DISSERTAÇÃO DE MESTRADO PROFISSIONAL

# Análise Abrangente de Vazamentos de Dados: Riscos, Conformidade e Estratégias de Prevenção

**Gabriel Arquelau Pimenta Rodrigues**

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

# Análise Abrangente de Vazamentos de Dados: Riscos, Conformidade e Estratégias de Prevenção

**Gabriel Arquelau Pimenta Rodrigues**

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

**Co-Orientador: Prof. Dr. Robson de Oliveira Albuquerque, ENE/UnB**

PUBLICAÇÃO: PPEE.MP.071

BRASÍLIA-DF

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

# Análise Abrangente de Vazamentos de Dados: Riscos, Conformidade e Estratégias de Prevenção

**Gabriel Arquelau Pimenta Rodrigues**

*Dissertação de mestrado profissional submetida ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica*

Banca examinadora

Dr. André Luiz Marques Serrano, EPR/UnB _____
*Presidente*

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

Dr. Rodrigo Bonacin,
Centro de Tecnologia da Informação Renato Archer _____
*Examinador Externo*

Dr. Clóvis Neumann, EPR/UnB _____
*Suplente*

**FICHA CATALOGRÁFICA**

RODRIGUES, GABRIEL ARQUELAU PIMENTA
Análise Abrangente de Vazamentos de Dados: Riscos, Conformidade e Estratégias de Prevenção [Distrito Federal] 2024.
xvi, 69 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2024).
Dissertação de mestrado profissional - Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia elétrica

1. Análise de risco
2. Conformidade
3. Controles de segurança
4. Vazamentos de dados

I. ENE/FT/UnB
II. Título (série)

**REFERÊNCIA BIBLIOGRÁFICA**

RODRIGUES, G.A.P (2024). *Análise Abrangente de Vazamentos de Dados: Riscos, Conformidade e Estratégias de Prevenção* . Dissertação de mestrado profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 69 p.



Gabriel Arquelau Pimenta Rodrigues
Departamento de Engenharia Elétrica (ENE) - FT
Universidade de Brasília (UnB)
Campus Darcy Ribeiro
CEP 70919-970 - Brasília - DF - Brasil

## DEDICATÓRIA

A Kyume Lopes.

## AGRADECIMENTOS

Agradeço especialmente à minha amada esposa, de quem sou admirador eterno, e quem mais me impulsiona a crescer.

Aos meus pais, familiares e amigos, que são os de mais elevada essência que eu poderia pedir.

Aos meus orientadores, por terem sido o suporte fundamental para o sucesso desta pesquisa; e aos demais professores e funcionários do Programa de Pós-Graduação Profissional em Engenharia Elétrica (PPEE).

À Universidade de Brasília, pelas oportunidades, aprendizagens e amizades que tem me agraciado desde que iniciei minha graduação.

Aos colegas de profissão que contribuíram tecnicamente com o estudo; e à minha Instituição de trabalho por ter amparado minha participação no PPEE.

À Confederação Nacional da Indústria (CNI) pelo apoio na condução desta pesquisa.

“Ser poeta é muito bom porque eu não tenho nenhuma obrigação de veracidade. Eu posso mentir à vontade, cientista é que não pode.”

---

Ariano Suassuna

## RESUMO

A transformação digital das empresas, impulsionada pela Internet das Coisas, sistemas de informação e serviços em nuvem, gerou um aumento na quantidade de dados disponíveis. Essa revolução, porém, trouxe consigo o crescente desafio dos vazamentos de dados, que impactam milhões de indivíduos anualmente e resultam em perda de informações, que podem ser sensíveis e privadas. Esse tipo de incidente de segurança afeta clientes, partes interessadas, organizações e empresas, com potencial de comprometer a segurança física dos indivíduos e de causar perdas econômicas substanciais em diferentes setores da indústria. Este estudo visa fornecer uma maior compreensão sobre a ocorrência de vazamentos de dados em um contexto global, buscando fortalecer a segurança contra esse tipo de incidente. Para tal, dois conjuntos de dados são analisados: (i) um com violações de dados em todo o mundo entre 2018 e 2019, e (ii) outro com violações de dados que afetaram empresas listadas na NYSE e na NASDAQ, nos EUA, entre 2005 e 2015. A análise estatística e dos aspectos de conformidade, correlacionando elementos-chave da legislação com os resultados, juntamente com a avaliação qualitativa de risco e a discussão dos controles de segurança, visa contribuir para uma compreensão abrangente das exposições de dados e capacitar as organizações a se protegerem de forma proativa. O estudo também contribui para a discussão sobre leis de proteção de dados, apoiando, por exemplo, o processo de decisão sobre a localização de armazenamento de dados na nuvem. Esta pesquisa mostra que o Brasil tem uma significativa incidência de vazamentos de dados no setor público e indica, com base numa análise qualitativa de risco, quais vetores de ataque devem ser prioritariamente mitigados, com indicações de controles de segurança.



**Palavras-chave:** Análise de riscos, Conformidade, Controles de Segurança, Privacidade, Vazamentos de dados

## ABSTRACT

The digital transformation of companies, driven by the Internet of Things, information systems, and cloud services, has generated an increase in the amount of data available. However, this revolution has also brought the growing challenge of data breaches, which impact millions of individuals annually and result in the loss of information, that can be sensitive and private. These security incidents affect customers, stakeholders, organizations and companies, with potential to compromise the physical security of individuals and to cause substantial economic losses in different industrial sectors. This study aims to provide a greater understanding of data breaches in a global context, seeking to strengthen the security against this type of incident. To achieve this, two datasets are analyzed: (i) one with data breaches worldwide between 2018 and 2019, and (ii) another with data breaches that affected companies listed on the NYSE and NASDAQ in the US between 2005 and 2015. The statistical analysis, along with the evaluation of compliance aspects, correlating key elements of legislation with the results, together with the qualitative risk assessment and discussion of security controls, aim to contribute to a comprehensive understanding of data breaches and empower organizations to proactively safeguard themselves. The study also contributes to the discussion on data protection and laws, supporting, for example, the decision-making process on the location of cloud data storage. This research shows that Brazil has a significant incidence of data leaks in the public sector and indicates, based on a qualitative risk analysis, which attack vectors should be prioritized for mitigation, with recommendations for security controls.



**Keywords: Compliance, Data breach, Privacy, Risk assessment, Security controls**

# SUMÁRIO

# LISTA DE FIGURAS

# LISTA DE TABELAS

# LISTA DE ABREVIAÇÕES

ANPD Autoridade Nacional de Proteção de Dados

APP *Australian Privacy Principles*

APPI *Act on the Protection of Personal Information*

BYOD *Bring Your Own Device*

CAC *Cyberspace Administration of China*

CARD Vazamento de dado de cartão de crédito/débito

CCPA *California Consumer Privacy Act*

CIS *Center for Internet Security*

CNP *Card Not Present*

CPA *Colorado Privacy Act*

CPRA *California Privacy Rights Act*

CPTED *Crime Prevention Through Environmental Design*

CSF *Cyber Security Framework*

CTDPA *Connecticut Data Privacy Act*

DISC Vazamento por negligência

DLP *Data Loss Prevention*

DPA *Data Privacy Act/Data Protection Authority*

DPDP *Digital Personal Data Protection*

DPDPA *Delaware Personal Data Privacy Act*

DPO *Data Protection Office*

EDPB *European Data Protection Board*

EMV *Europay, MasterCard, and Visa*

FADP *Federal Act on Data Protection*

FDBR *Florida Digital Bill of Rights*

FDPIC *Federal Data Protection and Information Commissioner*

FTC *Federal Trade Commission*

GDPR *General Data Protection Regulation*

GLBA *Gramm–Leach–Bliley Act*

HACK Vazamento por atividade de hacking

HD *Hard Drive*

HIPAA *Health Insurance Portability and Accountability Act*

ICDPA *Iowa Consumer Data Protection Act*

ICO *Information Commissioner's Office*

Indiana CDPA *Indiana Consumer Data Protection Act*

INSD Vazamento por *insider* malicioso

KVKK *Kişisel Verileri Koruma Kurumu (Personal Data Protection Authority)*

LGPD Lei Geral de Proteção de Dados

LPPD *Law on Protection of Personal Data*

MTCDPA *Montana Consumer Data Privacy Act*

NPC *National Privacy Commission*

NYSE *New York Stock Exchange*

OAIC *Office of the Australian Information Commissioner*

OCPA *Oregon Consumer Privacy Act*

OPC *Office of the Privacy Commissioner*

P2PE *Point-to-Point Encryption*

PA *Privacy Act*

PCI-DSS *Payment Card Industry Data Security Standard*

PCPD *Privacy Commissioner for Personal Data*

PDP *Personal Data Protection*

PDPA *Personal Data Protection Act*

PDPC *Personal Data Protection Commission*

PDPL *Personal Data Protection Law*

PDPO *Personal Data Privacy Ordinance*

PHYS Vazamento por documento em papel

PII *Personal Identifiable Information*

PIPEDA *Personal Information Protection and Electronic Documents Act*

PIPL *Personal Information Protection Law*

PORT Vazamento por ativo portável

PPA *Privacy Protection Authority*

PPC *Personal Information Protection Commission*

PPL *Protection of Privacy Law*

PRC *Privacy Rights Clearinghouse*

SETA *Security Education Training and Awareness*

SIC *Superintendence of Industry and Commerce*

SJR *Scimago Journal Rank*

SOF *Superintendence of Finance*

SOX *Sarbanes–Oxley Act*

SSD *Solid-State Drive*

SSS *Scientific Significance Score*

STAT Vazamento por ativo estacionário

TDPSA *Texas Data Privacy and Security Act*

TIPA *Tennessee Information Protection Act*

TTP *Tactics, Techniques, and Procedures*

UCPA *Utah Consumer Privacy Act*

UKGDPR *United Kingdom General Data Protection Regulation*

UNKN Vazamento por vetor desconhecido

VCDPA *Virginia Consumer Data Protection Act*

# 1 INTRODUÇÃO

A segurança da informação se tornou uma preocupação crítica para organizações de todos os setores, considerando a grande quantidade de dados que armazenam [8]. Esses dados propiciam um maior conhecimento acerca do mercado, clientes e operações. Com isso, as instituições se capacitam a tomar decisões mais informadas, a personalizar suas ofertas, a otimizar processos e a antecipar tendências [9]. Em consequência, as empresas podem alcançar vantagens competitivas significativas. Portanto, empresas como Facebook e Uber desenvolveram projetos de análise de dados, como Prophet e Orbit [10], respectivamente, para antecipar tendências e identificar padrões em seus dados.

Faz-se necessário, no entanto, resguardar esses dados de maneira a conferir-lhes confidencialidade conforme seus graus de sigilo. Isso porque o acesso indevido a esses dados pode gerar graves consequências para empresas e indivíduos, incluindo perdas financeiras [11], danos à reputação [12], roubo de identidade [13] e até mesmo danos físicos [14]. Especificamente no caso de empresas de capital aberto, incidentes de segurança cibernética podem impactar o preço de suas ações e acionistas, além de diminuir o apelo da empresa para potenciais investidores [15].

A segurança cibernética é também relevante no contexto industrial, já que alguns setores, como o energético, são considerados de infraestrutura crítica para uma nação. Assim, é fundamental resguardar os ativos de informação da indústria, para que possa se desenvolver sem incidentes que afetariam seus clientes e suas finanças.

Considerando esses impactos aos clientes e às instituições, inúmeras regulamentações e padrões estão em vigor globalmente para proteger a privacidade dos dados e prevenir crimes cibernéticos dirigidos a ela [16]. No Brasil, por exemplo, a Lei Geral de Proteção de Dados (LGPD) foi publicada em 2018 [17]. Empresas operando sob a jurisdição dessas regulamentações são obrigadas a cumprir tais leis.

Alguns países regulam o tema mesmo sem uma lei geral federal. É o caso, por exemplo, dos Estados Unidos que, embora não tenha uma legislação federal de proteção de dados em vigor, exige conformidade com leis em nível estadual [18] e leis específicas para determinados tipos de dados, como a *Health Insurance Portability and Accountability Act* (HIPAA), que regula a proteção de dados médicos e de saúde [19].

Além de demonstrar conformidade a essas leis e normas, as organizações devem comprovar diligência e cuidados adequados com seus ativos (em inglês, *due diligence* e *due care*, para protegê-los de atores maliciosos, incluindo os melhores esforços da organização para prevenir incidentes de segurança cibernética [20].

A mitigação desses incidentes, no entanto, deve ser pautada de forma a otimizar os recursos e priorizar a implementação de mecanismos de segurança conforme o apetite de riscos da organização. Assim, para uma estratégia de prevenção mais eficaz, é fundamental gerenciar os riscos cibernéticos, considerando sua probabilidade e potencial de impacto [21].

Par avançar essa discussão, este estudo se fundamenta em dois conjuntos de dados: o primeiro, desen-

volvido por [22],registra vazamentos de dados globalmente entre 2018 e 2019. O segundo, publicado por [23], contém informações sobre incidentes de violação de dados que afetaram empresas de capital aberto dos EUA entre 2005 e 2015.

De acordo com [15], o cenário dinâmico dos crimes cibernéticos torna difícil para empresas e profissionais de segurança antecipar os tipos, magnitude e gravidade das futuras violações. Considerando isso, a visualização de métricas relacionadas a esses conjuntos de dados pode beneficiar o gerenciamento de riscos e, consequentemente, a prevenção de incidentes. Além disso, a visualização de estatísticas acerca desses incidentes também possibilita a discussão acerca da legislação de proteção de dados a nível global e de controles de segurança que podem ser implementados para reduzir a frequência ou impacto desses incidentes.

Este trabalho avalia as estatísticas das exposições de dados listadas publicamente e utiliza os padrões observados e estudos de caso como base para a discussão sobre (i) conformidade com regulamentos de proteção de dados nos escopos geográfico e setorial; (ii) controles técnicos e administrativos de segurança aplicáveis para proteger os dados em uso, em trânsito e em repouso para diferentes vetores de ataque e ao longo de todo o ciclo de vida: criação, armazenamento, uso, compartilhamento, arquivamento e destruição; (iii) diretrizes para responder adequadamente a esse tipo de incidente; e (iv) as consequências de uma exfiltração tanto para o sujeito dos dados quanto para o proprietário dos dados.

## 1.1 MOTIVAÇÃO

As organizações levaram, em média, 204 dias para identificar um vazamento de dados em 2023, demonstrando mudanças mínimas em relação aos anos anteriores [5]. Além disso, o custo médio de uma violação de dados em organizações com poucos recursos de segurança é de 5,36 milhões de dólares [5]. Essas preocupações são agravadas pelo aumento na coleta de dados de várias fontes sensíveis como equipamentos médicos e transações online [24].

Em um caso específico, ocorrido em janeiro de 2024, 26 bilhões de registros foram expostos em um único incidente, marcando-o como o maior vazamento de dados conhecido. Esse incidente foi denominado de Mãe de Todos os Vazamentos e incluiu 12 TB de dados vazados de plataformas, afetando, por exemplo, redes sociais como LinkedIn e Twitter [25].

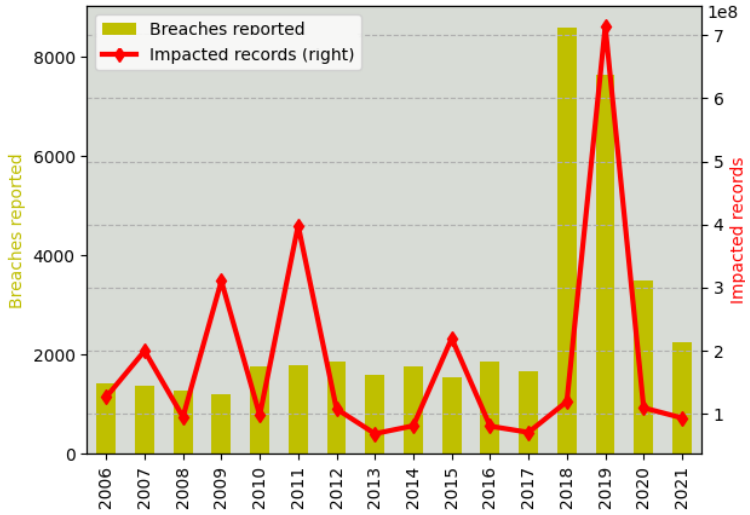



Figura 1.1: Séries temporais da quantidade de incidentes ocorridos (barras) e quantidade de registros vazados (linhas) nos EUA, de acordo com a *Privacy Rights Clearinghouse*

A Figura 1.1 mostra a quantidade histórica de incidentes registrados nos Estados Unidos pela Privacy Rights Clearinghouse (PRC), uma organização que fornece recursos sobre questões relacionadas à privacidade de dados e segurança da informação.

Conforme observado na Figura 1.1, entre 2006 e 2021, o mínimo de informação vazada anualmente, somente nos EUA, se aproxima de 100 milhões de registros. Em 2019, ano que teve a maior quantidade de dados violados, mais de 700 milhões de registros foram expostos. Esse número, que está restrito apenas aos Estados Unidos, representa aproximadamente 10% da população mundial. Além disso, de acordo com (22), há ainda mais ocorrências públicas de vazamentos de dados na Europa do que nos EUA.

Faz-se necessário, portanto, mitigar a ocorrência desses incidentes. Para desenvolver estratégias eficazes de proteção de dados e mitigação de riscos, é fundamental compreender as vulnerabilidades que podem viabilizar vazamentos de dados. Além do exposto, é também necessário conhecer o cenário regulatório para um planejamento estratégico mais eficaz.

## 1.2 OBJETIVOS

Este estudo tem como objetivo geral proporcionar uma compreensão mais ampla dos vazamentos de dados e identificar soluções para mitigar esses incidentes, tanto por meio de regulamentação quanto por meio de abordagens técnicas.

Para isso, tem-se como objetivos específicos:

- Visualizar métricas acerca da ocorrência de violações de dados globalmente;
- Analisar as regulações de proteção de dados dos países afetados, além das leis estaduais e setoriais dos Estados Unidos;
- Realizar uma análise qualitativa de riscos acerca dos vetores de ataques que ensejam esses incidentes;

- Levantar controles de segurança que podem ser implementados para mitigar os riscos identificados.

## 1.3 CONTRIBUIÇÕES ACADÊMICAS

Este estudo contribui ampliando a compreensão acerca da ocorrência de vazamentos de dados, e fornecendo uma análise abrangente das estatísticas relacionadas a incidentes desse tipo. A visualização de variáveis, como as causas dos vazamentos de dados e a incidência geográfica e setorial, aprofunda a compreensão dos padrões desses eventos.

Ao examinar as leis e regulamentos de proteção de dados, o estudo oferece uma visão detalhada das obrigações legais que as organizações enfrentam em relação à proteção das informações dos indivíduos. Isso inclui a análise de leis como o Regulamento Geral de Proteção de Dados (GDPR) na União Europeia, a LGPD brasileira, leis estaduais e setoriais dos Estados Unidos e regulamentos similares em outras jurisdições. O estudo ajuda, ainda, a identificar lacunas nas regulamentações existentes, apontando áreas onde a legislação precisa ser aprimorada. A análise acerca dos cenários regulatórios em diferentes países é relevante, por exemplo, na decisão de armazenamento de dados no exterior, como em serviços de nuvem.

Essa parte do trabalho, mais focada na regulação da proteção dos dados, foi publicada em artigo da revista acadêmica *Data*, intitulado *Understanding Data Breach from a Global Perspective: Incident Visualization and Data Protection Law Review* [26].

Ademais, a análise qualitativa dos riscos envolvidos nos vazamentos de dados, por meio da avaliação da frequência de ocorrência e da quantidade de dados vazados por incidente, aprimora o entendimento dos fatores que contribuem para esses incidentes. Essa análise também evidencia quais vetores de ataque devem ser priorizados num plano de mitigação, com base numa matriz de riscos.

Os resultados desta discussão foram publicados na revista científica *Results in Engineering*, no artigo *Mapping of data breaches in companies listed on the NYSE and NASDAQ: Insights and Implications* [27].

Por fim, os controles de segurança disponíveis para mitigar os riscos associados aos vazamentos de dados são discutidos. Com isso, o estudo fornece orientações para a implementação de medidas preventivas. Outrossim, a discussão sobre o impacto desses vazamentos, incluindo seus efeitos nos preços das ações das empresas afetadas, enfatiza a importância da condução adequada de resposta a incidentes.

Esses resultados foram publicados na revista *Future Internet*, no artigo *Impact, compliance, and countermeasures of data breaches in publicly traded U.S. companies* [28].

## 1.4 ESTRUTURA DO TRABALHO

Os demais capítulos deste trabalho estão organizados como descrito a seguir. O Capítulo 2 discute os trabalhos relacionados ao estudo de vazamento de dados, seja na área regulatória, de análise de riscos, ou de implementação de controles de segurança. O Capítulo 3 ilustra a metodologia e materiais usados. O Capítulo 4 apresenta os resultados e discussões fundamentadas neles. O Capítulo 5 conclui este trabalho,

apresentado as limitações e propondo trabalhos futuros.

# 2 REFERENCIAL TEÓRICO

Neste capítulo, estudos relevantes que ajudam a compreender os vazamentos de dados sob diferentes perspectivas são apresentados. A Seção 2.1 apresenta um levantamento bibliométrico do assunto, avaliando métricas como contagens de citações e padrões de publicação em vazamentos de dados. Obras significativas são selecionadas, de acordo com duas métricas diferentes, e brevemente descritas na Seção 2.2. Por último, a Seção 2.3 indica trabalhos que tenham disponibilizado conjuntos de dados sobre ocorrências de vazamentos de dados.

Como referência para classificar o estudo de exposições de dados, uma taxonomia do estudo de vazamento de dados, que é apresentada na Figura 2.1, foi proposta por [1]. Este trabalho aborda todos os elementos de prevenção da taxonomia, enquanto foca na consequência aos preços das ações na parte de impacto aos ativos. A Tabela 2.1 compara as contribuições deste estudo com as de alguns trabalhos similares.

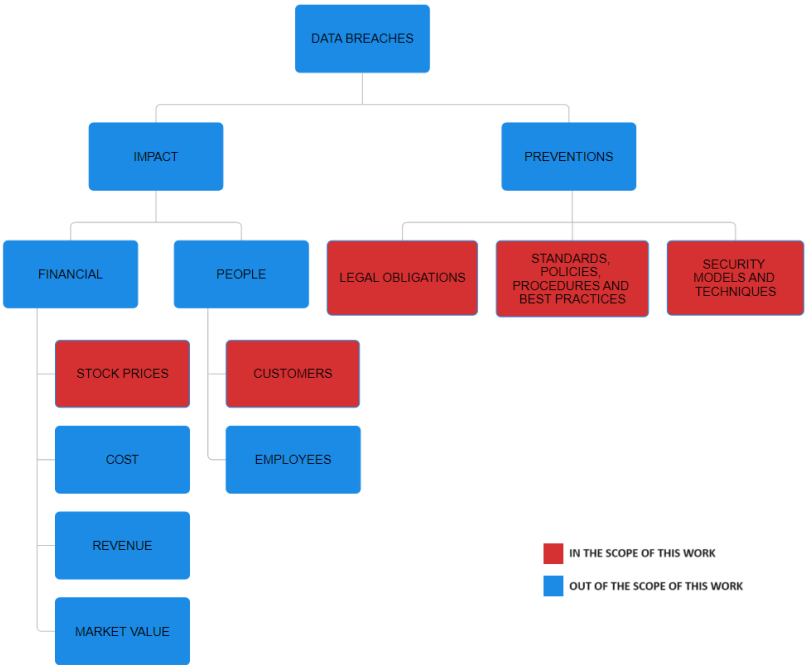



Figura 2.1: Taxonomia de violações de dados, proposta por [1]

Tabela 2.1: Contribuições de trabalhos similares.

| Referência | Conformidade | Risco | Impacto | Contramedidas |
|---|---|---|---|---|
| [29] | ✓ | | ✓ | |
| [30] | | ✓ | | |
| [31] | | | | ✓ |
| [32] | | | | ✓ |
| Este trabalho | ✓ | ✓ | ✓ | ✓ |

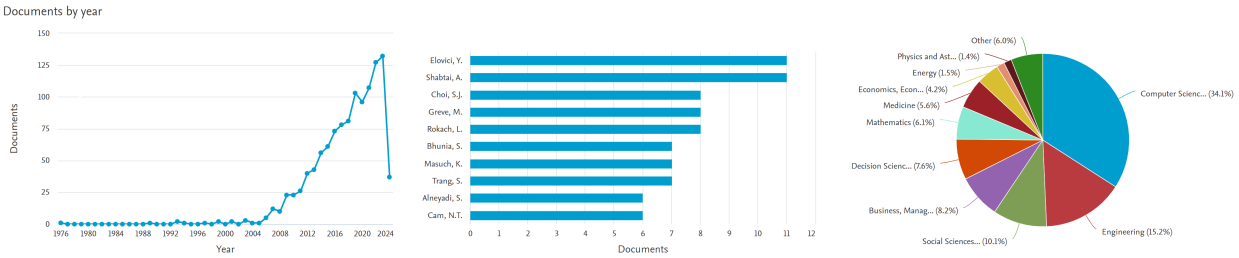
## 2.1 BIBLIOMETRIA

Para pesquisa bibliométrica acerca da temática de vazamento de dados, a base de documentos publicados da Scopus foi usada. Essa base foi selecionada devido à sua relevância no contexto acadêmico. O termo booleano da busca, que foi aplicado no título dos documentos, é apresentado na Listagem 2.1. A busca foi feita no dia 13 de abril de 2024.

Código 2.1: Termo de busca usados na Scopus

```
"data breach" OR "data exposure" OR "data leak" OR "data leakage" OR "data spill"
    OR "data spillage" OR "information breach" OR "privacy breach"
```

Essa busca resultou em 1.148 publicações. Algumas estatísticas acerca dessas publicações, fornecidas pela própria Scopus, são apresentadas na Figura 2.2. Pela quantidade de produções por ano, apresentada na Figura 2.2a, nota-se um crescimento na quantidade de documentos publicados sobre vazamento de dados a partir de 2005. Ressalta-se que a queda brusca de publicações no ano de 2024 é causada pela incompletude do ano.




(a) Publicações por ano (b) Autores mais frequentes (c) Áreas de pesquisa mais frequentes

Figura 2.2: Estatísticas da aba de análise da Scopus

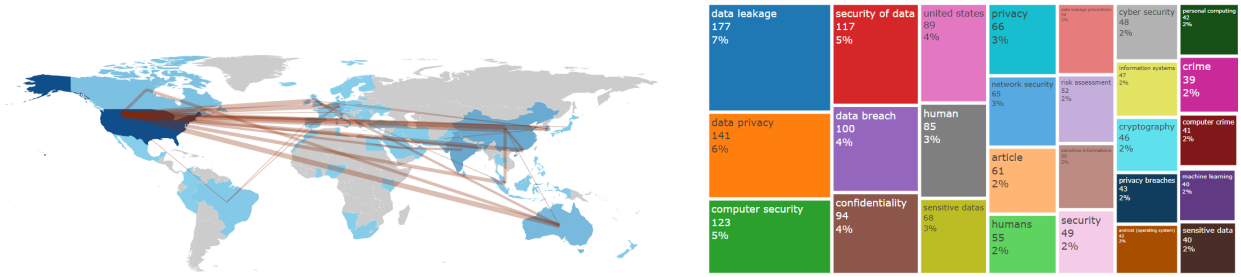
Considerando os autores que mais publicam nessa temática, conforme a Figura 2.2b, o trabalho de A. Shabtai, Y. Elovici, L. Rokach, e outros, [33] focou no uso de Prevenção de Perda de Dados (em inglês, *Data Loss Prevention* - DLP) para mitigação de vazamentos de dados. Além disso, S.J. Choi et al. [34] avaliou a ocorrência de violações de dados em hospitais dos EUA, e a relação entre a frequência desses incidentes e a qualidade dos serviços de saúde prestados. Já M. Greve et al. [35] avaliou o efeito de pedidos de desculpas e das compensações das empresas aos seus clientes após um incidente que infringiu a privacidade de seus dados.

