DISSERTAÇÃO DE MESTRADO PROFISSIONAL

Modelagem Integrada de Violações de Dados: Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais

EVANEI GOMES DOS SANTOS

Orientador: Prof. Dr. André Luiz Marques Serrano

Programa de Pós-Graduação Profissional em Engenharia Elétrica

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Modelagem Integrada de Violações de Dados: Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais

Evanei Gomes dos Santos

Orientador: Prof. Dr. André Luiz Marques Serrano, EPR/UnB

PUBLICAÇÃO: PPEE.MP.104

BRASÍLIA-DF

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

Modelagem Integrada de Violações de Dados: Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais

EVANEI GOMES DOS SANTOS

Orientador: Prof. Dr. André Luiz Marques Serrano

Dissertação de Mestrado Profissional submetida ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica

Banca Examinadora

Prof. Dr. André Luiz Marques Serrano, EPR/UnB _____
Orientador

Prof. Dr, Vinícius Pereira Gonçalves, ENE/UnB _____
Examinador Interno

Prof. Dr. Rodrigo Bonacin, Centro Universitário Campo Limpo (UNIFACCAMP) _____
Examinador externo

Prof. Dr. Geraldo Pereira Rocha Filho, DCET/UESB _____
Suplente

FICHA CATALOGRÁFICA

SANTOS, EVANEI GOMES
Modelagem Integrada de Violações de Dados:Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais [Distrito Federal] 2025.
xvi, 53 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2025).
Dissertação de Mestrado Profissional - Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia Elétrica

1. Vazamento de Dados
2. Privacidade de Dados
3. Aprendizado de Máquina
4. Aprendizagem Profunda

I. ENE/FT/UnB
II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

SANTOS, EVANEI GOMES (2025). *Modelagem Integrada de Violações de Dados:Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais*. Dissertação de Mestrado Profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 53 p.

CESSÃO DE DIREITOS

AUTOR: EVANEI GOMES DOS SANTOS
TÍTULO: Modelagem Integrada de Violações de Dados:Previsão de Incidentes, Análise de Sobrevivência e Privacidade Diferencial em Séries Temporais.
GRAU: Mestre em Engenharia Elétrica ANO: 2025

EVANEI GOMES DOS SANTOS
Depto. de Engenharia Elétrica (ENE) - FT
Universidade de Brasília (UnB)
Campus Darcy Ribeiro
CEP 70919-970 - Brasília - DF - Brasil

DEDICATÓRIA

Dedico este trabalho à minha esposa, Zilmara S. Soares dos Santos. Ela foi minha fonte constante de inspiração, encorajamento e equilíbrio. Enfrentamos esta jornada no meio de uma gestação e, depois, com a chegada do nosso pequeno Matias. Mesmo em meio a tantos desafios, ela permaneceu firme — não desistiu, insistiu e ainda me fortaleceu quando minhas forças pareciam poucas.

Ela é, sem dúvidas, a mulher mais forte que conheço. Dou graças a Deus por tê-la como esposa.

AGRADECIMENTOS

Agradeço, em primeiro lugar, a Deus. Sem Ele, para mim teria sido impossível chegar até aqui.

Agradeço à minha família, especialmente à minha esposa Zilmara e aos meus filhos, Elias e Matias, que foram minha fonte de motivação diária. Agradeço também à minha mãe, dona Maria, e ao meu pai, seu Cecílio, que, mesmo sem saber ler ou escrever, sonharam este dia por mim e plantaram as bases que me trouxeram até aqui.

Registro minha profunda gratidão ao meu orientador, Prof. André Luiz Marques Serrano, cuja orientação foi fundamental para a realização deste trabalho. Sua forma de ensinar — prática, instigante e sempre voltada ao desenvolvimento real do aprendizado — demonstra o quanto precisamos de professores com esse nível de comprometimento. O professor André é uma pessoa incrível e por quem cultivarei permanente admiração e sincera gratidão pelos ensinamentos e orientações ao longo desta jornada.

Agradeço ao grupo de pesquisa Lincex, em especial ao professor Geraldo, ao professor Vinícius, ao colega Gabriel Arquelau, cuja parceria e apoio foram fundamentais ao longo do desenvolvimento deste trabalho.

Estendo meus agradecimentos a Gustavo Henrique, Rodrigo Lange e João Paulo V., bem como ao Ministério da Justiça e Segurança Pública, ao Ibama e à Universidade de Brasília, cujas contribuições e apoio foram essenciais para a realização deste trabalho.

RESUMO

Esta dissertação investiga violações de dados sob a ótica da segurança da informação e da proteção de dados pessoais, articulando previsão temporal, análise de recorrência e privacidade diferencial. Parte-se de uma análise bibliométrica sobre privacidade e segurança cibernética, seguida da exploração da base *Data Breach Chronology* da *Privacy Rights Clearinghouse*, com estratificação por setor organizacional. No eixo preditivo, comparam-se modelos estatísticos (SARIMA e Prophet), modelos de aprendizado de máquina baseados em árvores de decisão (XGBoost) e redes neurais profundas (TCN e LSTM) para a previsão mensal de incidentes, avaliando o desempenho por meio das métricas *MAPE*, *MAE* e *RMSE*. Em seguida, aplica-se a análise de Sobrevivência (Kaplan–Meier) para estimar o tempo até a reincidência do próximo incidente por setor, evidenciando perfis diferenciados de risco temporal. Por fim, avalia-se o impacto da privacidade diferencial, por meio do mecanismo de *Laplace*, em dados tabulares de incidentes sintéticos, quantificando o equilíbrio entre orçamento de privacidade, deslocamento distributivo e perda de acurácia. Os resultados indicam melhor desempenho preditivo de redes neurais em séries mais complexas, janelas de recorrência menores em setores críticos e faixas de parâmetros de privacidade que preservam utilidade analítica. Em conjunto, os achados compõem um arcabouço aplicado para apoiar planejamento, definição de *SLAs* e divulgação responsável de estatísticas em conformidade com a Lei Geral de Proteção de Dados Pessoais.

Palavras-chave: Violação de dados; Segurança da Informação; Modelagem Preditiva; Análise de Sobrevivência; Privacidade Diferencial; LGPD.

ABSTRACT

This dissertation investigates data breaches from the perspective of information security and personal data protection, articulating temporal prediction, recurrence analysis, and differential privacy. It begins with a bibliometric analysis of privacy and cybersecurity, followed by an exploration of the Privacy Rights Clearinghouse's Data Breach Chronology database, stratified by organizational sector. In the predictive axis, statistical models (SARIMA and Prophet), machine learning models based on decision trees (XGBoost), and deep neural networks (TCN and LSTM) are compared for monthly incident prediction, evaluating performance using the metrics MAPE, MAE, and RMSE. Subsequently, Survival analysis (Kaplan–Meier) is applied to estimate the time until the next incident recurrence by sector, highlighting differentiated temporal risk profiles. Finally, the impact of differential privacy is evaluated, using the Laplace mechanism, on tabular synthetic incident data, quantifying the balance between privacy budget, distributional shift, and accuracy loss. The results indicate better predictive performance of neural networks in more complex series, smaller recurrence windows in critical sectors, and privacy parameter ranges that preserve analytical utility. Taken together, the findings comprise an applied framework to support planning, SLA definition, and responsible disclosure of statistics in accordance with the General Data Protection Law.

Keywords: Data Breaches; Information Security; Predictive Modeling; Survival Analysis; Differential Privacy; LGPD.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

1 INTRODUÇÃO

A transformação digital acelerou a criação, o trânsito e o armazenamento de dados sensíveis em organizações públicas e privadas (Lopes e Amaral 2022). Embora a evolução tecnológica traga vantagens, ela resultou em uma significativa geração de informações pessoais e de clientes (Belarmino, Ricarte e Motta 2024), intensificando a complexidade da regulação de segurança e privacidade de dados (Carvalho et al. 2023). Além de sistemas próprios, há hoje um ecossistema complexo de nuvem, *SaaS*, integrações por *Application Programming Interfaces (APIs)*, dispositivos móveis e *Internet of Things (IoT*) (Benzell et al. 2022), o que fragmenta perímetros e amplia a superfície de ataque (Rodrigues et al. 2024). Essa realidade expõe a sociedade a ameaças cibernéticas significativas (Bertoni et al. 2022). Nesse contexto, cyberataques continuam com frequência crescente em todo o mundo (Urooj et al. 2022), e violações de dados — oriundas de vazamentos (Bertoni 2020), *ransomware*, exploração de vulnerabilidades ou falhas de configuração — deixaram de ser eventos raros para se tornarem riscos operacionais recorrentes (Rodrigues et al. 2024).

Os incidentes de *ransomware*, por exemplo, evoluíram significativamente em complexidade e potência, visando cada vez mais companhias e organizações para obter resgates maiores (Urooj et al. 2022). As violações geram impactos financeiros, reputacionais e legais (Rodrigues et al. 2024), e podem afetar a continuidade do serviço. A perda de informações sensíveis, como registros médicos confidenciais, resulta em danos substanciais (Vainzof 2020). No setor público, tais incidentes podem interromper políticas públicas, afetar a confiança social e gerar encargos regulatórios adicionais (Rodrigues et al. 2024).

A Lei Geral de Proteção de Dados Pessoais (LGPD), formalizada pela Lei nº 13.709/2018 (Brasil 2018), é o principal marco regulatório brasileiro, inspirado no Regulamento Geral de Proteção de Dados (GDPR) da União Europeia (Belarmino, Ricarte e Motta 2024, Elger e Santander 2024, Fernandes e Nuzzi 2022), e tem como propósito fundamental proteger direitos essenciais como a privacidade, a liberdade e o livre desenvolvimento da personalidade dos titulares (Elger e Santander 2024, Fernandes e Nuzzi 2022). Para a Administração Pública, a LGPD estabelece que o tratamento de dados deve visar ao interesse público, permitindo a coleta e o uso compartilhado de informações quando estritamente necessário para a execução de políticas públicas (Art. 7º, III) (Brasil 2018, Fernandes e Nuzzi 2022). Este uso deve seguir rigorosamente o princípio da necessidade, que limita o tratamento ao mínimo indispensável, garantindo que os dados sejam pertinentes, proporcionais e não excessivos (Art. 6º, III) (Brasil 2018, Fernandes e Nuzzi 2022). Em síntese, a LGPD fixa diretrizes para o tratamento de dados pelo poder público, balizando-o pelo interesse público e pelo princípio da necessidade, com vistas à proteção dos direitos fundamentais dos titulares (Brasil 2018, Fernandes e Nuzzi 2022, Elger e Santander 2024).

À luz desse cenário, este trabalho adota uma perspectiva aplicada, orientada por evidências, com foco na análise e avaliação do panorama de violações de dados. A investigação organiza-se em frentes complementares: (i) previsão temporal da incidência de violações de dados; (ii) mensuração da recorrência (tempo-até-novo-incidente) por métodos de sobrevivência; e (iii) avaliação de estratégias de divulgação de evidências com salvaguardas formais de proteção de dados.

1.1 MOTIVAÇÃO

A crescente sofisticação dos ataques cibernéticos e o volume elevado de informações sensíveis armazenadas digitalmente intensificam as preocupações com violações de dados (Sun et al. 2023). Em 2023, organizações levaram, em média, 204 dias para identificar um vazamento, com pouca variação em relação a anos anteriores. Em paralelo, o custo médio global de uma violação em 2024 foi estimado em 4,88 milhões de dolares, podendo alcançar 5,36 milhões de dólares em ambientes com poucos recursos de segurança. Essas preocupações são agravadas pela expansão da coleta de dados a partir de fontes sensíveis — como registros de equipamentos médicos e transações *online* — que aumenta a superfície de exposição e a complexidade de proteção (Kumar e Gupta 2020).

Casos de grande escala ilustram esse cenário. Em janeiro de 2024, cerca de 26 bilhões de registros foram expostos em um único evento, conhecido como a "Mãe de Todos os Vazamentos", envolvendo aproximadamente 12 *terabytes* de dados e atingindo, entre outros, plataformas de redes sociais como LinkedIn e Twitter (Rodrigues et al. 2024). Em anos recentes, episódios de ampla repercussão — como o incidente nos hotéis Marriott, que expôs dados de 500 milhões de clientes — reforçam a materialidade do risco. Empresas como T-Mobile, Quora, Google, Orbitz e Facebook também reportaram incidentes que comprometeram populações superiores a 100 milhões de usuários, com efeitos diretos sobre conformidade regulatória e práticas de proteção (Yang e al. 2024).

Dados históricos da Privacy Rights Clearinghouse (PRC) indicam que, entre 2006 e 2021, o número anual mínimo de registros expostos nos Estados Unidos se aproximou de 100 milhões, atingindo mais de 700 milhões em 2019 (Autores do Artigo Healthcare 2020). Embora o recorte seja norte-americano, há relatos de maior número de ocorrências públicas na Europa em certos períodos, o que indica que a gravidade do fenômeno se estende a múltiplas jurisdições (Neto et al. 2021). Nesses eventos, a exposição de dados pessoais, de saúde e financeiros — por definição sensíveis e privados — configura incidentes de segurança em que informações confidenciais são acessadas por indivíduos não autorizados, com implicações de privacidade e proteção de dados (Rodrigues et al. 2024).

Para organizações públicas e privadas, as consequências incluem danos reputacionais (Perera et al. 2022), litígios e perdas financeiras diretas (Rodrigues et al. 2024), além de impactos indiretos significativos como aumento dos custos operacionais para investigação e remediação (Val et al. 2024), sanções administrativas regulatórias (Ainslie et al. 2023), interrupção ou degradação de serviços críticos (Val et al. 2024), perda de vantagem competitiva (Ainslie et al. 2023) e erosão da confiança de cidadãos, clientes e parceiros comerciais (Sharma e Bantan 2025).

Diante desse quadro, a preocupação com privacidade cresce de forma consistente, tornando essencial gerenciar riscos cibernéticos combinando probabilidade de ocorrência e potencial de impacto (International Organization for Standardization 2018). Assim, compreender o panorama de incidência, a recorrência temporal dos eventos e os limites para divulgação responsável de evidências desponta como requisito para um planejamento de proteção de dados mais eficaz e aderente às exigências regulatórias vigentes.

Nesse contexto, observa-se que, embora as violações de dados já sejam objeto de um conjunto consolidado de estudos e medidas de proteção — como controles de acesso, criptografia, políticas de segurança da

informação e marcos regulatórios alinhados à LGPD e a padrões internacionais —, a literatura indica que tais respostas permanecem, em grande medida, reativas, fragmentadas e orientadas à conformidade mínima (Rodrigues et al. 2024). Paralelamente, estudos recentes mostram que dados de clientes e cidadãos deixaram de estar confinados a infraestruturas *on-premises* e passaram a circular em ecossistemas distribuídos que combinam computação em nuvem, aplicações em modelo *Software as a Service (Saas)*, integrações baseadas em *APIs*, aplicativos móveis e dispositivos *IoT*, fragmentando o perímetro tradicional de segurança e ampliando a superfície de ataque (Zio e Miqueles 2024).

Diante desse cenário, torna-se imperativo reduzir a frequência e a severidade das violações de dados (Rodrigues et al. 2024), o que exige diagnosticar com rigor os diversos contextos que favorecem incidentes (Val et al. 2024), caracterizar estatisticamente o fenômeno (Mulla et al. 2025) — incluindo previsão temporal da incidência — e mensurar a recorrência por meio de abordagens apropriadas. Em paralelo, é indispensável situar essas ações no marco regulatório vigente, definindo critérios de divulgação responsável de dados que preservem, de forma equilibrada, a utilidade analítica e a privacidade (Ponte et al. 2024).

Nesse sentido, este trabalho não busca substituir arquiteturas, controles ou modelos de segurança já consolidados, mas sim complementá-los por meio de uma abordagem analítica voltada especificamente para violações de dados. Tal abordagem enfatiza a caracterização estatística do fenômeno, a antecipação de incidentes e a mensuração de sua recorrência, oferecendo insumos quantitativos adicionais para o planejamento, a tomada de decisão sobre medidas de proteção, resposta e monitoramento, bem como para a governança de dados pessoais. Ao fazê-lo, contribui para o uso mais eficiente de recursos organizacionais, o fortalecimento da segurança da informação e a conformidade regulatória.

1.2 OBJETIVOS

Este estudo tem como objetivo geral ampliar a compreensão sobre a mitigação de violações de dados a partir de uma abordagem integrada e proativa, que combina análise preditiva para antecipação de incidentes, modelagem de recorrência — voltada à estimativa do tempo até um novo incidente — e mecanismos de proteção de dados. Essa integração visa subsidiar decisões estratégicas de alocação de recursos, definir níveis de serviço (SLAs) adequados para resposta a incidentes e fortalecer a resiliência organizacional frente ao crescimento e à sofisticação das ameaças cibernéticas.

Para alcançar esse objetivo, são definidos os seguintes objetivos específicos:

- Comparar modelos de machine learning e deep learning, avaliando, em séries temporais de incidentes de violação de dados, a capacidade de prever a ocorrência de violação de dados por tipo de organização, de modo a orientar ações proativas para mitigação de riscos e fundamentar a definição de acordos de níveis de serviço (*SLAs*), fornecendo subsídios quantitativos que potencializam a eficiência operacional e fortalecem a governança em segurança da informação;
- Estimar a recorrência de incidentes cibernéticos, por meio da estimativa do tempo até a reincidência do evento em diferentes tipos de organização. Este objetivo específico visa subsidiar ações preditivas e o aprimoramento dos mecanismos de reporte institucional, fundamentando o gerenciamento proativo

de riscos segundo os padrões temporais e setoriais identificados na Análise de Sobrevivência;

- Quantificar e otimizar o *trade-off* entre privacidade e utilidade dos dados, mensurando o impacto da privacidade diferencial (ε) sobre a utilidade analítica e a acurácia preditiva, especialmente quando há injeção de ruído em atributos de alta entropia, e propondo diretrizes para escolha de ε que proteja dados sensíveis sem comprometer o desempenho de modelos; e
- Propor ações preventivas e melhorias operacionais, delineando e priorizando medidas para reduzir a incidência e a recorrência de violações de dados, com base nas evidências empíricas obtidas e em alinhamento com a governança de dados e os *SLAs* institucionais.

1.3 CONTRIBUIÇÕES ACADÊMICAS

Este estudo avança a compreensão das violações de dados ao integrar, de forma coesa, três frentes analíticas: previsão temporal, análise de recorrência e divulgação protegida de dados.

No eixo preditivo, a dissertação propõe um delineamento experimental unificado (Hou, Xue e Zhang 2020) para comparação de famílias de modelos de machine learning e deep learning (Janiesch, Zschech e Heinrich 2021), com protocolo padronizado de preparo, particionamento temporal (Mintarsih et al. 2023), calibração e avaliação (Fissler et al. 2020). Esse enquadramento metodológico fornece um caminho claro para seleção, operação e monitoramento de modelos em contextos setoriais (Jimenez et al. 2020), sem depender de escolhas *ad hoc* (Silva et al. 2020), e favorece sua aplicação em planejamento e gestão de capacidade em segurança da informação (Maddireddy e Maddireddy 2020).

Na dimensão temporal da recorrência, a pesquisa incorpora a análise de sobrevivência (Kaplan–Meier) (Papathanasiou, Demertzis e Tziritas 2023) para estimar probabilidade e janelas até um novo incidente por tipo de organização. Essa perspectiva complementa a previsão de volumes (Ansari et al. 2024) ao introduzir o "quando" como variável de apoio à decisão (Ponce et al. 2023), permitindo ajustar políticas de prevenção, ciclos de monitoramento (Zabierek et al. 2021) e gatilhos operacionais conforme o nível de recorrência de incidentes.

No campo da proteção e transparência, o trabalho estrutura um quadro quantitativo para aplicação de privacidade diferencial em dados tabulares de incidentes (Janiesch, Zschech e Heinrich 2021), relacionando parâmetros de privacidade à utilidade analítica (Saifuzzaman e al. 2024) por métricas de deslocamento distributivo e precisão (Ponte et al. 2024). Como resultado, apresenta diretrizes práticas para divulgação responsável de estatísticas (Ponte et al. 2024) em conformidade com a LGPD e boas práticas internacionais (Fernandes, Machado e Amaral 2023).

Por fim, o estudo disponibiliza um conjunto de artefatos — códigos, parâmetros, *scripts/notebooks*, visualizações e *pipelines* — que garantem rastreabilidade, auditabilidade e verificação independente dos resultados (Sharma e Bantan 2025). Esses materiais funcionam como referências técnicas estruturadas, apoiando a operacionalização dos modelos por equipes técnicas em etapas críticas, tais como seleção e monitoramento de desempenho, procedimentos sistemáticos de retreino, avaliação contínua de deriva temporal e calibragem dos mecanismos de privacidade (Lu et al. 2022). Em conjunto, as contribuições

resultam em um arcabouço coeso que traduz evidência quantitativa em governança e tomada de decisão (Ainslie et al. 2023) — incluindo políticas, *SLAs* e controles (Joint Task Force 2022) — preservando transparência e conformidade regulatória (Govindankutty e Goel 2024).

1.4 ESTRUTURA DO TRABALHO

Para melhor estruturar esse estudo, o trabalho está divido em 6 capítulos. O Capítulo 2 apresenta a revisão bibliográfica da pesquisa, fundamentado em uma análise bibliométrica sobre violações de dados, segurança da informação e privacidade, com foco na identificação e caracterização das principais fontes, periódicos, autores e temas recorrentes da literatura científica, além do mapeamento da estrutura conceitual do campo e das tendências temáticas atuais, proporcionando assim uma visão integrada e abrangente do panorama acadêmico sobre o tema.

O capítulo 3 detalha a modelagem preditiva de incidentes de violação de dados com base nos registros históricos da PRC. A metodologia inicia-se com a análise exploratória, na qual são definidos os critérios de seleção e preparação do conjunto de dados e explicitada a variável "tipo de organização" utilizada na estratificação setorial. Em seguida, aplicam-se diferentes famílias de modelos estatísticos e computacionais para previsão temporal dos incidentes e procede-se à comparação de desempenho entre as abordagens. Essa sequência estabelece uma base objetiva para a tomada de decisões estratégicas e orienta a definição de medidas de prevenção em contextos organizacionais.

