.1016/j.jik.2024.100598>
- Chirico, F., Naldi, L., Hitt, M. A., Sieger, P., Sirmon, D. G., & Xu, K. (2023). Orchestrating resources with suppliers for product innovation. *Journal of Product Innovation Management*, 41(4), 735–767. <https://doi.org/10.1111/jpim.12703>
- Cohen, J. (1992). *A power primer*. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cui, Y., Firdousi, S. F., Afzal, A., Awais, M., & Akram, Z. (2022). The influence of big data analytic capabilities building and education on business model innovation. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.999944>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fredriksson, A., & Hagberg, J. (2023). From strategy to execution: Bridging the gap between data strategy and data governance [Master's thesis, Chalmers University of Technology]. *Chalmers Publication Library*. <http://hdl.handle.net/20.500.12380/306275>
- Gür, I., Spiekermann, M., Arbter, M., & Otto, B. (2021). Data strategy development: A taxonomy for data strategy tools and methodologies in the economy. *Wirtschaftsinformatik Proceedings*. <https://aisel.aisnet.org/wi2021/YGeneralTrack/Track02/1>
- Hair, J. F., Babin, B. J., Ringle, C. M., Sarstedt, M., & Becker, J.-M. (2025). Covariance-based structural equation modeling (CB-SEM): a SmartPLS 4 software tutorial. *Journal of Marketing Analytics*. <https://doi.org/10.1057/s41270-025-00414-6>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2022). *Advanced Issues in Partial Least Squares Structural Equation Modeling* (2nd ed.). Sage.
- Hassani, A., & Marentes, J. (2024). Digital Transformation and Big Data Analytics Capabilities: A Systematic Literature Review. *AMCIS 2024 Proceedings*, 6. https://aisel.aisnet.org/amcis2024/sig_osra/sig_osra/6

- Huynh, M.-T., Nippa, M., & Aichner, T. (2023). Big data analytics capabilities: Patchwork or progress? A systematic review of the status quo and implications for future research. *Technological Forecasting and Social Change*, 197, 122884. <https://doi.org/10.1016/j.techfore.2023.122884>
- Karim, S., Adrees, M., & Khawaja, K. F. (2024). The influence of big data analytic capabilities on firm performance: Mediating role of value creation and moderating role of data governance. *Journal of Management & Social Science*, 1(4), 1–24. <https://www.rjmss.com/index.php/7/article/view/22>
- Korayim, D., Chotia, V., Jain, G., Hassan, S., & Paolone, F. (2024). How big data analytics can create competitive advantage in high-stake decision forecasting? The mediating role of organizational innovation. *Technological Forecasting and Social Change*, 199, 123040. <https://doi.org/10.1016/j.techfore.2023.123040>
- Langer, B. (2025). Understanding Data & Analytics Maturity: A Systematic Review of Maturity Model Composition. *Schmalenbach Journal of Business Research*. <https://doi.org/10.1007/s41471-024-00205-2>
- Liao, S., Hu, Q., & Wei, J. (2023). How to leverage big data analytic capabilities for innovation ambidexterity: A mediated moderation model. *Sustainability*, 15(5), 3948. <https://doi.org/10.3390/su15053948>
- Lin, J., Lin, S., Benitez, J., Luo, X., & Ajamieh, A. (2023). How to build supply chain resilience: The role of fit mechanisms between digitally-driven business capability and supply chain governance. *Information & Management*, 60(2). <https://doi.org/10.1016/j.im.2022.103747>
- Lis, D., Gelhaar, J., & Otto, B. (2023). Data strategy and policies: The role of data governance in data ecosystems. In I. Caballero & M. Piattini (Eds.), *Data governance* (pp. 1–13). Springer. https://doi.org/10.1007/978-3-031-43773-1_2