Pela Figura 2.2c, percebe-se que o estudo desses incidentes abrange diversas áreas do conhecimento. Como exemplo de um estudo aplicado às ciências sociais, Park et al. [36] buscam entender como a violação de dados de um varejista online afeta as percepções de risco à privacidade dos consumidores. Na área de gestão de negócios e contabilidade, Rezaee et al. [37] examinam a relação entre a irresponsabilidade social corporativa e incidentes de infração de privacidade, avaliando como as empresas afetadas respondem às violações de dados.

Em um estudo da área matemática e médica, H. Sun et al. [38] propuseram um modelo multivariado para modelar a frequência de incidentes e a quantidade de registros vazados na área de saúde.

### 2.1.1 Bilbiometrix

De maneira complementar à análise promovida pela Scopus, os mesmos 1.148 documentos são submetidos a uma análise por meio do pacote Bibliometrix, disponível pra linguagem R. A Figura 2.3 mostra os gráficos gerados por meio dessa ferramenta.




(a) Trabalhos e a colaboração internacional

(b) Termos frequentes nos documentos

Figura 2.3: Estatísticas dos documentos geradas pelo Bibliometrix

Pelo mapa apresentado na Figura 2.3a, percebe-se que os países com maior quantidade de publicações no assuntos são os Estados Unidos, Índia, China e Austrália, nessa ordem. Adicionalmente, nota-se também que o Brasil já publicou em parceria com pesquisadores dos EUA, Portugal e Espanha.

A Figura 2.3b mostra os termos mais frequentes nas publicações, reforçando a relação entre a área e a segurança de computadores, especialmente quanto à confidencialidade. Revela também uma maior quantidade de publicações voltadas aos Estados Unidos, possivelmente alavancada pela maior quantidade de dados disponíveis para o país. Evidencia também, pela presença significativa do termo Android, uma preocupação de vazamentos também nos dispositivos móveis. O único controle de segurança que aparece nessa lista de termos é a criptografia.

### 2.1.2 Publicações nas contribuições deste trabalho

Para avaliar as quantidades de publicações em cada uma das áreas de contribuições deste trabalho, os 1.148 documentos resultantes da busca da Listagem 2.1 foram novamente filtrados. Os filtros adicionais são apresentados nas Listagens 2.2 até 2.4. Esses termos de busca foram aplicadas ao título do documento, ao seu resumo e às palavras chave, por meio do operador lógico OU.

Código 2.2: Filtro adicional para documentos da área de conformidade

```
(compliance OR "data protection") AND (law OR regulation OR regulamentation)
```

Código 2.3: Filtro adicional para documentos da área de análise de riscos

```
(risk OR hazard) AND (assessment OR evaluation OR analysis)
```

Código 2.4: Filtro adicional para documentos da área de mitigação

```
countermeasure OR prevention OR remediation OR "security controls" OR mitigation
    OR "protective measures"
```

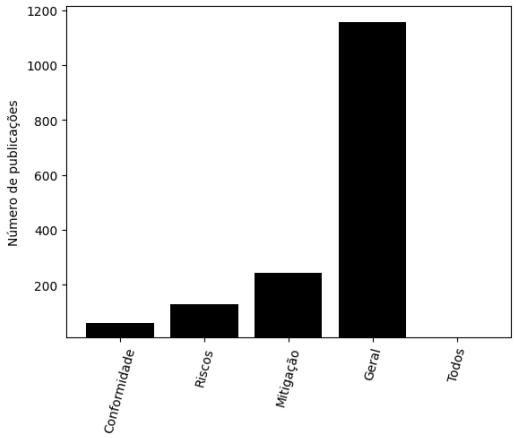



Figura 2.4: Número de publicações nas áreas de contribuição deste trabalho

Como resultado desses filtros subsequentes, foram retornados 53 documentos sobre conformidade em vazamentos de dados, 120 sobre análise de riscos e 234 sobre mitigação desses incidentes. Não foram encontradas publicações que combinassem todas as áreas concomitantemente, sugerindo uma novidade deste estudo. A Figura 2.4 mostra visualmente a quantidade de publicações nessas áreas.

## 2.2 TRABALHOS SIGNIFICATIVOS

Para identificar os trabalhos significativos entre os 1.148 documentos, são utilizadas duas métricas. A primeira, Z-score, é descrita na Equação 2.1, que é aplicada ao número de citações ($x$) de cada documento e serve como medida para avaliar a significância relativa de cada artigo em comparação com o número médio de citações de todo o conjunto de dados. A equação considera a média das citações ($\mu$) e seu desvio padrão ($\sigma$).

$$Z = \frac{x - \mu}{\sigma} \tag{2.1}$$

Em uma distribuição normal padrão, aproximadamente 99,7% dos dados estão dentro de três desvios padrão da média. Isso significa que os pontos de dados com Z-score maior que 3 são considerados no topo 0,3% da distribuição, representando eventos raros. Do total de documentos, apenas 23 têm Z-score > 3, o que corresponde a aproximadamente 2%.

Dentre esses trabalhos, Culnan e Williams [39] destacam os desafios enfrentados pelas organizações na proteção da privacidade das informações pessoais dos consumidores, argumentando que além da conformidade legal, as empresas têm uma responsabilidade ética de evitar danos e adotar medidas de precaução. Os autores também enfatizam a necessidade de programas de privacidade que promovam uma cultura de

integridade e responsabilidade gerencial.

Outro artigo que se destaca pela quantidade de citações é o de Sen e Borle [40], em que aplicam a teoria da oportunidade do crime para investigar o risco de violação de dados em decorrência da localização física de uma organização, sua indústria primária e o tipo de violação de dados que ela possa ter sofrido no passado.

No entanto, ao considerar somente a contagem de citações, as publicações mais antigas são privilegiadas. Para valorizar trabalhos mais recentes, utiliza-se a Pontuação de Significância Científica ($SSS$, do inglês, *Scientific Significance Score*), que é derivada da Equação 2.2 e adaptada de [41]. Ela SSS atribui maior peso a trabalhos mais recentes, reconhecendo sua contemporaneidade. Além disso, considera também o prestígio do periódico onde o documento foi publicado, conforme o *Scimago Journal Rank* ($SJR$). Os cinco trabalhos com maiores SSS são brevemente apresentados na Tabela 2.2.

$$SSS = \frac{\text{SJR} \times \text{Contagem\_citações}}{\text{Ano\_corrente} - \text{Ano\_publicação} + 1} \tag{2.2}$$

Tabela 2.2: Os documentos de mais elevados SSS

| Referência | SSS | Descrição |
|---|---|---|
| (42) | 36.74 | Sintetiza 43 artigos sobre os antecedentes e 83 sobre as consequências dos vazamentos de dados, destacando oito categorias para cada |
| (43) | 31.42 | Analisa postagens no Twitter relacionadas à violação de dados da Home Depot de 2014, avaliando o efeito na reputação da empresa |
| (44) | 28.90 | Investiga as consequências financeiras de vazamentos de dados, de acordo com os termos de empréstimos bancários |
| (45) | 28.77 | Examina a prática de roubo de credenciais, por meio de keyloggers e phishing, e seu impacto em milhões de usuários |
| (46) | 28.26 | Avaliou a ligação entre fatores, como a exposição da empresa, e a ocorrência de vazamentos de dados de saúde |

## 2.3 CONJUNTOS DE DADOS SIMILARES

Outros autores contribuíram com publicações de conjuntos de dados focados em vazamentos além dos dois utilizados neste trabalho [23, 22].

Como exemplo, Park [47] fornece uma base de dados acerca desse tipo de incidentes, focado nos anos entre 2012 e 2016 e no estado da Califórnia. Os dados incluem informação como ações judiciais que a empresa sofreu, a duração do monitoramento de crédito gratuito fornecido aos clientes afetados, o setor da economia em que a empresa opera, o vetor de ataque e o tamanho do vazamento.

Considerando a sensibilidade das violações de dados médicos, Ronquillo et al. [48] publicaram um conjunto de dados composto por violações de dados de saúde nos Estados Unidos, observando que, neste setor, as atividades de hacking foram responsáveis por cerca de 25% dos incidentes, mas comprometeram

quase 85% dos registros. Isso indica que as exposições de dados por hacking causam uma maior média de registros vazados por incidente do que outros vetores neste conjunto de dados.

Exposições de dados e ataques de ransomware ocorridos na Austrália entre 2004 e o início de 2020 são fornecidos por Tsen et al. [49], juntamente com informações sobre contramedidas técnicas e administrativas empregadas pelas organizações afetadas, como o uso de criptografia, políticas de segurança estabelecidas e segmentação inadequada de rede. Além disso, ainda na Austrália, pessoas foram entrevistadas com o intuito de melhor compreender as atitudes do público australiano em relação à governança de dados, incluindo o nível de preocupação com violações de dados [50].

Adicionalmente, Ikegami et al. [51] propõem um modelo probabilístico que estima o risco de uma violação de dados de uma determinada empresa. Eles fazem referência a dois conjuntos de dados associados a violações de dados no Japão de 2005 a 2018.

O *Information Commissioner's Officer* do Reino Unido publica informações sobre casos de vazamento de dados trimestralmente[1]. Da mesma forma, o site de dados abertos do governo dos Estados Unidos fornece um conjunto de dados cobrindo incidentes de violação de dados que afetaram pelo menos 500 residentes do estado de Washington[2].

Também nos Estados Unidos, a *Privacy Rights Clearinghouse* [3] compila informações acerca de vazamentos de dados registrados publicamente, com informações como o vetor de ataque utilizado e a quantidade de dado exposto no incidente. Esse conjunto de dados é usado como base para o trabalho de Rosati e Lynn (23), que o filtram para que abranja somente empresas listadas na bolsa e que é uma das bases deste estudo.

[1] <ico.org.uk/about-the-ico/our-information/complaints-and-concerns-data-sets/self-reported-personal-data-breach-cases/>, acessado em 07 de abril de 2024

[2] <catalog.data.gov/dataset/data-breach-notifications-affecting-washington-residents>, acessado em 07 de abril de 2024

[3] <https://privacyrights.org/data-breaches>, acesso em 14 de abril de 2024

# 3 METODOLOGIA

Este trabalho se fundamenta nos conjuntos de dados referentes a incidentes de vazamentos de dados que afetaram empresas listadas na bolsa dos EUA entre 2005 e 2015 [23] e a incidentes do mesmo tipo que ocorreram no mundo entre 2018 e 2019 [22].

Esses conjuntos de dados foram selecionados por serem de acesso aberto, o que promove a reproduzibilidade deste trabalho. Além disso, adotou-se os Estados Unidos como país objeto do estudo pela maior riqueza dos dados disponibilizados acerca dos incidentes nesse país.

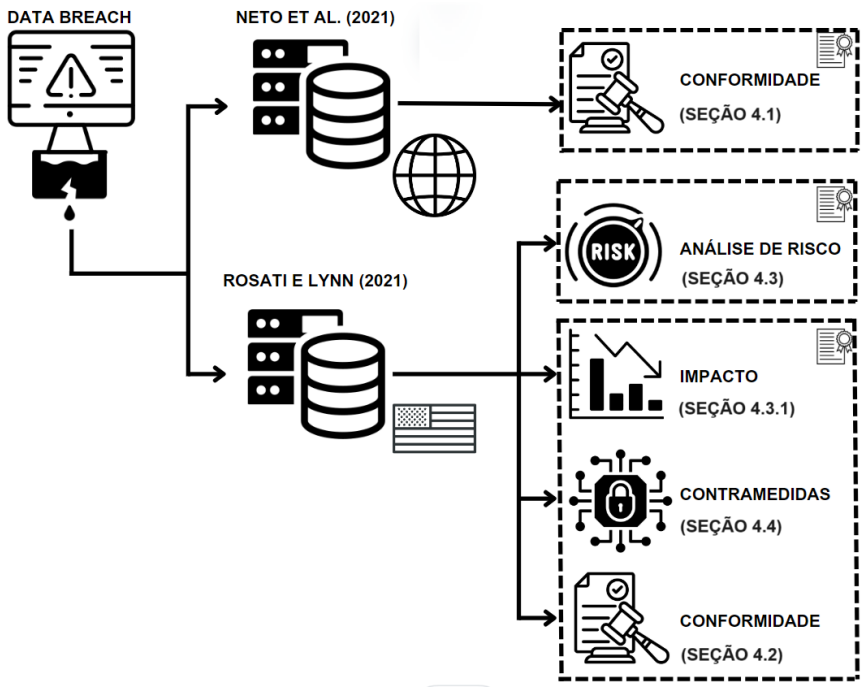



Figura 3.1: Diagrama representativo do estudo

A metodologia da análise apresentada neste trabalho é apresentada na Figura 3.1. Os conjuntos de dados são analisados com Python 3, para a geração dos gráficos apresentados e o avanço das discussões. Cada retângulo tracejado na Figura 3.1 indica um artigo científico publicado em decorrência deste trabalho.

As especificidades dos conjuntos de dados referentes aos incidentes dos EUA e globais são explicadas nas Seções 3.1 e 3.2, respectivamente.

## 3.1 CONJUNTO DE DADOS DOS EUA

Este conjunto de dados, publicado por Rosati e Lynn [23], foi originalmente obtido do repositório da Privacy Rights Clearinghouse, que contém estatísticas sobre vazamentos de dados publicamente relatadas nos Estados Unidos. Os autores posteriormente filtraram este conjunto de dados para abranger somente

eventos que impactaram empresas listadas na NYSE ou NASDAQ.

Para aprimorar a análise realizada, o conjunto de dados dos EUA foi enriquecido com informações sobre o setor econômico em que cada empresa opera, utilizando a API do Yahoo! Finance [52]. Esse enriquecimento permite uma análise mais esclarecedora dos fatores pertinentes aos incidentes. Como o conjunto de dados global já informa o setor da economia em que a instituição opera, não foi necessário utilizar a API do Yahoo! Finance.

De acordo com a PRC, as violações de dados foram obtidas primariamente dos Procuradores Gerais dos Estados Unidos e do Departamento de Saúde e Serviços Humanos, e não representam uma lista exaustiva de todos os vazamentos, refletindo apenas aquelas que foram relatadas e disponibilizadas publicamente nos Estados Unidos.

Esta base é ainda menos abrangente no caso deste trabalho, pois compreende exclusivamente empresas listadas na NYSE ou NASDAQ. No entanto, o conjunto de dados é considerado adequado para os objetivos deste estudo, que visa estimar o tamanho dos vazamentos e explorar vários aspectos de segurança relacionados a vazamentos de dados dentro de empresas listadas na bolsa estadunidense.

O conjunto de dados é apresentado em um formato tabular, em que cada coluna contém informações específicas sobre esses eventos. No total, a tabela compreende 506 linhas e 15 colunas. No entanto, para este artigo, apenas nove colunas são utilizadas e revisadas na Tabela 3.1, que revela a quantidade de informações que podem ser obtidas da fonte de dados.

Esse conjunto de dados, por informar o vetor de ataque utilizado, é adequado para fundamentar a discussão acerca de riscos e mitigação de vulnerabilidades. No entanto, por se restringir geograficamente aos Estados Unidos, não é suficiente para embasar uma comparação acerca da regulação de proteção de dados a nível global.

## 3.2 CONJUNTO DE DADOS GLOBAL

O outro conjunto de dados estudado neste artigo compreende vazamentos que ocorreram entre 2018 e 2019 no mundo, e que afetaram pelo menos 30.000 registros [22]. O conjunto de dados analisado neste estudo foi fornecido por Neto et al. [22]. De acordo com os autores, a criação do banco de dados foi baseada em fontes publicamente disponíveis de entidades governamentais, grupos de pesquisa em segurança, entidades de pesquisa e relatórios da mídia em vários idiomas. Um incidente só foi incluído no conjunto de dados se confirmado por múltiplas fontes ou se uma fonte fornecesse evidências de sua ocorrência.

Isso resultou em um conjunto de dados composto por 428 vazamentos de dados, acessível através de uma página web [1].

Para cada incidente, o conjunto de dados fornece o ano de ocorrência, a empresa afetada e seu setor, país, região geográfica, o número de registros vazados e a fonte da informação. O conjunto de dados abrange, ao todo, 37 países distribuídos pela América do Norte, América do Sul, Caribe, Europa, Ásia-Pacífico e África. Os setores das organizações incluem educação, governo e militar; médico e saúde;

[1] <databreachdb.com/>, acesso em 12 de abril de 2024

Tabela 3.1: Descrição dos campos originais no conjunto de dados dos EUA usado neste trabalho.

| Campo | Descrição | Valores possíveis |
|---|---|---|
| Event_ID | Identificador do evento | [1, 2 ,3 ,..., 506] |
| ticker | Símbolo do ticker da empresa afetada | Exemplos: AAPL, CAKE |
| event_date | Data de ocorrência do vazamento de dados | Exemplo: 21/06/2014 |
| confound_dum | Se a empresa afetada fez algum outro anúncio nos 7 dias anteriores ao anúncio da violação (53) | 0: nenhum anúncio<br>1: anúncio feito |
| confound_type | Tipo de anúncio (se houver) | Earnings: Anúncio de ganhos<br>Investigation: Investigação regulatória<br>IPO: Oferta pública inicial<br>M&A: Anúncio de fusão ou aquisição<br>Restatement: Retificação de demonstrações financeiras emitidas anteriormente<br>Statement: Divulgação de resultados financeiros trimestrais ou anuais<br>Other: Outro anúncio importante não incluído nas categorias acima |
| breach_size | Número de registros afetados pela violação (se disponível) | Exemplo: 930000 |
| breach_type | Vetor de ataque | CARD: Fraude envolvendo cartões de débito/crédito, não via hacking (por exemplo, dispositivos de skimming em terminais de ponto de venda)<br>HACK: Hackeado por uma parte externa ou infectado por malware<br>INSD: Interno (funcionário, contratado ou consumidor)<br>PHYS: Documentos em papel perdidos, descartados ou roubados<br>PORT: Dispositivo portátil perdido, descartado ou roubado<br>STAT: Perda de computador estacionário, acesso inadequado, descarte ou roubo<br>DISC: Divulgação não intencional (por exemplo, informações sensíveis divulgadas publicamente, manipuladas ou enviadas para a parte errada)<br>UNKN: Causa desconhecida |
| event_state | Estado dos EUA onde ocorreu a violação | Exemplo: Nova York |
| hq_state | Estado onde está localizada a sede da empresa afetada (pode estar fora dos EUA) | Exemplos: Texas, Tóquio |

Tabela 3.2: Descrição dos campos no conjunto de dados global usado neste trabalho.

| Campo | Descrição | Valores possíveis |
|---|---|---|
| Ano | Ano de ocorrência do vazamento de dados | [2018, 2019] |
| Empresa | Nome da instituição vitimada | Exemplos: Aadhaar, Microsoft |
| Região | Região geográfica atingida | Europa, América N., América S., Caribe, África e Ásia Pacífico |
| Setor | Área de atuação da empresa | Exemplos: Tecnologia, Saúde |
| Registros | Quantidade de registros vazados | Exemplos: 80000000, 56250000 |

negócios; entretenimento; tecnologia; e bancos, crédito e financeiro. A Tabela 3.2 sumariza esses campos.

Esse conjunto de dados não informa o vetor de ataque utilizado para infringir a confidencialidade dos dados, e, portanto, não embasa a discussão acerca de controles de segurança aplicáveis para proteger as informações. Por outro lado, ao abranger 37 países, é pertinente usar esta base para comparar a ocorrência de incidentes e as regulações de proteção de dados entre esses países e regiões.

# 4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados da análise estatística dos conjuntos de dados, em conjunto com a discussão das observações feitas. A Seção 4.1 apresenta os resultados referentes às avaliações de leis de proteção de dados, tanto globais quanto estadunidenses. A Seção 4.3 apresenta a análise qualitativa dos riscos associados a vazamentos de dados, enquanto a Seção 4.4 discute possíveis estratégias de mitigação aos riscos identificados.

## 4.1 REGULAÇÃO DE PROTEÇÃO DE DADOS GLOBAL

As regulamentações de proteção de dados estabelecem obrigações legais para as organizações no tratamento e proteção dos dados pessoais, incluindo medidas específicas para evitar vazamentos de dados e prevendo sanções no caso de descumprimento. Ainda, algumas regulamentações exigem que as organizações notifiquem autoridades reguladoras ou indivíduos afetados em caso de ocorrência desses incidentes,a fim de promover a transparência.

Considerando esta relação entre vazamentos de dados e os aspectos regulatórios de proteção de dados, esta seção busca avaliar o cenário legal dessa área em diferentes países, incluindo o Brasil.

### 4.1.1 Análise e visualização do conjunto de dados global

Esta Seção apresenta estatísticas das variáveis presentes no conjunto de dados global, disponibilizado por Neto et al. [22]. Um sumário do número de registros divulgados para esse conjunto de dados é apresentado na Tabela 4.1.

Tabela 4.1: Estatísticas descritivas do número de registros vazados no conjunto de dados global

| | Quantidade de dados vazados |
|---|---|
| Contagem | 428 |
| Média | 61.673.880 |
| Desvio padrão | 400.573.400 |
| Mínimo | 30.000 |
| 25% | 74.375 |
| 50% (mediana) | 422.548 |
| 75% | 6.000.000 |
| Máximo | 7.400.000.000 |

O incidente que resultou no maior vazamento de dados do conjunto de dados visado foi direcionado ao jornal francês Le Figaro, e expôs 7,4 bilhões de registros. Além disso, a soma de todos os registros vazados, 22 bilhões, supera a população mundial de 2019, que era de 7,7 bilhões de pessoas. Isso pode ser atribuído ao registro de dados de pessoas falecidas e à repetição do vazamento dos dados de uma mesma pessoa [22].

#### 4.1.1.1 Registros vazados por país e por região

Quando uma empresa decide adotar o modelo de computação em nuvem, por exemplo, um aspecto importante a ser considerado antes de contratar um provedor é onde os dados serão armazenados [54]. Esse modelo permite que os dados sejam armazenados fora das fronteiras do país do cliente, o que levanta desafios, especialmente relacionados à regulação e à conformidade [55].

As violações de dados em nuvem estão se tornando uma preocupação crescente [56], e, portanto, a escolha do local de armazenamento dos dados em nuvem deve levar em consideração a ocorrência desse tipo de incidente. Para fomentar a discussão sobre decisões relacionadas à nuvem e à localização dos dados, apresenta-se estatísticas sobre as ocorrências geográficas dos vazamentos de dados. A maioria dos vazamentos ocorreu nos Estados Unidos, como demonstrado na Figura 4.1, que exibe os dez países com o maior número de incidentes, desconsiderando o número de registros violados.

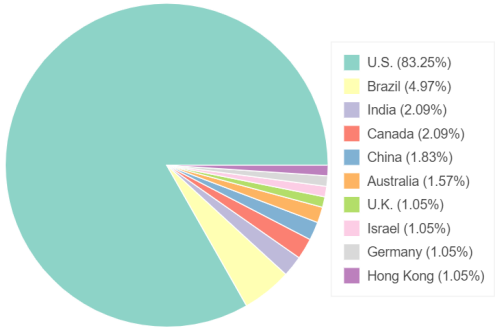



Figura 4.1: Os dez países mais frequentemente violados

Apesar de o Brasil aparecer em segundo lugar na Figura 4.1, não é um país significativo quando consideramos a magnitude dos vazamentos. A Figura 4.2 mostra o total de registros violados para cada país no conjunto de dados, evidenciando que, em termos dessa métrica, França e Estados Unidos foram os mais expressivos, seguidos por China e Índia.

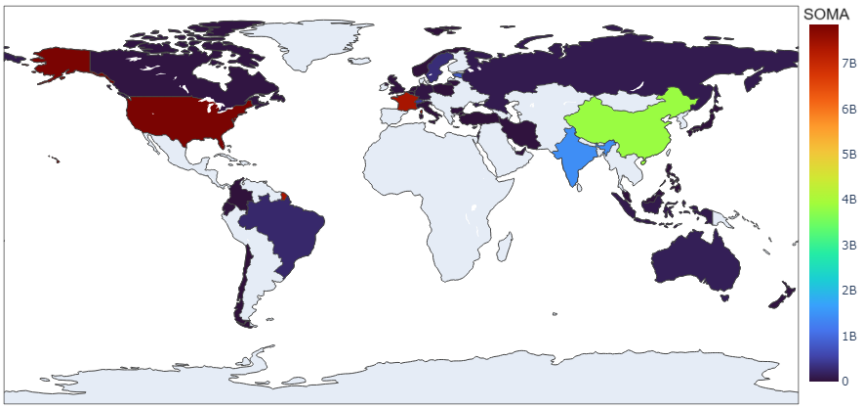



Figura 4.2: Soma de registros violados por país

O fato de a França estar entre os países com o maior número de registros vazados (Figura 4.2), enquanto não está entre os dez países mais frequentemente violados em termos de contagem de incidentes (Figura 4.1), sugere que teve menos incidentes, porém com maior volume, indicando uma maior quantidade de dados vazados por incidente.

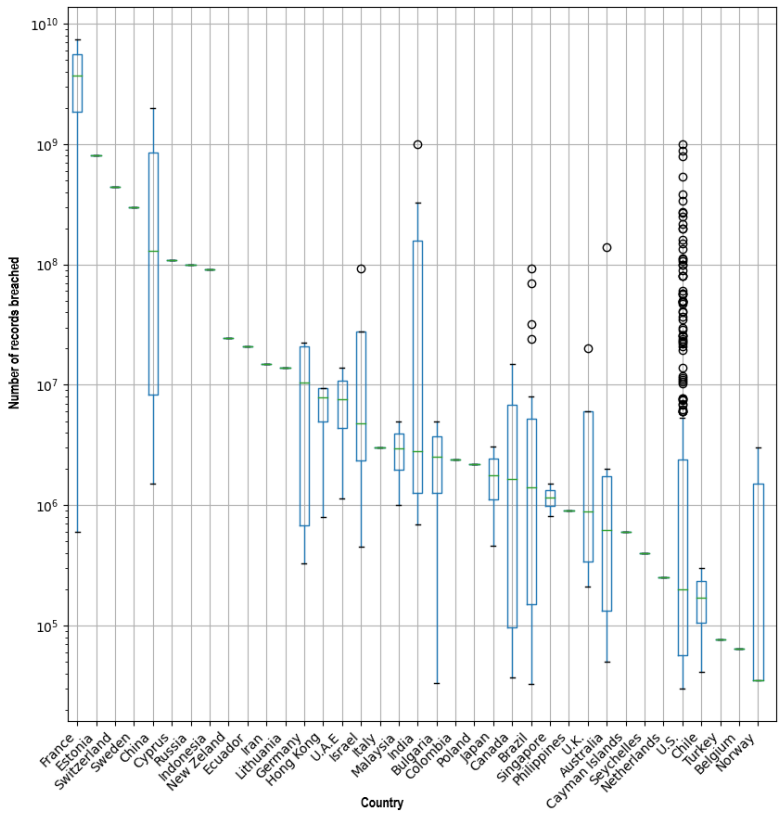



Figura 4.3: Boxplot de registros vazados por país

Esta conclusão é confirmada pela Figura 4.3, que indica que a França teve a maior mediana no tamanho dos registros violados por incidente entre todos os países no conjunto de dados. Países como Estônia, Suíça e Suécia, embora tenham violado um número relativamente alto de registros, foram violados apenas uma vez. Por outro lado, a mediana do número de registros violados nos Estados Unidos é menor do que a da Austrália, por exemplo, que não apresentou uma soma relevante de registros violados, como observado na Figura 4.2. Isso se deve ao fato de que os EUA relataram um grande número de incidentes, contribuindo para uma maior soma de registros vazados.