O capítulo 4 desenvolve a análise de sobrevivência de incidentes de violações de dados. A base *Data Breach Chronology* da PRC é novamente utilizada e, por meio do estimador de Kaplan–Meier, estima o tempo até a reincidência de um incidente em diferentes setores organizacionais. O capítulo descreve a padronização e preparação dos dados, a estratificação por tipo de organização e os procedimentos estatísticos para a construção das curvas de sobrevivência. Essas curvas permitem caracterizar perfis de recorrência por setor e fundamentar a priorização de recursos, o desenho de políticas preventivas e a definição de níveis de serviço (*SLAs*) proporcionais ao risco temporal observado.

O capítulo 5 trata da aplicação de privacidade diferencial na publicação de dados, com foco na categorização de atributos sensíveis, identificadores e quase-identificadores e na simulação de cenários no setor de reservas hoteleiras. Utiliza-se o mecanismo de *Laplace* para testar diferentes orçamentos de privacidade e níveis de sensibilidade, avaliando o impacto do ruído adicionado sobre a utilidade analítica dos dados. Métricas como entropia, distância Jensen–Shannon e erro percentual absoluto médio (*MAPE*) permitem quantificar o *trade-off* entre proteção e precisão. Os resultados orientam parâmetros para divulgação responsável de dados em conformidade com a LGPD, destacando benefícios, riscos e limitações da abordagem.

O capítulo 6 apresenta as principais conclusões do estudo, integrando os achados das análises preditivas, de sobrevivência e de privacidade diferencial aplicadas a incidentes de violação de dados. Destacam-se as implicações práticas para o planejamento estratégico e a alocação de recursos, além do reconhecimento das limitações metodológicas, como vieses de reporte e volatilidade setorial. Propõe-se, para pesquisas futuras, a ampliação com novas bases de dados, abordagens multivariadas, mecanismos explicáveis para proteção de

dados e aprimoramento dos métodos estatísticos, sempre alinhados às demandas regulatórias, em especial à LGPD, e à necessidade de fortalecer a resiliência organizacional frente aos riscos cibernéticos.

2 REVISÃO BIBLIOGRÁFICA

A ascensão da era digital e a proliferação de dados têm impulsionado a produção acadêmica em um ritmo sem precedentes, especialmente em temas críticos como privacidade e segurança cibernética. Neste contexto, a análise da literatura se torna uma atividade essencial para que pesquisadores possam mapear e identificar os artigos mais relevantes para suas investigações. No entanto, a vasta quantidade de publicações torna esse processo uma tarefa exaustiva e complexa, suscetível a vieses e demorada, o que destaca a necessidade de métodos automatizados para lidar com o volume crescente de estudos (Bispo et al. 2024).

Neste capítulo, o objetivo principal é apresentar uma análise bibliométrica da produção científica sobre privacidade e segurança cibernética, com base em uma metodologia estruturada para extração, processamento e análise de artigos. A coleta de dados foi realizada na plataforma *Scopus*, e o tratamento das informações ocorreu em ambiente R, utilizando os pacotes *bibliometrix* e *biblioshiny*. O pacote *bibliometrix* foi empregado para realizar as análises estatísticas, enquanto o *biblioshiny*, uma ferramenta baseada em interface gráfica, foi utilizado para facilitar a visualização e a exploração interativa dos dados.

A abordagem bibliométrica sistemática adotada segue princípios de estudos que empregam automação na revisão de literatura (Bispo et al. 2024) e articula-se a um conjunto de métricas alinhadas à estrutura analítica do capítulo. Para estruturar essa análise foram consideradas: Coleta e preparação dos dados (2.1); análise dos dados (2.2); principais fontes e periódicos (2.3); principais autores (2.4); tendências temáticas (2.5), estrutura conceitual e temas emergentes (2.6); rede de recorrência (2.7); e conclusão (2.8).

2.1 COLETA E PREPARAÇÃO DOS DADOS

A coleta de dados foi realizada na plataforma *Scopus*, importante base de dados de literatura acadêmica. Foram utilizadas palavras-chave estratégicas para identificar artigos relevantes nas áreas de privacidade e segurança cibernética. Após a busca, os metadados dos artigos selecionados foram exportados e salvos no formato *BibTeX*. O termo booleano de busca, aplicado ao campo de título dos documentos, é apresentado na Listagem 2.1. A busca foi realizada em 10 de agosto de 2025.

Código 2.1: Termo de busca usados na *Scopus*

```
(TITLE-ABS-KEY(privacy) AND TITLE-ABS-KEY(cyber security)) AND PUBYEAR
> 2020 AND PUBYEAR < 2025
```

A base de dados resultante foi submetida a um processo de pré-processamento, no qual artigos duplicados foram removidos para assegurar a exclusividade dos documentos e a integridade da análise. Após essa etapa, o conjunto final de dados foi preparado para importação no ambiente de análise.

2.2 ANÁLISE DA BASE DE DADOS

A base de dados analisada abrange o período de 2021 a 2024 e é composta por 4.262 documentos. A pesquisa foi conduzida por um total de 11.966 autores e utilizou 858 fontes de publicação. A média de idade dos documentos é de 2.03 anos, o que indica que a base é composta majoritariamente por estudos recentes. A Figura 2.1 apresenta uma visão geral das principais métricas da base de dados.

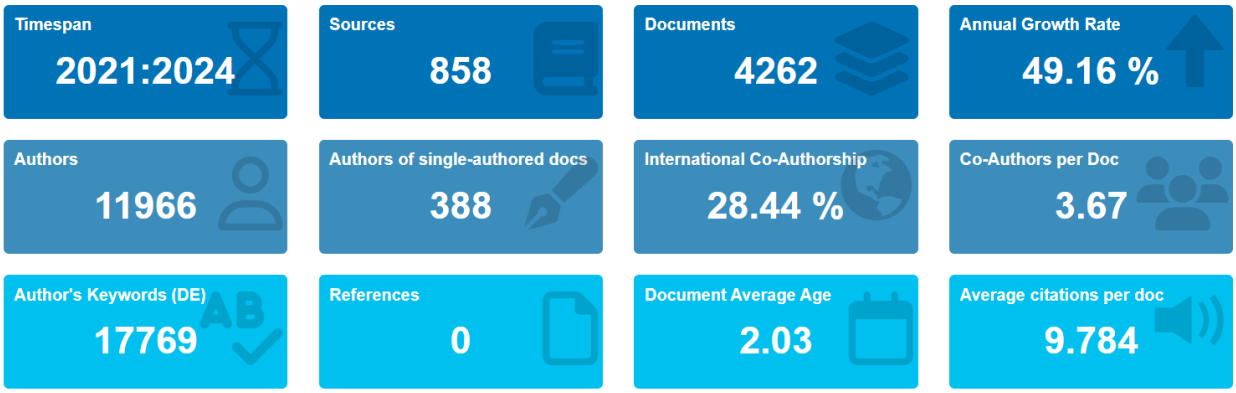

Figura 2.1: Visão Geral da Análise Bibliometrica

A análise do campo de estudo demonstra dinâmicas relevantes quanto ao volume de produção e aos padrões colaborativos. No período de 2021 a 2024, a amostra abrange 4.262 documentos. O corpo autoral totalizou 11.966 pesquisadores e a disseminação do conhecimento ocorreu por meio de 858 veículos distintos. A média de idade dos documentos situa-se em 2,03 anos, evidenciando que o acervo é predominantemente composto por trabalhos recentes.

A pesquisa também revela um forte padrão de colaboração na área, evidenciado pela taxa de crescimento anual de 49.16%. A média de coautores por documento é de 3.67, e a proporção de artigos com um único autor é pequena, totalizando 388 documentos. Um percentual notável de 28.44% dos artigos resulta de coautoria internacional, destacando a natureza global da pesquisa.

2.3 PRINCIPAIS FONTES E PERIÓDICOS

A análise das fontes de publicação revela os periódicos e anais de conferências mais relevantes para o campo de pesquisa. A figura 2.2 ilustra as principais fontes, ranqueadas pelo número de documentos publicados.

As fontes com o maior número de publicações são *IEEE Access*, com 105 documentos, seguidas por *Lecture Notes* in *Networks and Systems* com 101 documentos, e *Lecture Notes in Computer Science* com 88 documentos. A presença proeminente de séries de conferências, como as da *Lecture Notes* e da *ACM*, sugere que a pesquisa em privacidade e segurança cibernética é altamente dinâmica e frequentemente publicada em anais de eventos científicos.

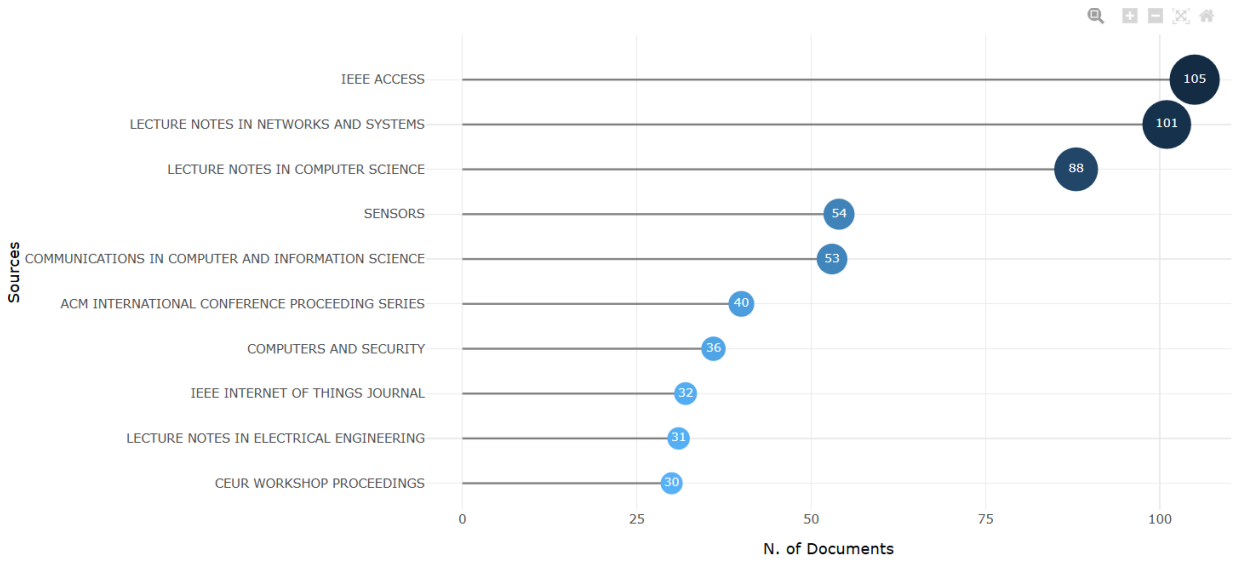

Figura 2.2: Fontes Relevantes

2.4 PRINCIPAIS AUTORES

A análise dos autores mais relevantes revela os pesquisadores mais produtivos no campo da privacidade e segurança cibernética. O gráfico 2.3 a seguir ranqueia os autores com base no número de publicações.

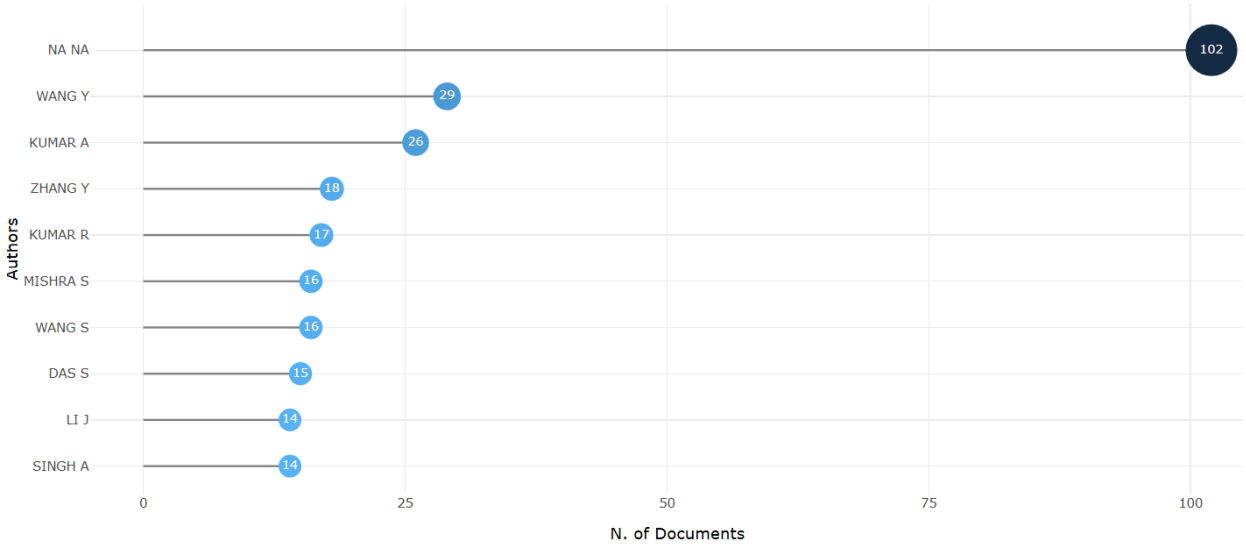

Figura 2.3: Principais Autores

Apesar da entrada anômala com 102 documentos, que provavelmente se refere a dados de autoria não indexados, a análise focada nos autores identificados mostra pesquisadores com contribuições substanciais. Wang Y. se destaca como o autor mais prolífico, com 29 documentos, seguido por Kumar A. (26 documentos) e Zhang Y. (18 documentos). Outros autores notáveis incluem Kumar R., Mishra S., Wang S., Das S., Li J. e Singh A., que publicaram entre 14 e 17 documentos. A predominância de nomes de origem asiática no topo da lista sugere a forte atuação de pesquisadores da Ásia neste campo de estudo.

A análise da estrutura social da pesquisa, representada pelo mapa de colaboração 2.4, evidencia a

natureza global do campo de estudo. A avaliação focada nos autores identificados revela pesquisadores com contribuições substanciais. A intensidade das linhas entre os países indica colaboração robusta e frequente, sugerindo uma troca dinâmica de conhecimento e recursos em escala global.

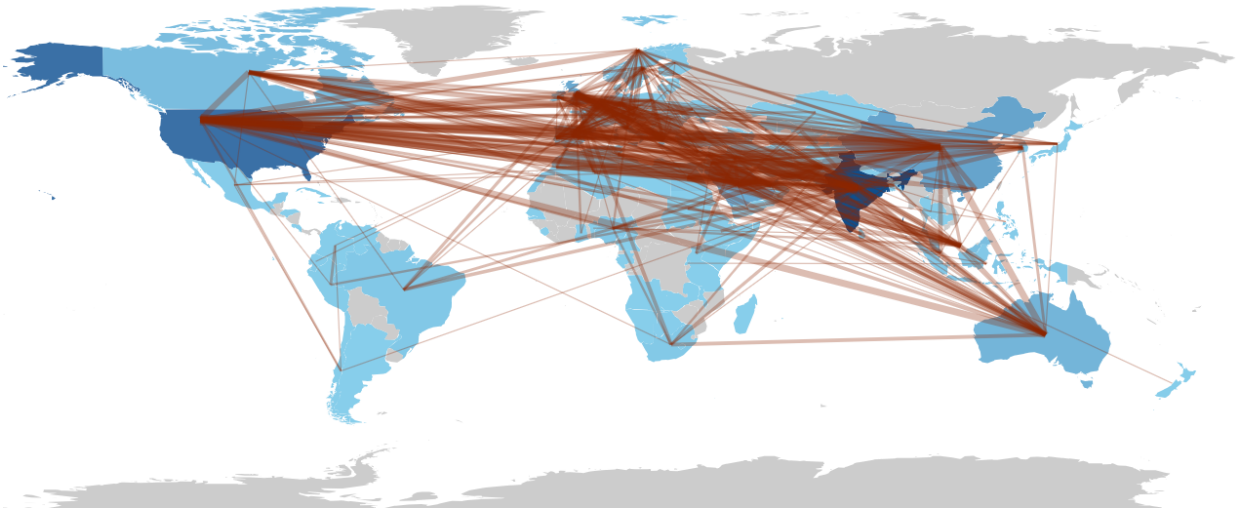

Figura 2.4: Mapa de Colaboração

2.5 TENDÊNCIAS TEMÁTICAS

A análise das palavras-chave mais frequentes é um excelente indicador das tendências e dos temas centrais do campo de pesquisa. O gráfico 2.5 a seguir ranqueia os termos por ocorrência, revelando os focos da literatura.

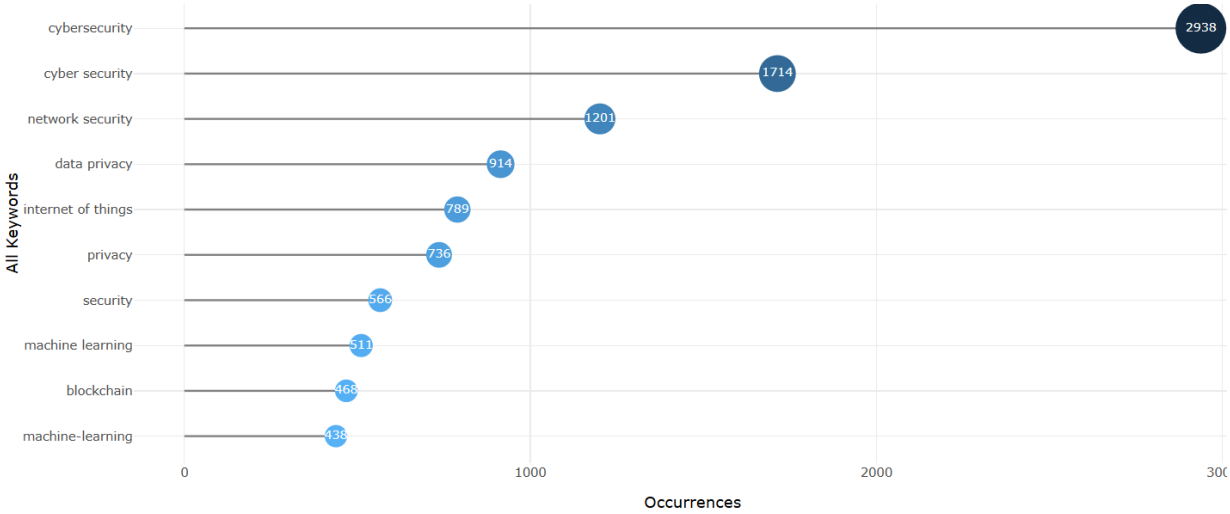

Figura 2.5: Palavras Chave

As palavras-chave mais proeminentes são aquelas diretamente relacionadas à segurança. Os termos *cybersecurity* (2938 ocorrências) e *cyber security* (1714 ocorrências), quando combinados, dominam o

cenário, seguidos por *network security* (1201 ocorrências). Isso confirma que a segurança cibernética é o tema central da pesquisa.

Além dos termos de segurança, a lista inclui palavras-chave que refletem a preocupação com a proteção de dados e tecnologias emergentes. *data privacy* (914 ocorrências) e *privacy* (736 ocorrências) são temas de alta relevância, mostrando que a proteção de informações é um foco principal da pesquisa. A presença de termos como internet of things, machine learning e *blockchain* demonstra que a comunidade acadêmica está explorando ativamente as implicações e as soluções de segurança e privacidade para novas tecnologias.

Além dos termos diretamente associados à segurança, a lista inclui palavras-chave que refletem a crescente preocupação com a proteção de dados e com o avanço de tecnologias emergentes. *Data privacy* (914 ocorrências) e *privacy* (736 ocorrências) aparecem como temas de alta relevância, evidenciando que a proteção de informações permanece um eixo central na produção científica. Da mesma forma, a presença de termos como *internet of things*, machine learning e *blockchain* indica que a comunidade acadêmica explora ativamente tanto as implicações quanto as soluções de segurança e privacidade relacionadas a novas tecnologias.

Complementando essa análise, a nuvem de palavras oferece uma representação visual clara dos temas predominantes na literatura examinada. Nela, o tamanho de cada termo reflete sua frequência na base de dados, permitindo identificar rapidamente os tópicos mais recorrentes. Conforme ilustrado na figura 2.6, o termo dominante é *cybersecurity*, confirmando sua posição como núcleo central das discussões científicas. Outros subtemas relevantes, como *network security* e *data privacy*, também surgem com destaque, reforçando áreas específicas de concentração. Além disso, a presença expressiva de conceitos como *privacy*, *internet of things*, machine learning e *blockchain* evidencia a interseção entre segurança, tecnologias emergentes e proteção de dados.

Figura 2.6: Nuvem de Palavras

A análise da frequência de palavras-chave ao longo do tempo evidencia a evolução e o dinamismo do campo de estudo. O gráfico de linhas 2.7, que apresenta as ocorrências cumulativas entre 2021 e 2024, ilustra um crescimento contínuo e consistente em todos os temas examinados. Nesse conjunto, o termo *cybersecurity* se destaca como líder absoluto, exibindo a curva de expansão mais acentuada e o maior

volume acumulado de ocorrências, o que reforça sua posição como tópico central da pesquisa. Outros temas, como *network security* e *cyber-security*, também apresentam trajetórias de crescimento robustas e estáveis, confirmando a relevância progressiva dessas áreas no período analisado.