- Marr, B. (2021). *Data strategy: How to profit from a world of big data, analytics and the internet of things* (2nd ed.). Kogan Page Ltd.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Mikalef, P., & Krogstie, J. (2018). Big data governance and dynamic capabilities: The moderating effect of environmental uncertainty. *PACIS 2018 Proceedings*, 206. <https://aisel.aisnet.org/pacis2018/206>
- Oliveira, M. I. S., & Lóscio, B. F. (2018). What is a data ecosystem? In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1–9). ACM. <https://doi.org/10.1145/3209281.3209335>
- Otto, B. (2011). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29. <https://doi.org/10.17705/1CAIS.02903>
- Salerno, F. F., & Maçada, A. C. G. (2025a). Data-driven culture and orchestrated data ecosystems: a conceptual model based on the resource-based view. *Revista de Gestão*, 32(2), 123–135. <https://doi.org/10.1108/rege-12-2024-0184>
- Salerno, F. F., & Maçada, A. C. G. (2025b). Developing a data orchestration scale: A validity and reliability study. *AMCIS 2025 Proceedings*, 3. https://aisel.aisnet.org/amcis2025/data_eco/data_eco/3

- Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal*, 27(3), 197–211. <https://www.sciencedirect.com/science/article/abs/pii/S1441358219301223>
- Sarwar, Z., Song, Z., Ali, S. T., Khan, M. A., & Ali, F. (2025). Unveiling the path to innovation: Exploring the roles of big data analytics management capabilities, strategic agility, and strategic alignment. *Journal of Innovation & Knowledge*, 10(1), 100643. <https://doi.org/10.1016/j.jik.2024.100643>
- Schrieck, M., Wiesche, M., & Krcmar, H. (2021). Capabilities for value co-creation and value capture in emergent platform ecosystems: A longitudinal case study of SAP's cloud platform. *Journal of Information Technology*, 36(4), 365–390. <https://doi.org/10.1177/02683962211023780>
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, 53(11), 2322–2347. <https://doi.org/10.1108/EJM-02-2019-0189>
- Sirmon, D. G., Hitt, M. A., Ireland, R. D., & Gilbert, B. A. (2011). Resource orchestration to create competitive advantage. *Journal of Management*, 37(5), 1390–1412. <https://doi.org/10.1177/0149206310385695>
- Spencer, D. (2009). *Card sorting: Designing usable categories*. Rosenfeld Media.
- Sumrit, D. (2025). An investigation of the impact of organizational big data analytics capabilities on healthcare supply chain resiliency. *Healthcare Analytics*, 7, 100393. <https://doi.org/10.1016/j.health.2025.100393>
- Weber, P., Hiller, S., Kurrle, S., Werling, M., & Werth, D. (2024). Evolutionary milestones in the development of data ecosystems. *AMCIS 2024 Proceedings*, 4. https://aisel.aisnet.org/amcis2024/data_eco/data_eco/4
- Xie, W., Zhang, Q., Lin, Y., Wang, Z., & Li, Z. (2023). The Effect of Big Data Capability on Organizational Innovation: a Resource Orchestration Perspective. *Journal of the Knowledge Economy*, 15(1), 3767–3791. <https://doi.org/10.1007/s13132-023-01208-w>
- Xu, D., Indulska, M., Asadi Someh, I., & Shanks, G. (2024). Time to reassess data value: The many faces of data in organizations. *The Journal of Strategic Information Systems*, 33(4), 101863. <https://doi.org/10.1016/j.jsis.2024.101863>
- Xu, J., & Pero, M. E. P. (2023). A resource orchestration perspective of organizational big data analytics adoption: Evidence from supply chain planning. *International Journal of Physical Distribution & Logistics Management*, 53(11), 71–97. <https://doi.org/10.1108/ijpdm-04-2022-0118>
- Zhang, D., Pee, L. G., Pan, S. L., & Cui, L. (2022). Big data analytics, resource orchestration, and digital sustainability: A case study of smart city development. *Government Information Quarterly*, 39(1), 101626. <https://doi.org/10.1016/j.giq.2021.101626>