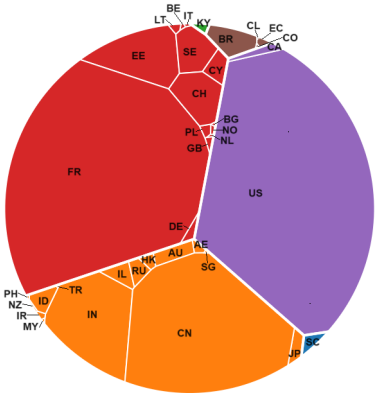



Figura 4.4: Proporção da soma de registros vazados por país em cada região

Para uma melhor visualização das proporções da soma de registros vazados para cada país em sua região, apresenta-se a Figura 4.4, que indica que a região europeia teve o maior número de registros vazados,

liderada pela França.

Da mesma forma, a América do Norte, a segunda região mais frequentemente violada, é liderada pelos Estados Unidos, com poucas ocorrências no Canadá. Na região Ásia-Pacífico, China e Índia representam os países com mais registros violados, com alguns outros países contribuindo com quantidades menores. América do Sul e África compreendem dados expostos majoritariamente no Brasil e nas Seychelles, respectivamente, enquanto a região do Caribe é representada exclusivamente pelas Ilhas Cayman.

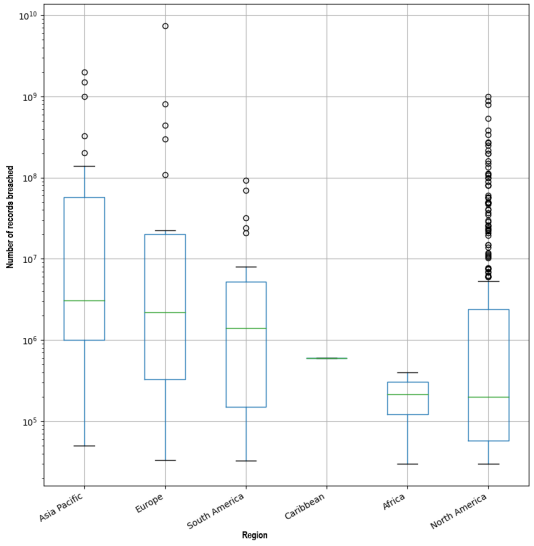



Figura 4.5: Boxplot de registros vazados por região

A região Ásia-Pacífico, apesar de ser apenas a terceira em soma de informações violadas, é a região com a maior mediana no tamanho das violações, como observado na Figura 4.5. A América do Norte foi a região com a menor mediana, principalmente devido às estatísticas dos Estados Unidos.

#### 4.1.1.2 Registros Vazados por Setor

Diferentes tipos de dados divulgados impactam de maneiras distintas a vida dos indivíduos afetados. Por exemplo, a divulgação de dados de saúde pode afetar informações médicas confidenciais [57], e vazamentos militares possivelmente expõem dados sensíveis das forças armadas de um país [58].

Além disso, a natureza dos dados armazenados determina também os aspectos regulatórios que a empresa deve cumprir [59]. Os aspectos legais das violações de dados são discutidos de forma mais aprofundada na Seção 4.1.2.

Conforme mostrado na Figura 4.6, o setor de tecnologia apresenta a maior mediana no número de registros vazados por incidente, seguido pelo governo/militar. No entanto, a violação mais volumosa no conjunto de dados ocorreu no setor empresarial. O setor de educação teve a menor mediana. No entanto, compreender as distribuições geográficas e temporais desses registros vazados pode revelar tendências e padrões na ocorrência desse tipo de incidente de segurança. A Figura 4.7 ilustra a soma de registros vazados por região para cada setor em 2018 e 2019.

A Figura mostra que a maioria dos incidentes ocorreu em 2019. Conforme afirmado por [22], isso

pode ser atribuído à entrada em vigor do GDPR em 2018 e às empresas estarem em conformidade com os requisitos de notificação de violação ao longo de 2019. Uma exceção a essa observação é o setor de saúde, com prevalência de registros violados em 2018. No entanto, o setor com o maior volume de informações vazadas em 2018 foi o de tecnologia, superado pelo setor empresarial em 2019.

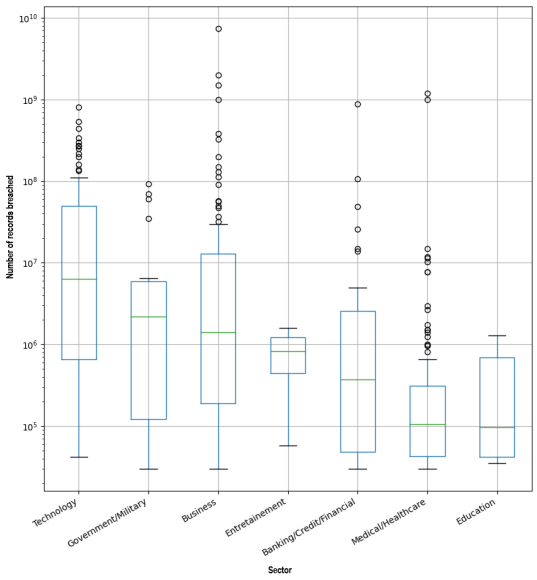



Figura 4.6: Boxplot de registros vazados por setor

Além disso, o setor empresarial foi o que teve mais registros violados quando consideramos ambos os anos, apesar de ter a terceira maior mediana de registros violados por incidente, indicando que as organizações nesse ramo sofreram mais vazamentos que vazaram menos registros do que a tecnologia e o governo, por exemplo.

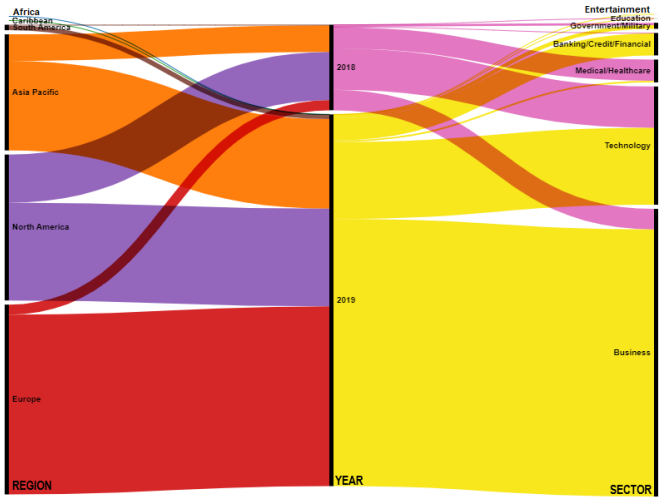



Figura 4.7: Soma de registros vazados por região por setor

É relevante, no entanto, interpretar o setor mais visado em cada país, pois pode evidenciar fraquezas técnicas em diferentes áreas para diferentes regiões e enfatizar a necessidade de regulamentação mais rigorosa e emprego de controles de segurança mais eficazes.

Com o intuito de viabilizar essa análise, a Figura 4.8 apresenta a relação entre os setores econômicos

nos quais as empresas violadas operam e os países onde o incidente ocorreu. Para esta análise, consideramos exclusivamente os dez países mais frequentemente violados (Figura 4.1), adicionando a França, pois representa uma parte significativa dos tamanhos de violação (conforme a Figura 4.4).

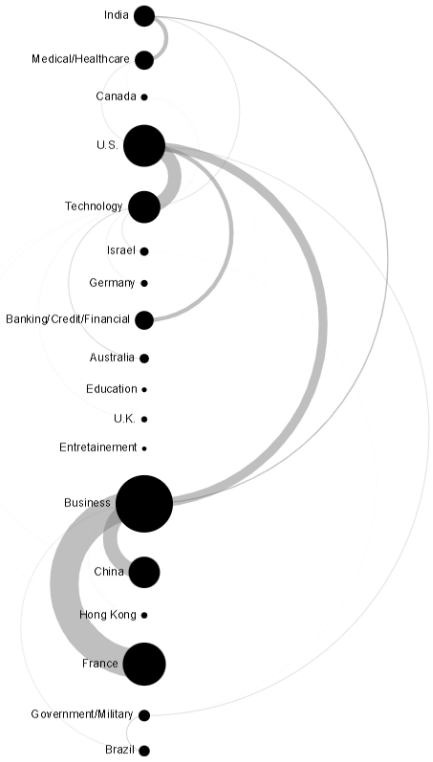



Figura 4.8: Setores mais explorados nos 10 países mais alvejados e a França

A partir da Figura 4.8, observa-se, por exemplo, que a França e a China foram predominantemente violadas no setor empresarial, que também compreende uma quantidade significativa de incidentes nos Estados Unidos. No entanto, as maiores violações nos EUA afetaram o setor de tecnologia, o que também pode ser uma consequência do grande número de empresas de tecnologia com base no país. Também observa-se que o setor bancário, de crédito e financeiro foi principalmente violado nos Estados Unidos, o que demonstra a importância da Lei Gramm-Leach-Bliley. Esse padrão demonstra uma mudança de tendência, pois de 2010 a 2017, os EUA registraram principalmente vazamentos de saúde [60].

Além disso, os incidentes de saúde afetaram principalmente organizações indianas. Churi et al. [61] reconhecem alguns problemas de privacidade relacionados ao setor de saúde na Índia, como a falta de tecnologia e infraestrutura, o que pode explicar a relação observada na Figura 4.8. Os autores relatam também ausência de confiança na relação entre médico e paciente, dados médicos armazenados na nuvem com preocupações com privacidade, controles de segurança fracos implementados, dados compartilhados sem consentimento do indivíduo, inadequação de políticas de segurança e aspectos culturais. Isso pode impactar a confiança e aceitação das tecnologias de saúde pelos cidadãos deste país [62].

Governo e militar divulgaram predominantemente registros do Brasil e dos Estados Unidos. No Brasil, [63] indicou uma falta de conformidade com a LGPD e imaturidade na área, com muitas organizações que ainda não haviam estabelecido um Oficial de Proteção de Dados (em inglês *Data Protection Officer*, DPO). Especificamente para o setor militar e de defesa, o Brasil tem demonstrado falta de atenção na formulação

de políticas de segurança nacional [64], o que pode ter efeitos nos incidentes de segurança cibernética neste setor.

As violações de dados do governo também podem surgir como consequência de conflitos geopolíticos. Como exemplos, [65] menciona quatro casos que afetaram pessoas políticas e figuras públicas nos EUA em operações de hacking e vazamento.

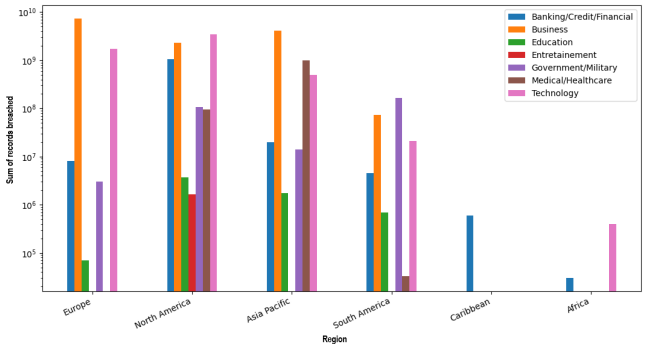



Figura 4.9: Distribuição da soma de registros expostos por setor e por região

Para um raciocínio mais abrangente globalmente dessa relação, a Figura 4.9 apresenta o número de registros violados por setor em cada região. O setor empresarial sofreu vazamentos significativas na Europa, América do Norte, Ásia-Pacífico e América do Sul, e a tecnologia também foi relevante na Europa e na América do Norte. No Ásia-Pacífico, incidentes médicos e de saúde também correspondem a porções significativas, principalmente devido aos incidentes na Índia, como visto na Figura 4.8. Governo e militar é o segundo setor principal afetado na América do Sul, com uma grande contribuição do Brasil.

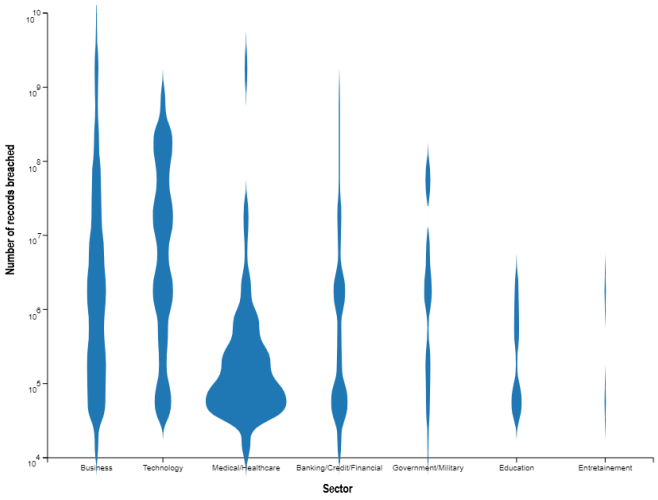



Figura 4.10: Distribuição do tamanho de vazamentos por setor

A única violação que ocorreu no Caribe afetou o Cayman National Bank, enquanto os países africanos sofreram vazamentos nas áreas financeira e tecnológica. Além disso, violações de entretenimento foram relatadas apenas na América do Norte, relacionadas aos incidentes envolvendo AMC Networks e MoviePass. A Europa, apesar de ser a região com a maioria dos registros de incidentes, não registrou publicamente nenhum incidente de vazamento de dados no setor de saúde e também foi a região com menos registros de educação divulgados entre as localidades que foram afetadas nesta área.

Os tamanhos das violações de dados nos setores industriais de tecnologia e negócios estão mais uniformemente distribuídos, especialmente na faixa entre $10^5$ e $10^9$ registros violados por incidente, como visto na Figura 4.10. As violações de dados médicos, diferentemente, estão mais concentradas em torno de $10^5$, com poucas ocorrências de $10^9$ registros vazados e nenhuma em torno de $10^8$.

### 4.1.2 Regulação de proteção de dados no mundo

Dada a sensibilidade dos dados pessoais, sua coleta, processamento, armazenamento, uso e destruição devem ser rigorosamente planejados. Para garantir esse planejamento, as leis de proteção de dados regulamentam a privacidade dos titulares de dados [66].

Uma legislação de proteção de dados visa ajudar na mitigação de vazamentos de dados por meio de mecanismos de segurança preventivos. Além disso, visa reduzir o impacto causado por vazamentos por meio de notificações oportunas às autoridades reguladoras, permitindo que ações apropriadas sejam tomadas [67].

Nesse último cenário, estruturas regulatórias auxiliam na mitigação de riscos tanto para empresas privadas quanto para o setor público. As regulamentações geralmente são fundamentadas em princípios como prevenção, responsabilidade e transparência, fornecendo medidas para evitar vazamentos e meios para identificar e responsabilizar as partes responsáveis. Assim, uma regulação adequada promove a mitigação riscos, mesmo que não seja robusta e falte disposições explícitas para proteção de dados, processamento ou notificações de violações.

Por outro lado, mesmo a legislação com disposições extensas, se não aplicada adequadamente, não será tão eficaz. A fiscalização e o papel da autoridade reguladora, com a conscientização, o monitoramento e a aplicação de penalidades, são fundamentais para atingir eficácia com as normas. Cada cenário envolve uma interpretação e adaptação distintas aos tipos de vazamentos ocorridos, tornando certas ferramentas mais adequadas para determinados países.

Para uma empresa localizada em uma região altamente regulamentada, a demonstração de conformidade é de importância fundamental [68]. Após uma violação de dados, além dos custos mencionados anteriormente, uma empresa pode incorrer em sanções se não conseguir demonstrar conformidade com a regulamentação pertinente. Por exemplo, 20% das empresas violadas pagaram pelo menos 250.000 dólares em multas [5].

### 4.1.3 Comparação dos níveis de regulamentação nos países afetados

Para identificar as regiões que são pouco e altamente regulamentadas quanto à proteção de dados e privacidade, comparamos brevemente aspectos da legislação pertinente de todos os países presentes nos dados. A Tabela 4.2 resume essa comparação, em que as linhas são classificadas em ordem decrescente pela soma dos registros vazados. A coluna 'Regulação e fiscalização' apresenta o nível de regulamentação do país, de acordo com [2], em que (++) representa uma regulamentação pesada, (+) uma regulamentação robusta, (-) moderada e (- -) limitada.

Uma possível análise para revelar se a legislação é eficaz na mitigação de incidentes é uma comparação

da frequência de vazamentos de dados antes e depois da aplicação da lei. Ou mesmo examinando como a autoridade de proteção de dados age em casos de vazamentos, a existência de canais de notificação fáceis, avaliando o impacto que esses vazamentos tiveram sobre os indivíduos e se as pessoas afetadas foram informadas, cumprindo os deveres de transparência e responsabilidade. No entanto, o conjunto de dados estudado neste trabalho é muito curto no tempo e não fornece informações suficientes para conduzir essas análises.

Portanto, nossa abordagem para avaliar a eficácia dessas regulamentações é comparar elementos-chave do cenário de normas de proteção de dados estabelecidas em países de diferentes níveis de regulamentação e fiscalização, ou seja, pesada, robusta, moderada e limitada.

O Regulamento Geral de Proteção de Dados, em vigor na Europa e considerado uma regulamentação pesada (++), é considerado um padrão global, devido, por exemplo, a seus direitos aprimorados para os indivíduos, requisitos de consentimento mais rigorosos, penalidades substanciais e requisito de demonstração de conformidade.

Ele estabelece um prazo máximo de 72 horas para notificação e outros mecanismos para relatar um vazamento de dados pessoais à autoridade supervisora, também exigindo a comunicação de uma violação de dados pessoais ao titular dos dados, conforme os Artigos 33 e 34. O Considerando 85 do GDPR também demanda a adoção de medidas de mitigação de riscos para danos físicos, materiais ou imateriais aos indivíduos [2]. Exemplos desses danos incluem a perda de controle sobre seus dados pessoais, limitação de seus direitos, discriminação, roubo ou usurpação de identidade, perdas financeiras, reversão não autorizada de pseudonimização, danos à reputação, perda de confidencialidade de dados pessoais protegidos por segredo profissional, ou qualquer outro desvantagem econômica ou social significativa para os indivíduos. O Artigo 4, subparágrafo 12, do GDPR define uma violação de dados como uma violação de segurança que leve à destruição, perda, alteração, divulgação ou acesso acidental ou ilícito de dados pessoais transmitidos, armazenados ou processados de outra forma. Portanto, segundo a lei da união Europeia, um vazamento constituiria uma violação por definição.

A Lei Japonesa foi selecionada como exemplo de regulamentação robusta (+). A Lei de Proteção de Informações Pessoais (APPI) foi inicialmente promulgada em 2003, mas foi alterada em 2017. Em 2020, foi aprovado um projeto de lei para alterar ainda mais a APPI, que entrou em vigor em 2022, integrando os setores público e privado, anteriormente separados [2]. A autoridade estabelecida é a Comissão de Proteção de Informações Pessoais (PPC). Sob a APPI alterada, operadores comerciais devem relatar incidentes de violação de dados à Comissão de Proteção de Informações Pessoais e aos titulares de dados afetados, se os dados do incidente de violação puderem prejudicar os direitos e interesses individuais. A PPC estabeleceu um limite concreto para obrigações de comunicação, onde o operador comercial precisa relatar à PPC e notificar os indivíduos afetados.

Os elementos definidores de um vazamento no Japão são: (i) informações pessoais sensíveis vazadas ou provavelmente vazadas; (ii) informações pessoais que causam danos financeiros devido ao uso não autorizado vazadas ou provavelmente vazadas; (iii) incidente de vazamento de dados devido a intenção maliciosa ocorreu ou provavelmente ocorreu; e (iv) incidente de vazamento de dados envolvendo mais de 1.000 titulares de dados ocorreu ou provavelmente ocorreu [2]. Além disso, as diretrizes da PPC sugerem que os operadores comerciais conduzam investigações necessárias, adotem medidas preventivas e

Tabela 4.2: Sumarização do cenário de proteção de dados pessoais nos países afetados por vazamento de dados [2], ordenado de maneira decrescente pela soma de registros vazados dividido pelo tamanho da população do país em 2019

| País | Regulação e fiscalização | Lei de proteção de dados | Ano de aprovação | Define dado pessoal | DPA | Requer registro | Requer DPO | Notificação de vazamento |
|---|---|---|---|---|---|---|---|---|
| CH | ++ | FADP | 2020 | ✓ | FDPIC | ✗ | ✗ | ✓ |
| EU | ++ | GDPR | 2016 | ✓ | EDPB | ✗ | ✓ | 72 horas |
| US | ++ | ✗ | - | ✓ | ✗ | ✗ | ✗ | ✓ |
| IL | ++ | PPL | 1981 | ✓ | PPA | ✓ | ✓ | imediatamente |
| KY | - | DPA | 2017 | ✓ | Ombudsman | ✗ | ✗ | 5 dias |
| AU | ++ | PA & APP | 1988 | ✓ | OAIC | ✗ | ✗ | 72 horas |
| NZ | + | PA | 2020 | ✓ | Privacy Commissioner | ✗ | ✓ | ✓ |
| SC | ++ | DPA | 2003 | ✓ | ✗ | ✓ | ✗ | ✗ |
| HK | ++ | PDPO | 1995 | ✓ | PCPD | ✗ | ✗ | ✗ |
| CN | ++ | PIPL | 2021 | ✓ | CAC | ✗ | ✓ | ✓ |
| AE | - | PDPL | 2021 | ✓ | ✗ | ✗ | ✓ | imediatamente |
| EC | - | PDPL | 2021 | ✓ | ✗ | ✓ | ✗ | 5 dias |
| BR | - | LGPD | 2018 | ✓ | ANPD | ✗ | ✓ | 2 dias úteis |
| IN | - - | DPDP | 2023 | ✓ | ✗ | ✗ | ✓ | ✓ |
| CA | ++ | PIPEDA | 2000 | ✓ | OPC | ✗ | ✓ | ✓ |
| RU | - | DPA | 2006 | ✓ | Roskomnadzor | ✓ | ✓ | 24 horas |
| SG | ++ | PDPA | 2012 | ✓ | PDPC | ✗ | ✓ | 3 dias corridos |
| ID | + | PDP | 2022 | ✓ | PDP Agency | ✗ | ✗ | 72 horas |
| GB | ++ | UKGDPR | 2018 | ✓ | ICO | ✓ | ✓ | 72 horas |
| MY | + | PDPA | 2010 | ✓ | PDPC | ✓ | ✗ | ✗ |
| IR | - - | ✗ | - | ✗ | ✗ | ✗ | ✗ | ✗ |
| CO | - | Law 1581 | 2012 | ✓ | SIC & SOF | ✓ | ✗ | 15 dias úteis |
| JP | + | APPI | 2003 | ✓ | PPC | ✗ | ✗ | ✓ |
| CL | - | PDPL | 1999 | ✓ | ✗ | ✗ | ✗ | ✗ |
| PH | - | DPA | 2012 | ✓ | NPC | ✓ | ✓ | 72 horas |
| TR | - | LPPD | 2016 | ✓ | KVKK | ✓ | ✗ | 72 horas |

divulguem a natureza do vazamento e as ações corretivas tomadas, se apropriado e necessário.

A Lei Geral de Proteção de Dados do Brasil (LGPD), considerada uma regulamentação moderada (-), lida com responsabilidade nos Artigos 31 e 32 e também com danos e reparação nos Artigos 42 a 45. A lei aborda a reparação pelo controlador ou processador que, devido ao processamento de dados pessoais, causa danos patrimoniais, morais, individuais ou coletivos em violação à legislação [2]. Assim, o não cumprimento das boas práticas, segurança e prevenção previstas pela lei exigiria reparação de danos em caso de vazamento de dados. O Artigo 48 aborda especificamente a notificação pelo controlador de incidentes de segurança à autoridade nacional (ANPD), que pode acarretar riscos ou danos relevantes aos titulares de dados [2]. A disposição esboça os requisitos mínimos: (i) descrição da natureza dos dados pessoais afetados; (ii) informações sobre os titulares de dados envolvidos; (iii) indicação das medidas técnicas e de segurança utilizadas para proteger os dados, observando segredos comerciais e industriais; (iv) riscos relacionados ao incidente; (v) razões para atraso se a comunicação não foi imediata; e (vi) medidas tomadas ou a serem tomadas para reverter ou mitigar os efeitos do dano. Não há prazo legal para notificações de incidentes à ANPD.

No entanto, a autoridade publicou diretrizes em 2021 afirmando que a comunicação deve ser feita dentro de dois dias úteis a partir da data de conhecimento do incidente. O site institucional [1] contém instruções para essa notificação, com diretrizes atualizadas sobre vazamentos. No caso de risco ou dano significativo aos titulares de dados, os indivíduos também podem precisar ser notificados. A notificação pode ser en-

[1] <www.gov.br/anpd/pt-br/canais_atendimento/agente-de-tratamento/comunicado-de-incidente-de-seguranca-cis>, acesso em 14 de abril de 2024

viada pelo Oficial de Proteção de Dados ou pelo representante legal, juntamente com a documentação ou autorização correspondente.

Uma recomendação adicional, não legalmente exigida, é a implementação de cláusulas contratuais estabelecendo obrigações de notificação entre controladores e processadores para agilizar a avaliação e minimizar os riscos aos titulares de dados. Embora não seja necessário fornecer uma lista de titulares de dados afetados, a ANPD pode solicitar que o controlador de dados apresente uma cópia da notificação aos titulares de dados sobre a violação [2]. Essa notificação ao titular dos dados deve ser feita individualmente, sempre que possível, e pode ser realizada por qualquer meio, como e-mail, carta ou mensagem eletrônica.

A Lei Indiana de Proteção de Dados Pessoais em 2019 (PDP), ainda passando por atualizações em 2022 e considerada uma regulamentação limitada (–), foi promulgada após a Suprema Corte Indiana em 2017 reconhecer a privacidade como um direito fundamental, consagrado no Artigo 21 da Constituição [2]. Ainda não há uma autoridade de proteção de dados estabelecida, o que pode comprometer a aplicabilidade da lei na construção e mitigação de vazamentos de dados.

Como observado, tanto a regulamentação europeia quanto a japonesa abrangem medidas de proteção de dados relacionadas a prevenção, auditoria e notificação. Elas também garantem a conformidade e a supervisão dos requisitos legais por meio de autoridades nacionais de proteção de dados. Em contraste, o Brasil, apesar de modelar sua lei nacional com base no GDPR e antecipar disposições semelhantes, permanece em um cenário regulatório relativamente imaturo. Isso é exemplificado pela subordinação aumentada e pela autonomia limitada na aplicação da lei [69]. Somente em 2022 a ANPD ganhou mais autonomia ao ser vinculada ao Ministério da Justiça e Segurança Pública do Brasil, embora sem subordinação.

Por outro lado, a lei de proteção de dados da Índia foi promulgada em 2023 e ainda não está totalmente implementada [70]. Além disso, a ausência de uma autoridade de proteção de dados destaca sua significativa imaturidade. Consequentemente, apesar da existência de leis e requisitos de segurança, a falta de mecanismos eficazes de aplicação, monitoramento e penalização dificulta a aplicação eficaz da proteção de dados. Essas questões de maturidade estão entre as possíveis causas da classificação das regulamentações de proteção de dados do Brasil e da Índia em níveis mais baixos quando comparadas às de países europeus e japoneses, por exemplo.