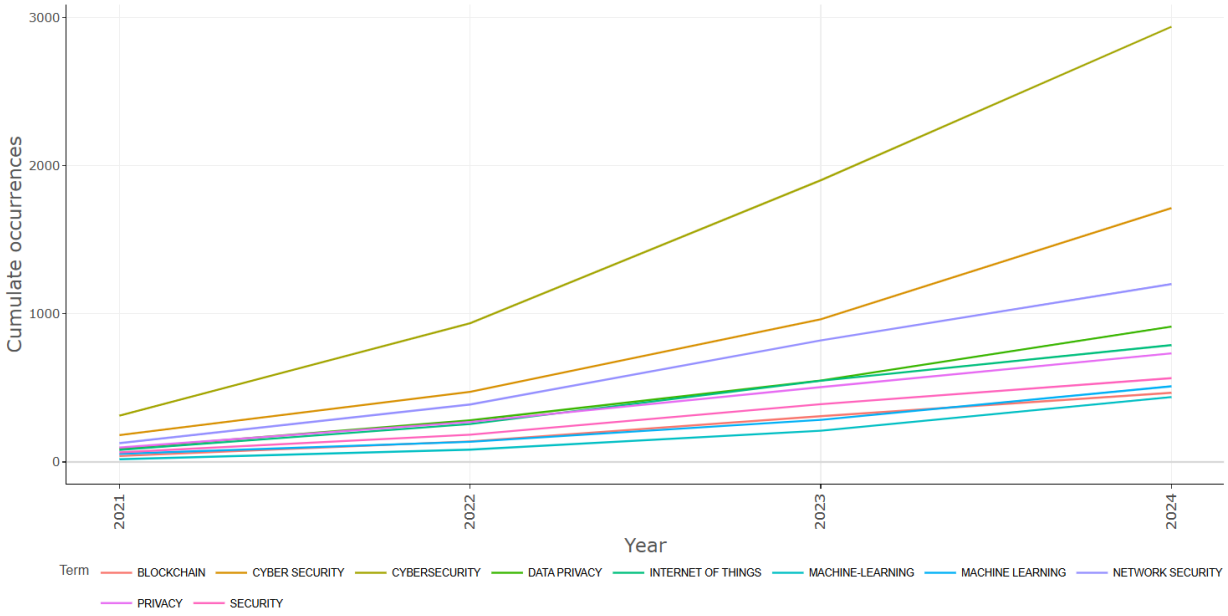

Figura 2.7: Frequência das Palavras ao Longo do Tempo

A ascensão de tópicos como *data privacy, internet of things*, machine learning e *blockchain* indica um interesse crescente e duradouro em suas interconexões com a privacidade e a segurança. O fato de todas as linhas do gráfico estarem em ascensão demonstra que o campo é dinâmico e está em contínua expansão, com todos os subtemas ganhando relevância de forma simultânea.

2.6 ESTRUTURA CONCEITUAL E TEMAS EMERGENTES

A análise da estrutura conceitual por meio do gráfico 2.8 oferece uma visão detalhada e hierárquica dos temas na pesquisa. A área de cada bloco representa a frequência de ocorrência da palavra-chave, fornecendo uma clara visualização da proporção de cada tema no campo.

O gráfico reforça a dominância de *cybersecurity* e *network security* como pilares da literatura, evidenciada pelas maiores áreas do *treemap*. Logo ao lado, o bloco correspondente a *data privacy* e *privacy* destaca a centralidade da proteção de dados no conjunto de pesquisas analisadas. A visualização também evidencia a relevância de áreas tecnológicas como *internet of things*, machine learning e *blockchain*, que aparecem como temas consolidados e de alta frequência na produção científica.

Além desses núcleos principais, a presença de blocos menores — como *intrusion detection, cryptography* e *authentication* — indica que o campo se ramifica em tópicos mais específicos e técnicos, refletindo a diversificação e a especialização natural da área.

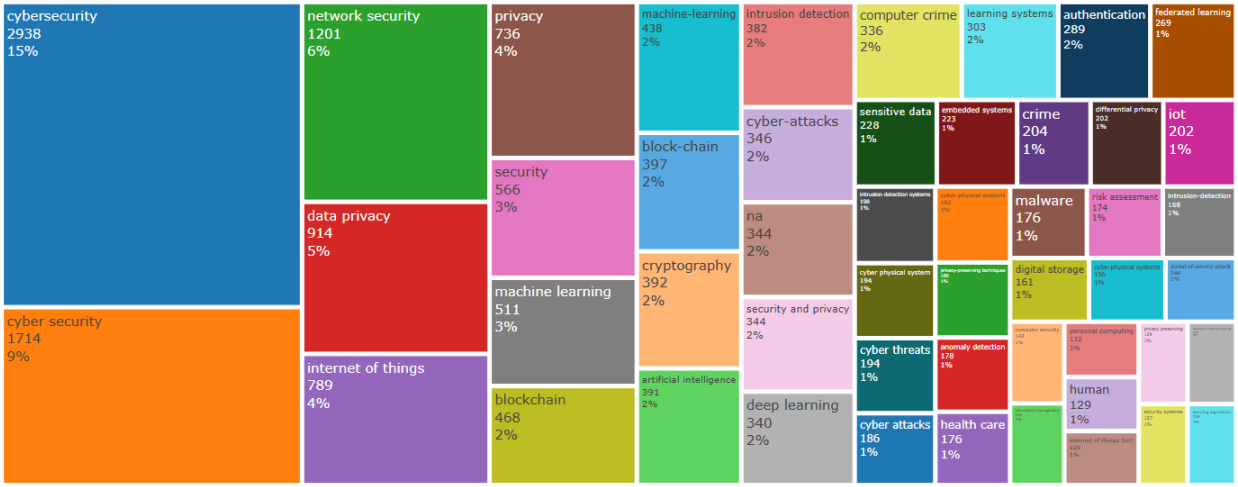

Figura 2.8: Hierarquia dos Temas de Pesquisa

2.7 REDE DE COOCORRÊNCIA

A rede de coocorrência de palavras-chave, apresentada na figura 2.9, revela a estrutura conceitual do campo de pesquisa, evidenciando como os temas se conectam e se agrupam. O gráfico mostra três *clusters* principais, indicando a organização do campo em subáreas especializadas.

O primeiro *cluster*, em vermelho, representa a interseção entre privacidade e segurança, com ênfase em *cybersecurity* e *data privacy*, conectando-se a temas como *internet of things*, *security and privacy* e *authentication*. O segundo cluster, em azul, agrupa abordagens técnicas relacionadas à proteção de sistemas, destacando termos como *cyber security*, *network security*, machine learning, deep learning e *intrusion detection*. O terceiro cluster, em verde, concentra-se em sistemas *ciber*-físicos, com foco em *cyber physical systems* e *embedded systems*.

A presença de conexões entre os clusters indica que, embora especializados, esses subcampos são interdependentes e contribuem para um ecossistema de pesquisa integrado no domínio da privacidade e da segurança cibernética.

2.8 CONCLUSÃO

Este capítulo apresentou uma análise bibliométrica sistemática da literatura em privacidade e segurança cibernética, a partir de dados extraídos da plataforma *Scopus* e processados com os pacotes *bibliometrix* e *biblioshiny* no ambiente R. O mapeamento revelou um campo dinâmico e em clara expansão, caracterizado por um volume expressivo de 4.262 documentos publicados entre 2021 e 2024, média de idade de 2,03 anos e crescimento anual consistente, o que evidencia a atualidade e a vitalidade da produção científica na área. A estrutura analítica adotada — contemplando coleta e preparação dos dados, análise descritiva da base, principais fontes, autores, padrões de colaboração, tendências temáticas, estrutura conceitual e redes de coocorrência — permitiu organizar de forma sistemática os resultados obtidos.

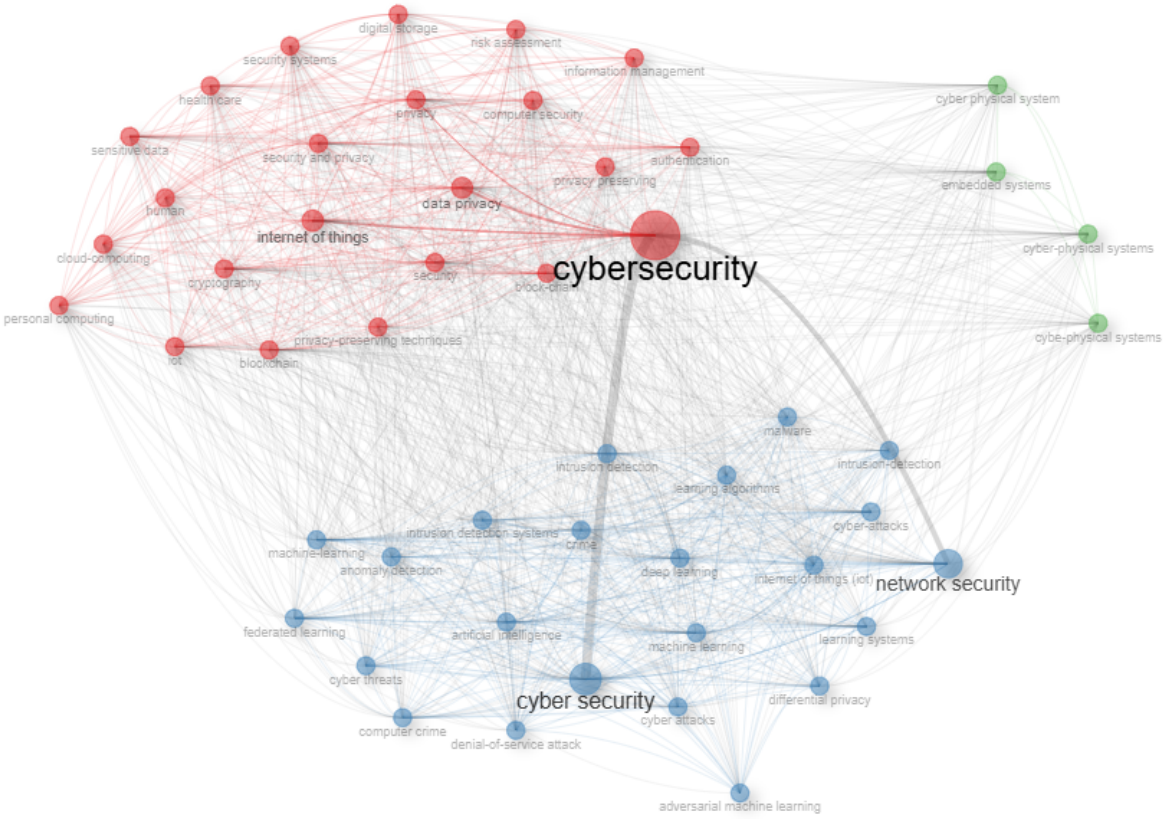

Figura 2.9: Rede de Coocorrência

Os achados indicam a consolidação de um ecossistema de pesquisa globalizado, com forte padrão de colaboração científica, refletido na elevada média de coautores por documento, na baixa proporção de trabalhos de autoria única e na taxa relevante de coautoria internacional. Do ponto de vista temático, observou-se a centralidade de *cybersecurity* e *network security*, acompanhada pela recorrência de tópicos como privacidade de dados, *Internet of Things* (*IoT*), aprendizado de máquina e *blockchain*. As análises de palavras-chave, nuvens de termos, evolução temporal e redes de coocorrência e colaboração evidenciaram ainda a organização do campo em *clusters* conceituais inter-relacionados, que articulam segurança, privacidade e tecnologias emergentes em sistemas distribuídos e *ciber*-físicos.

Em síntese, este capítulo ofereceu uma visão estruturada das principais fontes, autores, padrões de colaboração, temas centrais e frentes emergentes que moldam a produção científica em privacidade e segurança cibernética. Esses resultados não apenas caracterizam o estado da arte da área, mas também fornecem subsídios para a delimitação de lacunas, a formulação de questões de pesquisa e o alinhamento dos estudos empíricos desenvolvidos nos capítulos subsequentes desta dissertação.

3 MODELAGEM PREDITIVA DE INCIDENTES DE VIOLAÇÃO DE DADOS

No eixo preditivo, investiga-se a capacidade de diferentes famílias de modelos de previsão temporal em representar a dinâmica setorial dos incidentes de violação de dados, a partir de séries mensais derivadas da base *Data Breach Chronology* da PRC.

Em um delineamento experimental unificado, comparam-se abordagens estatísticas clássicas, métodos baseados em árvores de decisão e redes neurais profundas, quantificando o desempenho por meio de múltiplas métricas de erro, com ênfase no MAPE (Mulla et al. 2025).

O objetivo central é fornecer subsídios quantitativos para decisões relacionadas à prevenção e mitigação de violações de dados, por meio da avaliação comparativa do desempenho de diferentes famílias de modelos na previsão mensal de incidentes (Petropoulos et al. 2022). Essa avaliação considera tanto o total geral de violações de dados quanto as subdivisões por setor, buscando identificar quais modelos são mais adequados para capturar as distintas características e dinâmicas temporais presentes em cada série, dada a heterogeneidade e a complexidade dos dados ao longo do tempo.

Essa análise, derivada do artigo (Santos et al. 2025), previamente publicado no SBSeg 2025: XXV Simpósio Brasileiro de Cibersegurança, e incorporada a esta dissertação como parte de um estudo mais amplo, com pequenos ajustes e melhorias pontuais para ampliar sua aplicabilidade ao contexto desta pesquisa. Essa integração fortalece a interpretação dos resultados preditivos e orienta sua utilização em cenários de gestão de riscos de segurança da informação, permitindo que gestores e especialistas em cibersegurança tomem decisões mais precisas e estratégicas quanto à alocação de recursos e à implementação de medidas preventivas (Mulla et al. 2025).

3.1 PROCEDIMENTOS METODOLÓGICOS APLICADOS

O conjunto de dados utilizado da PRC reúne informações detalhadas sobre incidentes de violação de dados, incluindo o tipo de ataque, o número de pessoas afetadas, o tipo de organização, a data da ocorrência, o estado e a cidade, entre outros atributos. Neste estudo, o termo setor é empregado especificamente para designar o tipo de organização registrada na base de dados. Assim, para esta pesquisa, foram selecionadas as variáveis data da violação e tipo de organização, que refletem diretamente o setor de atuação correspondente. A Figura 3.1 apresenta uma visão geral da arquitetura proposta.

Inicialmente, conduziu-se uma análise exploratória para caracterizar o conjunto de dados e fundamentar as etapas subsequentes de modelagem. Na fase de preparação dos dados, foram realizados procedimentos como padronização e formatação dos campos, filtragem por tipo de organização e por período, agregação mensal das séries, cálculo do expoente de Hurst — indicador de dependência ou aleatoriedade temporal — e tratamento de valores extremos com base no intervalo interquartil (*IQR*).

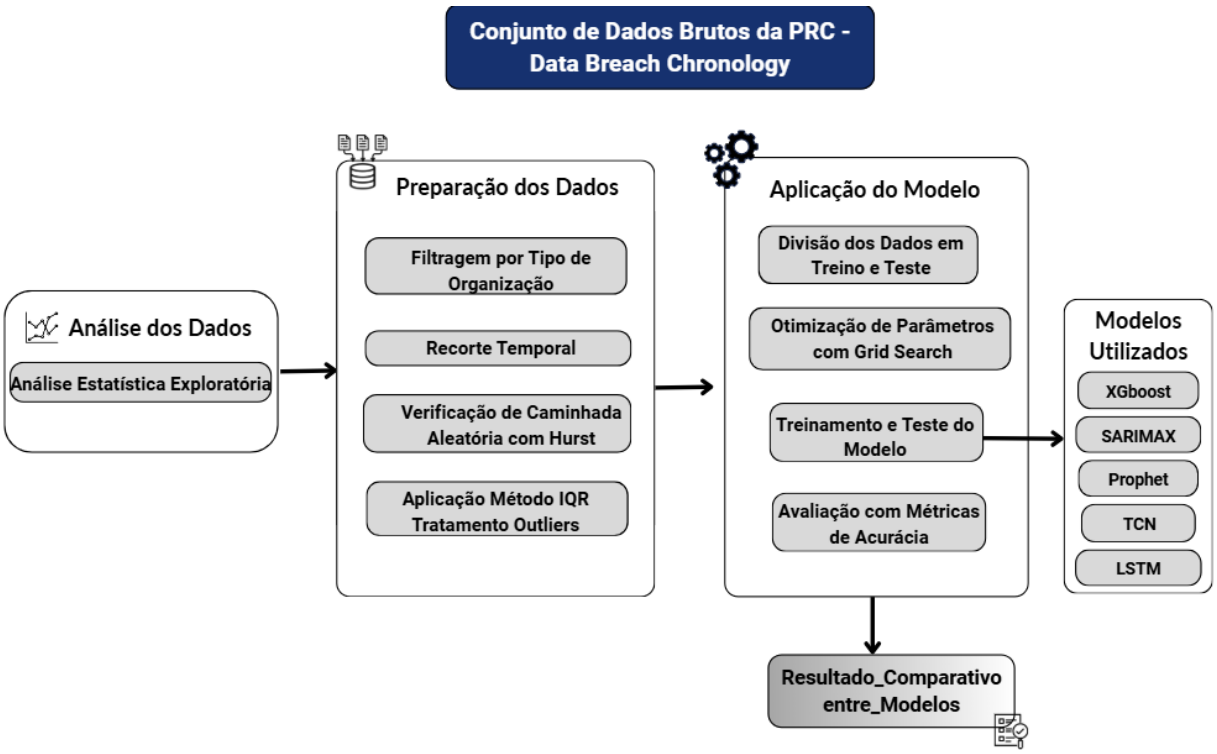

Figura 3.1: Arquitetura Metodológica para Aplicaçao dos Modelos Preditivos

Com as séries consolidadas, foram comparadas três famílias de modelos — estatísticos, métodos de combinação baseados em árvores e redes neurais profundas — calibradas por meio de busca em grade (*grid search*) e avaliadas a partir de uma partição temporal, composta por aproximadamente 85% dos dados para treino e 15% para teste. O desempenho foi mensurado pelas métricas *MAPE*, *MAE* e *RMSE*, possibilitando uma análise comparativa entre os diferentes setores e em relação ao agregado total.

Todo o processo de análise, modelagem e previsão foi realizado em *Python* 3.12, utilizando bibliotecas como *statsmodels*, *fbprophet*, *xgboost*, *tensorflow*, *keras*, *pandas*, *numpy*, *matplotlib* e *scikit-learn*. Os artefatos completos do projeto, incluindo scripts e documentação, estão disponíveis em: <https://github.com/evaneigomes/Modelos-Preditivos-para-Deteccao-de-Viola-oes-de-Dados-Uma-Abordagem-Comparativa>.

3.1.1 Análise e Preparação dos dados

Com o objetivo de compreender o comportamento das violações de dados ao longo do tempo, realizou-se uma análise exploratória. Esta etapa possibilitou a identificação de padrões, sazonalidades e variações significativas (Mulla et al. 2025), fornecendo subsídios para a construção de modelos preditivos mais consistentes e sensíveis às dinâmicas temporais observadas.

A análise da série histórica revelou um crescimento acentuado nas violações de dados a partir de 2010, evidenciando o aumento no uso de serviços digitais e a consequente ampliação da exposição das organizações a riscos de segurança (Ainslie et al. 2023). A Figura 3.2 apresenta essa evolução ao longo do período analisado e evidencia, a partir de 2024, uma queda brusca no número de violações registradas, que não reflete uma redução real no volume de dados vazados, mas decorre de atrasos nas notificações e na

consolidação dos incidentes reportados.

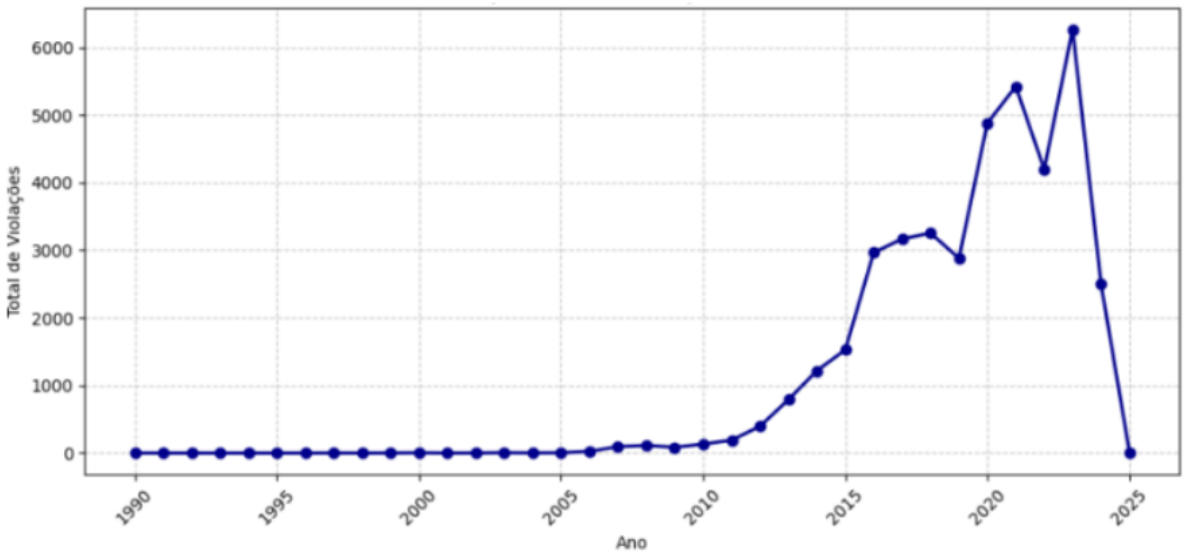

Figura 3.2: Evolução Anual das Violacões de Dados

A figura 3.3 apresenta a evolução anual das violações de dados classificadas por setor, permitindo a identificação de tendências históricas específicas para cada setor. A análise evidencia picos e quedas na ocorrência de incidentes, sugerindo possíveis relações com fatores externos, como mudanças regulatórias, avanços tecnológicos e alterações no perfil das ameaças cibernéticas. A descrição detalhada de cada setor considerado para este estudo encontra-se na tabela 3.1.

