

DOI: https://doi.org/10.48009/4_iis_2026_127

Trends and patterns in organizational data breaches: Evidence from a multi-year incident dataset

Olumide O. Malomo, Virginia State University, omalomo@vsu.edu

Emmanuel Omojokun, Virginia State University, emojokun@vsu.edu

Venkatapparao Mummalaneni, Virginia State University, vmummalaneni@vsu.edu

Christine Shikutwa, Saint John's University, cshikutwa001@csbsju.edu

Moses Garuba, Hampton University, moses.garuba@hamptonu.edu

Abstract

Organizational data breaches continue to pose significant operational, financial, and reputational risks across industries. Although prior research has examined breach costs and disclosure effects, comparative analyses characterizing breach magnitude and patterns across sectors and over time remain limited, often due to restricted access to proprietary datasets. This study addresses this gap with an empirical analysis of a multi-year dataset comprising 488 publicly documented organizational data-breach incidents across diverse sectors. Using the reported number of records exposed as a proxy for breach impact, the study examines temporal trends, sectoral distributions, and differences across breach-method categories. Descriptive and comparative analyses reveal notable variation in breach magnitude across sectors and incident types, with specific organizational contexts exhibiting disproportionately high-impact breaches. Temporal analysis indicates persistent large-scale breaches rather than a steady decline in impact over time. Rather than asserting causal relationships, the findings highlight patterns and empirical signals that can inform comparative risk assessment, benchmarking, and prioritization of cybersecurity controls. By leveraging a replicable, openly sourced incident dataset, this paper contributes an empirical foundation for understanding organizational breach risk across sectors and provides a basis for future theory-driven and decision-focused cybersecurity research.

Keywords: Data breaches, cybersecurity analytics, risk assessment, breach magnitude, empirical security studies

Introduction

Organizational data breaches have become a persistent feature of the modern digital landscape, affecting organizations across sectors and geographic regions. High-profile incidents involving hundreds of millions of records have heightened concerns regarding data protection, regulatory compliance, and organizational resilience (Malomo et al., 2025; Singh, 2025). Despite sustained investment in cybersecurity technologies and controls, large-scale breaches continue to occur, raising questions about how breach risk is distributed and whether meaningful patterns can be identified across time and industries (Malomo et al., 2022; SecurityBrief UK, 2025). Much of the existing breach literature focuses on financial impact, market reactions, or firm-specific consequences following breach disclosures. While valuable, such studies often rely on proprietary datasets or narrow sectoral samples, limiting replicability and cross-sector comparison (Muktadir-Al-Mukit & Ali, 2025; Smith et al., 2023). There remains a need for empirical studies that

characterize breach patterns using openly documented incidents, enabling broader benchmarking and comparative risk analysis. This study addresses this need by examining a multi-year dataset of publicly documented organizational data breaches. Rather than seeking to establish causality, the analysis focuses on identifying temporal trends, sectoral differences, and variations in impact associated with breach-method categories. By doing so, the paper provides an empirical baseline to support comparative risk assessment and inform cybersecurity prioritization efforts.

To guide the empirical investigation and systematically evaluate these objectives, the study is organized around the following research questions.

Research Questions

Guided by the study objectives and the available incident-level variables, this research seeks to answer the following questions regarding organizational data breaches across time, sectors, and incident types:

1. How do organizational data-breach incidents vary temporally in terms of frequency and breach magnitude over time?
2. To what extent do organizational sectors differ in breach magnitude and exposure patterns?
3. How do breach-method categories differ in terms of organizational exposure and breach severity?
4. To what extent is organizational breach impact concentrated among a small subset of high-impact incidents?

Contributions

This paper makes three primary contributions to the organizational cybersecurity and data-breach literature:

1. Reproducible empirical characterization of organizational data breaches: The study provides a reproducible empirical characterization of breach frequency and magnitude using a large, multi-year dataset of 488 publicly documented and independently verifiable organizational data-breach incidents spanning multiple sectors. Unlike many prior empirical studies that rely on proprietary or restricted datasets, this work demonstrates the value of openly sourced incident data for transparent cybersecurity research, comparative analysis, and future replication.
2. Comparative analysis of organizational breach characteristics: The study examines how readily available incident descriptors—particularly organizational sector and breach method—correspond with observable differences in breach magnitude. Rather than treating these descriptors as causal predictors, the analysis demonstrates their usefulness for comparative cybersecurity benchmarking, exposure profiling, and organizational risk prioritization.
3. Evidence of concentration in organizational breach impact: The study demonstrates that breach impact is highly concentrated across organizational sectors, with a relatively small number of sectors accounting for a disproportionate share of publicly reported records exposed. This distribution-aware perspective extends beyond traditional descriptive reporting by providing empirical evidence that can inform cybersecurity investment decisions, resilience planning, and strategic risk management.