Os Estados Unidos, apesar de também não possuírem uma lei de proteção de dados em nível federal promulgada nem uma autoridade de proteção de dados, são considerados um país de regulamentação pesada, o que pode ser devido às razões declaradas na Seção 4.1.3. A regulação dos EUA é discutida com mais detalhes na Seção 4.2

Essas diferenças nos níveis de regulamentação em diferentes regiões também podem surgir de diferenças culturais. Como exemplo, observou-se que as pessoas da América do Norte estão mais dispostas a abrir mão da privacidade do que as da Europa, que também estão mais preocupadas com vazamentos de dados e transparência sobre como os dados são usados [71].

Uma comparação dos níveis de maturidade entre as regiões geográficas, expressa pelos níveis de regulamentação e aplicação, então, é apresentada na Figura 4.11. Observa-se que todos os países da Europa, América do Norte e África aplicam regulamentações pesadas, enquanto todos os países da América do Sul e do Caribe são moderadamente regulamentados. A Ásia-Pacífico é a região mais diversificada, já que

também compreende um maior número de países no conjunto de dados.

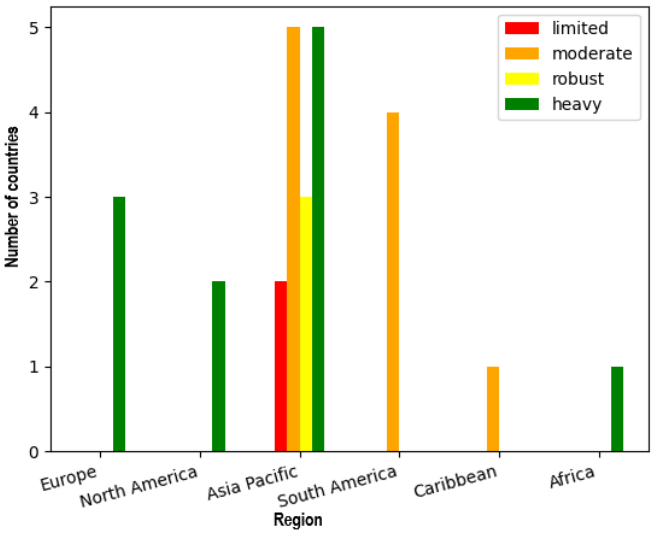



Figura 4.11: Nível de rigor da regulamentação e da fiscalização por país por região [2]

Também é observado que a Ásia-Pacífico, a única região que possui um país com regulamentação limitada no conjunto de dados, o Irã, é também a região com a maior mediana de registros violados por incidente, de acordo com a Figura 4.5, com um valor próximo à mediana na Europa, uma região fortemente regulamentada. A América do Norte, que também é fortemente regulamentada, apresenta o menor valor de mediana. Ao analisar os valores medianos nos países, da Figura 4.3, observa-se que os países mais bem classificados são fortemente regulamentados.

A partir disso, pode-se inferir que os níveis de regulamentação e o tamanho das violações não estão fortemente correlacionados. Na verdade, a quantidade de dados vazados está mais fortemente relacionada à quantidade de dados armazenados pela organização. A natureza dos dados armazenados também pode influenciar o tamanho da violação, já que as organizações podem aplicar mais esforços para proteger dados mais sensíveis, que poderiam ser mais severamente penalizados pelas leis de proteção de dados.

Como exemplo, dados relacionados a serviços médicos apresentaram os segundos menores valores de mediana, como visto na Figura 4.6. Após uma análise da Tabela 4.2, que está ordenada pela soma dos tamanhos das violações dividida pelos tamanhos das populações em 2019, observa-se que os países com maiores níveis de regulamentação e aplicação estão mais concentrados nas linhas superiores, o que sugere que esses países violam um número maior de registros por habitante do que aqueles de níveis mais baixos. Isso pode ser devido a vários fatores.

Um deles é que países desenvolvidos, como Estados Unidos e aqueles da Europa, são mais propensos a oferecer mais serviços relacionados a dados, como armazenamento em nuvem, mesmo para não residentes, o que aumenta as chances de vazamentos. Também pode ser uma consequência direta de uma maior aplicação da notificação de violações. No entanto, ainda evidencia a necessidade de controles de segurança mais eficazes para mitigar essas ocorrências.

Legislações de proteção de dados são regulamentações projetadas para proteger a privacidade e segurança dos dados individuais. Embora essas leis difiram em alguns aspectos para diferentes países, os princípios gerais e a base legal incluem consentimento, limitação de finalidade, integridade, confidenciali-

dade, responsabilidade e transparência no manuseio de dados.

Ainda que a maioria dos países dentro do escopo deste estudo tenha promulgado uma lei de proteção de dados, os Estados Unidos e o Irã, no momento desta escrita, não o fizeram. Esses países, no entanto, têm um projeto de lei pendente ou iniciativa de lei. Cabe ressaltar que, enquanto o Irã é classificado como tendo regulamentação e aplicação limitadas em relação à proteção de dados e privacidade, os Estados Unidos são categorizados como um país fortemente regulamentado, possivelmente devido à existência das leis estaduais e setoriais do país, que são discutidas na Seção 4.2.

Adicionalmente, alguns Estados Membros da União Europeia tinham uma lei nacional de proteção de dados em vigor antes da promulgação da GDPR. Algumas dessas ainda coexistem com a regulação europeia, como a francesa *Loi Informatique et Libertés* e a Alemã *Bundesdatenschutzgesetz* [72]. No entanto, esses Estados Membros seguem o GDPR e, com o objetivo de simplificar a análise, esses países foram grupados por região na Tabela 4.2.

Observa-se uma influência do GDPR nas regulamentações de proteção de dados em regiões fora da Europa, como consequência do Efeito Bruxelas [73].

Um cronograma dos anos em que as leis de proteção de dados foram promulgadas é apresentado na Figura 4.12. Os anos de promulgação das leis nos países presentes na Tabela 4.2 seguem [2], enquanto para os países que constituem a União Europeia, as datas são consonantes com o Conselho da Europa [2].

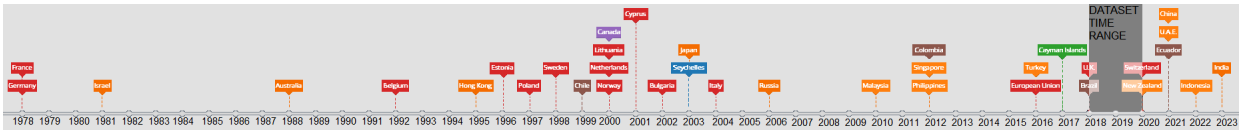



Figura 4.12: Cronograma de promulgação das leis de proteção de dados

Para a devida salvaguarda legal de dados, e para que os indivíduos compreendam quais dados estão sendo protegidos, uma definição clara de dados pessoais e sensíveis é necessária. Essa definição também promove a transparência e a fiscalização e aplicação da lei [74].

Como exemplo, o GDPR define dados pessoais como qualquer informação relacionada a uma pessoa natural identificada ou identificável, como nome, data de nascimento, endereço de e-mail e endereço de cobrança [2]. O LGPD do Brasil não classifica dados anonimizados como informações pessoais, a menos que possam ser revertidos mediante esforços razoáveis [2].Dados sensíveis também podem ser definidos. O APPI do Japão, por exemplo, define como qualquer informação que possa causar discriminação à pessoa, como raça, histórico médico e registro criminal [2]. Como visto na Tabela 4.2, o Irã é o único país que não define informações pessoais ou sensíveis.

Uma Autoridade de Proteção de Dados (em inglês, *Data Protection Authority*, DPA) é uma entidade pública responsável por supervisionar e aplicar as regulamentações de proteção de dados e também pode fornecer diretrizes e aumentar a conscientização sobre proteção de dados [75]. Na Europa, o Comitê Europeu para a Proteção de Dados (em inglês, *European Data Protection Board*, EDPB) padroniza a proteção de dados em cada Estado Membro, que também institui DPAs federais [76].

Algumas leis podem exigir que qualquer controlador de dados que pretenda processar pessoal notifique

[2]<coe.int/en/web/data-protection/>, acesso em 12 de abril de 2024

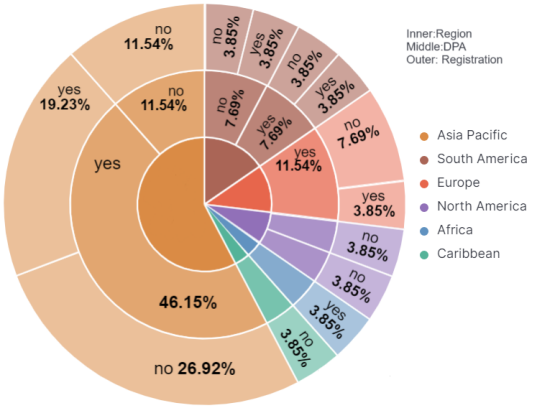



Figura 4.13: Distribuição de países que possuem uma DPA e exigem registro

a autoridade pública competente. Na Tabela 4.2, essa informação é apresentada na coluna 'exige registro'. A Rússia, por exemplo, exige que o registro mencione, por exemplo, o nome completo e o endereço do controlador de dados, o propósito do processamento, as categorias de dados em processamento, medidas de proteção implantadas e a ocorrência de transferência transfronteiriça de dados pessoais [2].

A distribuição de países que possuem uma DPA e exigem o registro de um controlador de dados entre as regiões é mostrada na Figura 4.13. Países que aparecem como 'não' na exigência da DPA, mas 'sim' no registro, têm uma previsão legal para estabelecer uma Autoridade de Proteção de Dados e exigir registro, mas ainda não a constituíram.

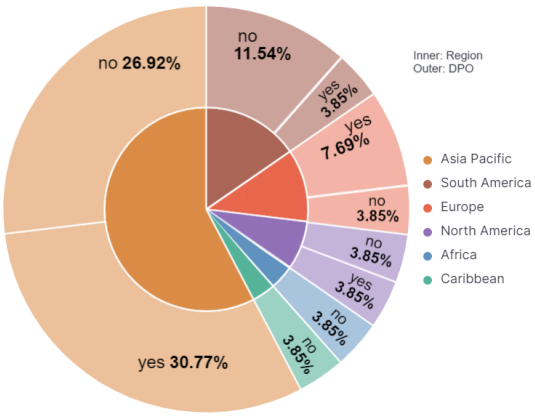



Figura 4.14: Distribuição de países que exigem um DPO

Por outro lado, as empresas podem ser obrigadas a indicar um ponto de contato para assuntos relacionados à proteção de dados. Essa função é identificada como o Oficial de Proteção de Dados, e algumas de suas atribuições incluem monitoramento de conformidade e aconselhamento ao empregador [77]. No Canadá, essa posição é ocupada por padrão pela maior autoridade dentro da organização, e suas responsabilidades também incluem responder e relatar vazamentos de segurança [2]. A distribuição de países que exigem um DPO nas empresas é apresentada na Figura 4.14.

A definição de vazamento de dados pode variar de acordo com as leis. Como exemplo, a Nova Zelândia a define como qualquer acesso não autorizado ou acidental, ou divulgação, alteração, perda ou destruição de informações pessoais, ou qualquer ação que impeça a agência de acessar as informações de forma

temporária ou permanente [2].

A lei da Indonésia exige que qualquer vazamento seja notificada por escrito dentro de 72 horas após a tomada de conhecimento do incidente, e a notificação deve ser direcionada tanto à autoridade nacional quanto aos usuários afetados, incluindo informações como a descrição dos dados violados, quando e como o incidente ocorreu e os esforços empreendidos para mitigá-lo [2].

Na Tabela 4.2, a coluna 'notificação de violação' especifica os requisitos de notificação para cada país. Nela, o valor ✓indica uma lei que exige notificação, mas não especifica um prazo. A Figura 4.15 mostra a distribuição de países que exigem notificação de violação de dados e, entre aqueles que o fazem, se especificam ou não um prazo. Também é digno de nota que países como Chile e Seychelles, embora estejam entre os três países com menos registros violados, não exigem notificação de violação de dados e, portanto, podem estar subnotificando incidentes [78].

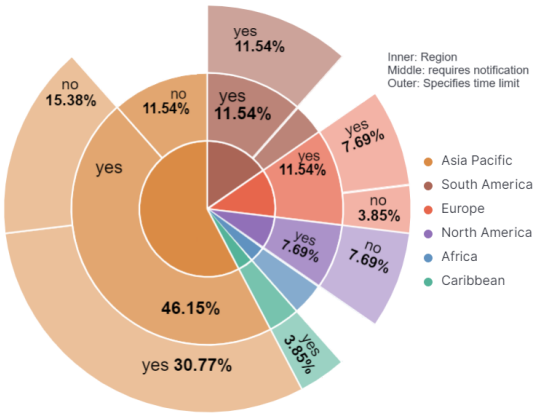



Figura 4.15: Distribuição de países que exigem notificação de vazamentos de dados e determinam um prazo para tal

A proteção de dados é uma questão complexa, e este trabalho não pretende apresentar uma comparação exaustiva da regulamentação nesses países, mas sim discutir alguns elementos-chave delas. Assim, os requisitos de transferência determinam os aspectos legais com os quais uma organização deve cumprir ao transferir dados nacional e internacionalmente. Especialmente ao transferir dados para o exterior, obrigações adicionais devem ser cumpridas. Nesse cenário, a lei israelense exige que as leis do país de destino forneçam um nível de proteção de dados não menos rigoroso do que o garantido pela lei israelense [2]. Caso contrário, pelo menos um dos outros critérios deve ser atendido, como o consentimento do titular dos dados, se a transferência for vital para a segurança pública ou outros [2].

Para a aplicação da lei, a entidade competente pode aplicar sanções a organizações não conformes, como interrupção da coleta de dados, destruição dos dados pessoais coletados e penalidades financeiras em Cingapura [2]. O Código Penal Turco também pune criminalmente com prisão uma pessoa que coleta, transfere, publica ou exclui dados ilegalmente [2].

As leis de proteção de dados também podem se aplicar ao marketing eletrônico. Um exemplo dessa aplicação é o PDPA da Malásia, que afirma que qualquer titular de dados pode exigir que seus dados cessem ou não comecem a ser processados para fins de marketing direto [2].

## 4.2 REGULAÇÃO DE PROTEÇÃO DE DADOS NOS ESTADOS UNIDOS

Existem, em alguns países, regulamentações específicas para tipos específicos de dados. Nos Estados Unidos, por exemplo, a Lei Sarbanes–Oxley (SOX) regulamenta a segurança de dados em empresas listadas publicamente [79], a Lei Gramm–Leach–Bliley (GLBA) em organizações financeiras [80], a Lei de Portabilidade e Responsabilidade de Seguro Saúde (em inglês, *Health Insurance Portability and Accountability Act*, HIPAA) [81].

Vale ressaltar que os EUA têm leis adicionais que regulam a proteção de dados, mas estas se referem a categorias de dados que estão fora do escopo do conjunto de dados em estudo. Por exemplo, a Lei de Direitos Educacionais e Privacidade da Família (em inglês, *Family Educational Rights and Privacy Act*, FERPA) em organizações educacionais [82]; a Lei de Proteção à Privacidade do Motorista (em inglês, *Driver's Privacy Protection Act*, DPPA) relaciona-se aos registros dos motoristas; a Lei de Modernização da Segurança da Informação Federal (em inglês, *Federal Information Security Modernization Act*, FISMA) em agências do governo federal dos EUA [83]; e a Lei de Proteção à Privacidade Online das Crianças (em inglês, *Children's Online Privacy Protection Rule*, COPPA) relacionada a crianças [84]. No entanto, o conjunto de dados utilizado neste trabalho não fornece informações suficientes para determinar se essas leis foram infligidas nas violações relatadas.

Há, ainda, o Padrão de Segurança de Dados da Indústria de Cartões de Pagamento (em inglês, *Payment Card Industry Data Security Standard*, PCI-DSS), que estabelece um conjunto de requisitos de segurança desenvolvido para garantir a proteção de dados de cartões de pagamento [85].

É importante destacar que os Estados Unidos atualmente não possuem uma lei de proteção de dados promulgada em nível federal. No entanto, há discussões em curso no Congresso sobre um projeto de lei proposto conhecido como a Lei Americana de Privacidade e Proteção de Dados (em inglês, *American Data Privacy and Protection Act*, ADPPA) [4].

A Tabela 4.3 detalha o arcabouço regulatório dos Estados Unidos conforme o tipo de dado protegido, além do número de incidentes associados às diferentes naturezas de dados. Esses diferentes tipos de dados estão fortemente atrelados ao setor da indústria que a empresa opera, e mostra diferentes necessidades regulatórias em função do segmento industrial.

Tabela 4.3: Normas de proteção de dados e sua aplicabilidade às violações no conjunto de dados por tipo de dado.

| Lei/Norma | Aplicável a | # (%) do conjunto de dados |
|---|---|---|
| SOX | Empresas listadas publicamente | 506 (100%) |
| GLBA | Financeiro | 131 (25.89%) |
| Lei de Telecomunicações | Comunicação | 35 (6.92%) |
| PCI-DSS | Dados de cartão de crédito | 25 (4.94%) |
| HIPAA | Saúde | 18 (3.56%) |

Em adição a essas normas setoriais, todos os 50 estados estabeleceram leis que determinam que empresas privadas devem ser obrigadas a informar os indivíduos sobre violações de segurança que comprometam informações pessoalmente identificáveis [86]. A maioria deles também inclui entidades governamentais

no compromisso.

No entanto, algumas dessas leis não estavam em vigor durante o período observado no conjunto de dados, entre 2005 e 2015. Por exemplo, a Lei de Notificação de Vazamento do Alabama, a lei estadual mais recente, foi promulgada em 2018. Da mesma forma, Dakota do Sul (2018) e Novo México (2017) introduziram suas respectivas leis após o período temporal do conjunto de dados. Portanto, as áreas desses estados são representadas em cinza na Figura 4.16, que exibe a ordem cronológica do início da vigência das leis de notificação de violação.

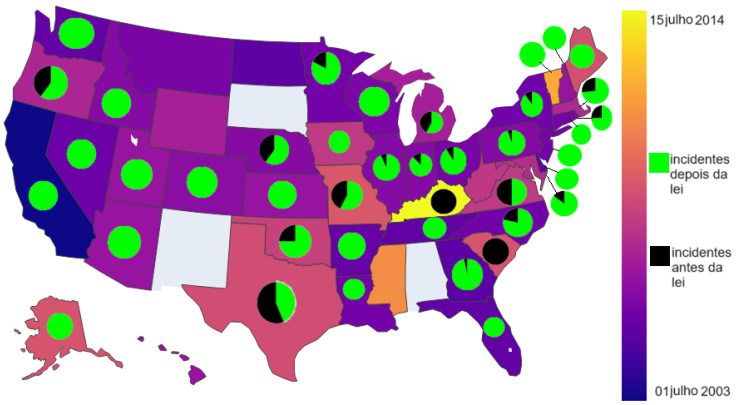



Figura 4.16: Status das leis de notificação de vazamento nos estados dos EUA. O mapa de cores indica o momento em que a lei começou a vigorar [3], e os gráficos de pizza indicam a distribuição das violações que ocorreram antes (preto) e depois (verde) do início da vigência da lei

Nesta figura, os gráficos de pizza denotam o número de incidentes de violação que ocorreram antes (preto) e depois (verde) do início da vigência da lei. Estados sem um gráfico de pizza ou não tiveram vazamentos ou não promulgaram leis de notificação dentro do período de tempo do conjunto de dados. A Califórnia foi o primeiro estado dos EUA a ter uma lei de notificação de violação em vigor a partir de 1º de julho de 2003, enquanto Kentucky se tornou o estado mais recente a fazer isso em 15 de julho de 2014 [4].

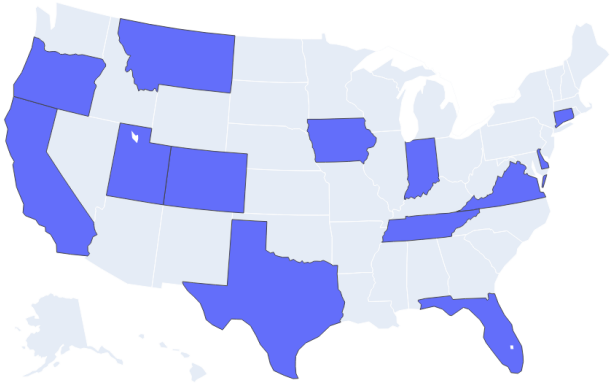


Figura 4.17: Estados dos EUA que promulgaram uma lei de proteção de dados [4]. Nenhum desses atos estava em vigor no intervalo de tempo do conjunto de dados

O trabalho de Coie [3] oferece uma análise abrangente dos detalhes de cada lei estadual. o autor detalha aspectos críticos, como suas definições de vazamentos de dados e informações pessoais, o momento e a estrutura das notificações de violação, sua aplicabilidade e outras disposições vitais que formam a base

dessas leis estaduais de notificação de vazamento de dados. Este recurso é relevante no entendimento das variações dentro dessas leis em diferentes estados.

Como evidenciado na Figura 4.22, a Califórnia teve o maior número de violações relatadas no conjunto de dados. Mas também foi o primeiro estado dos EUA a promulgar uma lei de proteção de dados, estabelecendo um precedente para outros estados. Nem todos os estados promulgaram legislação nesse sentido, e alguns ainda estão para iniciar seus efeitos, como visto na Tabela 4.4. Esses dados também podem ser visualizados geograficamente na Figura 4.17, ajudando a ilustrar os diferentes graus de cobertura legislativa entre os diferentes estados.

Tabela 4.4: Leis estaduais abrangentes de proteção de dados e sua aplicabilidade às violações no conjunto de dados por tipo de dados, com base em [4]. Nenhum desses atos estava em vigor no intervalo de tempo do conjunto de dados

| Lei | Estado | Data de vigência | # (%) do conjunto Fig 4.22 |
|---|---|---|---|
| CCPA | CA | 01/01/2020 | 79 (15.61%) |
| CPRA | CA | 01/01/2023 | 79 (15.61%) |
| VCDPA | VA | 01/01/2023 | 8 (1.58%) |
| CPA | CO | 07/01/2023 | 2 (0.4%) |
| CTDPA | CT | 07/01/2023 | 10 (1.98%) |
| UCPA | UT | 31/12/2023 | 2 (0.4%) |
| OCPA | OR | 01/07/2024 | 5 (0.99%) |
| TDPSA | TX | 01/01/2024 | 30 (5.93%) |
| FDBR | FL | 01/07/2024 | 19 (3.75%) |
| MTCDPA | MT | 01/10/2024 | 0 (0%) |
| ICDPA | IA | 01/01/2025 | 1 (0.2%) |
| DPDPA | DE | 01/01/2025 | 3 (0.59%) |
| TIPA | TN | 01/07/2025 | 3 (0.59%) |
| Indiana CDPA | IN | 01/01/2026 | 8 (1.58%) |

É importante destacar que o conjunto de dados examinado cobre o período entre 2005 e 2015, e a primeira lei de proteção de dados que entrou em vigor começou em 2020. Portanto, nenhum dos dados violados analisados neste artigo estava sujeito à regulamentação por uma lei de proteção de dados. Outra observação é que Nova York, o segundo estado mais violado no conjunto de dados, não promulgou nenhuma lei de proteção de dados. Em contraste, Montana não teve nenhum vazamento relatada nos dados analisados e promulgou a Lei de Privacidade de Dados do Consumidor de Montana (em inglês, *Montana Consumer Data Protection Act*, MTCDPA). No entanto, ainda não iniciou seus efeitos.

Elementos da Lei de Privacidade do Consumidor da Califórnia (em inglês, *Consumer California Privacy Act*, CCPA) incluem os direitos dos titulares de dados, abrangendo o direito de ser informado sobre a natureza dos dados coletados e suas práticas de venda; solicitar exclusão de dados, optar por não vender dados; acessar seus dados e não ser discriminado em serviço e preços ao exercer seus direitos de privacidade [87].

## 4.3 ANÁLISE QUALITATIVA DE RISCOS E IMPACTOS

Uma avaliação de risco qualitativa avalia a probabilidade e o impacto potencial de um evento ocorrer. Para possibilitar essa avaliação, considera-se a Figura 4.20d como o impacto, em termos do número de registros comprometidos. Para avaliar a probabilidade, consideramos a Figura 4.18, que apresenta a distribuição do número de incidentes de vazamento de dados relatados por vetor de ataque.

Ao comparar as Figuras 4.20d e 4.18, pode-se observar que as atividades de hacking não apenas corresponde ao vetor de ataque mais frequente, mas também ao com o maior impacto, em termos de mediana do tamanho das violações. Isso enfatiza a importância da adoção de controles de segurança técnica para mitigar os riscos relacionados ao hacking. De acordo com a correlação entre essas Figuras, o restante dos vetores de ataque inicial, exceto causas desconhecidas, são categorizados por sua probabilidade e impacto na Tabela 4.5, na qual as cores representam uma escala da gravidade de cada risco, sua probabilidade e seu impacto.