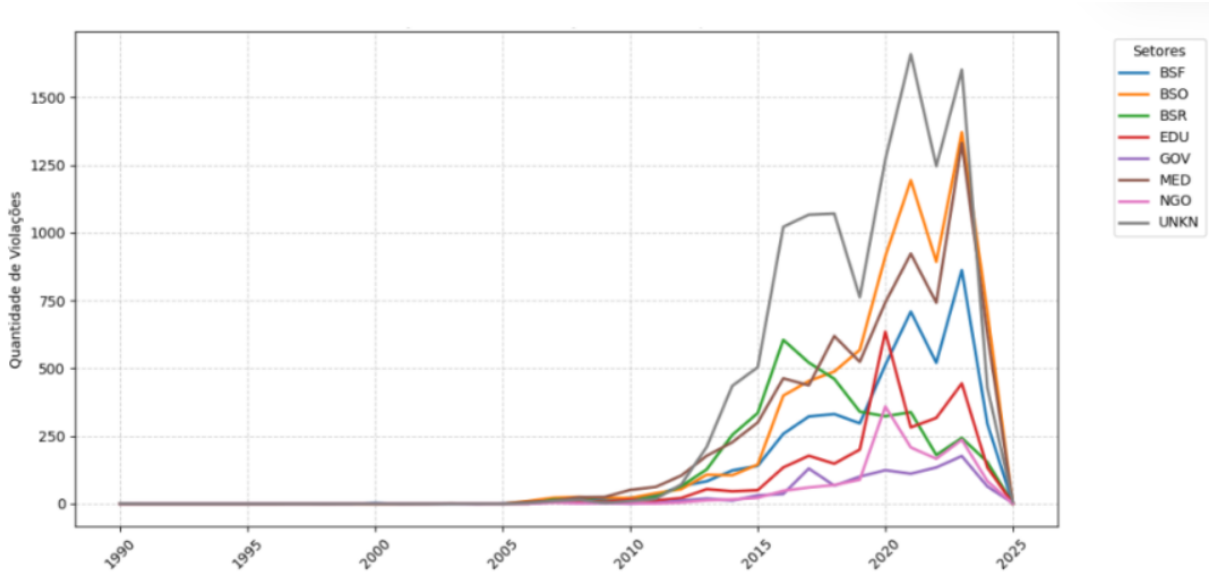

Figura 3.3: Evolução Anual das violações de Dados por Setor

Dessa forma, a análise exploratória realizada constituiu etapa fundamental para a definição da modelagem, ao revelar padrões, sazonalidades e variações específicas dos dados ao longo do tempo. Esses resultados orientaram a escolha dos métodos preditivos mais indicados para capturar as particularidades temporais e setoriais das séries, como será explorado nas subseções seguintes.

Na etapa de preparação dos dados, dos 72.553 registros iniciais de violações, 40.142 foram validados após a limpeza e a exclusão de inconsistências. O processo envolveu ajustes, filtragem por tipo de organização/setor e recorte temporal, estruturando as séries temporais de forma organizada. O expoente de Hurst foi empregado para avaliar se as séries temporais se aproximam de um comportamento aleatório ou se apresentam dependência de longo prazo em sua dinâmica, enquanto os valores atípicos foram tratados com base no intervalo interquartil (*IQR*), visando aprimorar a qualidade preditiva dos dados.

A filtragem dos dados permitiu organizar as informações de forma consistente por setor, assegurando maior homogeneidade ao longo do período analisado. Os setores organizacionais considerados estão apresentados e descritos na tabela 3.1.

Tabela 3.1: Descrição do Setor

Setor	**Descrição**
BSF	Serviços financeiros (bancos, corretoras, seguradoras não-sanitárias)
BSO	Outros negócios (tecnologia, manufatura, utilidades, serviços profissionais)
BSR	Varejo (lojas físicas e online)
EDU	Instituições educacionais (escolas, universidades, serviços educacionais)
GOV	Governo e militares (agências públicas e forças armadas)
MED	Saúde (hospitais e clínicas)
NGO	Organizações sem fins lucrativos (ONGs, igrejas, grupos de advocacia)
UNKN	Setor desconhecido devido a informações insuficientes para classificar

Para garantir maior consistência na análise, o período de estudo foi delimitado entre 2010 e 2023. Os registros anteriores a 2010 apresentavam volume reduzido, possivelmente em razão da menor maturidade dos mecanismos de detecção e da ausência de regulamentações específicas (Rodrigues et al. 2024). Embora o conjunto de dados inclua registros até 2025, optou-se por considerar apenas as informações até 2023, a fim de evitar distorções decorrentes de possíveis atrasos nas notificações de violações de dados, que poderiam comprometer a representatividade temporal da série (Petropoulos et al. 2022).

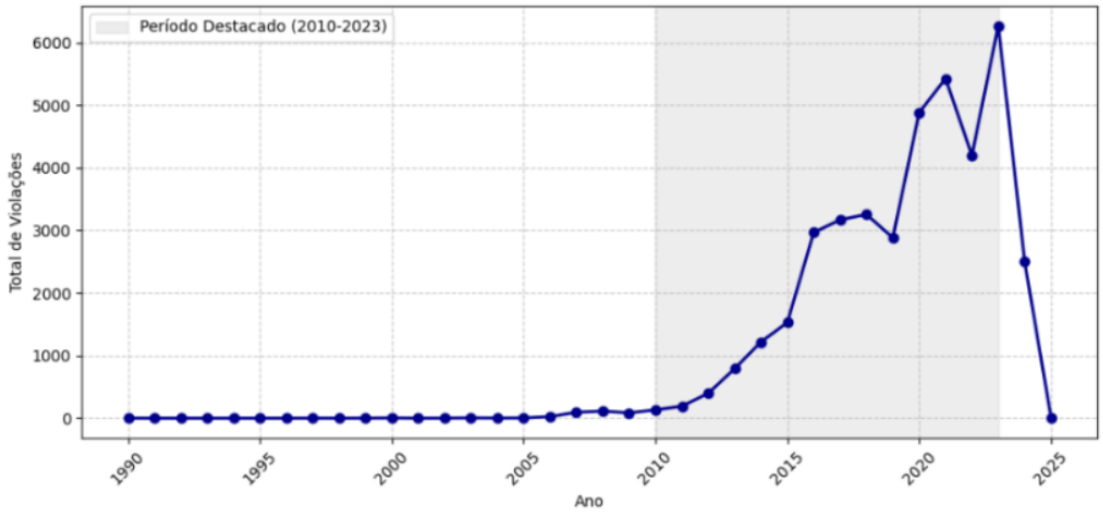

Figura 3.4: Delimitação Temporal

O expoente de Hurst foi utilizado para quantificar o comportamento dinâmico das séries temporais, funcionando como uma medida estatística que avalia a dependência de longo prazo e a previsibilidade da série, permitindo identificar se os dados apresentam persistência ou antipersistência (Zou et al. 2019). Esse

indicador permite classificar a série como tendencial (H>0,5), aleatória ($H \approx 0{,}5$) ou revertente à média (H<0,5). Essa abordagem contribui para uma compreensão mais aprofundada da estrutura temporal dos dados e subsidia a seleção de modelos preditivos mais adequados às características específicas de cada série. Os resultados estão apresentados na Tabela 3.2.

Tabela 3.2: Expoente de Hurst por Setor

Setor	**Expoente de Hurst (H)**
BSF	0.5237
BSO	0.5042
BSR	0.5773
EDU	0.5939
GOV	0.6508
MED	0.5703
NGO	0.6027
UNKN	0.4629
Total Geral	0.5269

Durante a análise, identificaram-se valores discrepantes (*outliers*) ao longo das séries, os quais poderiam comprometer a precisão e a capacidade de generalização dos modelos preditivos (Petropoulos et al. 2022). Para mitigar esse impacto, aplicou-se o método do Intervalo Interquartil (*IQR*), técnica reconhecida pela eficácia na detecção de *outliers*, particularmente em distribuições não paramétricas (Carvalho et al. 2023).

Para aplicar o método do Intervalo Interquartil (*IQR*), deve-se inicialmente calcular o primeiro quartil (Q_1), que corresponde ao valor que separa os 25% menores dados, e o terceiro quartil (Q_3), que separa os 75% menores dados. Em seguida, determina-se o intervalo interquartil, dado por:

$$IQR = Q_3 - Q_1 \tag{3.1}$$

A partir do valor do IQR, é possível identificar possíveis *outliers* no conjunto de dados. Para isso, calculam-se os limites inferior e superior, que definem o intervalo de variação aceitável:

$$\text{Limite inferior} = Q_1 - 1.5 \times IQR \tag{3.2}$$

$$\text{Limite superior} = Q_3 + 1.5 \times IQR \tag{3.3}$$

Qualquer valor abaixo do limite inferior ou acima do limite superior é considerado um *outlier*. Essa abordagem permite lidar com valores extremos sem removê-los, preservando a estrutura dos dados. De acordo com (Kumar, Kaur e Kumar 2023), a combinação do método do intervalo interquartil (*IQR*) com a *winsorização* possibilita a identificação e o ajuste eficiente de *outliers*. A *winsorização* substitui valores extremos pelos valores dos percentis-limite, reduzindo a influência de observações atípicas sem necessidade de exclusão (Martinez, Castle e Hendry 2021), sendo especialmente útil em análises sensíveis como séries temporais (Abraham e Box 1979).

Em síntese, esta subseção estabeleceu uma análise da base de dados e o protocolo de preparação necessários às etapas subsequentes: definição do escopo temporal (2010–2023) e das variáveis-chave (data da violação e tipo de organização); padronização e filtragem por setor e período; agregação mensal;

diagnóstico de dependência temporal pelo expoente de Hurst; e tratamento de valores extremos via *IQR*, com *winsorização* quando aplicável. Esses procedimentos produzem um painel setorial temporalmente consistente, com qualidade e rastreabilidade adequadas, servindo de alicerce para as análises que se seguem.

3.1.2 Aplicação do Modelo

A aplicação dos modelos foi realizada individualmente para cada setor, com o objetivo de avaliar comparativamente seu desempenho na previsão da quantidade de violações de dados ao longo do tempo. Para tanto, os dados foram particionados sequencialmente em conjuntos de treino e teste, de forma a preservar a ordem temporal da série — um requisito essencial em análises de séries temporais (Song et al. 2024).

Cada modelo considerado (*LSTM*, *TCN*, *FBPROPHET*, *SARIMA* e *XGBOOST*) (Petropoulos et al. 2022) foi treinado com aproximadamente 85% dos dados, correspondentes ao período de 2010 a 2021, permitindo o aprendizado dos padrões históricos e tendências relacionados às violações de dados. Os 15% restantes, referentes ao intervalo de 2022 a 2023, compuseram o conjunto de teste, possibilitando uma avaliação imparcial da capacidade preditiva dos modelos em relação a eventos não observados durante o treinamento (Petropoulos et al. 2022).

A implementação setorial dessa estratégia de particionamento garantiu uma comparação equitativa entre os diferentes modelos, permitindo identificar qual abordagem apresenta melhor adequação às especificidades temporais e características de cada segmento organizacional.

Nessa etapa, realizou-se inicialmente o ajuste de hiperparâmetros, procedimento essencial para otimizar o desempenho dos modelos preditivos (Gayam, Yellu e Thuniki 2021). Empregou-se a técnica de busca em grade (*grid search*), método sistemático que explora combinações de parâmetros dentro de um espaço predefinido (Thakkar e Lohiya 2021). Essa abordagem possibilitou testar uma ampla variedade de configurações na grade, avaliando o desempenho de cada modelo por meio das métricas de validação: erro médio absoluto (*MAE*), erro quadrático médio (*RMSE*) e erro percentual absoluto médio (*MAPE*) (Mulla et al. 2025).

A busca em grade foi aplicada em conjunto com validação cruzada, a fim de mitigar o risco de sobreajuste e garantir maior generalização dos resultados (Petropoulos et al. 2022). Esse processo foi executado individualmente para cada modelo e para cada setor organizacional analisado, respeitando as particularidades das séries temporais envolvidas (Petropoulos et al. 2022). Como resultado, foi possível identificar as configurações mais adequadas aos dados específicos de cada segmento, contribuindo diretamente para a acurácia das previsões obtidas (Carvalho et al. 2023).

Com a preparação concluída e as séries temporais devidamente estruturadas, procedeu-se à aplicação dos modelos preditivos. Foram selecionadas abordagens com naturezas e níveis de complexidade distintos — modelos estatísticos clássicos (*SARIMA*), modelos aditivos com heurísticas específicas para séries temporais (*fbprophet*), métodos baseados em árvores de decisão em regime de *boosting* (*xgboost*) e redes neurais profundas voltadas a dados sequenciais (*LSTM* e *TCN*). Essa diversidade de modelos, sintetizada na Tabela 3.3, permite explorar diferentes formas de capturar tendência, sazonalidade, não linearidades e dependências de longo prazo, oferecendo múltiplas perspectivas sobre a dinâmica dos incidentes de violação de dados.

Para mensurar o desempenho dos modelos aplicados à previsão de violações de dados organizacionais,

Tabela 3.3: Descrição dos Modelos Preditivos Utilizados

Modelo	**Tipo**	**Descrição e Características Relevantes**
SARIMA	Modelo estatístico clássico	Lida com componentes de tendência, sazonalidade e resíduos. Requer séries estacionárias e pode apresentar alta complexidade para ajuste de parâmetros.
fbprophet	Modelo estatístico aditivo com heurísticas	Desenvolvido para dados de séries temporais com fortes efeitos sazonais e de tendência. Automatiza a detecção de tendências e sazonalidades e é tolerante a falhas ou lacunas nos dados. Pode, contudo, superestimar tendências e ter menor desempenho com dados com ruídos ou *outliers* extremos.
xgboost	Ensemble de árvores de decisão (*Boosting*)	Algoritmo baseado em árvores de decisão que constrói um modelo preditivo de forma aditiva. Oferece alta performance com dados estruturados e boa explicabilidade de variáveis. Contudo, não modela diretamente a sequência temporal e requer engenharia de features para dados de séries temporais.
LSTM	Rede neural recorrente (Deep Learning)	Capaz de aprender dependências de longo prazo em dados sequenciais. É eficaz com dados não lineares e sazonais. Suas desvantagens incluem a necessidade de muitos dados e tempo de treinamento, além de ser sensível ao ajuste de hiperparâmetros.
TCN	Rede neural convolucional temporal (Deep Learning)	Utiliza convoluções para modelar dependências em séries temporais. Apresenta melhor paralelismo que *LSTM* e é capaz de captar padrões de longo prazo de forma mais estável. No entanto, é mais complexo para configurar e ainda menos difundido na literatura de séries temporais.

foram utilizadas três métricas estatísticas: o erro médio absoluto (*MAE*), a raiz do erro quadrático médio (*RMSE*) e o erro percentual absoluto médio (*MAPE*) (Mulla et al. 2025).

O *MAE* (*Mean Absolute Error*) representa a média dos erros absolutos entre os valores reais e previstos, fornecendo uma estimativa direta do desvio médio na mesma unidade da variável analisada, o que facilita sua interpretação (Carvalho et al. 2023). O *RMSE (Root Mean Squared Error)* corresponde à raiz quadrada da média dos quadrados dos erros, atribuindo maior peso a desvios elevados em função do termo quadrático (Rahimpour et al. 2024), de modo que penaliza erros grandes de forma mais intensa que o *MAE* (Song et al. 2024) e se torna, por isso, mais sensível à presença de valores atípicos (Ding et al. 2025). O *MAPE* (*Mean Absolute Percentage Error*), por sua vez, mede o erro percentual médio em relação aos valores reais, expressando-se em porcentagem e permitindo comparações equitativas entre diferentes séries ou modelos, o que reforça sua utilidade prática pela alta interpretabilidade (Fildes, Ma e Kolassa 2019).

Neste estudo, o *MAPE* foi adotado como métrica principal de avaliação da acurácia preditiva dos modelos (Mulla et al. 2025), em razão de sua capacidade de comparação entre diferentes séries temporais e de sua interpretação direta em termos percentuais (Fildes, Ma e Kolassa 2019). Seguindo os critérios propostos por (Lewis 1982), os valores de *MAPE* foram classificados conforme apresentado na Tabela 3.4.

Tabela 3.4: Classificação da Precisão das Previsões com base no MAPE

Intervalo do *MAPE* (%)	**Classificação**	**Descrição**
< 10	Previsão Altamente Precisa	Indica previsões com erro percentual muito baixo; trata-se de um desempenho excelente para fins analíticos e operacionais.
10 – 19,99	Boa Previsão	Indica modelos com boa acurácia, confiáveis para aplicações práticas, embora com margem de erro perceptível.
20 – 49,99	Previsão Razoável	As previsões possuem erro moderado; podem ser utilizadas em contextos exploratórios, mas com cautela nas decisões.
50 ou mais	Previsão Imprecisa	Reflete alto grau de erro percentual, tornando o modelo inadequado para aplicações que exigem confiabilidade nas estimativas.

3.2 RESULTADOS

Esta seção apresenta os principais resultados da análise da base de dados da PRC, com ênfase na comparação do desempenho preditivo dos modelos a partir do *MAPE* como métrica principal. Ao final,

sintetizam-se as considerações, aplicações práticas e teóricas.

3.2.1 Avaliação Comparativa da Precisão entre os Modelos

A análise dos resultados obtidos com os modelos *LSTM*, *fbprophet*, *SARIMA*, *TCN*, e *xgboost* foi realizada com base nas métricas *MAE*, *RMSE* e *MAPE*, considerando a previsão de violações de dados em diferentes tipos de organização, tendo o *MAPE* como métrica principal de avaliação. Foi incluída a figura 3.5 com resultados do *MAPE* por setor e por modelo aplicado:

Figura 3.5: MAPE por Modelo com Comparação entre Setores Organizacionais

Os resultados obtidos para cada setor, com base na aplicação dos diferentes modelos preditivos, foram avaliados segundo a classificação de (Lewis 1982), utilizando o *MAPE* como principal métrica de acurácia.

No setor Business–Financial (BSF), o modelo *LSTM* apresentou o melhor equilíbrio entre erros relativos e absolutos, com *MAPE* de 21,94%. O *xgboost* exibiu o menor erro relativo (*MAPE* 19,96%), embora a análise de *MAE* e *RMSE* não esteja representada na figura. Os demais modelos apresentaram *MAPE* superior: *SARIMA* (55,43%), *fbprophet* (39,47%) e *TCN* (34,46%).

No setor Business–Other (BSO), o *xgboost* apresentou o menor *MAPE* (11,40%), seguido de *LSTM* (16,64%) e *TCN* (19,44%). *fbprophet* (32,13%) e *SARIMA* (34,77%) mostraram desempenho inferior, refletindo a heterogeneidade do setor.

No setor Business–Retail (BSR), todos os modelos apresentaram resultados insatisfatórios, com *MAPE* de 68,00% (*LSTM*), 73,13% (*fbprophet*), 80,64% (*SARIMA*), 68,78% (*TCN*) e 83,34% (*xgboost*). A alta volatilidade e os eventos atípicos comprometeram a precisão das previsões.

No setor de Educação (EDU), o modelo *LSTM* registrou *MAPE* de 32,99%, enquanto *fbprophet* e *SARIMA* obtiveram desempenho semelhante (34,40% e 34,90%, respectivamente). O *xgboost* apresentou *MAPE* de 35,47%, e o *TCN*, 42,62%, indicando que a presença de tendência e sazonalidade beneficiou modelos estatísticos.

No setor Governamental (GOV), o *LSTM* destacou-se com *MAPE* de 31,01%, seguido de *xgboost* (35,03%) e *fbprophet* (39,47%). *SARIMA* (42,57%) e *TCN* (49,37%) apresentaram previsões menos precisas.

No setor Médico (MED), os modelos exibiram *MAPE* entre 19,44% (*TCN*) e 32,19% (*SARIMA*). O *LSTM* obteve 23,76%, *fbprophet* 26,54% e *xgboost* 21,53%, refletindo séries relativamente estáveis.

Para Organizações Não Governamentais (NGO), o menor *MAPE* foi do *LSTM* (37,10%), enquanto *SARIMA* (53,38%), *fbprophet* (48,57%), *TCN* (42,50%) e *xgboost* (42,10%) apresentaram desempenho inferior.

No Total Geral, o *xgboost* apresentou o menor *MAPE* (5,97%), seguido de *TCN* (7,82%) e *LSTM* (12,21%). *fbprophet* e *SARIMA* apresentaram 17,63% e 27,61%, respectivamente, indicando boa estabilidade preditiva no agregado.

Nos setores desconhecidos (*UNKN*), o *xgboost* foi o modelo mais preciso (10,03%), seguido de *LSTM* (12,78%) e *TCN* (12,88%). *SARIMA* (32,76%) e *fbprophet* (22,54%) apresentaram desempenho inferior.

De modo geral, os modelos *LSTM* e *TCN* mostraram desempenho consistente entre os setores, com destaque para o *xgboost*, que apresentou os menores valores de *MAPE* em vários contextos, especialmente no agregado e nos setores BSO e *UNKN*. *SARIMA* e *fbprophet* tiveram desempenho inferior nos setores mais voláteis, como o BSR, onde todas as abordagens registraram baixa acurácia.

3.2.2 Considerações Finais

Os resultados revelam que o desempenho preditivo varia de forma significativa entre setores, modelos e características estruturais das séries. O *xgboost* apresentou os menores valores de *MAPE* no agregado e em setores como BSO e *UNKN*, demonstrando elevada capacidade preditiva quando as séries são mais estáveis ou apresentam menor ruído. Por outro lado, *LSTM* e *TCN* mantiveram desempenho consistente em múltiplos setores, destacando-se em contextos com não linearidade e padrões mais complexos, embora não tenham sido sempre superiores em todos os cenários. Modelos estatísticos, como *SARIMA* e *fbprophet*, apresentaram limitações mais evidentes em setores voláteis, com destaque para o varejo (BSR), onde todos os modelos registraram *MAPE* elevado, reforçando o desafio inerente à previsão em ambientes altamente instáveis.