The following section reviews prior research on organizational data breaches and positions the present study within the broader literature on cybersecurity risk, breach analytics, and organizational resilience.

Related Work

Prior research on organizational data breaches has largely focused on firm-level economic and disclosure consequences. A substantial stream of work examines how breach announcements affect stock prices, market activity, and firm value using event-study designs. Studies show that data breach announcements can trigger negative abnormal returns and changes in trading volume and bid–ask spreads for publicly

traded firms, highlighting breaches as significant service crises and sources of asymmetric information for investors (Foerderer & Schuetz, 2022; Ford et al., 2021; Rasoulilian et al., 2023). More recent work extends this line of inquiry to repeated or multiple breaches, finding that repeated incidents can have distinct and sometimes compounding effects on firm value relative to one-off events (Peng et al., 2023; Schatz & Bashroush, 2016). Related research explores broader financial consequences, including increased risk of stock price crashes, shifts in Research and Development (R&D) spending, and changes in disclosure behavior following cybersecurity incidents, further underscoring breaches as material financial and strategic events (Garg et al., 2025; Singh et al., 2025; Zhang, 2024). These findings are consistent with emerging evidence that cyber events, cyberterrorism, and security incidents more generally affect valuation and risk perceptions in capital markets (Day & Booker, 2024; Liu & Babar, 2024). Related studies examining stock market responses to cyber-attack news and cyber terrorism incidents further reinforce the growing body of evidence linking cybersecurity events to organizational and market outcomes (Amani et al., 2025).

Another branch of research has examined reputational and governance responses to data breaches. For example, Akey and co-authors show that breaches can erode corporate reputation and prompt firms to increase corporate social responsibility (CSR), political contributions, wages, and IT investment as part of a broader repair strategy (Akey et al., 2023). Other studies analyze how cyber risk and breach exposure shape voluntary disclosure, mandatory reporting, and governance practices, including the use of CSR or narrative risk disclosure to manage stakeholder perceptions (Abdelraouf & Hussainey, 2025; Vo & Pham, 2025). This aligns with regulatory developments, such as enhanced breach disclosure obligations and evolving cybersecurity reporting regimes, that have intensified scrutiny of breach transparency and corporate responses (Newman et al., 2025). In parallel, work such as the HuCRAF paper has begun to foreground the human-centric and governance dimensions of cybersecurity risk, emphasizing decision environments and organizational factors that shape breach outcomes (Cremer et al., 2022; Malomo et al., 2025).

A third cluster of studies focuses on breach data sources and taxonomies. Many empirical analyses rely on proprietary or curated datasets derived from the Privacy Rights Clearinghouse (PRC) breach chronology or similar sources that aggregate notifications primarily for U.S. public companies or specific sectors (Rosati & Lynn, 2021). Systematic reviews of cyber risk and cybersecurity data highlight the uneven availability of incident-level data and the heavy reliance on a small number of breach repositories, often with access restrictions, incomplete coverage, or limited metadata for cross-sector comparison (Ran & Murphy, 2025). In parallel, several taxonomies and classification schemes have been proposed to categorize cyber incidents by attack vector, threat actor, target, or impact, providing conceptual structure but often focusing on technical attributes rather than cross-sector empirical pattern discovery (Rabitti et al., 2025). Studies such as Molitor et al. use PRC data to explore breach types, legal disputes, and narrative themes, illustrating the value of these datasets while also pointing to underexplored dimensions of breach impact and context (Molitor et al., 2023).

Taken together, this literature has advanced the understanding of breach consequences and mechanisms, particularly for publicly traded firms and financially measurable outcomes. However, it also reveals several limitations that motivate the present study. First, many empirical analyses rely on proprietary or restricted datasets or focus narrowly on listed firms, which constrains replication, extension, and cross-sectoral comparison (Rosati & Lynn, 2021). Second, a substantial portion of the breach literature emphasizes firm-level outcomes—such as stock price reactions, litigation, or disclosure strategies—rather than systematic characterization of breach magnitude patterns across sectors and over time (Rasoulilian et al., 2023). Third,

while taxonomies provide useful conceptual lenses, many are not empirically operationalized at scale to describe how breach impact is distributed across industries, methods, and years (Rabitti et al., 2025).