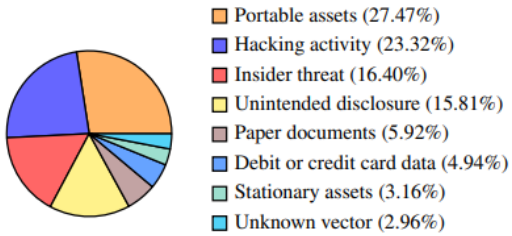



Figura 4.18: Número de violações relatadas por vetor de ataque

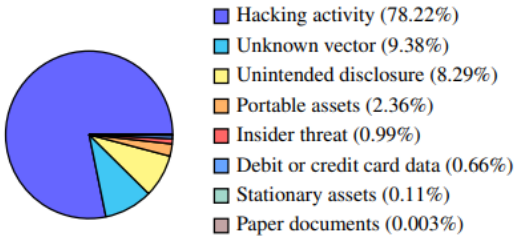
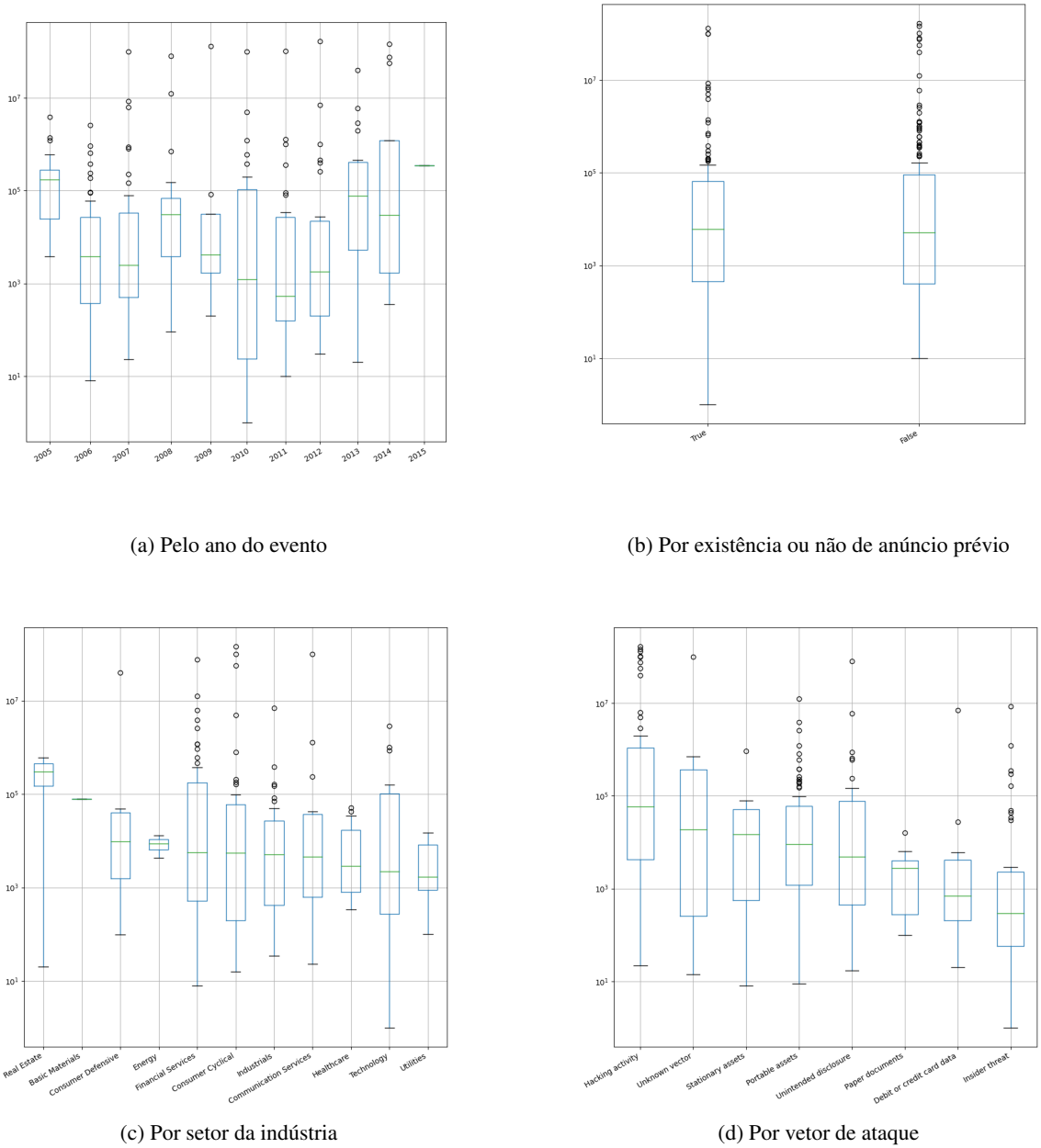



Figura 4.19: Proporções da soma de registros violados por vetor de ataque

Tabela 4.5: Análise qualitativa de risco para os diferentes vetores de ataque conhecidos

| Vetor de ataque | Probabilidade | Impacto |
|---|---|---|
| HACK | Alto | Alto |
| PORT | Alto | Médio |
| DISC | Médio | Médio |
| INSD | Médio | Baixo |
| CARD | Baixo | Baixo |
| STAT | Baixo | Baixo |
| PHYS | Baixo | Baixo |

Além disso, correlacionando as Figuras 4.19, que mostra a distribuição da soma de registros vazados por vetor de ataque, e 4.18 reforça o impacto médio das violações de dados em dispositivos portáteis, já que um grande número de ocorrências não resultou em uma quantidade significativa de registros vazados, quando comparado a outros vetores. Isso também enfatiza a importância de mitigar as atividades de hacking, já que, com uma correspondência de apenas 23,47% dos casos, resultou em 78,22% dos dados expostos. Essas conclusões sobre o impacto são confirmadas na Figura 4.20d.




(a) Pelo ano do evento

(b) Por existência ou não de anúncio prévio

(c) Por setor da indústria

(d) Por vetor de ataque

Figura 4.20: Boxplots da quantidade de dados vazados por incidente

Como observado na Figura 4.19, o tipo de violação mais prevalente é o de hacking, representando 78,22% dos registros violados no conjunto de dados. Em relação ao volume de informações vazadas, esse achado enfatiza a importância desse tipo de violação no cenário geral de vazamentos de dados.

Vários fatores podem influenciar o impacto de um incidente de vazamento de dados, como o tipo de dados, a origem da violação e o uso consequente dos dados comprometidos. Como o conjunto de dados usado neste estudo não fornece essas informações, nossa avaliação considera exclusivamente o número de registros violados. Considerando esses outros fatores, [30] obtiveram uma matriz diferente da apresentada na Tabela 4.5.

Contrariamente a este trabalho, os autores concluíram que as violações em ativos estacionários apresentam o maior impacto, apesar da baixa probabilidade, com divulgações não intencionais representando os vetores de ataque com o menor impacto. As conclusões convergentes indicam que vazamentos de hacking e ativos portáteis constituem os vetores mais significativos quando se considera tanto a probabilidade quanto o impacto do incidente.

Para uma estratégia de prevenção eficaz e para a implementação eficiente de controles de mitigação, é fundamental primeiro entender, em conjunto com a análise de riscos, como diferentes aspectos de um incidente cibernético contribuem para seus impactos, em termos da quantidade de dados vazados. Portanto, para aprimorar essa compreensão, a Figura 4.20 apresenta o resumo de cinco números do tamanho das violações de dados para diferentes anos de ocorrência (a), se a empresa vítima fez alguma divulgação anteriormente (b), e em função do setor da indústria (c) e do vetor de ataque utilizado (d).

A Figura 4.20a indica que o ano de 2005 teve a maior mediana de número de registros vazados por incidente, enquanto 2011 teve a menor. Da semelhança entre os quadros na Figura 4.20b, pode-se inferir que as divulgações feitas pelas organizações não impactam significativamente o número de registros violados. Isso pode ser devido ao curto período de 7 dias entre a divulgação e o evento de violação de dados, que pode não ser suficiente para um ator malicioso decidir agir com base na declaração e efetivamente roubar os dados. Por outro lado, as Figuras 4.20c e 4.20d mostram que o setor imobiliário e as atividades de Hacking alcançaram as maiores medianas em tamanho de vazamentos, respectivamente.

Além disso, o setor imobiliário alcançou a maior mediana em tamanhos de violação, apesar da baixa ocorrência de incidentes, enquanto serviços financeiros e cíclicos ao consumidor, que registraram mais exposições, apresentam medianas intermediárias de registros violados entre os setores.

### 4.3.1 Impacto no mercado de ações

Vazamentos de dados podem ter implicações significativas para o mercado de acoes, frequentemente levando a flutuações nos preços das acoes das empresas afetadas. Quando ocorre um incidente, a confiança dos investidores pode diminuir devido a preocupações sobre a capacidade da empresa de proteger informações sensíveis, resultando em uma queda no valor das acoes.

No entanto, o impacto das exposições de dados no mercado de acoes nem sempre é direto. Além disso, praticas de gerenciamento inadequadas apos uma violação de dados, como respostas inadequadas ou medidas de segurança de dados comprometidas, podem agravar ainda mais a situação e contribuir para a queda nos preços das ações.

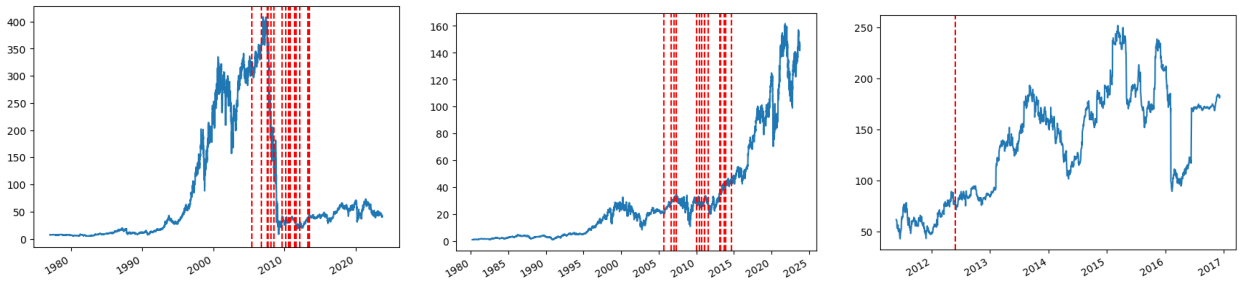
Como exemplos, menciona-se as trajetórias históricas dos preços das ações de três empresas notáveis dentro do conjunto de dados: Citigroup Inc., que sofreu o maior número de vazamentos; JPMorgan Chase,

a segunda empresa mais afetada; e LinkedIn.com, que sofreu a violação mais volumosa no conjunto de dados. Eles são representados na Figura 4.21, e é possível observar diferentes reações do mercado a essas exposições de dados, permitindo a observação de respostas variadas dos preços das ações a essas exposições de dados.

A queda no preço das ações de uma empresa pode resultar de vários fatores, incluindo vazamentos de dados, crises financeiras, modificações regulatórias e dinâmicas macroeconômicas mais amplas. Por causa disso, é essencial enfatizar que a conexão entre vazamentos de dados e a depreciação do preço das ações nem sempre é direta. Em certos casos, uma violação de dados pode desencadear uma queda no valor das ações ao diminuir a confiança dos investidores na empresa. Isso pode ser devido a preocupações de que a empresa seja incapaz de proteger adequadamente as informações sensíveis de seus clientes e funcionários ou devido à apreensão de que a violação possa causar danos financeiros ou de reputação.

Por outro lado, existem cenários em que uma queda no preço das ações pode ser atribuída a outros fatores, como crises financeiras ou alterações regulatórias. Por exemplo, os investidores podem optar por vender suas participações durante uma crise financeira como estratégia de mitigação de risco. Da mesma forma, se uma empresa enfrentar novos requisitos regulatórios, os investidores podem optar por vender suas ações diante de preocupações de que possam ter dificuldades para cumprir as regras atualizadas. Compreender a natureza dessas relações é essencial ao avaliar o impacto das violações de dados na situação financeira de uma empresa e seu efeito nos preços das ações.

Além disso, é vital considerar que práticas de gestão inadequadas podem estabelecer um ciclo prejudicial que agrava crises e intensifica a queda nos valores das ações. Por exemplo, quando uma empresa enfrenta uma crise financeira, a má gestão pode levar a decisões que agravam o problema. Essas decisões, como implementar medidas de corte de custos que comprometem a segurança de dados, podem promover a ocorrência de novas violações de dados, agravando ainda mais o prejuízo financeiro e resultando em uma subsequente queda contínua nos preços das ações.




(a) Citigroup Inc (b) JPMorgan Chase & Co (c) LinkedIn.com

Figura 4.21: Preços históricos das ações (em USD) das duas empresas com o maior número de ocorrências de incidentes (a e b), e o incidente com o maior número de registros expostos (c). As linhas verticais tracejadas representam eventos de violação de dados.

Embora, em alguns casos, um vazamento de dados possa de fato contribuir para uma queda no preço das ações, outros fatores também devem ser considerados. As reações dos preços das ações após anúncios de violação de dados têm sido estudadas por outros autores, gerando conhecimento sobre essa inter-relação [88].

Como estudo de um caso concreto, tem-se o vazamento de dados que afetou a empresa Heartland Payment Systems, ocorrido em Nova Jersey em 2009. Este ataque envolveu um informante do Serviço Secreto dos Estados Unidos que voltou à sua vida criminal como hacker e explorou uma vulnerabilidade de injeção SQL [89]. A empresa esperou um ano para anunciar a divulgação das informações do cartão de crédito de 130 milhões de clientes, causando uma queda no preço das suas ações de quase 80% [90].

## 4.4 MITIGAÇÃO DE VULNERABILIDADES

Estratégias eficazes de mitigação podem reduzir a frequência de ocorrência e os impactos de um vazamento de dados, protegendo os ativos da organização e preservando sua reputação e as pessoas envolvidas. Além disso, a adoção de controles de segurança é essencial para garantir a conformidade com regulamentações de proteção de dados e evitar multas regulatórias, custos de remediação e danos à reputação. Com o uso desses controles, a resiliência cibernética da organização é fortalecida, ajudando também a melhorar sua capacidade de detectar, responder e se recuperar de vazamentos de dados de maneira rápida e eficaz.

### 4.4.1 Análise e visualização do conjunto de dados dos EUA

Esta seção apresenta várias observações sobre o conjunto de dados fornecido por Rosati e Lynn [23], que é restrito às empresas listadas nas bolsas NYSE e NASDAQ dos Estados Unidos. A Tabela 4.6 sumariza os tamanhos dos vazamentos ocorridos nos EUA.

Tabela 4.6: Descrição estatística da quantidade de registros vazados

| | breach_size |
|---|---|
| Contagem | 272 |
| Média | 3.951716e+06 |
| Desvio padrão | 2.007628e+07 |
| Mínimo | 1 |
| 25% | 4.000000e+02 |
| 50% (mediana) | 5.154000e+03 |
| 75% | 7.979000e+04 |
| Máximo | 1.670000e+08 |

#### 4.4.1.1 Visão Geográfica

Como observado na Figura 4.22, os estados que experimentaram o maior número de vazamentos de dados relatadas foram a Califórnia, com 79 incidentes, e Nova York, com 75. Por outro lado, durante o período coberto pelo conjunto de dados, não houve violações relatadas nos seguintes estados: Havaí, Mississippi, Montana, Novo México, Dakota do Norte, Dakota do Sul, Virgínia Ocidental e Wyoming.

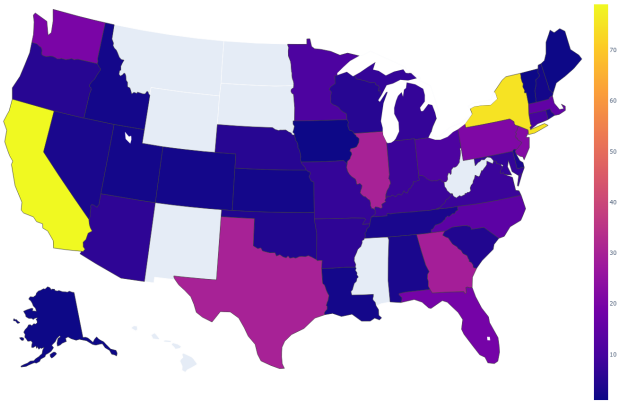


Figura 4.22: Distribuição geográfica da contagem de incidentes reportados por estado dos EUA

Entre os estados onde ocorreram vazamentos, Alasca, Iowa, Maine e Vermont registraram o menor número de incidentes, com apenas uma ocorrência cada. A exposição de Maine ocorreu em 2012 e visou a New York State Electric & Gas (NYSEG) e Rochester Gas and Electric (RG&E), subsidiárias da Iberdrola EUA, empresa que atua no ramo de energia elétrica. Este vazamento resultou na divulgação de 5.100 registros contendo números de Seguro Social, datas de nascimento e números de contas bancárias [91]. Em resposta ao incidente, a empresa posteriormente ofereceu uma associação de monitoramento de crédito assistencial[3].

Consistente com o mapa mostrado na Figura 4.22, a distribuição de tipos de violação entre os dez estados mais afetados, mostrada na Figura 4.23, é predominantemente liderada por Califórnia e Nova York. Este gráfico também mostra proporções semelhantes de tipos de vazamentos nos estados. Destaca-se, no entanto, que enquanto as violações de hacking representam uma parcela significativa das exposições nesses estados, Nova Jersey teve apenas um caso relatado desse tipo, ainda menos do que outros estados que registraram um pequeno número de incidentes. Esse caso se refere ao relatado na Seção 4.3.1.

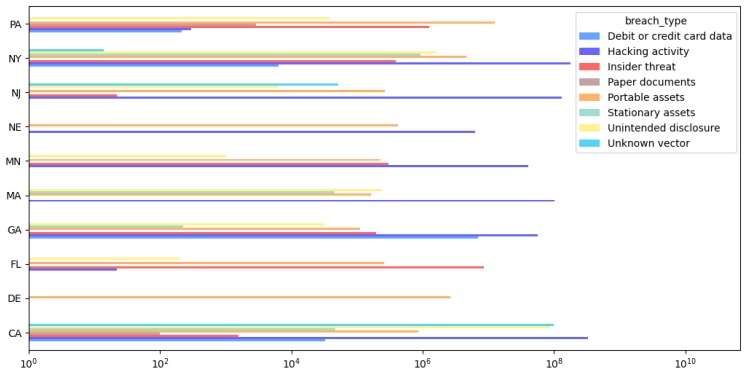



Figura 4.23: Distribuição de tipos de vazamento por estado

#### 4.4.1.2 Empresas

O conjunto de dados compreende 506 vazamentos de dados distribuídas entre 274 empresas únicas. A Figura 4.24 representa as dez empresas que experimentaram o maior número de vazamentos.

[3] <oag.ca.gov/ecrime/databreach/reports/sb24-22146> acesso em 12 de abril de 2024

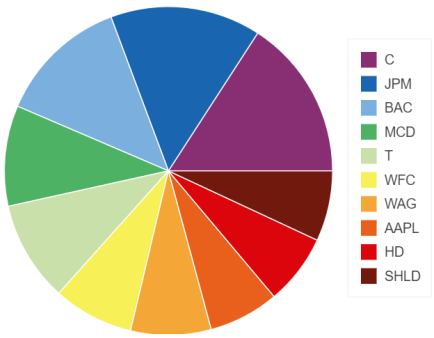



Figura 4.24: Contagem de vazamentos causado nas empresas mais afetadas

Uma tradução entre o ticker da empresa e seu nome é fornecida na Tabela 4.7, que também indica o setor da empresa. Entre as dez empresas que sofreram mais vazamentos, quatro operam dentro do setor financeiro. Como destacado na pesquisa, este setor econômico é o mais visado no conjunto de dados.

Com o objetivo de identificar áreas de vulnerabilidade aumentada dentro dessas empresas, a Figura 4.25 ilustra os vários tipos de incidentes para cada corporação. Consequentemente, vale ressaltar, por exemplo, que enquanto a Sears Holding teve uma distribuição um tanto uniforme de tipos, todas as violações de segurança na Apple foram atribuídas a atividades de hacking, enquanto uma proporção significativa das divulgações do McDonald's resultou de ações maliciosas de insiders. Além disso, incidentes de fraude com cartão de débito e crédito foram observados apenas em empresas financeiras; e perda de computadores estacionários, acessados de maneira inadequada, descartados ou roubados, tiveram sucesso em violar a Walgreens.

Tabela 4.7: Dicionário dos nomes das empresas mais vazadas e seus tickers

| Ticker | Nome | Setor |
|---|---|---|
| C | Citigroup Inc | Financeiro |
| JPM | JPMorgan Chase & Co | Financeiro |
| BAC | Bank of America Corp | Financeiro |
| MCD | McDonald's | Cíclico de consumo |
| T | AT&T Inc. | Comunicação |
| WFC | Wells Fargo & Co | Financeiro |
| WAG | Walgreens | - |
| AAPL | Apple | Tecnologia |
| HD | Home Depot | Cíclico de consumo |
| SHLD | Sears Holding Corp | - |

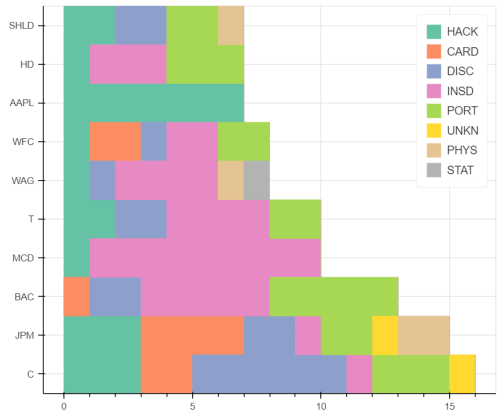



Figura 4.25: Tipos de vazamento nas empresas mais vazadas

#### 4.4.1.3 Tamanhos de Violação

A soma total de registros violados no conjunto de dados, que considera exclusivamente empresas de capital aberto, resulta em aproximadamente $1.07410 \times 10^9$, o que equivale aproximadamente a 3 vezes a população do país em 2013, cerca de $3.16128 \times 10^8$ pessoas [92]. De maneira semelhante ao observado no conjunto de dados global, isso pode ser devido à violação de dados de pessoas falecidas e aos dados da mesma pessoa sendo vazados várias vezes.

As dez violações mais significativas, em termos de registros vazados, no conjunto de dados e seus tipos correspondentes são mostrados na Figura 4.26.

No caso do incidente do LinkedIn, que constitui a exposição de dados mais volumosa examinada neste estudo, foi relatado que senhas com hash SHA-1 sem uso de *salt* foram vazadas. No entanto, há informações limitadas disponíveis sobre o método específico pelo qual os dados foram roubados [93]. Pesquisadores [94] conseguiram quebrar com sucesso aproximadamente 2,5% dessas senhas.

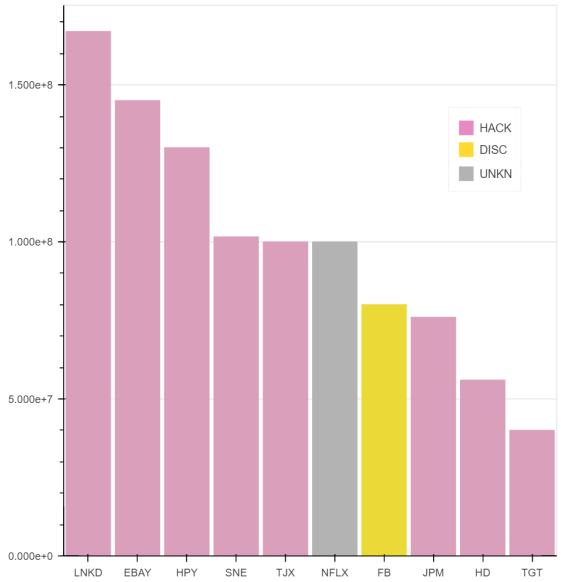



Figura 4.26: As dez maiores violações no conjunto de dados e seus tipos

Tabela 4.8: Dicionário de empresas com as maiores violações relacionando seus nomes, tickers e data de violação

| Ticker | Nome | Data | Estado |
|---|---|---|---|
| LNKD | LinkedIn.com | 29/05/2012 | CA |
| EBAY | eBay & Co | 21/05/2014 | CA |
| HPY | Heartland Payment | 20/01/2009 | NJ |
| SNE | Sony | 26/04/2011 | NY |
| TJX | TJ stores | 17/01/2007 | MA |
| NFLX | Netflix & Co | 01/01/2010 | CA |
| FB | Facebook | 17/07/2008 | CA |
| JPM | JPMorgan Chase & Co | 27/08/2014 | NY |
| HD | Home Depot | 02/09/2014 | GA |
| TGT | Target Corp | 13/12/2013 | MN |

Em um incidente de 2014 que ocorreu na Califórnia, uma campanha de spear-phishing efetivamente comprometeu aproximadamente 145 milhões de registros armazenados pelo eBay [95]. Como destacado pelo autor, essa violação expôs dados sensíveis, incluindo nomes de clientes, endereços de e-mail, endereços físicos, números de telefone e datas de nascimento, todos em texto não criptografado, resultando em um custo estimado para a empresa de U$300 milhões.

Os autores de [96] demonstraram a gravidade desse vazamento, destacando que, naquela época, o incidente permitiu que o atacante recuperasse o histórico completo de compras para um nome de usuário conhecido. Como consequência, facilitou a identificação de compradores de itens sensíveis, como armas de fogo e testes de gravidez e HIV.

O caso da Heartland Payment Systems já foi discutido na Seção 4.4.1.1. Os casos da Home Depot e Target são discutidos em 4.4.3.1.

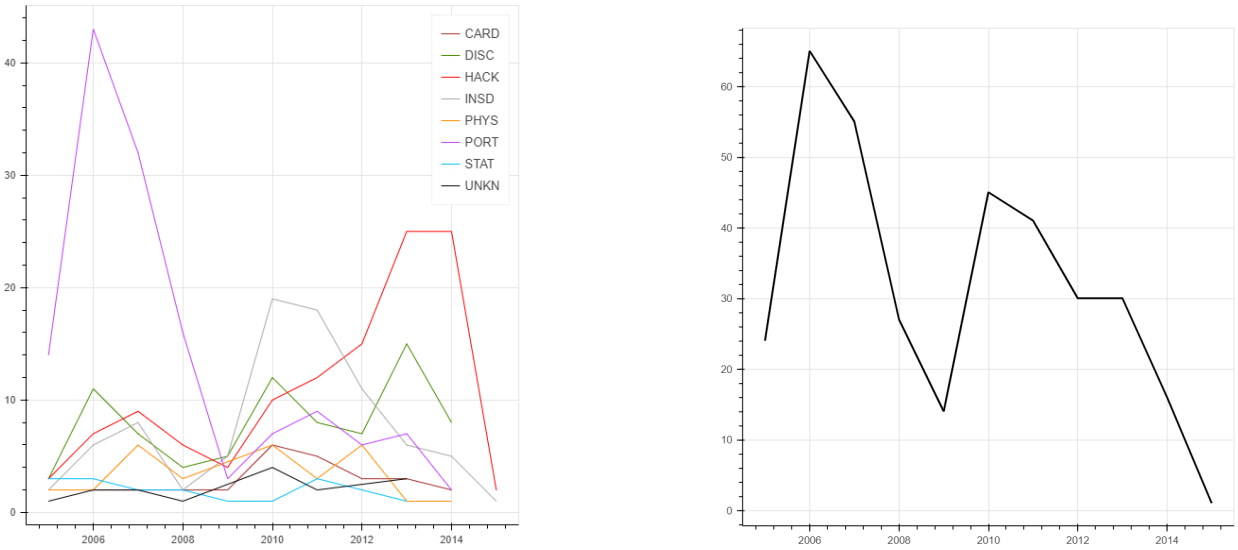
### 4.4.2 Vetores de ataque

Como observado na Seção 4.3, os tipos de violação mais prevalentes são PORT (139 eventos) e HACK (118), representando coletivamente 50,79% dos incidentes dentro do conjunto de dados. Essa compreensão reforça a importância desses tipos de violação no panorama geral das violações de dados. Casos com causas desconhecidas (UNKN) podem indicar uma investigação forense ineficaz ou falta de transparência.

Determinar se o problema das violações de dados está piorando, bem como identificar as tendências predominantes, é uma preocupação crítica. Os dados usados para gerar informações frequentemente apresentam uma variação significativa. É importante enfatizar que essas tendências nem sempre são imediatamente aparentes. Assim, há uma necessidade de análises de dados rigorosas e estatisticamente sólidas para determinar se existem tendências discerníveis. Além disso, quando possível, tais estudos podem ajudar a fazer previsões sobre a trajetória das violações de dados. Essa abordagem baseada em dados é relevante na compreensão do cenário de segurança de dados em constante evolução.

Para melhorar a compreensão das Táticas, Técnicas e Procedimentos (TTPs) empregados pelos atacantes em vazamentos de dados e suas tendências evolutivas ao longo do tempo, a Figura 4.27a fornece

uma visualização das violações totais relatadas por ano para cada tipo de violação. Ao correlacionar as Figuras 4.19 e 4.27a, torna-se evidente que, enquanto as violações PORT são a categoria mais frequente de exposições, os ataques HACK cresceram significativamente desde 2012.




(a) Por vetor de ataque

(b) Agrupando vetores de ataque de fator humano

Figura 4.27: Contagem de vazamentos por ano

A divulgação de dados como consequência das TTPs de PORT foi mais proeminente durante os primeiros anos do conjunto de dados, com destaque para 2006, especialmente com 43 vazamentos nessa categoria. Esse número foi quase quatro vezes maior que o segundo tipo mais relevante, DISC, que teve 11 vazamentos. Essas tendências oferecem informações sobre as táticas e prioridades em mutação dos atacantes ao longo do tempo.

A afirmação de que os seres humanos são frequentemente considerados o elo mais fraco na cibersegurança é bem documentada. Aprofundando essa questão, [97] concluíram que as violações decorrentes de fatores humanos estavam em declínio, possivelmente devido ao aumento da conscientização entre o pessoal. Uma análise semelhante foi realizada em empresas listadas na NYSE e NASDAQ para validar essa afirmação. Esta pesquisa investiga se há evidências que sustentem a tendência de diminuição das violações atribuíveis a fatores humanos nessas situações.