Na prática, recomenda-se uma estratégia diferenciada por setor: utilizar *xgboost* em séries com boa estabilidade ou como *baseline* de menor erro relativo no agregado; aplicar *LSTM* e *TCN* em contextos que exigem maior robustez a padrões não lineares e interações temporais mais complexas; e empregar modelos estatísticos como alternativas interpretáveis, especialmente quando sazonalidades regulares ou tendências bem definidas predominam. Métricas como *MAPE*, *MAE* e *RMSE* devem orientar definições de *SLAs*, intervalos de tolerância e políticas de alocação de recursos, considerando a previsibilidade específica de cada setor. A calibração periódica, com re-treinos e monitoramento de deriva temporal, torna-se fundamental diante de mudanças de comportamento nas séries.

Os achados reforçam que a escolha do modelo deve considerar a estrutura da série, o grau de volatilidade setorial e a finalidade operacional da previsão. Propõe-se a adoção de portfólios híbridos de modelos por setor, *SLAs* proporcionais à previsibilidade observada e integração sistemática das previsões ao ciclo de resposta organizacional. A incorporação de variáveis exógenas, aliada a monitoramento contínuo da

performance e ao ajuste dinâmico dos modelos, pode fortalecer a gestão de riscos e aprimorar a governança cibernética baseada em previsões.

4 ANÁLISE DE SOBREVIVÊNCIA DE INCIDENTES CIBERNÉTICOS

A partir da compreensão preditiva da frequência e da distribuição dos incidentes, o estudo avança para a análise temporal da recorrência das violações de dados por meio de técnicas de sobrevivência (Wang, Li e Reddy 2019), utilizando a base *Data Breach Chronology* da PRC como fonte empírica. Essa abordagem estima o tempo até a ocorrência de um novo incidente cibernético, incorporando informações parciais de observações que ainda não sofreram novo evento, e amplia a compreensão do comportamento das séries ao oferecer uma perspectiva complementar às técnicas tradicionais de previsão (Bradley, Alhajjar e Bastian 2023).

A partir da compreensão preditiva da frequência e da distribuição dos incidentes, o estudo avança para a dimensão temporal da recorrência dos incidentes de violação de dados por meio da análise de sobrevivência (Wang, Li e Reddy 2019), utilizando também a base *Data Breach Chronology* da PRC como fonte empírica. Essa abordagem estima o tempo até a ocorrência de um novo incidente cibernético e oferece uma perspectiva complementar às técnicas de predição, abordando uma nova perspectiva que complementa as técnicas tradicionais de previsão (Bradley, Alhajjar e Bastian 2023).

A Análise de Sobrevivência permite estimar, para cada setor, a função de sobrevivência e as respectivas probabilidades de permanecer sem novo incidente (Val et al. 2024), além de identificar fatores que influenciam o intervalo entre violações de dados (Papathanasiou, Demertzis e Tziritas 2023). Essa abordagem aprofunda a compreensão do ciclo de vida dos incidentes ao revelar padrões temporais e variabilidades intrínsecas que afetam diretamente a recorrência(Alvarez et al. 2025).

Essa abordagem permite que gestores desenvolvam políticas fundamentadas em estimativas temporais (Petropoulos et al. 2022), otimizem a alocação de recursos (Kotsias, Ahmad e Scheepers 2022) e implementem medidas preventivas ajustadas à vulnerabilidade específica de cada setor ao longo do tempo (Val et al. 2024). Com isso, fortalece-se a capacidade de antecipação e de resposta, ampliando a eficácia na mitigação da recorrência de incidentes de violação de dados (Rodrigues et al. 2024).

4.1 PROCEDIMENTOS METODOLÓGICOS APLICADOS À ANÁLISE DE SOBREVIVÊNCIA

A figura 4.1 sintetiza o encadeamento metodológico adotado na análise de sobrevivência de incidentes cibernéticos. O processo inicia-se na seleção da fonte de dados, a base PRC – *Data Breach* Chronology, e no recorte temporal de 2010–2023, escolhido por oferecer maior consistência e padronização dos registros. Em seguida, procede-se à preparação dos dados, etapa em que são realizadas a limpeza dos registros, a padronização dos formatos de data e a harmonização da taxonomia setorial, de modo a garantir comparabilidade entre os diferentes tipos de organização.

Na sequência, são construídas as variáveis analíticas centrais da modelagem de sobrevivência: a duração em dias entre incidentes (`duration_days`), o indicador de ocorrência de evento (`event`) e o código setorial padronizado (`sector_code`). Com esse conjunto estruturado, aplica-se o estimador de Kaplan–Meier, implementado em ambiente do *Google Colab* com a biblioteca *lifelines*, para obter as curvas de sobrevivência estratificadas por setor, bem como as medianas de tempo e as probabilidades $S(t)$ em diferentes horizontes. Por fim, as saídas analíticas são interpretadas sob a perspectiva de governança: os resultados subsidiam a priorização de setores mais críticos, a definição de acordos de nível de serviço (*SLAs*) alinhados ao risco temporal de reincidência e o ajuste de estratégias de mitigação de incidentes de violação de dados.

Figura 4.1: Fluxo metodológico da análise de sobrevivência de incidentes cibernéticos

O recorte temporal da base de dados da PRC, para o período de 2010–2023, foi definido por corresponder ao intervalo em que a cobertura dos registros e a padronização dos campos são mais consistentes (Ainslie et al. 2023). Essa delimitação busca mitigar vieses de subnotificação nos anos iniciais do cadastro e assegurar maior comparabilidade entre os diferentes tipos de organização ao longo do tempo.

As variáveis analíticas centrais desta pesquisa são a data da violação e o tipo de organização (ver Tabela 3.1), ambas fundamentais para caracterizar a dinâmica temporal dos incidentes. A data da violação estabelece o eixo cronológico das análises e orienta a construção das séries temporais (Mulla et al. 2025). Para este estudo, entende-se que o tipo de organização corresponde ao setor de atuação da respectiva entidade.

A variável temporal constitui a base para a decomposição das séries em componentes de tendência e de sazonalidade (Rhif e al. 2019), além de permitir a identificação de choques exógenos, como mudanças regulatórias ou eventos de grande impacto (Marczak e Proietti 2016). No contexto da análise de sobrevivência, essa mesma dimensão temporal é utilizada para derivar as durações entre incidentes e definir o tempo até o

evento, o que viabiliza a estimação de funções de sobrevivência e de risco associadas às violações de dados. Também sustenta a análise de rupturas estruturais, sinalizando transições relevantes no comportamento de incidência das violações. Esses aspectos temporais são fundamentais para compreender a evolução do fenômeno, interpretar a dinâmica de recorrência e contextualizar as variações entre períodos e grupos de observação (Song et al. 2024).

O tipo de organização é a variável que sustenta a segmentação por setor, necessária para controlar as heterogeneidades estruturais (Val et al. 2024). Neste estudo, o termo setor é empregado especificamente para designar o tipo de organização registrada na base de dados. Setores distintos apresentam variações relevantes na exposição, na maturidade de segurança, na regulação e no tipo de dado tratado. Essa segmentação permite contrastar as curvas de sobrevivência por grupo setorial (Val et al. 2024), testar diferenças significativas e derivar implicações práticas, como o ajuste de estratégias de mitigação e a priorização de controles (Val et al. 2024).

Combinadas, as variáveis data da violação e setor da organização permitem análises comparativas consistentes de recorrência, articulando uma base temporal uniforme (agregação mensal e ordenação dos eventos) à heterogeneidade setorial observada. Para assegurar essa uniformidade, as datas foram padronizadas, os formatos ajustados e o índice temporal definido, o que viabilizou a construção de históricos de eventos por setor, a derivação das durações (tempo entre incidentes) e a codificação do indicador de evento ou censura. Assim, obtém-se um painel (*setor* × *tempo*) apropriado às etapas de análise de sobrevivência, garantindo consistência temporal e comparabilidade.

O estimador de Kaplan–Meier foi empregado para obter a função de sobrevivência $S(t)$ (Val et al. 2024), interpretada como a probabilidade de não ocorrer um novo incidente até o tempo t após o evento anterior (Val et al. 2024). Em conformidade com o *pipeline* implementado *Google Colab*, derivou-se a variável de duração como o número de dias desde a última violação por setor (Alvarez et al. 2025), bem como o indicador de evento 1 para ocorrência observada e 0 para censura à direita (Alvarez et al. 2024). O estimador não paramétrico foi calculado pelo produto-limite

$$S(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{4.1}$$

(Papathanasiou, Demertzis e Tziritas 2023), com intervalos de confiança obtidos pela variância de Greenwood, mantendo o pressuposto de censura não informativa.

Para esta análise de sobrevivência, utilizou-se o estimador de Kaplan–Meier, uma ferramenta não paramétrica amplamente empregada para estimar a função de sobrevivência $S(t)$, que representa a probabilidade de que um evento de interesse, como a reincidência de um incidente, não ocorra até um instante de tempo t (Val et al. 2024). Essa técnica constrói uma função por partes, decrescente em degraus nos tempos observados dos eventos, incorporando tanto dados completos quanto censurados, isto é, observações nas quais o evento não foi observado até o final do período de análise (Val et al. 2024).

Neste estudo, utilizou-se o *pipeline* Colab combinado com a biblioteca *lifelines* para derivar a variável de duração, definida como o número de dias desde a última violação, por setor ou estrato. O indicador de evento foi codificado como 1 para ocorrência observada e 0 para censura à direita (Alvarez et al. 2025,

Alvarez et al. 2024). O estimador de Kaplan–Meier foi calculado pelo produto-limite:

$$S(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{4.2}$$

onde d_i representa o número de eventos ocorridos no tempo t_i e n_i o número de indivíduos em risco imediatamente antes de t_i (Papathanasiou, Demertzis e Tziritas 2023). Os intervalos de confiança foram obtidos pela variância de Greenwood, sob o pressuposto de censura não informativa.

Para capturar a heterogeneidade estrutural, as curvas foram estratificadas segundo o tipo de organização, permitindo comparações entre perfis de risco (Papathanasiou, Demertzis e Tziritas 2023) e, quando aplicável, a aplicação do teste log-rank. O processo de preparação dos dados envolveu a padronização de datas, a ordenação temporal, o tratamento mínimo de valores ausentes e a definição explícita do conjunto em risco por estrato (Papathanasiou, Demertzis e Tziritas 2023). A visualização foi configurada com escala temporal de 0–30 dias, rótulos e grade informativa, assegurando rastreabilidade entre os dados, o código de estimação e os gráficos gerados (Bradley, Alhajjar e Bastian 2023).

Foram geradas curvas de sobrevivência para cada setor, permitindo comparar os padrões de recorrência de incidentes entre setores. A análise foi conduzida utilizando *Python*, com bibliotecas como *pandas* para tratamento dos dados, *matplotlib* para visualização e para modelagem de sobrevivência (Val et al. 2024). Scripts e documentação estão disponíveis em repositório público no *GitHub*.

4.2 RESULTADOS

Conforme a Figura 4.2, as curvas de sobrevivência Kaplan–Meier por tipo de organização evidenciam padrões distintos de reincidência entre os setores analisados. Observa-se que todas as curvas apresentam queda acentuada nos primeiros dias após uma violação, indicando que a probabilidade de ocorrer um novo incidente é significativamente maior logo no período inicial. Entretanto, a intensidade desse decaimento varia de forma substantiva entre os grupos.

Setores como Saúde (MED) e Outros Negócios (BSO) apresentam declínios acentuados na função de sobrevivência nos períodos iniciais, evidenciando alta probabilidade de reincidência de incidentes em janelas temporais curtas. Esse padrão é compatível com ambientes de elevada criticidade operacional e regulatória, nos quais a exigência de disponibilidade contínua e a concentração de dados sensíveis ampliam a exposição a novos eventos adversos. De forma semelhante, os setores Business–Financial (BSF) e Business–Retail (BSR) também exibem quedas aceleradas nas curvas de sobrevivência, refletindo sua intensa exposição a ataques direcionados e a contextos fortemente transacionais.

O setor Desconhecido (*UNKN*) igualmente apresenta queda abrupta da função de sobrevivência, especialmente nos estágios iniciais, sugerindo alta reincidência de incidentes em curto prazo. Esse comportamento pode estar associado à heterogeneidade do agrupamento, que reúne organizações com diferentes perfis de risco e níveis de maturidade em segurança, concentrando eventos recorrentes em intervalos reduzidos. Em contraste, o setor de Organizações Sem Fins Lucrativos (NGO) exibe uma curva mais suavizada ao longo do tempo, indicando menor frequência relativa de reincidência ou maior variabilidade nos intervalos entre

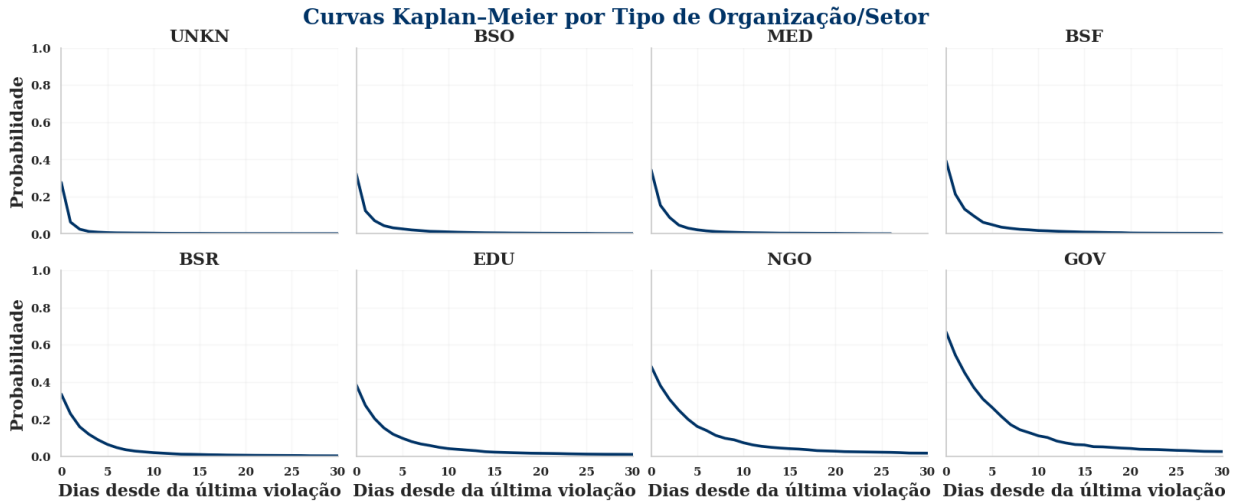

Figura 4.2: Kaplan-Meier

incidentes. Já o setor de Educação (EDU) apresenta um padrão intermediário, com declínio menos abrupto quando comparado aos setores de maior risco.

Esse conjunto de curvas evidencia que o risco temporal não é homogêneo entre setores, reforçando a importância de estratificar políticas de segurança cibernética de acordo com o perfil operacional de cada tipo de organização. A análise temporal permite ainda inferir intervalos médios distintos até a ocorrência de um novo incidente, gerando insumos diretos para a priorização de recursos, definição de janelas de monitoramento e ajuste de estratégias de prevenção. Setores com queda mais rápida na sobrevivência demandam ciclos de resposta mais curtos e mecanismos contínuos de mitigação, enquanto setores com curvas mais estáveis podem adotar abordagens de monitoramento com periodicidade mais espaçada, sem perda significativa de cobertura de risco.

4.2.1 Considerações Finais

A aplicação da Análise de Sobrevivência, especificamente por meio do estimador Kaplan-Meier, proporcionou uma visão detalhada da dinâmica temporal da reincidência de incidentes de violação de dados em diferentes setores organizacionais. Os resultados evidenciam que o perfil setorial é determinante para as estratégias de mitigação, ajustando o nível de resposta e recursos segundo o risco temporal identificado.

O estudo mostrou-se eficaz para incorporar as incertezas temporais e o tratamento da censura, características frequentes em bases reais de incidentes cibernéticos. Recomenda-se a continuidade da pesquisa com a inclusão de modelos de riscos proporcionais de Cox, técnica amplamente utilizada em análise de sobrevivência para investigar como diferentes fatores influenciam o risco de ocorrência de um novo incidente ao longo do tempo. Esses modelos permitem estimar o efeito de variáveis explicativas — como tipos específicos de ataque, atributos organizacionais ou características setoriais — sobre a probabilidade instantânea de uma nova violação, sem impor uma forma paramétrica ao tempo até o evento (Papathanasiou, Demertzis e Tziritas 2023).

Além disso, o monitoramento contínuo, a atualização periódica dos dados e a integração dessas análises

na governança corporativa são cruciais para antecipar violações futuras. Tal abordagem contribui para aprimorar políticas de segurança cibernética e alocação eficiente de recursos em ambientes complexos e dinâmicos.

5 PRIVACIDADE DIFERENCIAL APLICADA A DADOS DE INCIDENTES

Neste capítulo, investiga-se o impacto introduzido pela privacidade diferencial (*DP*) em dados tabulares associados a incidentes de segurança da informação. O estudo baseia-se no artigo (Rodrigues et al. 2025), previamente publicado no SBSeg 2025, do qual o autor desta dissertação é coautor, sendo aqui incorporado como componente de uma investigação mais abrangente. A ênfase recai sobre a relação entre o orçamento de privacidade (ε), a sensibilidade da função (Δf) e a preservação da utilidade analítica (Ponte et al. 2024). A problemática central consiste em compreender de que forma a aplicação de ruído controlado, inerente aos mecanismos de *DP*, altera a estrutura estatística dos dados e, consequentemente, influencia a validade de modelos preditivos ou inferenciais derivados desses conjuntos (Henderson e al. 2023).

Tal análise busca sustentar a transparência e a colaboração interorganizacional — por meio de divulgação de estatísticas agregadas ou geração de dados sintéticos — em conformidade com os princípios da Lei Geral de Proteção de Dados e com abordagens consagradas na literatura sobre proteção de dados sensíveis (Ponte et al. 2024).

O objetivo é investigar como mecanismos e parâmetros de *DP* influenciam a utilidade de análises e o risco de reidentificação em dados de incidentes (Ponte et al. 2024). As questões de pesquisa incluem: (i) quais impactos sobre métricas de utilidade são observados sob diferentes configurações de ε e Δf; (ii) em que medida atributos com maior entropia apresentam maior degradação quando submetidos a ruído (Dwork et al. 2025); e (iii) como estabelecer diretrizes práticas para divulgação responsável de estatísticas observando os limites impostos pelo orçamento de privacidade (Kifer, Messing e Roth 2020).

5.1 PROCEDIMENTOS METODOLÓGICOS APLICADOS AO ESTUDO DE PRIVACIDADE DIFERENCIAL

Parte-se de um conjunto de dados gerado sinteticamente que simula vazamentos em dezesseis cenários hipotéticos de divulgação adversarial (Sharma e Bantan 2025), abrangendo setores sensíveis como o bancário, comércio eletrônico, saúde, reservas de viagens, redes sociais e entretenimento digital (Wang et al. 2024). Cada domínio modela variáveis demográficas e comportamentais de modo estatisticamente coerente, assegurando utilidade analítica sem expor indivíduos (Alaa et al. 2022).

O delineamento empírico considera um cenário adversarial no domínio de reservas de hotel (Alhamad e Singh 2021), no qual se aplica o mecanismo de Laplace para avaliar, de forma sistemática, o impacto combinado dos parâmetros de privacidade ε (orçamento de privacidade) e Δf (sensibilidade da função) em 6.120 configurações experimentais. A análise quantitativa utiliza como métricas a *Jensen–Shannon Distance (JSD)*, para mensurar a divergência entre distribuições originais e perturbadas, e o erro percentual absoluto médio MAPE, para avaliar a degradação da acurácia analítica (Ponte et al. 2024).

A metodologia adotada nesta pesquisa é apresentada na Figura 5.1. As análises foram desenvolvidas em

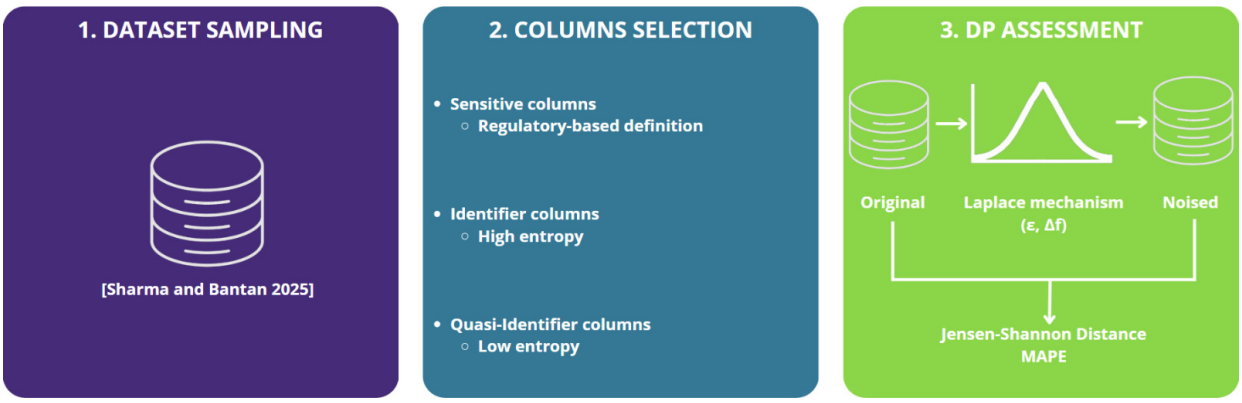

Figura 5.1: Fluxograma do Estudo de DP

ambiente *Python*, versão 3.11.11, assegurando a reprodutibilidade e a consistência nos procedimentos de processamento, modelagem e avaliação dos experimentos.

5.2 DADOS E VARIÁVEIS

Para o experimento, realiza-se uma seleção das colunas do conjunto de dados às quais será aplicada a técnica de privacidade diferencial. Para facilitar esse processo, as colunas do conjunto de dados são classificadas em três categorias distintas: atributos sensíveis, identificadores e quase-identificadores (QIs). Essa categorização permite direcionar a aplicação da privacidade diferencial de forma adequada, focando principalmente nos quase-identificadores, que são atributos que, isoladamente, não identificam um indivíduo, mas que em combinação podem representar uma ameaça à privacidade.