This paper complements and extends prior work by leveraging a large, openly verifiable dataset of publicly documented organizational breaches to examine impact patterns across sectors, incident types, and years, without relying on proprietary data sources. By focusing on distribution-aware descriptive and comparative analysis—for example, median versus mean breach magnitude and the concentration of impact among a small number of very large incidents—the study offers a replicable empirical perspective that supports cross-sector risk benchmarking and provides a quantitative foundation for future theory-driven and human-centric cybersecurity research, including frameworks such as HuCRAF (Malomo et al., 2025).

Methods

Dataset Construction and Preparation

This study employs a secondary analysis of a researcher-compiled dataset of publicly documented organizational data-breach incidents. The dataset was constructed by systematically collecting publicly reported breach incidents from Internet sources, including reputable cybersecurity news outlets, technology publications, organizational breach notifications, and other publicly accessible reports describing cybersecurity incidents. Only incidents that could be independently verified through publicly available information were retained, thereby enhancing the dataset's transparency and reproducibility.

The final dataset comprises 488 organizational data-breach incidents spanning 2004 to 2024. To ensure consistency and analytical reliability, incidents were included only when sufficient information was available to identify: (1) the affected organization, (2) the year the breach became publicly known, (3) an estimate of the number of records exposed, and (4) sufficient descriptive information to classify both the organizational sector and the primary reported breach method. Duplicate reports of the same breach were consolidated into a single incident record, and incomplete reports lacking sufficient information for classification were excluded from the analysis. Each incident record contains four principal analytical variables:

- Year of disclosure, representing the year in which the breach became publicly known.
- Organizational sector, classifying the affected organization into one of sixteen sectors (e.g., Web, Finance, Government, Healthcare, Retail, Technology, Telecommunications, and Academic).
- Breach method, identifying the primary reported mechanism associated with the incident (e.g., hacked, poor security, insider activity, lost device, or third-party compromise).
- Records exposed, representing the publicly reported estimate of the number of compromised records and serving as the operational measure of breach magnitude throughout the study.

Because the dataset was compiled from publicly reported incidents rather than proprietary breach repositories, it is intended to characterize observable organizational breach patterns rather than estimate the complete population of cybersecurity incidents. Consequently, the findings should be interpreted as empirical observations derived from documented breach events that are independently verifiable through public reporting. Table 1 summarizes the principal variables extracted from each publicly documented breach incident and used throughout the analysis.

Table 1. Variables Included in the Dataset

VARIABLE	DESCRIPTION
Year of Disclosure	Year in which the breach became publicly known.
Affected Organization	Name of the affected organization.
Organizational Sector	Sector classification assigned to the affected organization (e.g., Web, Finance, Government, Healthcare, Academic).
Breach Method	Primary reported breach mechanism (e.g., hacked, poor security, lost device, insider activity, third-party compromise).
Records Exposed	Publicly reported estimate of the number of compromised records used as the operational measure of breach magnitude.

Analytical Approach

The study employs descriptive and comparative statistical analyses to examine breach frequency and magnitude across time, organizational sectors, and breach-method categories. Breach magnitude is operationalized using the reported number of records exposed. Because breach-size data exhibit substantial positive skewness due to the presence of extremely large "mega breaches," median values are emphasized alongside arithmetic means to provide a more representative measure of typical organizational exposure, while maximum values illustrate the influence of exceptionally large incidents. Additional cumulative analyses are performed to examine the concentration of breach impact across organizational sectors. Throughout the analysis, organizational sector and breach method are treated as explanatory descriptors rather than causal predictors; therefore, the findings should be interpreted as empirical patterns that support comparative risk assessment, cybersecurity benchmarking, and prioritization rather than causal inference.

Results

Dataset Overview

The compiled dataset contains 488 publicly documented organizational data-breach incidents spanning the period 2004–2024 and representing sixteen organizational sectors. Before examining temporal and sector-specific patterns, Table 2 provides a statistical overview of the dataset, including its temporal coverage, sector diversity, and descriptive measures of breach magnitude.

Table 2. Dataset Summary Statistics

MEASURE	VALUE
Total breach incidents	488
Time span	Multi-year (2004–2024)
Number of sectors	16
Median records exposed	3,000,000

Mean records exposed	$\approx 35,361,049$
Maximum records exposed	1,000,000,000

Temporal Distribution of Breach Incidents

Analysis of breach frequency by year indicates that reported incidents persist across the entire observation period, with fluctuations rather than a clear monotonic increase or decrease. While some years show elevated reporting activity, large-scale breaches are not confined to any single time period, suggesting that high-impact incidents remain a persistent feature of the organizational cybersecurity landscape.