Nesse contexto, focamos nos tipos de violação INSD, PHYS, PORT, STAT e DISC, que estão associados a fatores humanos. A Figura 4.27b exibe o número cumulativo de violações relacionadas a esses tipos ao longo dos anos. Essa análise nos leva a concluir que o envolvimento do fator humano em vazamentos de dados está diminuindo em empresas de capital aberto. No entanto, é importante observar que essa redução nos casos também pode ser influenciada pelo declínio nos casos de PORT, como discutido anteriormente, que foram particularmente numerosos em 2006.

A Seção 4.4.3, explora mais a fundo as possíveis causas para a tendência decrescente em violações de dados relacionadas a fatores humanos. Além disso, apresentamos uma discussão mais abrangente de estratégias de mitigação específicas para cada tipo de violação. Essa análise fornece uma compreensão mais profunda dos fatores que influenciam a redução dessas violações, abordando estratégias de mitigações a

essas vulnerabilidades.

#### 4.4.2.1 Setor da empresa

Combinando o tipo de violação de dados com o setor em que uma empresa opera, é possível verificar se existe uma predisposição de uma TTP ao mirar em um setor econômico específico. Embora nenhuma preferência seja evidenciada pela Figura 4.28, que mostra uma distribuição semelhante de tipos nos setores das empresas, algumas outras observações são feitas. Por exemplo, nota-se que a indústria de saúde não foi alvo de nenhuma atividade de hacking no período coberto pelo conjunto de dados, nem as empresas nos setores industriais tiveram vazamentos por meio de documentos em papel (PHYS). Essas percepções fornecem contexto valioso sobre a distribuição de tipos de vazamentos dentro de diferentes setores econômicos e podem ajudar a informar estratégias de segurança para essas indústrias.

Além disso, uma relação mais robusta é encontrada entre fraudes envolvendo cartões de débito/crédito (CARD), que é o segundo tipo de violação menos frequente conhecido (Figura 4.19), e empresas financeiras. Essa correlação está alinhada com a natureza das operações financeiras, demonstrando uma concentração desse tipo específico de exposição no setor financeiro. Diferentemente, os ataques direcionados a dispositivos estacionários (STAT), o tipo de violação menos frequentemente conhecido, estavam principalmente presentes nos quatro setores mais violados, a saber, serviços financeiros, ciclo de consumo, industriais e tecnologia.

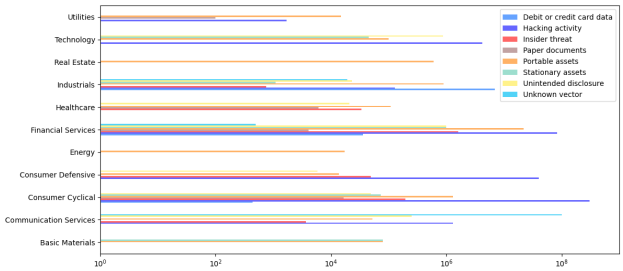



Figura 4.28: Distribuição do tipo de violação por setor da empresa

É interessante contrastar esses resultados com os de [97], que, em sua análise do conjunto de dados da PRC de 2005 a 2019, observaram que as empresas mais visadas estavam nos setores de saúde e manufatura/tecnologia/comunicações, devido à sensibilidade dos dados que possuem. No entanto, a Figura 4.28 mostra que, ao restringir essa análise às empresas listadas publicamente no conjunto de dados, esse cenário muda, e as empresas financeiras, que também possuem dados sensíveis, são as mais violadas.

A Figura 4.29 ilustra que o número substancial de vazamentos neste setor mostrado na Figura 4.28 não foi concentrado em um período específico, mas relativamente constante ao longo do tempo contemplado no conjunto de dados. Isso fornece informações valiosas sobre a persistência dos desafios de segurança enfrentados pelas empresas no setor financeiro. Na Figura 4.28, o hífen (-) representa empresas para as quais a API do Yahoo! Finance não conseguiu identificar o setor pertencente.

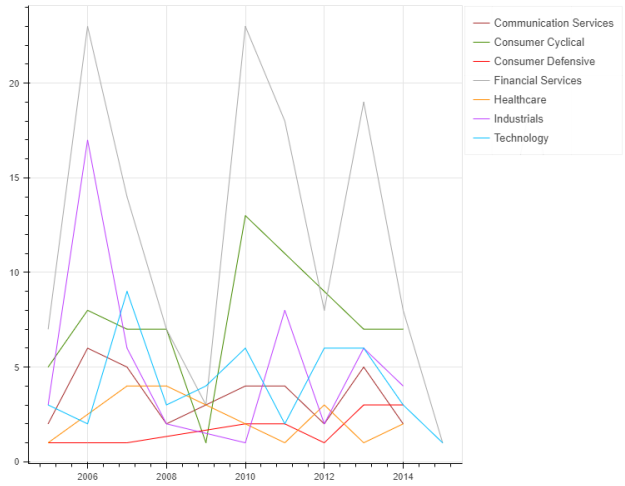



Figura 4.29: Número de vazamentos por setor da empresa por ano

É importante enfatizar que a diminuição no número de casos em 2015 se deve à incompletude dos dados para esse ano, que só foram até março. Por outro lado, os casos de violação para as empresas do setor industrial foram mais frequentes em 2006 e têm demonstrado uma tendência descendente desde então.

Uma associação entre o setor da empresa e os padrões de proteção de dados é pertinente. Essa conexão surge da relação próxima entre a indústria de uma empresa e o tipo de dados que ela armazena, o que, por sua vez, influencia as responsabilidades e obrigações do proprietário dos dados. Essa relação e uma visão geral da conformidade são exploradas mais detalhadamente na Seção 4.2.

### 4.4.3 Controles de segurança

Para garantir a proteção adequada dos dados dos clientes e a conformidade com essas regulamentações, uma empresa deve implementar efetivamente controles de segurança que mitiguem as vulnerabilidades que podem levar a um vazamento de dados. Diferentes tipos de incidentes requerem diferentes contramedidas de segurança como mitigação.

Como visto na Figura 4.30, as violações de dados em 2023 foram mais frequentemente causadas por ataques de phishing e as mais caras originaram-se de insiders maliciosos [5]. Esta seção revisa algumas das contramedidas de segurança adequadas que podem mitigar esses e outros vetores. Tais contramedidas poderiam ter reduzido a probabilidade e/ou o impacto das violações no escopo deste estudo.

Para melhorar a segurança geral, as empresas podem adotar um framework estruturado, como o NIST Cybersecurity Framework (CSF). Esse é um framework agnóstico que categoriza diversos controles de segurança em cinco núcleos, como mostrado na Tabela 4.9. Essas categorias principais fornecem uma abordagem organizada para melhorar as medidas de cibersegurança. Outro padrão significativo é o CIS Controls, que apresenta um conjunto de práticas recomendadas para a segurança cibernética.

Especialmente para mitigação de vulnerabilidades, a função principal Protect apresenta algumas recomendações valiosas divididas em categorias: Gerenciamento de Identidade, Autenticação e Controle de Acesso, Conscientização e Treinamento, Manutenção, Tecnologia Protetora, Processos e Procedimentos

de Proteção de Informações e Segurança de Dados.

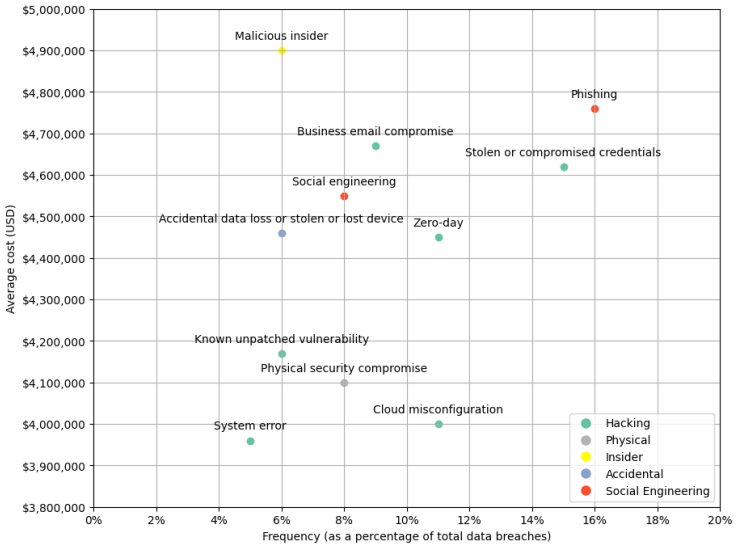



Figura 4.30: Frequência (como porcentagem do total de vazamentos de dados) e custo médio (medido em milhões de dólares) dos vetores de ataque inicial responsáveis por vazamentos de dados em 2023 [5]

Esta última, mais pertinente a este estudo, é então dividida em subcategorias, que são: Dados em repouso são protegidos; Dados em trânsito são protegidos; Ativos são formalmente gerenciados durante remoção, transferências e disposição; Capacidade adequada para garantir disponibilidade é mantida; Mecanismos de verificação de integridade são usados para verificar a integridade do software, firmware e informações; Os ambientes de desenvolvimento e teste são separados do ambiente de produção; Mecanismos de verificação de integridade são usados para verificar a integridade do hardware; e Proteções contra vazamentos de dados são implementadas.

Tabela 4.9: Funções principais do NIST CSF

| Função principal | Descrição |
|---|---|
| Identificar | Ajuda a determinar o risco atual de cibersegurança para a organização |
| Proteger | Usa salvaguardas para prevenir ou reduzir o risco de cibersegurança |
| Detectar | Encontra e analisa possíveis ataques e comprometimentos de cibersegurança |
| Responder | Tome medidas em relação a um incidente de cibersegurança detectado |
| Recuperar | Restaurar ativos e operações que foram impactados por um incidente de cibersegurança |

Mais uma vez, esta última subcategoria é mais relevante para este trabalho, e o NIST CSF faz referência ao Center for Internet Security (CIS) Critical Security Control, COBIT 5, ISA 62443-3-3:2013, ISO/IEC

27001:2013 e NIST SP 800-53.

Especificamente para o NIST SP 800-53, que revisa controles de segurança e privacidade para sistemas de informação e organizações, o framework menciona as seções sobre aplicação do fluxo de informação (AC-4), separação de funções (AC-5), o princípio do mínimo privilégio (AC-6), triagem de pessoal (PS-3), acordos de acesso (PS-6), proteção de limite (SC-7), confidencialidade e integridade da transmissão (SC-8), proteção criptográfica (SC-13), análise de canal oculto (SC-31), monitoramento do sistema (SI-4) e proteção contra vazamento de informações devido a emissão eletromagnética (PE-19). Em relação a este último, o TEMPEST é uma especificação valiosa sobre o blindagem de equipamentos contra vazamento não intencional de sinais de rádio ou elétricos, sons e vibrações [98].

Alguns desses controles são subsequentemente discutidos com mais profundidade. Como essas contramedidas estão intimamente relacionadas ao vetor de ataque, considerou-se o `breach_type` do conjunto de dados para revisá-las.

No entanto, é essencial observar que listar todas as medidas de segurança para mitigar vulnerabilidades relacionadas a violações de dados é impraticável devido ao grande número de vetores de ataque. Portanto, esta seção fornece diretrizes gerais e boas práticas contra vetores de ataque comuns. Ainda assim, não é uma lista exaustiva de controles de segurança nem um guia abrangente de prevenção contra vazamentos de dados.

#### 4.4.3.1 CARD

Conforme discutido na Seção 4.2, o PCI-DSS delimita requisitos específicos para cada objetivo. Esses requisitos são detalhados na Tabela 4.10, fornecendo uma visão abrangente dos padrões PCI-DSS e seus requisitos associados. O PCI exige que controles técnicos e operacionais sejam implementados por qualquer entidade que armazene, processe ou transmita dados de cartões de crédito [99].

Isso é aplicado por meio de três etapas contínuas: uma avaliação, identificação de todos os locais de dados do titular do cartão, um inventário de ativos e análise de vulnerabilidades que poderiam expor dados do titular do cartão. A etapa seguinte é corrigir as vulnerabilidades encontradas e, por último, relatar os detalhes da avaliação e remediação e enviar o documento resultante às entidades com as quais a empresa faz negócios.

Em relação à criptografia, o PCI-DSS exige conformidade com o padrão de Criptografia de Ponta a Ponta (P2PE) usando uma das soluções validadas listadas por eles. Dois vazamentos de dados relevantes envolvendo informações de cartões de débito e crédito referem-se aos vazamentos da Home Depot e Target Corp., exibidos na Figura 4.26 e apresentam os vazamentos mais volumosos.

A causa mais provável do vazamento no Target é uma infecção pelo malware Citadel [100]. Este malware, baseado em seu antecessor Zeus, executa um ataque de Man-in-the-Browser. Outro malware usado no ataque foi o BlackPOS, que visa dispositivos de Pontos de Venda (em inglês, *point of sales*, POS) [101]. A Seção 4.4.3.2 fornece mais informações sobre malware bancário.

Consequentemente, aproximadamente 40 milhões de registros de cartões de crédito e débito foram vazados, incluindo seus números de PIN criptografados e outras informações pessoais identificáveis.

Tabela 4.10: Requisitos resumidos do PCI-DSS

| Metas | Requisitos |
|---|---|
| Construir e Manter uma Rede e Sistemas Seguros | Instalar e manter um firewall |
| | Não usar as configurações padrão fornecidas pelo fornecedor para senhas do sistema e outros parâmetros de segurança |
| Proteger os Dados do Titular do Cartão | Proteger dados do titular do cartão armazenados |
| | Criptografar a transmissão de dados do titular do cartão através de redes abertas e públicas |
| Manter um Programa de Gerenciamento de Vulnerabilidades | Proteger todos os sistemas contra malware e atualizar regularmente o software antivírus ou programas |
| | Desenvolver e manter sistemas e aplicativos seguros |
| Implementar Medidas de Controle de Acesso Forte | Restringir o acesso aos dados do titular do cartão pela necessidade de negócios |
| | Identificar e autenticar acesso a componentes do sistema |
| | Restringir o acesso físico aos dados do titular do cartão |
| Monitorar e Testar Redes Regularmente | Rastrear e monitorar todo o acesso aos recursos da rede e dados do titular do cartão |
| | Testar regularmente sistemas e processos de segurança |
| Manter uma Política de Segurança da Informação | Manter uma política que aborda a segurança da informação para todo o pessoal |

De acordo com o conjunto de dados estudado, as lições não foram aprendidas, e a Home Depot também foi infectada pelo BlackPOS [101], vazando 56 milhões de registros de pagamento. Os autores de [43] estudaram os efeitos na reputação da empresa após esse evento por meio de postagens no Twitter (agora X) e concluíram que raiva, desgosto e tristeza eram emoções significativas entre os usuários de mídias sociais.

As duas empresas também estavam em conformidade com o PCI-DSS no momento do vazamento [102], embora a Tabela 4.10 mostre alguns requisitos que poderiam ter prevenido uma infecção por malware se implementados com sucesso, como atualizar regularmente o software antivírus e manter sistemas e aplicativos seguros. Além disso, o vazamento de dados da Home Depot poderia ter sido evitado usando P2PE e segregação de rede [103]. Portanto, observa-se que o PCI-DSS serve como uma base confiável para a segurança de cartões de crédito, mas para uma melhor segurança, não deve ser implementado exclusivamente.

Uma tecnologia adicional que pode ser usada para melhorar a segurança de transações é o chip micro-

processador EMV (Europay, MasterCard e Visa), aumentando a complexidade e os custos para a falsificação de cartões [104], conhecida como *skimming*.

Os casos de falsificação estão diminuindo acentuadamente em áreas onde o EMV é implementado, mas há um aumento consequente nos crimes de compra por Cartão Não Presente (CNP) [105]. Um crime CNP é o uso não autorizado dos detalhes de pagamento de outra pessoa para uma transação, principalmente por meio de meios online. As informações de pagamento podem ser obtidas após um vazamento de dados, por exemplo. [106] descrevem o script seguido por criminosos em um crime CNP, o que permite uma consideração mais fundamentada para estratégias de mitigação.

Além disso, tanto os vazamentos do Target quanto da Home Depot foram iniciados com um ataque de phishing [101], o que reforça a necessidade de Educação, Treinamento e Conscientização em Segurança, discutida na Seção 4.4.3.7.

#### 4.4.3.2 HACK

Como evidenciado nos casos examinados neste artigo, como eBay, Target e Home Depot, o phishing é um vetor de ameaça prevalente usado para iniciar vazamentos de dados. Uma abordagem prática para reduzir a taxa de sucesso desses ataques é implementar um programa robusto de Educação, Treinamento e Conscientização em Segurança, que é discutido com mais detalhes na Seção 4.4.3.7. Um programa desse tipo desempenha um papel fundamental em aprimorar a capacidade dos funcionários de reconhecer e frustrar ameaças à segurança, como tentativas de phishing, fortalecendo a postura de segurança da organização.

O estudo de Naqvi et al. [107] revisa a literatura sobre procedimentos de mitigação de phishing por meio de diferentes vetores, como e-mail e sites. A maioria das técnicas propostas baseia-se em aprendizado de máquina ou treinamento e conscientização. A Autenticação Multifator (MFA) pode proteger a conta do usuário mesmo após um phishing bem-sucedido, pois a identificação e a senha do usuário obtidas com a técnica não seriam suficientes para fazer login, exigindo um fator extra.

Como relacionado na Seção 4.4.3.1, após o phishing bem-sucedido nos casos da Target e da Home Depot, o invasor usou malware bancário para exfiltrar dados. Como o setor financeiro representa a parte de contribuição mais significativa no conjunto de dados, julgou-se conveniente discutir esse tipo de malware.

Alguns desses malwares, nomeadamente Zeus V2, Citadel (que foi usado no vazamento do Target), Carberp, Vawtrak, Dridex, Dyre e Rovnix, têm regras de detecção conhecidas [108]. Essas regras são conhecidas como Indicadores de Comprometimento (IoC) do Citadel, e exemplos incluem as chaves de registro que ele altera e seus comportamentos de rede.

No entanto, é relevante observar que certas cepas de malware têm foco regional, como Guildma, Grandoreiro e Javali, que visavam principalmente entidades brasileiras [109]. Um projeto de inteligência de ameaças pode ser necessário para identificar atividades maliciosas comuns dentro do domínio operacional da organização.

Mesmo após uma infecção bem-sucedida por malware, a violação de dados pode ser evitada se a empresa aplicar efetivamente outras medidas de segurança, como criptografia e controle de acesso. Não foi o caso do LinkedIn, por exemplo, que, como discutido na Seção 4.4.1.3, teve milhões de hashes de senhas

sem salt vazados, destacando a importância crítica de medidas de segurança abrangentes para proteger dados sensíveis.

Vários fatores agravaram as vulnerabilidades de segurança do LinkedIn. Em primeiro lugar, a empresa empregava o algoritmo de hash SHA-1, que demonstrou ser fraco e vulnerável a vários ataques [110]. A revisão 2 do NIST SP 800-131A proibiu o uso de SHA-1, permitindo-o apenas para aplicativos de assinatura não digital. Atualmente, SHA-2 e SHA-3 são algoritmos de digestão de mensagens seguros.

Em segundo lugar, a segurança do LinkedIn foi comprometida pela ausência de um algoritmo de "salt"para aumentar a segurança das senhas hash. Quando a mesma senha é processada usando o mesmo algoritmo de digestão de mensagem, ela gera consistentemente o mesmo valor de hash, o que aumenta a previsibilidade e a suscetibilidade a ataques de força bruta. Algoritmos de salt envolvem a adição de uma string única (o sal) à senha antes de hash, aumentando significativamente sua segurança [111]. Além disso, as bases de vazamento de senhas devem ser continuamente monitoradas em busca de credenciais em uso na organização.

A Seção 4.4.1.1 introduziu brevemente o vazamento do Heartland Payment System, que dependia de injeção de SQL. Para vulnerabilidades de nível de aplicativo, como a explorada na HPY, o Open Worldwide Application Security Project (OWASP) é uma referência notória [112] . Eles publicam regularmente as 10 principais vulnerabilidades no escopo de segurança de aplicativos, juntamente com suas estratégias de mitigação [113], como validação de entrada, Firewall de Aplicativos da Web (WAF) e testes de software [114].

Embora não esteja no conjunto de dados estudado, os invasores conseguiram invadir provedores de nuvem no caso do hack do SolarWinds em 2020, expondo e violando os dados de seus clientes [115]. Neste incidente, os adversários inseriram código arbitrário no código-fonte de um produto da empresa chamado Orion. Posteriormente, a SolarWinds distribuiu o código malicioso para seus clientes como parte do produto, infectando mais de 18.000, incluindo entidades governamentais e empresas privadas [116]. Este exemplo reforça a importância da Gestão da Cadeia de Suprimentos em cibersegurança [117].

Além disso, em empresas grandes e tecnologicamente complexas, manter os sistemas atualizados pode ser desafiador. Como consequência, os invasores podem explorar vulnerabilidades conhecidas nos sistemas. Portanto, é fundamental estabelecer um programa de gerenciamento de patches para atualizar e proteger oportunamente os ativos da organização [118].

Para ataques de dia zero, que exploram vulnerabilidades previamente desconhecidas, um patch de segurança ainda não foi publicado pelo desenvolvedor do produto, e a detecção de assinaturas é ineficaz [119]. Métodos de ML podem detectar essas intrusões com base na percepção de atividades suspeitas que diferem da linha de base esperada [120].

Soluções de DLP também contribuem para a segurança de dados e para evitar vazamentos de dados. Essas tecnologias detectam e impedem transferências de dados não autorizadas, incluindo a prevenção de vazamentos de dados PII [121]. No entanto, é importante notar que o DLP é ineficaz na detecção de exfiltração de dados por meio de técnicas esteganográficas [122].

Quando há necessidade de publicar métricas estatísticas relacionadas a um conjunto de dados, mas há preocupação em preservar a privacidade dos indivíduos dentro do conjunto de dados, a privacidade dife-

rencial pode ser uma abordagem adequada. Essa técnica viabiliza a divulgação de informações agregadas sobre um conjunto de dados, adicionando ruído aos dados para que os registros individuais permaneçam privados e indistinguíveis. Isso garante que as informações sensíveis estejam protegidas e que conhecimentos estatísticos possam ser derivados sem comprometer a privacidade dos sujeitos de dados [123].

#### 4.4.3.3 INSD

Uma ameaça interna é qualquer pessoa com acesso autorizado ou conhecimento dos recursos de uma organização [124]. A empresa confia nessa pessoa, que conhece os fundamentos da empresa e tem acesso aos seus ativos. Por causa disso, um agente interno malicioso (chamado de *insider*) pode causar danos significativos à empresa de forma imperceptível. O tempo médio de uma empresa para detectar as ações maliciosas de um insider é de 85 dias [125].

Um insider pode ser classificado como não intencional ou intencional. Como as ameaças internas não intencionais são mais adequadas no escopo deste trabalho para DISC, PHYS, PORT e START, nesta seção, discutimos principalmente insiders maliciosos intencionais.

As principais motivações para realizar um ataque interno são benefícios financeiros ou espionagem [126]. A ação do incidente interno é o abuso de privilégios, enquanto as ações são realizadas principalmente por meio do abuso de privilégios. Por causa disso, a política de privilégio mínimo e a tecnologia de Gerenciamento de Acesso Privilegiado (PAM) são ferramentas úteis para prevenir vazamentos internos.

As principais tecnologias adotadas para mitigar as ameaças internas são Prevenção de Perda de Dados (DLP), Gerenciamento de Acesso Privilegiado, Análise de Comportamento de Usuário e Entidade (UEBA), Gerenciamento de Informações e Eventos de Segurança (SIEM), Detecção e Resposta em Endpoints (EDR) e Gerenciamento de Ameaças Internas (ITM) [125].

Controles de segurança administrativos também podem ser implementados. Uma verificação de antecedentes e uma triagem de funcionários na contratação podem revelar um histórico malicioso para o candidato, permitindo que a empresa cancele o processo de contratação. Se uma pessoa passar por essa investigação, a imposição da assinatura de um Acordo de Não Divulgação (NDA) é uma contramedida adicional, pois constituirá legalmente sua responsabilidade.

Após a contratação de um funcionário, outras medidas de segurança ainda devem ser adotadas. Uma delas é avaliar a necessidade de saber para cada funcionário, aplicada por meio de um mecanismo de controle de acesso. Conceder mais acesso ao conhecimento e aos dados do que o funcionário precisa para realizar suas tarefas habituais expõe as informações desnecessariamente. Um controle semelhante é baseado no princípio do privilégio mínimo, que concede a um trabalhador os privilégios mínimos necessários, o que ajuda a minar uma possível colaboração.

A separação de funções é outra forma de mitigar ameaças internas intencionais, que divide tarefas críticas entre vários funcionários, de acordo com seus departamentos na organização, por exemplo [127]. Uma política de rotação de empregos, embora às vezes inviável, também pode ajudar a manifestar fraudes, sabotagens ou espionagem. Encerrar o contrato com o empregador é outro passo crítico na prevenção de vazamentos de dados, e a empresa deve garantir que as contas do usuário sejam desativadas, preferenci-

almente durante a entrevista de saída, em que qualquer equipamento pertencente à organização deve ser devolvido,. Após essa entrevista, o ex-funcionário deve ser escoltado para fora das instalações.

#### 4.4.3.4 PHYS, PORT, STAT: perdas

Um dispositivo móvel é significativamente mais fácil de ser perdido do que um dispositivo fixo, já que uma pessoa pode levá-lo para qualquer lugar e ser roubado ou perder o equipamento. Nisso, dados sensíveis armazenados no dispositivo podem ser vazados. Portanto, o conceito de *Bring Your Own Device* (BYOD) aumentou o potencial para tais ocorrências. Refere-se ao uso de dispositivos móveis de propriedade do funcionário para acessar o conteúdo ou redes empresariais [128] . Conceitos de portabilidade semelhantes são *Choose Your Own Device* (CYOD), *Company Owned and Personally Enabled* (COPE) e *Company Owned Business Only* (COBO), e todos levantam preocupações de segurança [129].

Os autores de [6] listam alguns dos desafios que esses dispositivos móveis trazem para hospitais, que também podem ser aplicáveis a empresas em geral. Eles categorizam esses desafios como relacionados à tecnologia, fatores humanos e políticas, conforme resumido na Tabela 4.11. Oportunamente, eles também fornecem possíveis soluções para esses desafios, como visto na Tabela 4.12, juntamente com alguns elementos-chave que devem estar presentes em uma Política de BYOD.

Tabela 4.11: Desafios de BYOD [6].

| Técnico | Pessoas | Política |
|---|---|---|
| Dispositivo inseguro | Comportamento inadequado | Falta de política |
| Ausência de bloqueio | Falta de consciência | Conformidade |
| Rede insegura | Experiência do usuário ruim | Sanções por vazamentos |
| Aplicativo suspeito instalado | Escassez de habilidades | |

Além do BYOD, o teletrabalho e os espaços de co-working podem representar uma ameaça à segurança das empresas. Esses modelos de trabalho, que aumentaram desde a pandemia de COVID-19 [130], também implicam novas lacunas de segurança semelhantes às relacionadas ao BYOD [131].

O *geofencing* é outro controle de segurança adequado para esse cenário, que se refere a ações acionadas em resposta a um dispositivo deixando uma geolocalização pré-definida [132]. Tais ações poderiam ser, por exemplo, desativar sua interface de rede ou limpar remotamente o dispositivo para evitar vazamentos de dados ao sair de uma área autorizada.

Sobre esse assunto, Uz [133] avaliou a eficácia da limpeza remota de dados, considerando que os dados excluídos podem ser recuperados forensicamente, conforme explicado na Seção 4.4.3.5.