5.2.1 Dados Sensíveis, Quase-Identificadores e Identificadores

Com base em requisitos regulatórios — em particular o *California Consumer Privacy Act (CCPA)* e o *California Privacy Rights Act (CPRA)*, que definem como sensível qualquer informação capaz de expor credenciais ou dados financeiros, e o *Payment Card Industry Data Security Standard (PCI-DSS)*, que estabelece regras específicas para o armazenamento seguro de dados de cartões de pagamento — as seguintes colunas são classificadas como atributos sensíveis: nome de usuário, endereço de e-mail, senha, número do cartão de crédito, CVV e data de validade do cartão. Como a proteção desses dados exige confidencialidade absoluta, essas colunas são excluídas da avaliação de privacidade diferencial. Os últimos quatro dígitos do cartão, entretanto, não são considerados sensíveis, pois o *PCI-DSS* permite seu armazenamento desde que estejam devidamente mascarados.

Além dos atributos sensíveis, as colunas são classificadas como identificadores ou quase-identificadores. Identificadores são atributos que identificam unicamente um indivíduo sem precisar de informações adicionais — por exemplo, nome completo, endereço físico ou número de telefone. Por outro lado, quase-identificadores (QIs) são atributos que não identificam unicamente um indivíduo isoladamente, mas podem fazê-lo quando combinados com outros QIs, como cidade, código postal, informações do dispositivo ou

métodos de pagamento.

Para distinguir atributos identificadores de quase-identificadores, calcula-se a entropia de Shannon de cada coluna como medida da incerteza associada à distribuição de seus valores. Colunas que apresentam entropia elevada, refletindo alto grau de unicidade entre os registros, são classificadas como identificadores e excluídas da avaliação de privacidade diferencial, uma vez que a adição de ruído nesses casos tenderia a comprometer a unicidade intrínseca dos dados e a reduzir de forma significativa sua utilidade analítica.

A entropia $H(X)$ é estimada a partir da distribuição empírica de cada coluna. Para atributos cujos valores são essencialmente únicos, assume-se uma distribuição aproximadamente uniforme sobre as 185.000 observações amostradas, o que resulta em uma entropia teórica de $\ln(185000) \approx 12{,}13$ nats. Com base nesse critério, colunas com entropia superior a 12 nats são classificadas como identificadoras, enquanto aquelas com entropia inferior, desde que não caracterizadas como sensíveis por definição regulatória, são tratadas como quase-identificadores.

Assim, atributos como nome, endereço físico, número de telefone e histórico de pagamentos são excluídos da aplicação de mecanismos de privacidade diferencial. Em contrapartida, os mecanismos de privacidade diferencial — em particular o mecanismo de Laplace — são aplicados exclusivamente aos quase-identificadores, cuja menor entropia está diretamente associada tanto ao risco potencial de reidentificação quanto ao impacto da injeção de ruído sobre as distribuições de contagem. Essa escolha permite mitigar riscos indiretos de identificação ao mesmo tempo em que preserva, na medida do possível, a utilidade analítica dos dados.

A categorização sistemática das colunas em atributos sensíveis, identificadores e quase-identificadores fundamenta uma estratégia direcionada de preservação da privacidade, na qual a confidencialidade é assegurada para os dados mais críticos, enquanto a privacidade estatística é aplicada aos atributos que representam ameaças indiretas à identificação. Essa abordagem contribui para o atendimento aos requisitos regulatórios e para o equilíbrio entre proteção da privacidade e utilidade analítica.

5.3 SELEÇÃO DO CENÁRIO DE DIVULGAÇÃO

Dentre os 16 cenários de vazamento simulados apresentados por (Sharma e Bantan 2025), selecionamos o cenário de vazamento associado a reservas de hotel, por ter apresentado o maior número de quase-identificadores (QIs). Essa escolha possibilita uma avaliação mais abrangente da aplicação da privacidade diferencial, visto que a presença de múltiplos QIs aumenta a complexidade do desafio de proteção dos dados (Carvalho et al. 2023). Essa seleção é sustentada por evidências da literatura que mostram que a indústria hoteleira, inserida no próprio setor de turismo, apresenta recorrentes fragilidades de segurança e elevada exposição a ameaças cibernéticas, o que torna esse contexto particularmente relevante para a avaliação de estratégias de proteção de dados (Li et al. 2010).

Ao optar por este cenário, é possível investigar com mais profundidade os impactos da privacidade diferencial em um conjunto de dados realista, rico em atributos que podem potencialmente levar à reidentificação de indivíduos quando combinados. Dessa forma, a análise realizada contribui tanto para o aprimoramento das técnicas de proteção quanto para a compreensão dos *trade-offs* entre privacidade e utilidade dos dados

na prática (Ponte et al. 2024).

O cenário de divulgação adversária selecionado consiste em aproximadamente 2,9 milhões de linhas (pré-amostragem), contendo as colunas classificadas como quase-identificadores apresentadas na tabela 5.1. Essa tabela apresenta as principais colunas consideradas na avaliação de privacidade diferencial, acompanhadas de descrições sucintas de cada atributo. As colunas classificadas como quase-identificadores (QI) são aquelas que possibilitam a identificação indireta dos indivíduos presentes no conjunto de dados, isto é, informações que, embora não revelem a identidade por si sós, apresentam potencial de reidentificação quando combinadas com outras variáveis.

Tabela 5.1: Colunas categorizadas como QI e consideradas na avaliação de DP

Coluna	**Descrição**
Informação de dispositivo	Sistema operacional do telefone (Android ou iOS) e sua versão, considerando aqueles lançados entre 2016 e 2022.
Hábitos de Viagem	Pode ser qualquer combinação de viagem de carro, viagem de trem, cruzeiros, voos domésticos ou internacionais.
Métodos de Pagamento	Pode ser qualquer combinação de dinheiro, carteira digital, pagamento móvel, transferência bancária, cartão de débito ou cartão de crédito.
Cidade	Cidade do endereço do titular (Estados Únidos da America - EUA).
Últimos 4 dígitos do cartão	Últimos quatro dígitos do número do cartão de crédito.
Código postal	Código postal (ZIP code) do endereço do titular (Estados Unidos da America - EUA).

A tabela 5.2 apresenta o resumo estatístico dos quase-identificadores, indicando para cada coluna o número total de entradas, a quantidade de valores distintos, o valor mais frequente e sua respectiva ocorrência. Essa organização facilita a compreensão da diversidade e da distribuição desses atributos, fornecendo subsídios para avaliar o risco de divulgação adversarial.

Tabela 5.2: Estatística resumida das tabelas de QI

Coluna	**Contar valores únicos**	**Mais frequentes**	**Frequência**
Informações do dispositivo	185.000 16	Android, Android 12	11.773
Hábitos de viagem	185.000 155	Viagem de trem	12.646
Métodos de pagamento	185.000 3.905	Cartão de crédito	7.590
Cidade	185.000 17.610	Washington	1.444
Últimos 4 dígitos do cartão	185.000 10.000	**** **** **** 5519 51342	36
Códigos Postais	185.000 39.359	–	16

5.4 AVALIAÇÃO DE PRIVACIDADE DIFERENCIAL

Para a avaliação da privacidade diferencial, utilizou-se a biblioteca *diffprivlib*, que oferece mecanismos consolidados para *DP*, além de permitir a configuração dos principais parâmetros de privacidade. A biblioteca possui interface simples, boa documentação e alertas integrados que ajudam a evitar violações acidentais de privacidade.

Neste trabalho, optou-se pelo uso do mecanismo de *Laplace*, amplamente empregado em análises de

dados com preservação de privacidade devido à sua simplicidade e eficiência. O mecanismo de *Laplace* atua adicionando ruído calibrado às consultas sobre os dados, controlado pelos parâmetros de sensibilidade e orçamento de privacidade (Anon. 2021), permitindo equilibrar o grau de proteção com a utilidade dos dados (Ponte et al. 2024). Tal abordagem possibilita que as análises e publicações estatísticas sejam realizadas minimizando os riscos de reidentificação dos indivíduos, sem comprometer excessivamente a acurácia das respostas (Seeman e Susser 2024).

5.5 MECANISMO DE LAPLACE PARA PRIVACIDADE DIFERENCIAL

O mecanismo de *Laplace (ML)*, aplicado à privacidade diferencial, garante a propriedade de ε-privacidade diferencial ao adicionar ruído calibrado ao resultado das consultas realizadas sobre os dados. Esse ruído é gerado a partir de uma distribuição de *Laplace*, cujo parâmetro de escala é inversamente proporcional ao orçamento de privacidade ε. Em outras palavras, quanto menor o valor de ε, maior será o ruído adicionado, resultando em uma proteção de privacidade mais forte, porém com maior impacto sobre a precisão dos resultados (Ponomareva et al. 2024). O mecanismo pode ser expresso matematicamente como:

$$M_L(f(x)) = f(x) + \text{Laplace}\left(\frac{\Delta f}{\varepsilon}\right) \tag{5.1}$$

onde $f(x)$ é a consulta de interesse (Ponomareva et al. 2024), Δf corresponde à sensibilidade global da função f, e Laplace$(\cdot)$ representa a distribuição de *Laplace* centrada em zero e ajustada pela escala apropriada (Ponomareva et al. 2024).

Dada uma consulta original $f(x)$, com sensibilidade L_1 igual a Δf, o mecanismo de *Laplace* gera como saída $M_L(x)$, conforme a expressão:

$$M_L(x) = f(x) + \text{Lap}(0, b) \tag{5.2}$$

onde $\text{Lap}(0, b)$ representa ruído amostrado de uma distribuição de *Laplace* com média zero e escala $b = \frac{\Delta f}{\varepsilon}$. O parâmetro de escala b é determinado pela sensibilidade Δf da função e pelo orçamento de privacidade ε.

A magnitude do ruído introduzido depende diretamente da sensibilidade da consulta: quanto maior Δf, maior será o ruído somado ao resultado, garantindo maior privacidade, porém diminuindo a precisão das respostas. Valores menores de Δf resultam em menor ruído, favorecendo a precisão mas reduzindo a proteção de privacidade.

O orçamento de privacidade ε atua como regulador deste equilíbrio: valores menores de ε aumentam o nível de privacidade ao adicionar mais ruído, enquanto valores maiores favorecem precisão, mas com menor proteção. Assim, tanto a sensibilidade Δf quanto ε são fundamentais para controlar o *trade-off* entre utilidade e privacidade na aplicação do mecanismo de *Laplace*.

Neste estudo, são analisados os impactos das variações na sensibilidade (Δf) e no orçamento de privacidade (ε) sobre os deslocamentos distribucionais dos resultados de consultas de contagem $f(x)$. Foram avaliadas alterações na perda de privacidade mediante a variação de ε no intervalo $\{0.1, 0.2, 0.3, \ldots, 12.0\}$

e de Δf em $\{0.1, 1.1, 2.1, \ldots, 50.1\}$. Para garantir a validade dos resultados, todas as saídas negativas, que não possuem significado na consulta de contagem, são truncadas para o valor mínimo de zero.

5.6 MÉTRICAS DE COMPARAÇÃO

Para comparar as distribuições de contagem de cada coluna de quase-identificadores (QIs) antes e após a aplicação de privacidade diferencial, utiliza-se a métrica *Jensen-Shannon Distance (JSD)*, definida como a raiz quadrada da divergência *Jensen-Shannon*.

A divergência *Jensen-Shannon* é derivada da divergência *Kullback-Leibler (KLD)*, definida conforme a equação (5.3), que também é denominada entropia relativa. A *KLD* quantifica a discrepância esperada na surpresa ao assumir a distribuição Q em vez da distribuição real P.

$$D_{KL}(P \parallel Q) = \sum_{i=1}^{n} P(x_i) \log \frac{P(x_i)}{Q(x_i)} \tag{5.3}$$

Trata-se de uma medida assimétrica e não limitada superiormente, tal que

$$D_{KL}(P \parallel Q) \in [0, +\infty) \tag{5.4}$$

sendo zero se e somente se

$$P = Q. \tag{5.5}$$

A presença de zeros na distribuição Q faz com que

$$D_{KL}(P \parallel Q) \tag{5.6}$$

tenda ao infinito devido à fração

$$\frac{P(x_i)}{Q(x_i)}. \tag{5.7}$$

A *Jensen-Shannon Distance*, conforme expressa na equação (5.8), quantifica a discrepância entre duas distribuições P e Q calculando a raiz quadrada da média das divergências *Kullback-Leibler* entre cada distribuição e sua distribuição média $M = \frac{P+Q}{2}$.

$$JSD(P \parallel Q) = \sqrt{\frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(Q \parallel M)} \tag{5.8}$$

Ao contrário da *KLD*, a *JSD* é uma medida simétrica e limitada, tal que

$$JSD(P \parallel Q) \in [0, \sqrt{\ln 2}] \tag{5.9}$$

assumindo valor zero somente quando

$$P = Q \tag{5.10}$$

e valor máximo

$$\sqrt{\ln 2} \tag{5.11}$$

quando P e Q são disjuntas. Por depender da distribuição média M, a *JSD* evita divergências infinitas da *KLD*, não sendo necessário adicionar constantes para lidar com zeros.

5.7 RESULTADOS

Esta seção apresenta os resultados obtidos nos experimentos, com ênfase na quantificação dos deslocamentos distribucionais introduzidos pela aplicação da privacidade diferencial. Como base para a análise e discussão subsequentes, são calculados os valores de entropia para cada coluna de quase-identificadores no conjunto de dados. Esses valores de entropia fornecem informações acerca da variabilidade de cada atributo QI, os quais influenciam o impacto da adição de ruído sob distintas configurações da privacidade diferencial.

5.7.1 Entropia das Colunas quase-identificadoras

A entropia para cada coluna de quase-identificadores selecionada foi calculada e está apresentada na tabela 5.3. Correlacionando essas informações com as estatísticas resumidas da tabela 5.2, observa-se que colunas com menor entropia tendem a possuir menos valores únicos. Essa relação decorre da definição de entropia como uma medida de imprevisibilidade. Quando o número de valores únicos em uma coluna é baixo, a diversidade de resultados possíveis é reduzida, tornando os dados mais previsíveis e, consequentemente, diminuindo sua entropia.

Tabela 5.3: Entropia dos atributos (em Nats)

Coluna	**Entropia (Nats)**
Informação de Dispositivo	2.78
Hábitos de Viagem	4.32
Método de Pagamento	6.42
Cidade	8.93
4 últimos dígitos Cartão	9.18
Código Postal	10.48

Colunas com menor entropia, como as relacionadas a informação de dispositivo, apresentam risco reduzido de reidentificação, devido ao fato de muitos registros compartilharem os mesmos valores. Essa uniformidade diminui a probabilidade de que um indivíduo seja identificado de forma única, possivelmente exigindo uma proteção de privacidade menos rigorosa. Contudo, colunas de baixa entropia também impõem desafios à aplicação da privacidade diferencial. Devido à limitada variabilidade nessas colunas, o ruído

adicionado pode não ser suficiente para ocultar efetivamente os valores excepcionais. Consequentemente, valores raros ou únicos podem permanecer evidentes mesmo após a adição do ruído, reduzindo o nível geral de proteção da privacidade nesses casos.

Por outro lado, colunas de alta entropia, como código postal, apresentam maior potencial de reidentificação devido à sua maior unicidade entre os registros. Contudo, dado o elevado grau de variabilidade nessas colunas, mesmo uma pequena quantidade de ruído pode ser eficaz para provocar um deslocamento distribucional.

A figura 5.2 apresenta as funções de distribuição acumulada empírica (ECDFs) dos atributos quase-identificadores, utilizadas para avaliar sua representatividade e capacidade de generalização em relação a distribuições típicas de bases reais. As ECDFs são construídas a partir do rank das categorias (eixo X, em escala logarítmica), ordenadas da mais frequente para a menos frequente, enquanto o eixo Y expressa a proporção acumulada de registros à medida que novas categorias são incorporadas. Curvas que apresentam crescimento rápido — como *Device Information* e *Travel Habits* — evidenciam forte concentração, indicando que poucas categorias dominam a maior parte dos registros e caracterizando atributos de baixa entropia, nos quais a incerteza sobre o valor observado é reduzida. Em contraste, atributos como *City, Zip Codes* e *Card Last4Digits* exibem crescimento gradual ao longo do eixo X, revelando alta dispersão e necessidade de milhares de categorias para abranger o conjunto completo de registros. Esses padrões refletem atributos de alta entropia, que carregam maior diversidade informacional e, consequentemente, maior potencial de quase-identificação.

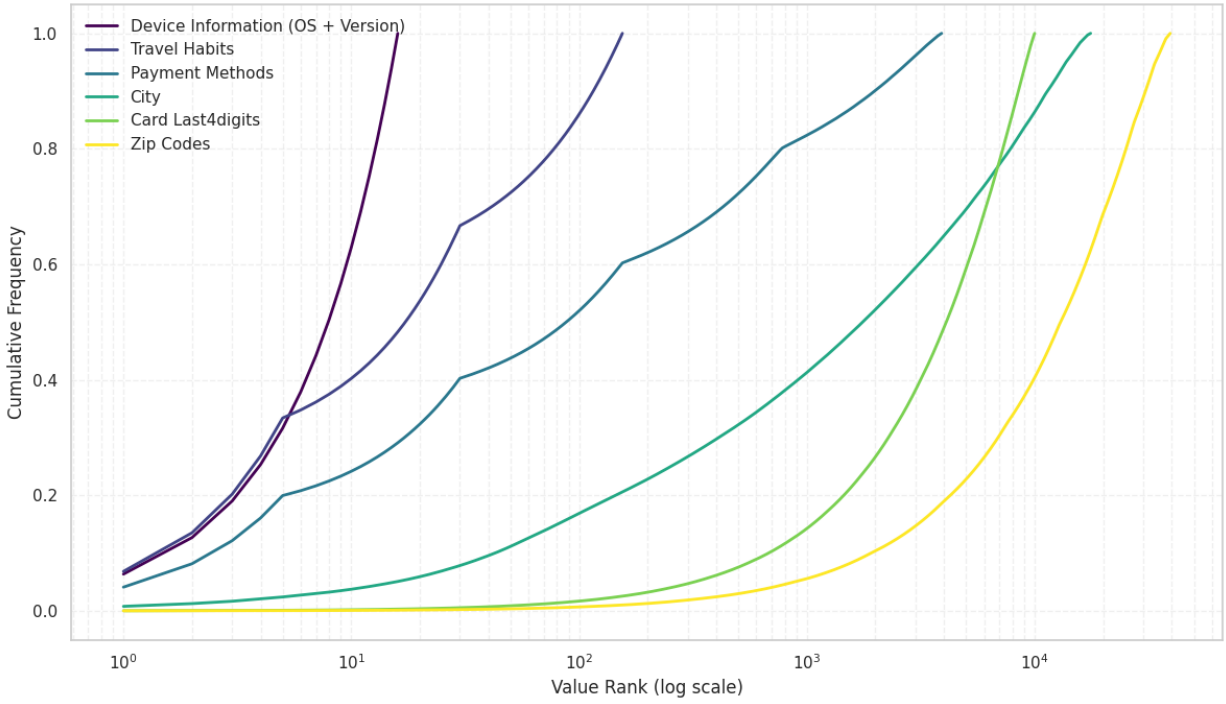

Figura 5.2: Distribuição de Classificação de Valores das Colunas QI

Dessa forma, a análise comparativa das curvas sintetiza o grau de diversidade ou concentração presente em cada atributo QI, oferecendo uma base empírica sólida para estimar riscos de reidentificação e orientar a seleção de mecanismos de proteção apropriados sob o arcabouço da privacidade diferencial.

5.7.2 Avaliação de mudança distributivas

Um deslocamento distribucional refere-se a uma mudança nas propriedades estatísticas dos dados de entrada. Em contextos como aprendizado de máquina, onde a privacidade diferencial é frequentemente empregada para proteger os sujeitos dos dados, tais deslocamentos podem acarretar consequências significativas. Especificamente, quando ocorre um descompasso entre a distribuição dos dados utilizada durante o treinamento do modelo e aquela encontrada na aplicação real, o desempenho do modelo pode se degradar. Isso acontece porque o modelo é otimizado com base na distribuição de treinamento e pode não generalizar adequadamente para dados com características alteradas, o que pode levar à redução da acurácia e confiabilidade durante a inferência.

Os deslocamentos distribucionais em função do orçamento de privacidade ε e da sensibilidade Δf, medidos pela distância *Jensen-Shannon (JSD)*, são apresentados na figura 5.3. Conforme esperado, observa-se que deslocamentos maiores ocorrem quando se adiciona mais ruído aos dados, o que ocorre para valores menores de ε e maiores de Δf. Essas condições resultam em maior perturbação e, consequentemente, aumento da divergência em relação à distribuição original.

Esse comportamento evidencia o *trade-off* fundamental da privacidade diferencial: a adição de mais ruído fortalece as garantias de privacidade, mas eleva o deslocamento distribucional, refletido nos maiores valores de *JSD*. A seleção dos valores adequados para ε e Δf requer um equilíbrio cuidadoso entre esses objetivos conflitantes. Em cenários de maior risco, valores menores de ε são preferíveis para assegurar maior proteção à privacidade, enquanto valores maiores podem ser tolerados para atributos menos sensíveis, preservando maior precisão. Assim, a escolha dos parâmetros deve ser orientada pelos requisitos de privacidade e pelos objetivos analíticos da aplicação específica.

Os padrões de variação observados na figura 5.3 podem ser diretamente relacionados aos níveis de entropia de cada coluna. Colunas com maior entropia, como *payment Methods, City, Card Last4digits* e *Zip Codes* (figuras 5.3c, 5.3f), apresentam valores mais elevados da distância *Jensen-Shannon*, mesmo quando o ruído adicionado é mínimo (isto é, para valores maiores de ε e menores de Δf). Esse comportamento reflete a maior variabilidade dessas colunas, tornando suas distribuições mais sensíveis mesmo a pequenas perturbações introduzidas pelos mecanismos de privacidade diferencial.