A criptografia também é recomendada para proteção de dados e, para dispositivos móveis, a Criptografia Baseada em Arquivos (FBE) é obrigatória no Android desde sua 10ª versão [134], e, para notebooks, BitLocker e VeraCrypt são algumas das opções disponíveis [135].

Um método de controle de acesso apropriado para BYOD é o Controle de Acesso Baseado em Atributos

Tabela 4.12: Soluções de BYOD [6]

| Técnico | Pessoas | Política |
|---|---|---|
| Gerenciamento de Dispositivos Móveis | Cultura de Segurança | Estratégia e Governança de BYOD |
| Containerização | Conscientização e Treinamento | Acordo do Usuário |
| Gerenciamento de Identidade e Acesso | Melhoria de Habilidades | Política de BYOD |
| Ferramentas de Segurança para Endpoints | | |
| Plataformas de Comunicação Segura | | |

(ABAC). Segundo esse paradigma, a empresa pode negar e conceder acesso a uma identidade com base em atributos da solicitação, como localização, hora do dia e objeto acessado.

#### 4.4.3.5 PHYS, PORT, STAT: descartes

Ao descartar dados sensíveis, é preciso estar ciente da possibilidade de um adversário revirar a lixeira, o que é conhecido como *dumpster diving*, um ataque de engenharia social [136].

Com esse método, o atacante pode acessar qualquer objeto que a empresa tenha descartado, como equipamentos, documentos, anotações e contas. Supondo que haja dados sensíveis entre esse material descartado, como senhas anotadas de funcionários ou dados de clientes. Nesse caso, o ator malicioso terá mais informações para realizar o ataque ou possuir os dados violados.

Cabe ressaltar que, desde o caso California v. Greenwood em 1988, a legalidade da busca e apreensão sem mandado do lixo deixado em áreas públicas foi estabelecida nos EUA [137]. A busca em lixo em local público também não é tipificada no Brasil. Por causa disso, para mais uma camada de segurança contra vazamentos de dados, as empresas devem manter suas lixeiras trancadas em áreas privadas.

Como uma medida adicional contra essa abordagem, uma empresa deve, ao final do ciclo de vida dos dados, realizar uma adequada disposição das informações. Para isso, Políticas de Classificação de Dados e Disposição de Ativos devem ser implementadas e divulgadas para aumentar a conscientização dos funcionários. Para auxiliar na edição adequada dessas e de outras políticas, várias organizações de segurança respeitadas fornecem modelos de políticas, como o Instituto SANS[4] e CIS[5].

Antes do descarte, a mídia deve ser sanitizada, ou seja, ter seus dados tornados inacessíveis para um determinado nível de esforço do atacante, dependendo da classificação dos dados. Técnicas adequadas de sanitização de mídia são apresentadas pelo NIST SP 800-88, para diferentes tipos de mídia.

Para documentos em papel, por exemplo, o padrão estabelece que eles devem ser triturados em pedaços pequenos o suficiente para haver uma garantia razoável de que os dados não podem ser reconstruídos em

[4] <sans.org/information-security-policy/> acessado em 07 de abril de 2024
[5] <cisecurity.org/> acessado em 07 de abril de 2024

proporção à confidencialidade dos dados. Para dificultar ainda mais uma reconstrução maliciosa, documentos sensíveis podem ser misturados com papel público na entrada do triturador. Em relação ao tamanho dos pedaços triturados para cada nível de classificação, o padrão alemão DIN 66399 fornece diretrizes, algumas das quais estão resumidas na Tabela 4.13.

Tabela 4.13: Tamanhos de trituração de papel da norma alemã DIN 66399 de acordo com a sensibilidade dos dados.

| Nível de Classificação | Área máxima do pedaço ($mm^2$) |
|---|---|
| P-1 (menos sensível) | 2000 |
| P-2 | 800 |
| P-3 | 320 |
| P-4 | 160 |
| P-5 | 30 |
| P-6 | 10 |
| P-7 (mais sensível) | 5 |

Abordagens de disposição similares devem ser aplicadas a dispositivos digitais, como HDs, SSDs, pen drives e CD/DVD. Apesar de a destruição física e a trituração ainda serem possíveis para esses tipos de mídia, e de fato serem recomendadas para casos mais sensíveis, a natureza desses dispositivos permite outros mecanismos de apagamento, especialmente para os dados menos sensíveis.

É sabido que simplesmente excluir arquivos via sistema operacional não é uma maneira eficaz de purgar dados, pois técnicas de *carving* de dados podem recuperar esses arquivos [138]. Outras técnicas, como o preenchimento com zeros, em que todos os dados são sobrescritos com zeros, são eficazes contra mecanismos de recuperação de dados comumente disponíveis, de acordo com o NIST SP 800-88. Rondas adicionais de preenchimento podem ser realizadas para aumentar a segurança.

Especificamente para HDs magnéticos, a técnica de desmagnetização pode ser usada. Essa abordagem consiste em aplicar um campo magnético ao HD, o que muda os padrões magnéticos no dispositivo, consequentemente destruindo os dados [139].

A abordagem de desmagnetização não é eficaz para SSDs, que não são magnéticos. Para este tipo de mídia, uma maneira segura de lidar com a remanência de dados é a cripto-trituração, também chamada de cripto-apagamento (em inglês, *crypto shredding*). Neste procedimento, os dados armazenados no dispositivo são criptografados com um algoritmo seguro e, em seguida, a chave de descriptografia é descartada, tornando os dados irrecuperáveis [140].

Além das disposições, essas técnicas de sanitização também devem ser aplicadas ao doar ou vender os dispositivos, se a sensibilidade dos dados permitir a transferência da propriedade da mídia.

#### 4.4.3.6 PHYS, PORT, STAT: furtos e acessos inadequados

Esta seção discute principalmente aspectos de segurança física que podem ser implementados em uma instalação da empresa para evitar uma violação de dados. Entendemos que as abordagens de mitigação relacionadas aos furtos dos ativos da empresa em posse de um funcionário fora das instalações da empresa

são abordadas na Seção 4.4.3.4.

O projeto de segurança física em uma empresa começa na fase de arranjo arquitetônico da construção da instalação [141] . Estratégias de Prevenção ao Crime por Meio do Projeto Ambiental (em inglês, *Crime Prevention Through Environmental Design*, CPTED) podem ser empregadas durante esta fase. Por meio dessa abordagem, criminosos são dissuadidos e mais facilmente detectados pelo layout físico do espaço [142].

Os principais princípios de CPTED são vigilância natural, controle de acesso, reforço territorial e manutenção [143]. Como exemplo, é recomendado que cercas tenham pelo menos 1 metro de altura para dissuadir invasores casuais e pelo menos 2,5 metros de altura para dissuadir infiltradores intencionais [144].

Outros controles físicos devem ser implementados para prevenir incidentes, especialmente em áreas mais sensíveis, como centros de dados. Exemplos incluem o uso de um controle de acesso físico por PIN (controle preventivo), guardas (dissuasores), câmeras de segurança (detectivas) e alarmes (corretivos).

No entanto, todas essas medidas de segurança serão inúteis se o fator humano for explorado com sucesso. Um adversário pode, por exemplo, se infiltrar furtivamente por uma porta aberta por pessoal autorizado, uma prática conhecida como *tailgating*, ou podem convencer alguém a deixá-los entrar, por exemplo, dizendo que esqueceram o crachá e estão com pressa. Este último é uma tática de engenharia social conhecida como *piggybacking*. Uma contramedida eficaz para essas intrusões é o uso de uma câmara de acesso (em inglês, *mantrap*).

No entanto, os intrusos nem sempre perpetrarão incidentes físicos. Visitantes autorizados, por exemplo, podem realizar atividades não autorizadas e, nesse caso, medidas físicas adicionais devem ser implementadas.

Uma possível lacuna é a observação direta do teclado dos dispositivos. Nesses casos, um adversário pode obter dados sensíveis, como senhas, através de técnica conhecida como *sholder surffing*. Para dificultar essa atividade, pode ser necessário relocar os dispositivos.

Outra medida de segurança em relação ao local de trabalho do funcionário é a implementação de uma Política de Mesa Limpa, que obriga que todas as mesas dentro da empresa estejam livres de objetos e documentos. Após a implementação bem-sucedida dessa política, um intruso não pode roubar um documento sensível da mesa de um trabalhador.

O Capítulo 15 do NIST SP 800-12 revisa outras práticas de segurança física. A segurança física, que visa proteger a integridade física, vida e saúde das pessoas, é outro tópico relevante nesta discussão. No entanto, como esses incidentes geralmente não resultam em vazamentos de dados, que são o foco deste trabalho, não os incluímos na discussão.

#### 4.4.3.7 DISC

A divulgação não intencional pode ser classificada como resultado de uma ameaça interna não intencional, seja devido a negligência ou imprudência [124].

Um programa eficaz de Educação, Treinamento e Conscientização em Segurança (em inglês, *Security*

Tabela 4.14: Elementos do SETA [7].

| Segurança | Propósito | Método de entrega exemplar | Público-alvo | Nível |
|---|---|---|---|---|
| Educação | Capacitar funcionários com aprendizado profundo sobre conhecimento e habilidades de segurança | Simulações de ataques cibernéticos | Pessoal de TI | Alto |
| Treinamento | Construir conhecimento e habilidades de segurança dos funcionários | Aulas e webinars | Todos os funcionários | Intermediário |
| Conscientização | Chamar a atenção dos funcionários para a segurança | Pôsteres, banners | Todos os funcionários | Básico |

*Education, Training and Awareness*, SETA) pode ser capaz de reduzir a incidência desses casos e promover a conformidade em uma organização [145] . Cada elemento do SETA é melhor caracterizado na Tabela 4.14.

O estudo de Alyami et al. [146] avaliou os fatores críticos para implantar um programa de SETA bem-sucedido com base em uma pesquisa com 65 respondentes. Eles produziram uma lista classificada dos fatores essenciais de sucesso. A gamificação também é vista como uma maneira razoável de melhorar o engajamento no programa [147].

A partir da Tabela 4.14, pode-se observar que os programas de conscientização em segurança são úteis para mitigar os riscos associados à utilização geral de recursos tecnológicos pelo usuário comum, como comprometimento de credenciais e ataques de engenharia social, como phishing [148] Por outro lado, o treinamento e a educação em segurança focam na prevenção de incidentes cibernéticos baseados em vulnerabilidades técnicas, como falhas originadas de má configuração, e devem ser direcionados ao pessoal de TI [149].

De acordo com PCR (Tabela 3.1), a categorização do tipo DISC inclui informações publicadas publicamente enviadas para a parte errada. Além do SETA, um controle de duas pessoas também pode reduzir a probabilidade dessas divulgações. Com essa abordagem, duas pessoas devem autorizar uma ação antes de efetivar sua execução [150].

### 4.4.4 Contenção, recuperação e resposta

A empresa afetada deve estudar uma estratégia de resposta após um atacante e uma violação de dados terem contornado os controles de segurança. Do ponto de vista técnico, a empresa deve conter rapidamente o vazamento de dados para minimizar os danos potenciais e, em seguida, identificar e erradicar os componentes do incidente. O NIST SP 800-61 fornece um guia mais detalhado para o tratamento de incidentes de segurança de computadores.

Esta publicação do NIST divide o processo de resposta a incidentes em cinco etapas: Preparação, que ocorre antes de um incidente e corresponde a medidas de segurança preventivas; Detecção e Análise, na qual o vetor de ataque e as TTPs são identificados; Contenção, erradicação e recuperação, que compreendem uma restrição inicial da atividade maliciosa, uma subsequente limpeza de artefatos maliciosos (embora mantendo-os para análise forense), seguida pela Restauração da operação dos sistemas. Finalmente, as atividades pós-incidente incluem discutir e documentar o incidente para entendê-lo melhor e prevenir intrusões futuras semelhantes.

Para uma melhor compreensão das causas do incidente e das eventuais modificações no sistema feitas pelo intruso, ferramentas e técnicas forenses podem ser úteis [151]. O NIST SP 800-86 fornece diretrizes para integrar técnicas forenses na resposta a incidentes, incluindo coleta, exame e relatório de dados de diferentes fontes, como arquivos, sistemas operacionais, redes e aplicativos. Ao realizar a perícia digital, é importante manter uma cadeia de custódia e preservar a integridade da evidência [152].

Especificamente no domínio de vazamento de dados, Rabello et al. [153] propuseram uma metodologia de resposta a vazamentos de dados com base na ISO 27035 e NIST 800-61. Seu estudo enfatiza a importância de automatizar esse processo, especialmente devido ao curto tempo necessário para notificar uma violação para a conformidade legislativa.

Hillmann et al. [154] realizaram doze entrevistas com clientes sobre suas expectativas em relação a uma resposta a vazamentos de dados. Eles concluíram que a expectativa varia com vários fatores, como gravidade da violação, tipo de vazamento de dados e setor da empresa. Portanto, uma empresa deve adaptar sua estratégia de resposta ao cenário específico para maximizar a chance de corresponder às expectativas de seus clientes.

Como guia mais geral, a Comissão Federal de Comércio estadunidense (em inglês, *Federal Trade Commission*, FTC) apresenta recomendações para uma resposta adequada a vazamentos de dados[6], como montar uma equipe de resposta a incidentes, corrigir as vulnerabilidades e remover informações publicadas incorretamente na web.

Para notificação, especialmente para cumprir a legislação mencionada na Seção 4.2, a FTC menciona a importância de notificar a polícia e empresas e indivíduos afetados, especificando o que aconteceu, quais informações foram roubadas, como os atacantes usaram as informações, que medidas de remediação foram tomadas e como os clientes podem entrar em contato com a organização em relação à violação. Eles também fornecem um modelo de carta para notificação de violação de dados.

[6] <www.ftc.gov/business-guidance/resources/data-breach-response-guide-business> acessado em 07 de abril de 2024

# 5 CONCLUSÃO

Vazamentos de dados são uma ameaça crescente para organizações de todos os tamanhos e setores. A análise abrangente desses incidentes melhora a compreensão do cenário em evolução das ameaças de segurança cibernética globalmente. Esta pesquisa revelou observações que aprimoram o entendimento sobre os padrões e implicações dos vazamentos de dados.

O estudo identificou que o setor financeiro dos EUA emergiu como o principal alvo para atores maliciosos. Isso enfatiza a necessidade crítica de implementar medidas robustas de segurança cibernética dentro da indústria. Compreender as vulnerabilidades que tornam o setor financeiro suscetível a vazamentos é primordial para proteger os dados sensíveis manipulados dentro dessas organizações. Assim, também foram mostradas as causas mais comuns para vazamentos nesse setor econômico.

Não apenas no setor financeiro, identificou-se também que incidentes relacionados a dispositivos portáteis (PORT) e a atores maliciosos externos (HACK) foram os tipos mais prevalentes de vazamentos. Isso destaca a importância de as organizações adotarem contramedidas proativas para proteger seus dados e mitigar o risco associados a esses vetores.

Esta investigação identificou, também, várias relações fortes entre países e setores econômicos afetados por vazamentos de dados, como a Índia e o setor médico ou de saúde; e o Brasil e o setor governamental ou militar. Uma correlação com os aspectos regulatórios desses países e regiões pode promover o reconhecimento das causas raiz das vazamentos de dados e das estratégias de mitigação.

Os resultados também apoiam o processo de gestão de risco de vazamentos de dados, avaliando a probabilidade e os impactos desses incidente.

É relevante mencionar, no entanto, que essas descobertas dependem muito do conjunto de dados utilizado. Apesar da metodologia aparentemente robusta na coleta de dados, as ameaças potenciais à sua validade devem ser consideradas. Essas limitações podem prejudicar a representação da realidade. Além disso, a discussão acerca dos controles de segurança não tem a intenção de apresentar os mecanismos de mitigação de maneira exaustiva, uma vez que diversas estratégias de segurança podem ser adotadas.

## 5.1 LIMITAÇÕES E AMEAÇAS À VALIDADE

As conclusões deste estudo enfrentam ameaças à sua validade, que podem ser classificadas em quatro tipos, de acordo com [155].

**Validade externa**. Os conjuntos de dados podem ter sido gerados com foco em incidentes que ocorreram em regiões específicas, como países ou estados com economias maiores ou que possam ser mais representativos na mídia. Violações menos divulgadas que atendem aos critérios do conjunto de dados podem ter sido omitidas. Embora o conjunto de dado global incorpore fontes de vários idiomas [22], a barreira linguística também representa uma potencial ameaça externa à validade. Essa limitação pode

resultar em uma cobertura menos abrangente de idiomas menos frequentemente falados, levando a uma sub-representação potencial de vazamentos nesses cenários linguísticos.

**Validade interna**. Os padrões observados podem estar especificamente relacionados aos intervalos de tempo limitados abrangidos pelos conjuntos de dados e não representar o cenário geral de vazamentos de dados nos países estudados. Mudanças nas leis de proteção de dados também podem alterar os padrões desses incidentes ao longo do tempo. Além disso, a relação causal entre o nível de regulamentação em um país e a frequência e impacto de uma violação de dados dentro de sua jurisdição nem sempre é clara, pois outros fatores, como tamanho da população, também podem influenciar as métricas desses incidentes.

**Validade de construção**. Os países têm definições diferentes para dados pessoais e o que constitui uma violação de dados. Isso ameaça a consistência dos incidentes entre regiões. Além disso, uma regulamentação de proteção de dados é rica em detalhes que não foram explorados profundamente, o que pode afetar a comparabilidade entre os aspectos regulatórios discutidos e as violações de dados relacionadas.

**Validade de conclusão**. O conjunto de dados pode sofrer de subnotificação, especialmente em países onde a notificação de vazamento de dados não era legalmente exigida entre 2018 e 2019.

Apesar dessas ameaças, este estudo promove a discussão sobre conformidade e proteção de dados nos países afetados durante o intervalo de tempo dos conjuntos de dados.

## 5.2 TRABALHOS FUTUROS

Como trabalho futuro, sugere-se a previsão, por meio de modelos de séries temporais, da ocorrência de incidentes de vazamento de dados e da quantidade de dados vazados nesses conjuntos de dados específicos. Esse estudo poderia evidenciar padrões e tendências desses incidentes ao longo do tempo, auxiliando na melhor compreensão, planejamento estratégico e prevenção de segurança cibernética.

Adicionalmente, este estudo focou a análise exploratória dos dados referente aos incidentes de uma maneira geral no mundo e filtrada nos EUA para empresas da bolsa de valores. Trabalhos futuros poderiam explorar a incidência de incidentes cibernéticos com outros filtros, como em setores da indústria de infraestrutura crítica, por exemplo no setor de energia elétrica. Esse trabalho futuro pode incluir um dashboard configurável, com visualização de estatísticas referentes aos incidentes cibernéticos na indústria com filtros customizáveis pelo usuário. O uso de técnicas de Aprendizado de Máquina, como o Aprendizado Federado, também pode enriquecer a análise desses dados.

Sugere-se, ainda, como trabalho futuro, uma análise exploratória dos dados vazados de um incidente específico, com a finalidade de explorar o potencial de violação à privacidade desse tipo de exploração cibernética.

Por fim, é proposto um aprofundamento da análise do impacto de vazamentos de dados na bolsa de uma empresa, com uso de técnicas econométricas, como análise de volatilidade e mudança de regimes de Markov, e também avaliar a variação do impacto na bolsa em função de diferentes respostas dadas pelas empresas afetadas.

# REFERÊNCIAS


1 ASLAM, M.; ABBASI, M. A. K.; KHALID, T.; SHAN, R. U.; ULLAH, S.; AHMAD, T.; SAEED, S.; ALABBAD, D. A.; AHMAD, R. Getting smarter about smart cities: Improving data security and privacy through compliance. *Sensors*, MDPI, v. 22, n. 23, p. 9338, 2022.

2 PIPER, D. *DATA PROTECTION LAWS OF THE WORLD Full Handbook*. [S.l.], 2023.

3 COIE, P. *Security Breach Notification Chart*. 2014.

4 POTTER, A.; CAMPBELL, K.; BALDIN, A.; CHAMBERS, H.; TOTO, B.; SATURNINO, F.; PRESCOTT, V. *COMPARING COMPREHENSIVE US PRIVACY LAWS: A GUIDE TO COMPLIANCE*. [S.l.], 2023.

5 IBM. *Cost of a Data Breach Report*. [S.l.], 2023.

6 WANI, T. A.; MENDOZA, A.; GRAY, K. Hospital bring-your-own-device security challenges and solutions: systematic review of gray literature. *JMIR mHealth and uHealth*, JMIR Publications Inc., Toronto, Canada, v. 8, n. 6, p. e18175, 2020.

7 HU, S.; HSU, C.; ZHOU, Z. Security education, training, and awareness programs: Literature review. *Journal of Computer Information Systems*, Taylor & Francis, v. 62, n. 4, p. 752–764, 2022.

8 RASHID, F. Y. The rise of confidential computing: Big tech companies are adopting a new security model to protect data while it's in use-[news]. *IEEE Spectrum*, IEEE, v. 57, n. 6, p. 8–9, 2020.

9 MIKALEF, P.; PAPPAS, I. O.; KROGSTIE, J.; PAVLOU, P. A. *Big data and business analytics: A research agenda for realizing business value*. [S.l.]: Elsevier, 2020. 103237 p.

10 DESETTI, S. S.; GHOSH, I. Prediction and deeper analysis of market fear in pre-covid-19, covid-19 and russia-ukraine conflict: A comparative study of facebook prophet, uber orbit and explainable ai. In: SPRINGER. *International Conference on Computational Intelligence in Communications and Business Analytics*. [S.l.], 2023. p. 213–227.

11 JUMA'H, A. H.; ALNSOUR, Y. The effect of data breaches on company performance. *International Journal of Accounting & Information Management*, Emerald Publishing Limited, v. 28, n. 2, p. 275–301, 2020.

12 KUIPERS, S.; SCHONHEIT, M. Data breaches and effective crisis communication: a comparative analysis of corporate reputational crises. *Corporate Reputation Review*, Springer, v. 25, n. 3, p. 176–197, 2022.

13 BURNES, D.; DELIEMA, M.; LANGTON, L. Risk and protective factors of identity theft victimization in the united states. *Preventive medicine reports*, Elsevier, v. 17, p. 101058, 2020.

14 ISSA, W. B.; AKOUR, I. A.; IBRAHIM, A.; ALMARZOUQI, A.; ABBAS, S.; HISHAM, F.; GRIFFITHS, J. Privacy, confidentiality, security and patient safety concerns about electronic health records. *International nursing review*, Wiley Online Library, v. 67, n. 2, p. 218–230, 2020.

15 ALI, S. E. A.; LAI, F.-W.; DOMINIC, P.; BROWN, N. J.; LOWRY, P. B. B.; ALI, R. F. Stock market reactions to favorable and unfavorable information security events: A systematic literature review. *Computers & Security*, Elsevier, v. 110, p. 102451, 2021.

16 RUSTAD, M. L.; KOENIG, T. H. Towards a global data privacy standard. *Fla. L. Rev.*, HeinOnline, v. 71, p. 365, 2019.

17 MARTINS, A. D. F.; BARROS, P. V. da S.; MONTEIRO, J. M.; MACHADO, J. de C. Lgpd: a formal concept analysis and its evaluation. In: SBC. *Anais do XXXV Simpósio Brasileiro de Bancos de Dados*. [S.l.], 2020. p. 259–264.

18 BLANKE, J. M. Protection for 'inferences drawn': A comparison between the general data protection regulation and the california consumer privacy act. *Global Privacy Law Review*, v. 1, n. 2, 2020.

19 SHUAIB, M.; ALAM, S.; ALAM, M. S.; NASIR, M. S. Compliance with hipaa and gdpr in blockchain-based electronic health record. *Materials Today: Proceedings*, Elsevier, 2021.

20 COCO, A.; DIAS, T. de S. 'cyber due diligence': A patchwork of protective obligations in international law. *European Journal of International Law*, Oxford University Press UK, v. 32, n. 3, p. 771–806, 2021.

21 ALAHMARI, A.; DUNCAN, B. Cybersecurity risk management in small and medium-sized enterprises: A systematic review of recent evidence. In: IEEE. *2020 international conference on cyber situational awareness, data analytics and assessment (CyberSA)*. [S.l.], 2020. p. 1–5.

22 NETO, N. N.; MADNICK, S.; PAULA, A. M. G. D.; BORGES, N. M. Developing a global data breach database and the challenges encountered. *Journal of Data and Information Quality (JDIQ)*, ACM New York, NY, USA, v. 13, n. 1, p. 1–33, 2021.

23 ROSATI, P.; LYNN, T. A dataset for accounting, finance and economics research on us data breaches. *Data in Brief*, Elsevier, v. 35, p. 106924, 2021.

24 VARSHNEY, S.; MUNJAL, D.; BHATTACHARYA, O.; SABOO, S.; AGGARWAL, N. Big data privacy breach prevention strategies. In: IEEE. *2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*. [S.l.], 2020. p. 1–6.

25 PETKAUSKAS, V. *Mother of all breaches reveals 26 Billion Records*. 2024. Disponível em: <https://cybernews.com/security/billions-passwords-credentials-leaked-mother-of-all-breaches>.

26 RODRIGUES, G. A. P.; SERRANO, A. L. M.; LEMOS, A. N. L. E.; CANEDO, E. D.; MENDONÇA, F. L. L. d.; ALBUQUERQUE, R. de O.; OROZCO, A. L. S.; VILLALBA, L. J. G. Understanding data breach from a global perspective: Incident visualization and data protection law review. *Data*, MDPI, v. 9, n. 2, p. 27, 2024.

27 RODRIGUES, G. A. P.; SERRANO, A. L. M.; ALBUQUERQUE, R. de O.; SAIKI, G. M.; RIBEIRO, S. S.; OROZCO, A. L. S.; VILLALBA, L. J. G. Mapping of data breaches in companies listed on the nyse and nasdaq: Insights and implications. *Results in Engineering*, Elsevier, p. 101893, 2024.

28 RODRIGUES, G. A. P.; SERRANO, A. L. M.; VERGARA, G. F.; ALBUQUERQUE, R. d. O.; NZE, G. D. A. Impact, compliance, and countermeasures in relation to data breaches in publicly traded us companies. *Future Internet*, Multidisciplinary Digital Publishing Institute, v. 16, n. 6, p. 201, 2024.

29 OBAYDIN, I.; XU, L.; ZURBRUEGG, R. The unintended cost of data breach notification laws: Evidence from managerial bad news hoarding. *Available at SSRN 3926962*, 2021.

30 ZADEH, A.; LAVINE, B.; ZOLBANIN, H. M.; HOPKINS, D. A cybersecurity risk quantification and classification framework for informed risk mitigation decisions. *Decision Analytics Journal*, Elsevier, p. 100328, 2023.

31 BENZELL, S.; HERSH, J. S.; ALSTYNE, M. W. V.; LAGARDA, G. How apis create growth by inverting the firm. *Available at SSRN 3432591*, 2022.

32 NIYONZIGIRA, F. *Exploring Nonprofit Organizations' Successful Compliance Strategies Against Cyber Threats: A Qualitative Study Inquiry*. Tese (Doutorado) — Capella University, 2023.

33 SHABTAI, A.; ELOVICI, Y.; ROKACH, L.; SHABTAI, A.; ELOVICI, Y.; ROKACH, L. *Data leakage detection/prevention solutions*. [S.l.]: Springer, 2012.

34 CHOI, S. J.; JOHNSON, M. E.; LEHMANN, C. U. Data breach remediation efforts and their implications for hospital quality. *Health services research*, Wiley Online Library, v. 54, n. 5, p. 971–980, 2019.

35 MASUCH, K.; GREVE, M.; TRANG, S. What to do after a data breach? examining apology and compensation as response strategies for health service providers. *Electronic Markets*, Springer, v. 31, n. 4, p. 829–848, 2021.

36 PARK, J.; SHIN, W.; KIM, B.; KIM, M. Spillover effects of data breach on consumer perceptions: evidence from the e-commerce industry. *Internet Research*, Emerald Publishing Limited, 2024.

37 REZAEE, Z.; ZHOU, G.; BU, L. L. Corporate social irresponsibility and the occurrence of data breaches: A stakeholder management perspective. *International Journal of Accounting Information Systems*, Elsevier, v. 53, p. 100677, 2024.

38 SUN, H.; XU, M.; ZHAO, P. A multivariate frequency-severity framework for healthcare data breaches. *The Annals of Applied Statistics*, Institute of Mathematical Statistics, v. 17, n. 1, p. 240–268, 2023.

39 CULNAN, M. J.; WILLIAMS, C. C. How ethics can enhance organizational privacy: lessons from the choicepoint and tjx data breaches. *MIS quarterly*, JSTOR, p. 673–687, 2009.

40 SEN, R.; BORLE, S. Estimating the contextual risk of data breach: An empirical approach. *Journal of Management Information Systems*, Taylor & Francis, v. 32, n. 2, p. 314–341, 2015.

41 BISPO, G. D.; VERGARA, G. F.; SAIKI, G. M.; MARTINS, P. H. d. S.; COELHO, J. G.; RODRIGUES, G. A. P.; OLIVEIRA, M. N. d.; MOSQUÉRA, L. R.; GONÇALVES, V. P.; NEUMANN, C. et al. Automatic literature mapping selection: Classification of papers on industry productivity. *Applied Sciences*, MDPI, v. 14, n. 9, p. 3679, 2024.

42 SCHLACKL, F.; LINK, N.; HOEHLE, H. Antecedents and consequences of data breaches: A systematic review. *Information & Management*, Elsevier, v. 59, n. 4, p. 103638, 2022.

43 SYED, R. Enterprise reputation threats on social media: A case of data breach framing. *The Journal of Strategic Information Systems*, Elsevier, v. 28, n. 3, p. 257–274, 2019.

44 HUANG, H. H.; WANG, C. Do banks price firms' data breaches? *The Accounting Review*, American Accounting Association, v. 96, n. 3, p. 261–286, 2021.

45 THOMAS, K.; LI, F.; ZAND, A.; BARRETT, J.; RANIERI, J.; INVERNIZZI, L.; MARKOV, Y.; COMANESCU, O.; ERANTI, V.; MOSCICKI, A. et al. Data breaches, phishing, or malware? understanding the risks of stolen credentials. In: *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. [S.l.: s.n.], 2017. p. 1421–1434.

46 MCLEOD, A.; DOLEZEL, D. Cyber-analytics: Modeling factors associated with healthcare data breaches. *Decision Support Systems*, Elsevier, v. 108, p. 57–68, 2018.

47 PARK, S. Why information security law has been ineffective in addressing security vulnerabilities: Evidence from california data breach notifications and relevant court and government records. *International Review of Law and Economics*, Elsevier, v. 58, p. 132–145, 2019.

48 RONQUILLO, J. G.; WINTERHOLLER, J. E.; CWIKLA, K.; SZYMANSKI, R.; LEVY, C. Health it, hacking, and cybersecurity: national trends in data breaches of protected health information. *JAMIA open*, Oxford University Press, v. 1, n. 1, p. 15–19, 2018.

49 TSEN, E.; KO, R.; SLAPNICAR, S. Dataset of data breaches and ransomware attacks over 15 years from 2004. *The University of Queensland*, 2020.

50 BIDDLE, N.; EDWARDS, B.; GRAY, M.; MCEACHERN, S. Anu poll 2018: Data governance. *ADA Dataverse*, 2020. Disponível em: <http://dx.doi.org/10.26193/XHORAI>.

51 IKEGAMI, K.; KIKUCHI, H. Modeling the risk of data breach incidents at the firm level. In: SPRINGER. *Innovative Mobile and Internet Services in Ubiquitous Computing: Proceedings of the 14th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS-2020)*. [S.l.], 2021. p. 135–148.

52 LEE, J.; LEE, C.-F. Data collection, presentation, and yahoo! finance. In: *Essentials of Excel VBA, Python, and R: Volume I: Financial Statistics and Portfolio Analysis*. [S.l.]: Springer, 2023. p. 19–80.

53 ROSATI, P.; LYNN, T. Corrigendum to "a dataset for accounting, finance and economics research on us data breaches"[data in brief 35 (2021) 1–6/106924]. *Data in Brief*, Elsevier, v. 40, 2022.

54 WU, E. Sovereignty and data localization. *Belfer Center for Science and International Affairs, Harvard Kennedy School, Cambridge, MA.*, 2021.

55 GEORGE, D. A. S.; GEORGE, A. H. Potential risk: Hosting cloud services outside the country. *International Journal of Advanced Research in Computer and Communication Engineering*, v. 11, n. 4, p. 5–11, 2022.

56 SAMPSON, D.; CHOWDHURY, M. M. The growing security concerns of cloud computing. In: IEEE. *2021 IEEE International Conference on Electro Information Technology (EIT)*. [S.l.], 2021. p. 050–055.

57 THANTILAGE, R. D.; LE-KHAC, N.-A.; KECHADI, M.-T. Healthcare data security and privacy in data warehouse architectures. *Informatics in Medicine Unlocked*, Elsevier, p. 101270, 2023.

58 KOCH, R. Hidden in the shadow: The dark web-a growing risk for military operations? In: IEEE. *2019 11th International Conference on Cyber Conflict (CyCon)*. [S.l.], 2019. v. 900, p. 1–24.

59 HABER, M. J.; CHAPPELL, B.; HILLS, C. Regulatory compliance. In: *Cloud Attack Vectors: Building Effective Cyber-Defense Strategies to Protect Cloud Resources*. [S.l.]: Springer, 2022. p. 297–373.

60 MCCOY, T. H.; PERLIS, R. H. Temporal trends and characteristics of reportable health data breaches, 2010-2017. *Jama*, American Medical Association, v. 320, n. 12, p. 1282–1284, 2018.

61 CHURI, P.; PAWAR, A.; MORENO-GUERRERO, A.-J. A comprehensive survey on data utility and privacy: Taking indian healthcare system as a potential case study. *Inventions*, MDPI, v. 6, n. 3, p. 45, 2021.

62 DHAGARRA, D.; GOSWAMI, M.; KUMAR, G. Impact of trust and privacy concerns on technology acceptance in healthcare: an indian perspective. *International journal of medical informatics*, Elsevier, v. 141, p. 104164, 2020.

63 FERRÃO, S. É. R.; CARVALHO, A. P.; CANEDO, E. D.; MOTA, A. P. B.; COSTA, P. H. T.; CERQUEIRA, A. J. Diagnostic of data processing by brazilian organizations—a low compliance issue. *Information*, MDPI, v. 12, n. 4, p. 168, 2021.

64 LIMA, R. C.; SILVA, P. F.; RUDZIT, G. No power vacuum: national security neglect and the defence sector in brazil. *Defence Studies*, Taylor & Francis, v. 21, n. 1, p. 84–106, 2021.

65 SHIRES, J. The simulation of scandal: Hack-and-leak operations, the gulf states, and us politics (fall 2020). *Texas National Security Review*, Texas National Security Review, 2020.

66 HOOFNAGLE, C. J.; SLOOT, B. V. D.; BORGESIUS, F. Z. The european union general data protection regulation: what it is and what it means. *Information & Communications Technology Law*, Taylor & Francis, v. 28, n. 1, p. 65–98, 2019.

67 SHASTRI, S.; WASSERMAN, M.; CHIDAMBARAM, V. The seven sins of {Personal-Data} processing systems under {GDPR}. In: *11th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 19)*. [S.l.: s.n.], 2019.

68 CHATTERJEE, C.; SOKOL, D. D. Data security, data breaches, and compliance. *Cambridge Handbook on Compliance (D. Daniel Sokol & Benjamin van Rooij editors, Cambridge University Press, forthcoming)*, 2019.

69 SARLET, G. B. S.; RODRIGUEZ, D. P. Alternatives for an adequate structuring of the national data protection authority (anpd) in its independent profile: proposals to overcome the technological challenges in the age of digital governance. *International Cybersecurity Law Review*, Springer, p. 1–15, 2023.

70 SRINIVASAN, S.; SINHA, V.; MODI, S. Drafting a pro-antitrust and data protection regulatory framework. *Indian Public Policy Review*, v. 4, n. 5 (Sep-Oct), p. 35–56, 2023.

71 SHETH, S.; KAISER, G.; MAALEJ, W. Us and them: a study of privacy requirements across north america, asia, and europe. In: *Proceedings of the 36th International Conference on Software Engineering*. [S.l.: s.n.], 2014. p. 859–870.

72 ETTELDORF, C. Germany revisited: The second data protection adaption and implementation act. *Eur. Data Prot. L. Rev.*, HeinOnline, v. 5, p. 397, 2019.

73 MAHIEU, R.; ASGHARI, H.; PARSONS, C.; HOBOKEN, J. van; CRETE-NISHIHATA, M.; HILTS, A.; ANSTIS, S. Measuring the brussels effect through access requests: Has the european general data protection regulation influenced the data protection rights of canadian citizens? *Journal of Information Policy*, Pennsylvania State University Press, v. 11, p. 301–349, 2021.

74 FINCK, M.; PALLAS, F. They who must not be identified—distinguishing personal from non-personal data under the gdpr. *International Data Privacy Law*, Oxford University Press, v. 10, n. 1, p. 11–36, 2020.

75 SEVİNÇ, İ.; KARABULUT, N. A review on the personal data protection authority of turkey. *Akademik Hassasiyetler*, Hüzeyfe Süleyman ARSLAN, v. 7, n. 13, p. 449–472, 2020.

76 BOTTA, M.; WIEDEMANN, K. The interaction of eu competition, consumer, and data protection law in the digital economy: the regulatory dilemma in the facebook odyssey. *The Antitrust Bulletin*, SAGE Publications Sage CA: Los Angeles, CA, v. 64, n. 3, p. 428–446, 2019.

77 CICLOSI, F.; MASSACCI, F. The data protection officer: a ubiquitous role that no one really knows. *IEEE Security & Privacy*, IEEE, v. 21, n. 1, p. 66–77, 2022.

78 AMIR, E.; LEVI, S.; LIVNE, T. Do firms underreport information on cyber-attacks? evidence from capital markets. *Review of Accounting Studies*, Springer, v. 23, p. 1177–1206, 2018.

79 SEBASTIAN, G. Could incorporating cybersecurity reporting into sox have prevented most data breaches at us publicly traded companies? an exploratory study. *International Cybersecurity Law Review*, Springer, v. 3, n. 2, p. 367–383, 2022.

80 PANG, M.-S.; TANRIVERDI, H. Strategic roles of it modernization and cloud migration in reducing cybersecurity risks of organizations: The case of us federal government. *The Journal of Strategic Information Systems*, Elsevier, v. 31, n. 1, p. 101707, 2022.

81 MOORE, W.; FRYE, S. Review of hipaa, part 2: limitations, rights, violations, and role for the imaging technologist. *Journal of nuclear medicine technology*, Soc Nuclear Med, v. 48, n. 1, p. 17–23, 2020.

82 COHEN, B.; HU, A.; PATINO, D.; COFFMAN, J. Educational data in the cloud legal implications and technical recommendations. In: IEEE. *2022 IEEE/ACM 15th International Conference on Utility and Cloud Computing (UCC)*. [S.l.], 2022. p. 181–182.

83 RYLE, P.; YAN, J.; GARDINER, L. R. Gramm-leach-bliley gets a systems upgrade: What the ftc's proposed safeguards rule changes mean for small and medium american financial institutions. *EDPACS*, Taylor & Francis, v. 65, n. 2, p. 6–17, 2022.

84 SKOWRONSKI, D. S. Coppa and educational technologies: The need for additional online privacy protections for students. *Georgia State University Law Review*, v. 38, n. 4, p. 12, 2022.

85 ROBINSON, P. Can pci dss 4.0 reverse the decline in compliance? *Computer Fraud & Security*, MA Business London, v. 2022, n. 6, 2022.

86 GREENWOOD, B. N.; VAALER, P. M. Do us state breach notification laws decrease firm data breaches? *Minnesota Legal Studies Research Paper*, 2023.

87 BAIK, J. S. Data privacy against innovation or against discrimination?: The case of the california consumer privacy act (ccpa). *Telematics and Informatics*, v. 52, 2020.

88 ISLAM, R. The impact of data breaches on stock performance. *Glucksman Inst. for Res. in Securities Markets, Leonard N. Stern School of Bus., New York Univ. New York, USA*, 2020.

89 REIDENBACH, M.; WANG, P. Heartland payment systems: cybersecurity impact on audits and financial statement contingencies. *Issues in Accounting Education*, American Accounting Association, v. 36, n. 2, p. 93–109, 2021.

90 KLAUS, T.; ELZWEIG, B. The impact of data breaches on corporations and the status of potential regulation and litigation. *Law and Financial Markets Review*, Taylor & Francis, v. 14, n. 4, p. 255–260, 2020.

91 STEVENS, G. M. *Data security breach notification laws*. [S.l.]: Congressional Research Service Washington, DC, 2012.

92 COHEN, D. T.; HATCHARD, G. W.; WILSON, S. G. et al. *Population trends in incorporated places: 2000 to 2013*. [S.l.]: US Department of Commerce, Economics and Statistics Administration, US ..., 2015.

93 LAYTON, R.; WATTERS, P. A. A methodology for estimating the tangible cost of data breaches. *Journal of Information Security and Applications*, Elsevier, v. 19, n. 6, p. 321–330, 2014.

94 POORNACHANDRAN, P.; NITHUN, M.; PAL, S.; ASHOK, A.; AJAYAN, A. Password reuse behavior: How massive online data breaches impacts personal data in web. In: SPRINGER. *Innovations in Computer Science and Engineering: Proceedings of the Third ICICSE, 2015*. [S.l.], 2016. p. 199–210.

95 ROBERTS, S. Learning lessons from data breaches. *Network Security*, MA Business London, v. 2018, n. 11, p. 8–11, 2018.

96 MINKUS, T.; ROSS, K. W. I know what you're buying: Privacy breaches on ebay. In: SPRINGER. *Privacy Enhancing Technologies: 14th International Symposium, PETS 2014, Amsterdam, The Netherlands, July 16-18, 2014. Proceedings 14*. [S.l.], 2014. p. 164–183.

97 HAMMOUCHI, H.; CHERQI, O.; MEZZOUR, G.; GHOGHO, M.; KOUTBI, M. E. Digging deeper into data breaches: An exploratory data analysis of hacking breaches over time. *Procedia Computer Science*, Elsevier, v. 151, p. 1004–1009, 2019.

98 KUBIAK, I.; BOITAN, A.; HALUNGA, S. Assessing the security of tempest fonts against electromagnetic eavesdropping by using different specialized receivers. *Applied Sciences*, MDPI, v. 10, n. 8, p. 2828, 2020.

99 SULISTYOWATI, D.; HANDAYANI, F.; SURYANTO, Y. Comparative analysis and design of cybersecurity maturity assessment methodology using nist csf, cobit, iso/iec 27002 and pci dss. *JOIV: International Journal on Informatics Visualization*, v. 4, n. 4, p. 225–230, 2020.

100 PLACHKINOVA, M.; MAURER, C. Security breach at target. *Journal of Information Systems Education*, v. 29, n. 1, p. 11–20, 2018.

101 SHU, X.; TIAN, K.; CIAMBRONE, A.; YAO, D. Breaking the target: An analysis of target data breach and lessons learned. *arXiv preprint arXiv:1701.04940*, 2017.

102 ROSENBLUM, P. *Lessons From Home Depot: Expect Hackers To Crack More Retailers This Holiday Season*. 2014. <https://www.forbes.com/sites/paularosenblum/2014/11/06/lessons-from-home-depot-expect-hackers-to-crack-more-retailers-this-holiday-season/?sh=1f6436ea68bc>. [Online; accessed 22-October-20023].

103 HAWKINS, B. *Case Study: The Home Depot Data Breach*. [S.l.], 2021.

104 WILKINS, S. The evolution of europay, mastercard, and visa (emv) and the impacts on credit card fraud. *La Salle University Digital Commons*, 2018.

105 FROUD, D. The global implications of us emv adoption. *Computer Fraud & Security*, Elsevier, v. 2016, n. 2, p. 5–7, 2016.

106 BODKER, A.; CONNOLLY, P.; SING, O.; HUTCHINS, B.; TOWNSLEY, M.; DREW, J. Card-not-present fraud: using crime scripts to inform crime prevention initiatives. *Security Journal*, Springer, p. 1–19, 2022.

107 NAQVI, B.; PEROVA, K.; FAROOQ, A.; MAKHDOOM, I.; OYEDEJI, S.; PORRAS, J. Mitigation strategies against the phishing attacks: A systematic literature review. *Computers & Security*, Elsevier, p. 103387, 2023.

108 BLACK, P.; GONDAL, I.; LAYTON, R. A survey of similarities in banking malware behaviours. *Computers & Security*, Elsevier, v. 77, p. 756–772, 2018.

109 BHARDWAJ, A.; KAUSHIK, K.; MAASHI, M. S.; ALJEBREEN, M.; BHARANY, S. Alternate data stream attack framework to perform stealth attacks on active directory hosts. *Sustainability*, MDPI, v. 14, n. 19, p. 12288, 2022.

110 CHEVAL, V.; CREMERS, C.; DAX, A.; HIRSCHI, L.; JACOMME, C.; KREMER, S. Hash gone bad: Automated discovery of protocol attacks that exploit hash function weaknesses. In: *32nd USENIX Security Symposium*. [S.l.: s.n.], 2023. p. 1–18.

111 FARAWN, A. A.; RJEIB, H. D.; ALI, N. S.; AL-SADAWI, B. Secured e-payment system based on automated authentication data and iterated salted hash algorithm. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, v. 18, n. 1, p. 538–544, 2020.

112 KIRUBA, B.; SARAVANAN, V.; VASANTH, T.; YOGESHWAR, B. Owasp attack prevention. In: IEEE. *3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*. [S.l.], 2022. p. 1671–1675.

113 ALJABRI, M.; ALDOSSARY, M.; AL-HOMEED, N.; ALHETELAH, B.; ALTHUBIANY, M.; ALOTAIBI, O.; ALSAQER, S. Testing and exploiting tools to improve owasp top ten security vulnerabilities detection. In: IEEE. *2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)*. [S.l.], 2022. p. 797–803.

114 FAISAL, F.; ELSHOUSH, H. T. Input validation vulnerabilities in web applications: Systematic review, classification, and analysis of the current state-of-the-art. *IEEE Access*, IEEE, 2023.

115 MARELLI, M. The solarwinds hack: Lessons for international humanitarian organizations. *International Review of the Red Cross*, Cambridge University Press, v. 104, n. 919, p. 1267–1284, 2022.

116 MARTÍNEZ, J.; DURÁN, J. M. Software supply chain attacks, a threat to global cybersecurity: Solarwinds' case study. *International Journal of Safety and Security Engineering*, v. 11, n. 5, p. 537–545, 2021.

117 HASSIJA, V.; CHAMOLA, V.; GUPTA, V.; JAIN, S.; GUIZANI, N. A survey on supply chain security: Application areas, security threats, and solution architectures. *IEEE Internet of Things Journal*, IEEE, v. 8, n. 8, p. 6222–6246, 2020.

118 DISSANAYAKE, N.; JAYATILAKA, A.; ZAHEDI, M.; BABAR, M. A. Software security patch management-a systematic literature review of challenges, approaches, tools and practices. *Information and Software Technology*, Elsevier, v. 144, p. 106771, 2022.

119 AHMAD, R.; ALSMADI, I.; ALHAMDANI, W.; TAWALBEH, L. Zero-day attack detection: a systematic literature review. *Artificial Intelligence Review*, Springer, p. 1–79, 2023.

120 KUMAR, R.; SUBBIAH, G. Zero-day malware detection and effective malware analysis using shapley ensemble boosting and bagging approach. *Sensors*, MDPI, v. 22, n. 7, p. 2798, 2022.

121 FUGKEAW, S.; WORAPALUK, K.; TUEKLA, A.; NAMKEATSAKUL, S. Design and development of a dynamic and efficient pii data loss prevention system. In: SPRINGER. *International Conference on Computing and Information Technology*. [S.l.], 2021. p. 23–33.

122 CÁRDENAS, J. M. G. Steganography and data loss prevention: an overlooked risk? *International Journal of Security and Its Applications*, Science and Engineering Research Support Society, 2017.

123 HASSAN, M. U.; REHMANI, M. H.; CHEN, J. Differential privacy techniques for cyber physical systems: a survey. *IEEE Communications Surveys & Tutorials*, IEEE, v. 22, n. 1, p. 746–789, 2019.

124 CISA. *Insider Threat Mitigation Guide*. [S.l.]: Cybersecurity and Infrastructure Security Agency, 2020.

125 INSTITUTE, P. *Cost of Insider Threats Global Report*. [S.l.]: Proofpoint, 2020.

126 ENISA. *Threat Landscape Report 2016*. [S.l.]: European Union Agency For Network and Information Security, 2016.

127 THEIS, M.; TRZECIAK, R. F.; COSTA, D. L.; MOORE, A. P.; MILLER, S.; CASSIDY, T.; CLAYCOMB, W. R. Common sense guide to mitigating insider threats. *Carnegie Mellon University*, Software Engineering Institute, 2019.

128 AGUBOSHIM, F. C.; UDOBI, J. I. Security issues with mobile it: A narrative review of bring your own device (byod). *Information Technology (IT)*, v. 8, n. 1, 2019.

129 OGUNYEMI, R.; IDOWU, A. Data security concerns raised by 'bring your own device'in corporate organisations' hybrid and remote work environments in nigeria. *The Commonwealth Cybercrime Journal*, p. 111, 2023.

130 MARIOTTI, I.; CEINAR, I. M. et al. Teleworking in post-pandemic times: May local coworking spaces be the future trend? *Romanian Journal of Regional Science*, v. 15, n. 1, p. 52–76, 2021.

131 EVANGELAKOS, G. Keeping critical assets safe when teleworking is the new norm. *Network security*, Elsevier, v. 2020, n. 6, p. 11–14, 2020.

132 PANDE, P.; MEDATIYA, A. K.; MALLAIAH, K.; GANDHI, R. K.; SRINIVASACHARY, S. Mandatory enforcement of geofenced security in android. In: IEEE. *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*. [S.l.], 2021. p. 1031–1035.

133 UZ, A. *The effectiveness of remote wipe as a valid defense for enterprises implementing a BYOD policy*. Tese (Doutorado) — Université d'Ottawa/University of Ottawa, 2014.

134 GROSS, T.; BUSCH, M.; MÜLLER, T. One key to rule them all: Recovering the master key from ram to break android's file-based encryption. *Forensic Science International: Digital Investigation*, Elsevier, v. 36, p. 301113, 2021.

135 SHARMA, R.; DANGI, S.; MISHRA, P. A comprehensive review on encryption based open source cyber security tools. In: IEEE. *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*. [S.l.], 2021. p. 614–619.

136 SALAHDINE, F.; KAABOUCH, N. Social engineering attacks: A survey. *Future internet*, MDPI, v. 11, n. 4, p. 89, 2019.

137 HERDRICH, M. A. California v. greenwood: The trashing of privacy. *Am. UL Rev.*, HeinOnline, v. 38, p. 993, 1988.

138 AZEEM, E. A. et al. The data carving-the art of retrieving deleted data as evidence. *International Journal for Electronic Crime Investigation*, v. 6, n. 2, p. 8–8, 2022.

139 XU, Z.; ZHANG, Z.; LI, P.; LIU, X.; TANG, J. Review of research on degaussing technology of magnetic storage media. In: IEEE. *2020 Chinese Automation Congress (CAC)*. [S.l.], 2020. p. 3400–3405.

140 SHUKLA, M. K.; DUBEY, A. K.; UPADHYAY, D.; NOVIKOV, B. Group key management in cloud for shared media sanitization. In: IEEE. *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*. [S.l.], 2020. p. 117–120.

141 MIHINJAC, M.; SAVILLE, G. Third-generation crime prevention through environmental design (cpted). *Social Sciences*, MDPI, v. 8, n. 6, p. 182, 2019.

142 PIROOZFAR, P.; FARR, E. R.; ABOAGYE-NIMO, E.; OSEI-BERCHIE, J. Crime prevention in urban spaces through environmental design: A critical uk perspective. *Cities*, Elsevier, v. 95, p. 102411, 2019.

143 TAN, W. H.; ABAS, H. Systematic literature review crime prevention through environmental design (cpted) in physical security for it organization. *Open International Journal of Informatics*, v. 10, n. 1, p. 68–83, 2022.

144 FENNELLY, L. J.; PERRY, M. A. Encompassing effective cpted solutions in 2020 and beyond: concepts and strategies. In: *Handbook of Loss Prevention and Crime Prevention*. [S.l.]: Elsevier, 2020. p. 45–77.

145 MUBARKOOT, M.; ALTMANN, J.; RASTI-BARZOKI, M.; EGGER, B.; LEE, H. Software compliance requirements, factors, and policies: A systematic literature review. *Computers & Security*, Elsevier, p. 102985, 2022.

146 ALYAMI, A.; SAMMON, D.; NEVILLE, K.; MAHONY, C. Critical success factors for security education, training and awareness (seta) programme effectiveness: an empirical comparison of practitioner perspectives. *Information & Computer Security*, Emerald Publishing Limited, 2023.

147 SILIC, M.; LOWRY, P. B. Using design-science based gamification to improve organizational security training and compliance. *Journal of management information systems*, Taylor & Francis, v. 37, n. 1, p. 129–161, 2020.

148 ALDAWOOD, H.; SKINNER, G. Reviewing cyber security social engineering training and awareness programs—pitfalls and ongoing issues. *Future Internet*, MDPI, v. 11, n. 3, p. 73, 2019.

149 LATEș, I.; BOJA, C. Cyber range as a competency based education instrument in cyber security. In: *International Conference on New Trends in Sustainable Business and Consumption*. [S.l.: s.n.], 2022. p. 703–710.

150 EBAD, S. A. Exploring how to apply secure software design principles. *IEEE Access*, IEEE, v. 10, p. 128983–128993, 2022.

151 FERNANDO, V. Cyber forensics tools: A review on mechanism and emerging challenges. In: IEEE. *2021 11th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. [S.l.], 2021. p. 1–7.

152 D'ANNA, T.; PUNTARELLO, M.; CANNELLA, G.; SCALZO, G.; BUSCEMI, R.; ZERBO, S.; ARGO, A. The chain of custody in the era of modern forensics: From the classic procedures for gathering evidence to the new challenges related to digital data. In: MDPI. *Healthcare*. [S.l.], 2023. v. 11, n. 5, p. 634.

153 RABELLO, A.; GOULART, J.; KARAM, M.; PITANGA, M.; FILHO, R. G. B.; RICIONI, R. Proposed incident response methodology for data leakage. *ICSEA 2021*, p. 60, 2021.

154 HILLMANN, F.; KLAUENBERG, T.; SCHROEDER, L.; DIESTERHÖFT, T. O. A user-centric view on data breach response expectations. *CIISR*, p. 19, 2023.

155 WOHLIN, C.; RUNESON, P.; HÖST, M.; OHLSSON, M. C.; REGNELL, B.; WESSLÉN, A. *Experimentation in software engineering*. [S.l.]: Springer Science & Business Media, 2012.