Como esses atributos possuem um número maior de valores distintos e distribuições menos uniformes, apresentam maior sensibilidade à injeção de ruído. Consequentemente, as perturbações decorrentes dos mecanismos de privacidade diferencial acarretam deslocamentos mais acentuados em suas distribuições de probabilidade.

Em contraste, colunas como *device information* e *travel habits* (figuras 5.3a, 5.3b) apresentam consistentemente baixos valores de divergência. Essas características exibem menor entropia, conforme demonstrado na tabela 5.3, devido ao menor número de valores únicos, evidenciado na tabela 5.3. Isso resulta em impacto mínimo da injeção de ruído, conduzindo a deslocamentos distribucionais reduzidos e, consequentemente, menor divergência, refletido nas superfícies quase planas dos gráficos correspondentes.

Essas colunas, portanto, requerem ruído significativamente maior (isto é, ε mais baixo e/ou Δf mais alto) para alcançar o mesmo nível de proteção de privacidade que colunas de maior entropia. Observa-se que a magnitude da divergência correlaciona-se com a entropia do atributo, sugerindo que atributos com

maior diversidade são mais suscetíveis à distorção quando mecanismos de preservação de privacidade são aplicados.

5.7.3 Avaliação da precisão

Para avaliar o impacto do ruído sobre a utilidade dos dados, a acurácia foi medida por meio do Erro Percentual Absoluto Médio (*MAPE*) entre os valores originais e os valores com ruído. Como esperado, orçamentos de privacidade menores e sensibilidades maiores resultam consistentemente em valores elevados de *MAPE*, refletindo o *trade-off* entre a preservação da privacidade e a fidelidade dos dados para análise.

A degradação da acurácia segue um padrão previsível, possibilitando a seleção dos parâmetros ε e Δf que equilibram os requisitos de privacidade e o nível aceitável de precisão dos dados para cada caso específico. A figura 5.4 apresenta o *MAPE* entre os dados originais e os dados protegidos por privacidade diferencial para cada quase-identificador, considerando $\varepsilon \leq 2$ para melhor visualização. Em conjunto com a tabela 5.3, que reporta os valores de entropia, observa-se uma relação entre entropia e erro sob privacidade diferencial.

Atributos de baixa entropia, como *device Information* (2,78 nats), apresentam consistentemente menores valores de *MAPE* ao longo dos diferentes níveis de privacidade (figura 5.3a). Colunas com entropia média, como *Payment Methods* (6,42 nats), exibem *MAPE* moderado (figura 5.3c), enquanto colunas de alta entropia, como *Zip Codes* (10,48 nats), apresentam maiores valores de *MAPE*, especialmente para valores baixos de ε (figura 5.3f). Esses atributos são mais suscetíveis a degradação significativa da acurácia, requerendo menor injeção de ruído para induzir erros substanciais. Observa-se, portanto, que maior entropia e sensibilidade resultam em distorção mais acentuada para o mesmo orçamento de privacidade.

Nas configurações testadas, observa-se que valores de ε na faixa de [1,5, 4,0] geralmente resultam em injeção moderada de ruído, equilibrando a privacidade com níveis razoáveis de degradação da acurácia (*MAPE* < 15%) para a maioria dos quase-identificadores. Da mesma forma, valores de sensibilidade na faixa [5,0, 20,0] preservaram a acurácia sem comprometer excessivamente a privacidade para atributos de entropia média a alta.

Embora essas faixas não sejam definitivas, elas representam pontos iniciais para implantação em contextos reais com estruturas de dados similares. Importa ressaltar que atributos de maior entropia indicam a necessidade de ajuste dos parâmetros de privacidade diferencial conforme as características específicas de cada atributo, ao invés da aplicação de valores uniformes.

5.7.4 Considerações finais

A aplicação de *DP* no contexto de incidentes possibilita transparência responsável com risco mensurável, desde que mecanismos e parâmetros sejam calibrados ao uso pretendido. A sensibilidade dos atributos e a presença de contagens baixas demandam cautela adicional na interpretação dos resultados, enquanto aspectos de implementação — como potenciais canais laterais decorrentes de operações temporais ou de ponto flutuante — requerem práticas seguras de execução e publicação (Jin et al. 2021). Do ponto de vista de governança, o uso de *DP* alinha-se aos princípios da LGPD — minimização, finalidade, segurança e

prestação de contas — e favorece análises reprodutíveis com salvaguardas formais (Brasil 2018).

As principais ameaças identificadas incluem: (i) viés de reporte e cobertura nos dados de incidentes; (ii) dependência do desenho das consultas e da hipótese de sensibilidade (Δf); (iii) efeitos de deslocamento distribucional; e (iv) riscos de implementação capazes de degradar garantias formais quando não mitigados. Estratégias de mitigação incluem reporte transparente de parâmetros, rastreabilidade das transformações aplicadas e testes de robustez (Jin et al. 2021).

Em síntese, o estudo observa os princípios da LGPD e adota transparência responsável na divulgação de resultados, privilegiando agregações e proteção por *DP* quando pertinente. As evidências obtidas oferecem diretrizes para a publicação de estatísticas fortalecendo a *accountability* e apoiando práticas seguras de colaboração interinstitucional.

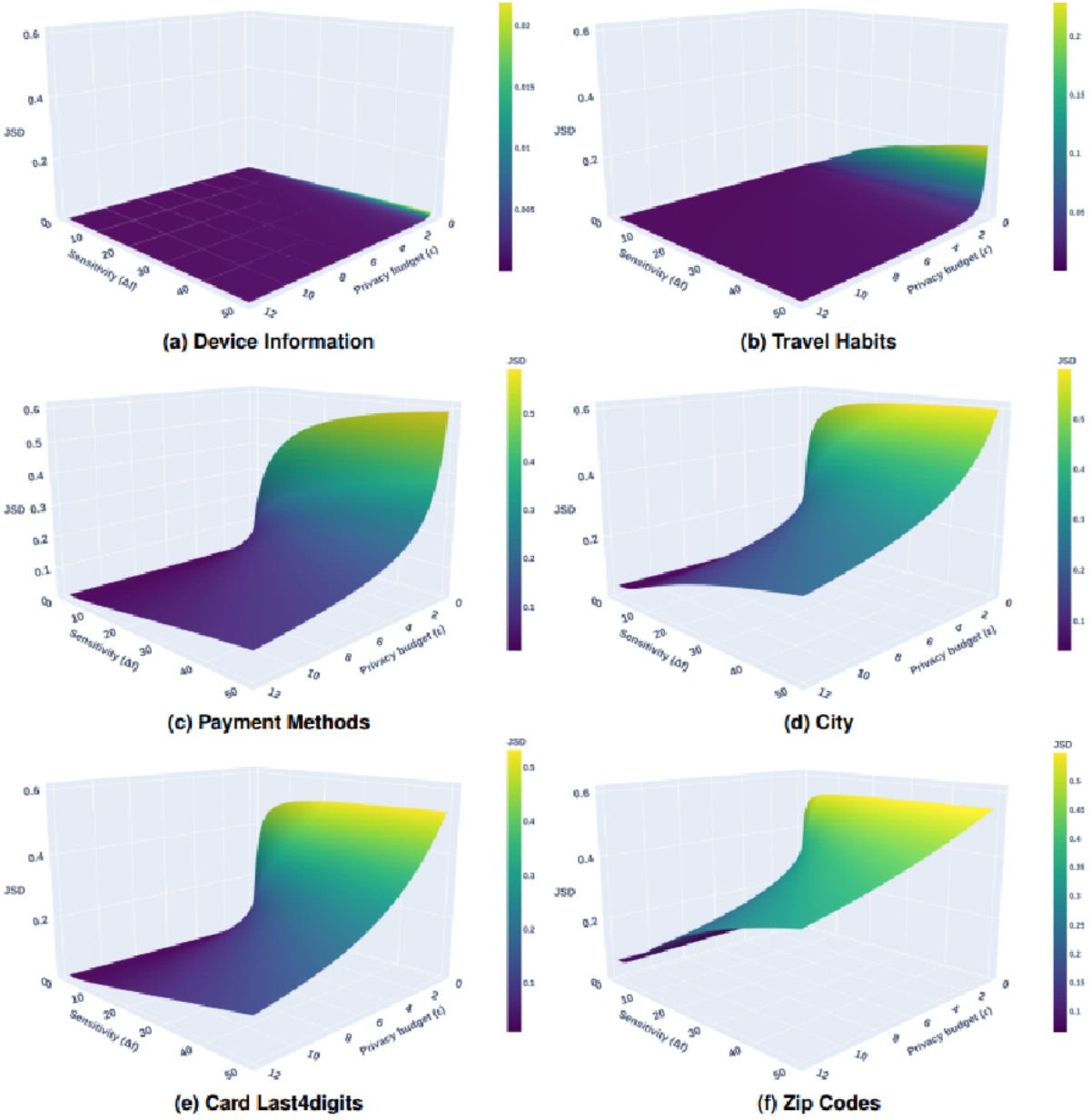

Figura 5.3: Distância Jansen-Shanon entre o sinal original e o sinal com ruído

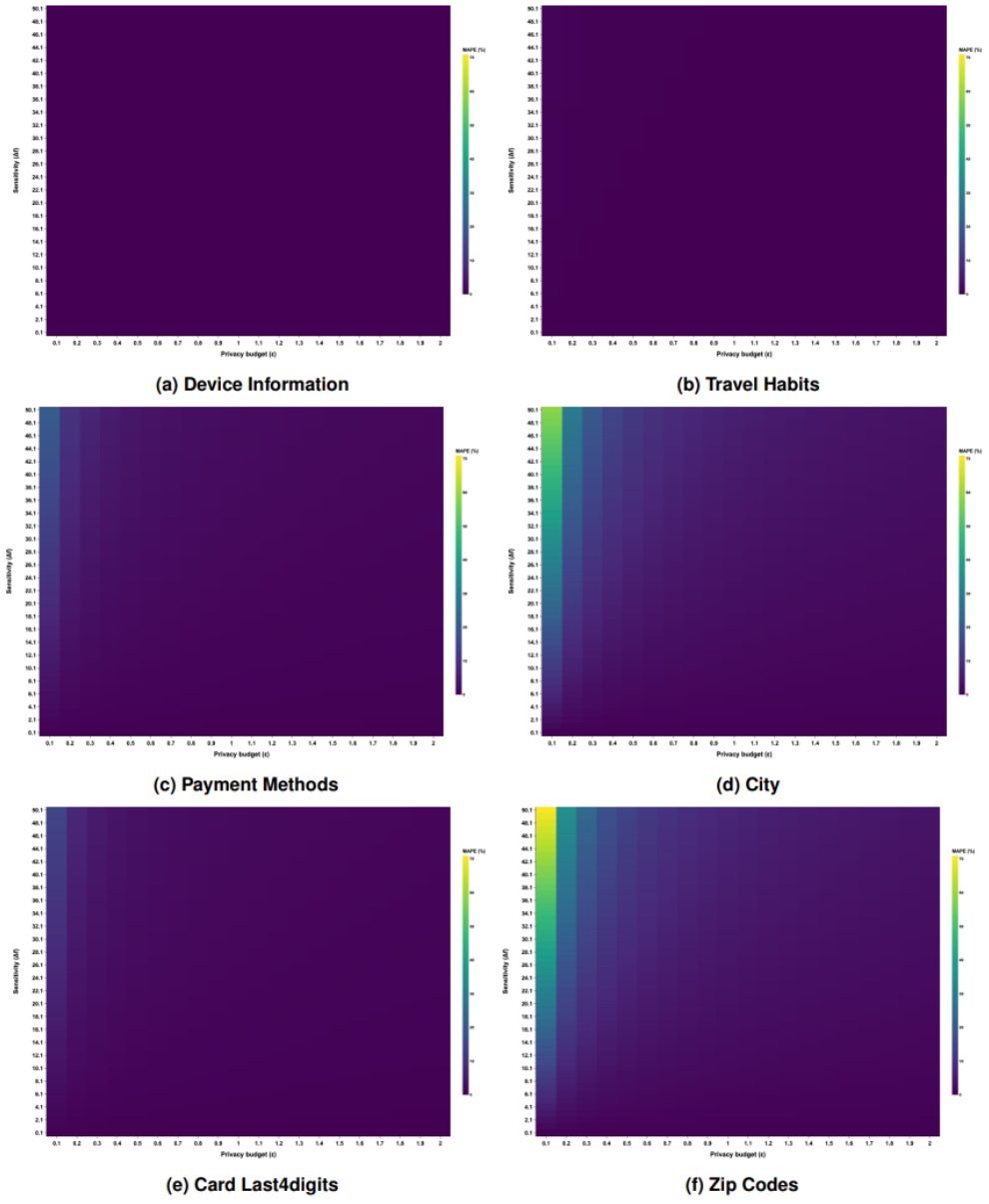

Figura 5.4: MAPE entre as distribuições com ruído e original

6 CONCLUSÃO

Este capítulo apresenta a síntese integrada dos achados obtidos nos três eixos analíticos da pesquisa, bem como suas principais implicações práticas e acadêmicas, limitações e possibilidades de continuidade. A dissertação é estruturada como um estudo empírico único, desenvolvido em três frentes articuladas: (i) modelagem preditiva da incidência de violações de dados; (ii) análise de recorrência por métodos de sobrevivência; e (iii) aplicação de privacidade diferencial (*DP*) em dados de incidentes. Em conjunto, esses eixos formam um arcabouço coerente de governança baseada em evidências, alinhado às exigências regulatórias contemporâneas, especialmente às diretrizes da LGPD (Brasil 2018).

6.1 SÍNTESE DOS ACHADOS

O primeiro eixo analítico concentrou-se na modelagem preditiva da incidência mensal de violações de dados por setor, a partir de séries temporais derivadas da base *Data Breach Chronology* da PRC. Foram comparadas famílias de modelos estatísticos, métodos baseados em árvores de decisão e redes neurais profundas, em um delineamento experimental unificado, com protocolo padronizado de preparo dos dados, particionamento temporal, calibração e avaliação. Os resultados indicaram desempenho relativo superior das redes *LSTM* e *TCN* no agregado, com *MAPE* médio mais baixo e maior estabilidade entre setores, sobretudo na presença de padrões não lineares, dependências de longo prazo e heterogeneidades setoriais. Modelos estatísticos (*SARIMA*, *fbprophet*) mantiveram desempenho competitivo em contextos com sazonalidade mais clara e menor volatilidade, enquanto o *xgboost* se destacou por reduzir erros absolutos em determinadas configurações, embora com maior variação relativa em cenários de baixa contagem e alta variabilidade.

O segundo eixo deslocou o foco do volume de incidentes para a dimensão temporal da recorrência, por meio da análise de sobrevivência com o estimador de Kaplan–Meier. Utilizando o mesmo recorte temporal (2010–2023) e a mesma base da PRC, estimou-se o tempo-até-novo-incidente por setor, com tratamento explícito da censura à direita e estratificação setorial. As curvas de sobrevivência obtidas evidenciaram heterogeneidade significativa na probabilidade de permanecer sem novo incidente ao longo do tempo. Setores como Saúde (MED) e Governo (GOV) apresentaram declínios mais rápidos nas funções de sobrevivência, sugerindo janelas críticas de reincidência mais curtas e, portanto, maior urgência na renovação de controles, testes e auditorias. Outros setores, como organizações sem fins lucrativos, exibiram curvas mais suaves, indicando intervalos médios mais longos entre incidentes. Essas evidências complementam o eixo preditivo ao introduzir explicitamente a dimensão do "quando", permitindo associar níveis de risco temporal a perfis setoriais e a padrões empíricos de recorrência.

O terceiro eixo investigou a aplicação de privacidade diferencial a dados tabulares associados a incidentes, com foco em um cenário sintético de reservas de hotel, selecionado dentre dezesseis cenários adversariais por sua maior concentração de quase-identificadores (Sharma e Bantan 2025). Com base no mecanismo de *Laplace* e na variação sistemática do orçamento de privacidade (ε) e da sensibilidade da função (Δf), foram executadas 6.120 configurações experimentais, avaliando-se deslocamentos distributivos por meio da

Jensen–Shannon Distance (JSD) e degradação de acurácia por meio do *MAPE*. Os resultados mostraram que atributos com maior entropia (por exemplo, códigos postais, cidades e combinações de hábitos e métodos de pagamento) são mais sensíveis à injeção de ruído, apresentando maiores *JSD* e *MAPE* mesmo em cenários com ε moderado. Em contraste, atributos de baixa entropia, com menor diversidade de valores, sofrem deslocamentos mais discretos e demandam ruído mais intenso para alcançar níveis comparáveis de proteção.

Tomados em conjunto, os três eixos configuram um ciclo analítico integrado: (i) a modelagem preditiva fornece estimativas quantitativas de "quanto" tende a ocorrer, em termos de volume de incidentes por setor; (ii) a análise de sobrevivência explicita "quando" novas ocorrências são mais prováveis, revelando janelas críticas de tempo-até-novo-incidente; e (iii) a privacidade diferencial oferece meios formais para divulgar estatísticas, preservando a utilidade analítica sob restrições de proteção de dados. Essa integração sustenta uma abordagem operacional em que previsão, recorrência e transparência responsável se reforçam mutuamente, fornecendo insumos robustos para a governança de violações de dados em ambientes regulados pela LGPD (Brasil 2018).

6.2 IMPLICAÇÕES PRÁTICAS

Do ponto de vista aplicado, os achados desta dissertação fornecem instrumentos concretos para apoiar o planejamento, a operação e a prestação de contas em segurança da informação e proteção de dados, especialmente em organizações públicas.

Em primeiro lugar, o eixo preditivo demonstra que modelos de séries temporais, em particular arquiteturas neurais profundas como *LSTM* e *TCN*, podem ser incorporados a fluxos de monitoramento contínuo para antecipar a incidência de violações por setor. Ao produzir previsões mensais com *MAPE* em faixas consideradas boas ou razoáveis na maioria dos setores, esses modelos oferecem base quantitativa para decisões como dimensionamento de equipes, priorização de auditorias, renegociação de contratos de serviços de segurança e definição de janelas de manutenção preventiva. A utilização combinada de *MAPE*, *MAE* e *RMSE* permite calibrar expectativas sobre precisão e volatilidade setorial, facilitando a tradução dos resultados em metas e indicadores operacionais.

Em segundo lugar, o eixo de análise de recorrência fornece um eixo temporal adicional para orientar a definição de *SLAs* e políticas de mitigação. As curvas de Kaplan–Meier por setor permitem estabelecer horizontes temporais nos quais a probabilidade de reincidência se torna crítica, sustentando, por exemplo, a intensificação de monitoramento em períodos subsequentes a incidentes em setores com alta taxa de recorrência. No contexto da administração pública, onde recursos são escassos e as demandas são concorrentes, essa diferenciação temporal por setor apoia a alocação seletiva de esforços, alinhando níveis de serviço a perfis de risco empiricamente observados, em consonância com práticas de gestão baseada em risco.

Em terceiro lugar, o eixo de privacidade diferencial demonstra que é possível conciliar transparência e proteção de dados ao divulgar estatísticas ou dados sintéticos sobre incidentes. Ao quantificar o efeito de diferentes combinações de ε e Δf sobre a *JSD* e o *MAPE*, a dissertação oferece subsídios para a escolha de parâmetros que preservem a utilidade de análises agregadas sem expor indevidamente indivíduos ou organizações. Esse resultado é particularmente relevante para órgãos públicos sujeitos à LGPD, que precisam

equilibrar obrigações de transparência, cooperação interinstitucional e prestação de contas com deveres de confidencialidade, minimização e segurança no tratamento de dados pessoais (Brasil 2018).

Assim, do ponto de vista prático, o trabalho aponta para a possibilidade de construir *pipelines* de governança em que previsões, perfis de recorrência e camadas de *DP* são integrados a sistemas de apoio à decisão, alimentando painéis gerenciais, relatórios regulatórios e fluxos operacionais de resposta a incidentes.

6.3 LIMITAÇÕES

Apesar da robustez metodológica e da reprodutibilidade dos resultados, algumas limitações devem ser explicitadas, delimitando o escopo de generalização da pesquisa.

Uma primeira limitação diz respeito à própria base empírica utilizada. A PRC concentra incidentes reportados em um contexto específico (principalmente Estados Unidos) e está sujeita a vieses de subnotificação, atrasos de comunicação e mudanças de política de divulgação ao longo do tempo. Isso impacta tanto a modelagem preditiva quanto a análise de sobrevivência, podendo produzir estimativas conservadoras ou distorcidas para determinados setores ou períodos. A extrapolação dos resultados para outras jurisdições, como o contexto brasileiro, requer, portanto, adaptações e validações adicionais.

Uma segunda limitação relaciona-se à não estacionariedade das séries temporais e às mudanças de regime. Incidentes de violação de dados são influenciados por fatores tecnológicos (novas vulnerabilidades, adoção de nuvens e *APIs*), regulatórios (novas normas, sanções) e socioeconômicos (crises, guerras, novas modalidades de fraude) que podem alterar abruptamente o padrão de incidência. Embora a dissertação tenha utilizado métricas como o expoente de hurst e procedimentos rigorosos de preparação, os modelos treinados em um regime podem perder desempenho quando o contexto se altera de forma significativa, exigindo re-treinos periódicos, monitoramento de desempenho e ajustes dinâmicos.

Uma terceira limitação refere-se aos recortes com baixa contagem de incidentes. Em alguns setores e intervalos, o número de eventos observados é reduzido, elevando a incerteza das estimativas de sobrevivência e a sensibilidade a *outliers*. Essa limitação é intrínseca à natureza dos dados e demanda cautela ao interpretar curvas de sobrevivência e estatísticas derivadas em segmentos com amostras pequenas.

Por fim, o eixo de privacidade diferencial depende fortemente do desenho das consultas (contagens, agregações), da classificação dos atributos (sensíveis, identificadores, quase-identificadores) e da hipótese adotada para a sensibilidade global Δf (Jin et al. 2021). Pequenas mudanças nesses elementos podem alterar o balanço utilidade–privacidade. Adicionalmente, riscos de implementação, tais como canais laterais de tempo ou imprecisões de ponto flutuante, podem comprometer garantias formais se não forem mitigados por práticas de engenharia seguras e auditorias independentes.

Essas limitações não invalidam os resultados, mas indicam que eles devem ser lidos como evidências condicionadas a um determinado contexto de dados, hipóteses e escolhas de modelagem, o que reforça a importância de replicação, extensão e validação cruzada em ambientes distintos.

6.4 TRABALHOS FUTUROS

Os eixos desenvolvidos nesta dissertação abrem um amplo conjunto de possibilidades para aprofundamento teórico e aplicação prática em contextos organizacionais, regulatórios e operacionais diversos. Uma das primeiras direções de continuidade envolve a ampliação e integração de bases de incidentes, incorporando informações provenientes de autoridades reguladoras, órgãos setoriais e repositórios nacionais. A harmonização de taxonomias de tipos de incidente, setores e níveis de severidade permitirá reduzir vieses de reporte, fortalecer a robustez das estimativas e viabilizar análises comparativas entre países e domínios distintos.

Outra frente relevante diz respeito à exploração de portfólios híbridos e modelos explicáveis. A combinação de arquiteturas neurais, como *LSTM* e *TCN*, com modelos estatísticos e ensembles baseados em árvores pode articular desempenho preditivo e interpretabilidade de maneira mais equilibrada. A incorporação de técnicas de explicabilidade, incluindo métodos baseados em *SHAP* e diferentes métricas de importância de variáveis, aproxima a modelagem das exigências de auditoria, transparência e governança típicas da administração pública.

O tratamento da não estacionariedade também representa uma oportunidade de expansão. O aprofundamento em técnicas de detecção de mudanças de regime *change point detection*, validação temporal mais estrita e estratégias de re-treino contínuo permitiria alinhar o ciclo de vida dos modelos a processos recorrentes de revisão de políticas de segurança e renegociação de contratos de serviços tecnológicos. Esse aprimoramento metodológico contribui para maior resiliência em ambientes em constante mutação.

A análise de sobrevivência pode igualmente ser estendida mediante o uso de modelos multivariados, como o modelo de riscos proporcionais de Cox e outras abordagens que incorporam covariáveis de forma explícita. Quando houver disponibilidade de atributos explicativos em nível organizacional e operacional, tais modelos permitirão quantificar efeitos de variáveis como tipo de ataque, vetor de intrusão, porte institucional e maturidade em segurança sobre o tempo até um novo incidente, ampliando o valor analítico das estimativas.

No eixo de privacidade diferencial, uma vertente promissora envolve a comparação entre mecanismos alternativos — como o Gaussiano e o *staircase* — e diferentes estratégias de composição do orçamento de privacidade. Avaliações específicas para cenários de alta entropia ou atributos com baixa contagem podem gerar recomendações operacionais mais refinadas para órgãos reguladores e equipes técnicas responsáveis pela publicação de estatísticas sensíveis.

A consolidação de boas práticas e ferramentas de implementação também merece destaque. A elaboração de guias técnicos, verificações independentes de código, *toolkits* e fluxos de auditoria voltados à aplicação segura de privacidade diferencial em ambientes produtivos poderá fortalecer a mitigação de canais laterais, aprimorar o reporte de parâmetros e padronizar a documentação de garantias (Jin et al. 2021). A criação de referenciais direcionados à administração pública tem potencial para acelerar a adoção responsável dessas técnicas nos órgãos governamentais.

Por fim, recomenda-se aprofundar a investigação sobre o impacto da privacidade diferencial em modelos preditivos. A análise sistemática dos efeitos da proteção de dados no desempenho de algoritmos de previsão e classificação, tanto em nível agregado quanto em recortes setoriais, permitirá documentar de forma robusta

o *trade-off* entre utilidade e privacidade em diferentes cenários operacionais. Essa linha de continuidade conecta diretamente o eixo preditivo ao eixo de proteção de dados, aproximando a pesquisa de casos concretos de uso no setor público e privado.

6.5 CONSIDERAÇÕES FINAIS

A dissertação atingiu o objetivo de articular, em um único estudo, três eixos complementares — previsão temporal, análise de recorrência e privacidade diferencial — em uma abordagem integrada de governança de violações de dados, com aderência explícita aos princípios e diretrizes da LGPD (Brasil 2018). Ao combinar modelagem preditiva, análise de sobrevivência e experimentos controlados com *DP*, o trabalho oferece uma contribuição tanto conceitual quanto aplicada para o campo de segurança da informação e proteção de dados.

Do ponto de vista acadêmico, a pesquisa consolida um delineamento experimental reprodutível para comparação de modelos preditivos, introduz a análise de sobrevivência como eixo relevante para compreender recorrência de incidentes e quantifica, em termos de deslocamento distributivo e degradação de acurácia, o impacto de parâmetros de *DP* sobre dados tabulares sensíveis. Sob a perspectiva prática, os resultados delineiam um caminho factível para que organizações públicas e privadas avancem de uma postura predominantemente reativa para uma abordagem proativa, baseada em evidências e alinhada a requisitos regulatórios, na gestão de violações de dados.

Em síntese, ao integrar técnicas de previsão, análise temporal e mecanismos formais de proteção de dados em um mesmo arcabouço, esta dissertação contribui para a construção de uma cultura de decisão informada, na qual resiliência cibernética, conformidade regulatória e transparência responsável deixam de ser objetivos concorrentes e passam a ser dimensões complementares de uma mesma estratégia de governança.

REFERÊNCIAS BIBLIOGRÁFICAS

Abraham e Box 1979 ABRAHAM, B.; BOX, G. E. P. Bayesian analysis of some outlier problems in time series. *Biometrika*, v. 66, n. 2, p. 229–236, 1979.

Ainslie et al. 2023 AINSLIE, S. et al. Cyber-threat intelligence in practice: A systematic literature review and framework for future research. *Computers & Security*, v. 132, p. 103352, 2023.

Alaa et al. 2022 ALAA, A. et al. How faithful is your synthetic data? sample-level metrics for evaluating and auditing generative models. In: *Proceedings of the 39th International Conference on Machine Learning*. [S.l.: s.n.], 2022. v. 162, p. 290–306.

Alhamad e Singh 2021 ALHAMAD, I. A.; SINGH, H. P. Predicting key factors impacting online hotel ratings using data mining approach: A case study of the makkah city of saudi arabia. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, v. 12, n. 2, p. 12A2N, 1–12, 2021.

Alvarez et al. 2024 ALVAREZ, C. A. R. et al. In survival analysis, there are three types of censorship: right, left, and interval (right censoring in survival analysis). *Reliability Engineering and System Safety*, v. 246, p. 110040, 2024.

Alvarez et al. 2025 ALVAREZ, C. A. R. et al. Predicting iot device vulnerability fix times with survival and failure time models. In: *2025 Australasian Computer Science Week (ACSW 2025)*. [S.l.: s.n.], 2025.

Anon. 2021 ANON., A. Differential privacy based location privacy protection schemes in location-based services: A survey. *Computers & Security*, v. 111, p. 102464, 2021.

Ansari et al. 2024 ANSARI, A. F. et al. Chronos: Learning the language of time series. *arXiv preprint*, 2024.

Autores do Artigo Healthcare 2020 Autores do Artigo Healthcare. Artigo sobre análise de violações de dados de saúde nos eua (inferred from context). *Healthcare*, v. 8, n. 133, 2020.

Belarmino, Ricarte e Motta 2024 BELARMINO, G. S.; RICARTE, D. R. D.; MOTTA, G. H. M. B. A lei geral de proteção de dados do brasil à luz do regimento europeu: Um exame comparativo e prospectivo através de uma revisão sistemática. 2024.

Benzell et al. 2022 BENZELL, S. et al. How apis create growth by inverting the firm. *SSRN*, n. 3432591, 2022.

Bertoni et al. 2022 BERTONI, A. P. S. et al. Internet das coisas de saúde: Aplicando iot, interoperabilidade e aprendizado de máquina com foco no paciente. 2022.

Bertoni 2020 BERTONI, E. O novo vazamento de dados na saúde e suas consequências. *Nexo*, 2020.

Bispo et al. 2024 BISPO, G. D. et al. Automatic literature mapping selection: Classification of papers on industry productivity. *Applied Sciences*, v. 14, n. 9, p. 3679, 2024.

Bradley, Alhajjar e Bastian 2023 BRADLEY, T.; ALHAJJAR, E.; BASTIAN, N. D. Novelty detection in network traffic: Using survival analysis for feature identification. In: *2023 IEEE International Conference on Assured Autonomy (ICAA)*. [S.l.: s.n.], 2023. p. 11–18.

Brasil 2018 BRASIL. Lei nº13.709, de 14 de agosto de 2018: Lei geral de proteção de dados pessoais (lgpd). 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>.

Carvalho et al. 2023 CARVALHO, T. et al. Towards a data privacy-predictive performance trade-off. *Expert Systems With Applications*, v. 223, p. 119785, 2023.

Carvalho et al. 2023 CARVALHO, T. et al. On the impact of privacy-preserving techniques on re-identification risk and predictive performance. *Expert Systems With Applications*, v. 223, p. 119785, 2023.

Ding et al. 2025 DING, J. et al. Non-specified title (related to forecasting and remote sensing). *International Journal of Applied Earth Observation and Geoinformation*, p. 104473, 2025. A informação exata da fonte: "RMSE is more sensitive to outliers than MAE."[4-6].

Dwork et al. 2025 DWORK, C. et al. Calibrating noise to sensitivity in private data analysis. *International Journal of Intelligent Engineering and Systems*, v. 18, n. 9, 2025.

Elger e Santander 2024 ELGER, E.; SANTANDER, V. A. A engenharia de requisitos e a lei geral de proteção de dados (lgpd): Uma revisão sistemática da literatura. *Contextual Reference (Engenharia de Software/Requisitos)*, 2024.

Fernandes, Machado e Amaral 2023 FERNANDES, J.; MACHADO, C.; AMARAL, L. Towards a readiness model derived from critical success factors, for the general data protection regulation implementation in higher education institutions. *Strategic Management*, v. 28, n. 1, p. 4–19, 2023.

Fernandes e Nuzzi 2022 FERNANDES, M. E.; NUZZI, A. P. E. Fundamentos da lei geral de proteção de dados (lgpd): Uma revisão narrativa. *Research, Society and Development*, v. 11, n. 12, p. e310111234247, 2022.

Fildes, Ma e Kolassa 2019 FILDES, R.; MA, S.; KOLASSA, S. Retail forecasting: research and practice. *International Journal of Forecasting*, 2019.

Fissler et al. 2020 FISSLER, T. et al. Forecast evaluation of quantiles, prediction intervals, and other set-valued functionals. *arXiv:1910.07912*, 2020.

Gayam, Yellu e Thuniki 2021 GAYAM, S. R.; YELLU, R. R.; THUNIKI, P. Artificial intelligence for real-time predictive analytics: Advanced algorithms and applications in dynamic data environments. *Distrib Learn Broad Appl Sci Res*, v. 7, p. 18–37, 2021.

Govindankutty e Goel 2024 GOVINDANKUTTY, S.; GOEL, P. Enhancing transparency and compliance in ai systems. *Journal of Quantum Science and Technology (JQST)*, v. 1, p. 504–511, 2024.

Henderson e al. 2023 HENDERSON, S.; AL. et. Statistical data privacy. *Annual Review of Statistics and Its Application*, 2023.

Hou, Xue e Zhang 2020 HOU, K.; XUE, C.; ZHANG, L. Replicating anomalies. *Review of Financial Studies*, v. 33, n. 7, p. 2979–3033, 2020.

International Organization for Standardization 2018 INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Iso/iec 27005:2018—information technology—security techniques—information security risk management. 2018. Conforme referenciado em Alzahrani et al. (2023).

Janiesch, Zschech e Heinrich 2021 JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. *Electronic Markets*, v. 31, n. 3, p. 685–695, 2021.

Jimenez et al. 2020 JIMENEZ, J. J. M. et al. Towards multi-model approaches to predictive maintenance. *Journal of Manufacturing Systems*, v. 55, p. 149–164, 2020.

Jin et al. 2021 JIN, J. et al. Are we there yet? timing and floating-point attacks on differential privacy systems. In: IEEE. *IEEE Symposium on Security and Privacy (SP)*. [S.l.], 2021.

Joint Task Force 2022 Joint Task Force. *Assessing Security and Privacy Controls in Information Systems and Organizations*. [S.l.], 2022. (NIST Special Publication 800-53A).

Kifer, Messing e Roth 2020 KIFER, D.; MESSING, S.; ROTH, A. Guidelines for implementing and auditing differentially private systems. *arXiv*, 2020.

Kotsias, Ahmad e Scheepers 2022 KOTSIAS, J.; AHMAD, A.; SCHEEPERS, R. Adopting and integrating cyber-threat intelligence in a commercial organization. *European Journal of Information Systems*, 2022.

Kumar e Gupta 2020 KUMAR, R.; GUPTA, H. A study on healthcare data breaches. *Healthcare*, v. 8, n. 2, p. 133, 2020.

Kumar, Kaur e Kumar 2023 KUMAR, S.; KAUR, A.; KUMAR, R. A hybrid oversampling approach to deal with data imbalance and outliers for credit card fraud detection. *Applications in Computing and Mathematics for Engineering*, Elsevier, v. 2, n. 1, p. 100004, 2023. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2772662223000048>.

Lewis 1982 LEWIS, C. D. *Métodos de Previsão Industrial e Empresarial: Um Guia Prático para Suavização Exponencial e Ajuste de Curvas*. Oxford, Reino Unido: Butterworth Scientific, 1982. Disponível em: Google Scholar.

Li et al. 2010 LI, X. et al. A location-aware recommender system for tourism mobile commerce. In: *Proceedings of the International Conference on Information Science and Engineering*. [S.l.: s.n.], 2010. p. 1709–11.

Lopes e Amaral 2022 LOPES, F.; AMARAL, M. A. A. Implementação da lei geral de proteção de dados pessoais (lgpd) em uma instituição sem fins lucrativos, atuante na Área da educação básica. *Projetos e Relatórios de Estágios*, v. 4, n. 1, p. 21–21, 2022.

Lu et al. 2022 LU, F. et al. A general framework for auditing differentially private machine learning. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2022.

Maddireddy e Maddireddy 2020 MADDIREDDY, B. R.; MADDIREDDY, B. R. Proactive cyber defense: Utilizing ai for early threat detection and risk assessment. *International Journal of Advanced Engineering Technologies and Innovations*, v. 1, n. 1, p. 37–53, 2020.

Marczak e Proietti 2016 MARCZAK, M.; PROIETTI, T. Outlier detection in structural time series models: The indicator saturation approach. *International Journal of Forecasting*, v. 32, n. 1, p. 180–202, 2016.

Martinez, Castle e Hendry 2021 MARTINEZ, A. B.; CASTLE, J. L.; HENDRY, D. F. Smooth robust multi-horizon forecasts. *Advances in Econometrics*, Forthcoming, 2021. Forthcoming.

Mintarsih et al. 2023 MINTARSIH, F. et al. Lstm variants comparison for exchange rate idr/usd forecasting with rolling window cross validation. In: *Eighth International Conference on Informatics and Computing (ICIC)*. [S.l.: s.n.], 2023.

Mulla et al. 2025 MULLA, S. M. et al. Exploring the variations among arima models for time series forecasting of data breaches. *Data Intelligence*, v. 7, p. 40–69, 2025.

Neto et al. 2021 NETO, N. N. et al. Developing a global data breach database and the challenges encountered. *J. Data Inf. Qual.*, v. 13, p. 1–33, 2021.

Papathanasiou, Demertzis e Tziritas 2023 PAPATHANASIOU, D.; DEMERTZIS, K.; TZIRITAS, N. Machine failure prediction using survival analysis. *Future Internet*, v. 15, n. 5, p. 153, 2023.

Perera et al. 2022 PERERA, S. et al. Factors affecting reputational damage to organisations due to cyberattacks. *Informatics*, v. 9, n. 1, p. 28, 2022.

Petropoulos et al. 2022 PETROPOULOS, F. et al. The theory and practice of forecasting. *International Journal of Forecasting*, v. 38, n. 2, p. 705–871, 2022.

Ponce et al. 2023 PONCE, L. M. et al. Um arcabouço para processamento escalável de vulnerabilidades e caracterização de riscos à conformidade da lgpd. In: *Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg)*. [S.l.: s.n.], 2023. p. 15–28.

Ponomareva et al. 2024 PONOMAREVA, N. et al. The laplace distribution centered at zero. *Unknown*, 2024.

Ponomareva et al. 2024 PONOMAREVA, N. et al. The query is a function f that takes a dataset as input and outputs a quantity of interest. *Unknown*, 2024.

Ponomareva et al. 2024 PONOMAREVA, N. et al. Smaller ϵ's typically lead to lower utility and more protection. *Unknown*, 2024.

Ponte et al. 2024 PONTE, G. R. et al. Balancing data utility and privacy risk in marketing analytics using differential privacy. *International Journal of Research in Marketing*, v. 41, p. 529–546, 2024.

Rahimpour et al. 2024 RAHIMPOUR, H. et al. The rmse calculates the square root of the mean squared differences, emphasising larger errors due to the squared term. *Unpublished Preprint (ssrn-5244222.pdf)*, 2024. A informação exata da fonte: "The RMSE calculates the square root of the mean squared differences, emphasising larger errors due to the squared term."[1].

Rhif e al. 2019 RHIF, M.; AL., E. Non-stationary time series (ts) analysis has gained an explosive interest over the recent decades in different applied sciences... *Applied Sciences*, v. 9, n. 1345, 2019.

Rodrigues et al. 2025 RODRIGUES, G. A. P. et al. Balancing privacy and utility: Evaluating distributional shifts and accuracy in differentially private synthetic breached data. In: *Anais do Simpósio Brasileiro de Cibersegurança (SBSEG 2025)*. Foz do Iguaçu, PR, Brasil: Sociedade Brasileira de Computação, 2025.

Rodrigues et al. 2024 RODRIGUES, G. A. P. et al. Impact, compliance, and countermeasures in relation to data breaches in publicly traded u.s. companies. *Future Internet*, v. 16, n. 6, p. 201, 2024.

Saifuzzaman e al. 2024 SAIFUZZAMAN, M.; AL. et. A systematic literature review on wearable health data publishing under differential privacy. *A systematic literature review*, 2024.

Santos et al. 2025 SANTOS, E. et al. Modelos preditivos para detecção de violações de dados: Uma abordagem comparativa entre técnicas clássicas e de aprendizado profundo. In: *Anais do XXV Simpósio Brasileiro de Cibersegurança*. Porto Alegre, RS, Brasil: SBC, 2025. p. 626–642. ISSN 0000-0000. Disponível em: <https://sol.sbc.org.br/index.php/sbseg/article/view/36648>.

Seeman e Susser 2024 SEEMAN, J.; SUSSER, D. Between privacy and utility: On differential privacy in theory and practice. *ACM J. Responsible Comput.*, v. 1, p. 1–18, 2024.

Sharma e Bantan 2025 SHARMA, A.; BANTAN, M. Simulating data breaches: Synthetic datasets for depicting personally identifiable information through scenario-based breaches. *Data in Brief*, v. 58, p. 111207, 2025.

Silva et al. 2020 SILVA, A. De Melo e et al. A methodology to evaluate standards and platforms within cyber threat intelligence. *Future Internet*, v. 12, n. 12, p. 218, 2020.

Song et al. 2024 SONG, X. et al. Deep learning-based time series forecasting. *Non-specified Journal (s10462-024-10989-8.pdf)*, 2024.

Sun et al. 2023 SUN, N. et al. Cyber threat intelligence mining for proactive cybersecurity defense: A survey and new perspectives. *IEEE Communications Surveys & Tutorials*, v. 25, p. 1748–1774, 2023.

Thakkar e Lohiya 2021 THAKKAR, A.; LOHIYA, R. A review on machine learning and deep learning perspectives of ids for iot: recent updates, security issues, and challenges. *Archives of Computational Methods in Engineering*, v. 28, p. 3211–3243, 2021.

Urooj et al. 2022 UROOJ, U. et al. Ransomware detection using the dynamic analysis and machine learning: A survey and research directions. *Applied Sciences*, v. 12, n. 1, p. 172, 2022.

Vainzof 2020 VAINZOF, R. Lei 13.709, de 14 de agosto de 2018: Capítulo i – disposições preliminares. In: MALDONADO, V. N.; BLUM, R. O. (Ed.). *LGPD: Lei Geral de Proteção de Dados Comentada*. São Paulo: Thomson Reuters Brasil, 2020.

Val et al. 2024 VAL, O. O. et al. Strengthening cybersecurity measures for the defense of critical infrastructure in the united states. *Asian Journal of Research in Computer Science*, v. 17, n. 11, p. 25–45, 2024.

Wang, Li e Reddy 2019 WANG, P.; LI, Y.; REDDY, C. K. Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)*, v. 51, n. 6, p. 1–36, 2019.

Wang et al. 2024 WANG, S. et al. Data privacy and cybersecurity challenges in the digital transformation of the banking sector. *Computers & Security*, v. 147, p. 104051, 2024.

Yang e al. 2024 YANG, M.; AL. et. Reported incidents of data exposure on social networks. *Computer Standards & Interfaces*, v. 89, p. 103827, 2024.

Zabierek et al. 2021 ZABIEREK, L. et al. Toward a collaborative cyber defense and enhanced threat intelligence structure foreword and select discussion by. *Report/Paper*, 2021.

Zio e Miqueles 2024 ZIO, E.; MIQUELES, L. Methodology, challenges, and solutions related to data integration and cybersecurity in digital twins. *Reliability Engineering and System Safety*, v. 246, p. 110040, 2024.

Zou et al. 2019 ZOU, Y. et al. Nonlinear time series analysis based on complex networks: A survey. *Physics Reports*, v. 787, p. 1–97, 2019.