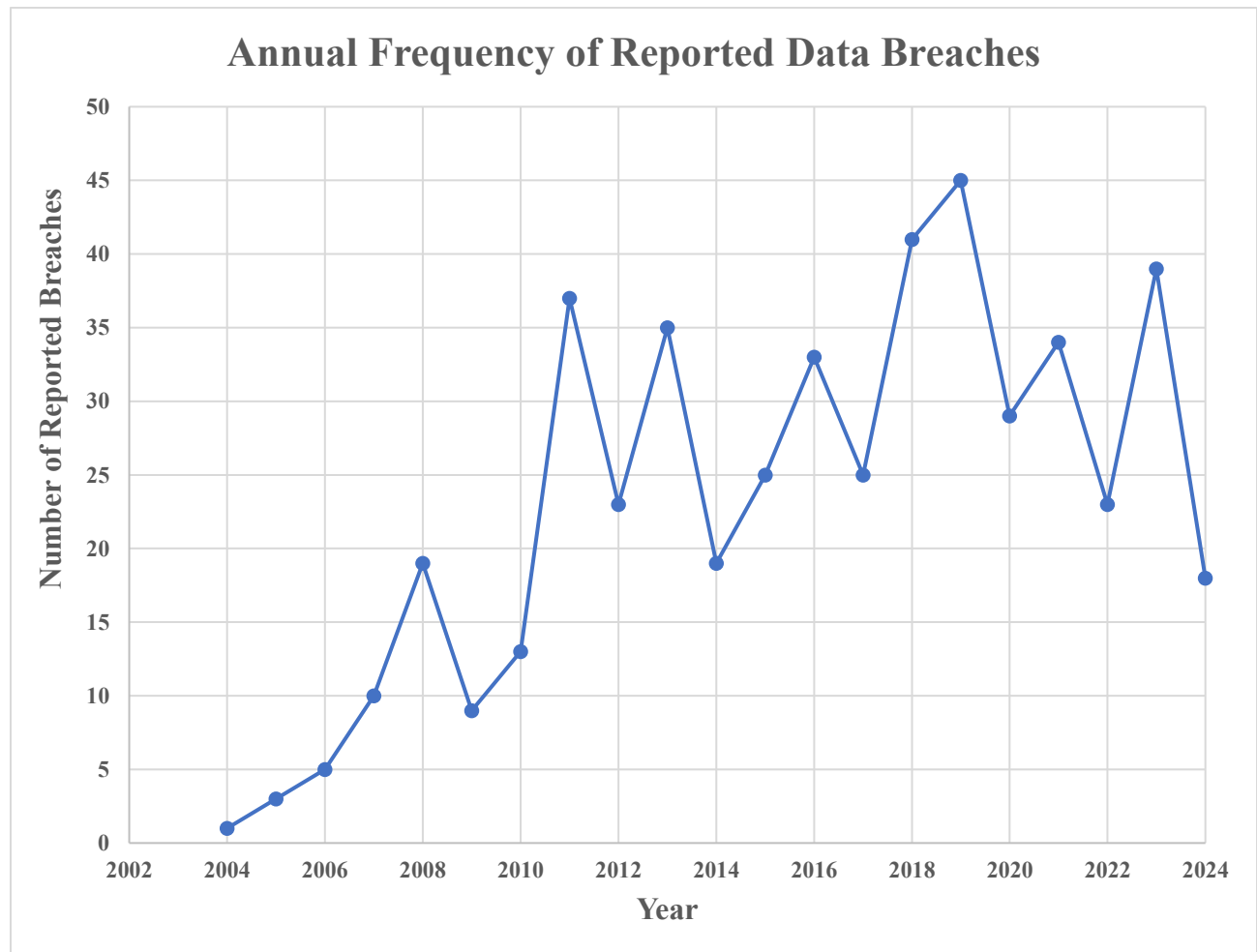


Figure 1. Annual Frequency of Reported Data Breaches. Number of Publicly Reported Organizational Data Breach Incidents Per Year in the Dataset.

Temporal Trends in Breach Magnitude

Breach magnitude, measured by the number of records exposed, exhibits substantial variability across years. As illustrated in Figure 2, median breach size remains relatively stable over time, while maximum breach values fluctuate considerably due to intermittent extreme cases—often referred to as “mega breaches.” This pattern suggests that improvements in security controls have not eliminated the risk of very large-scale data exposures.

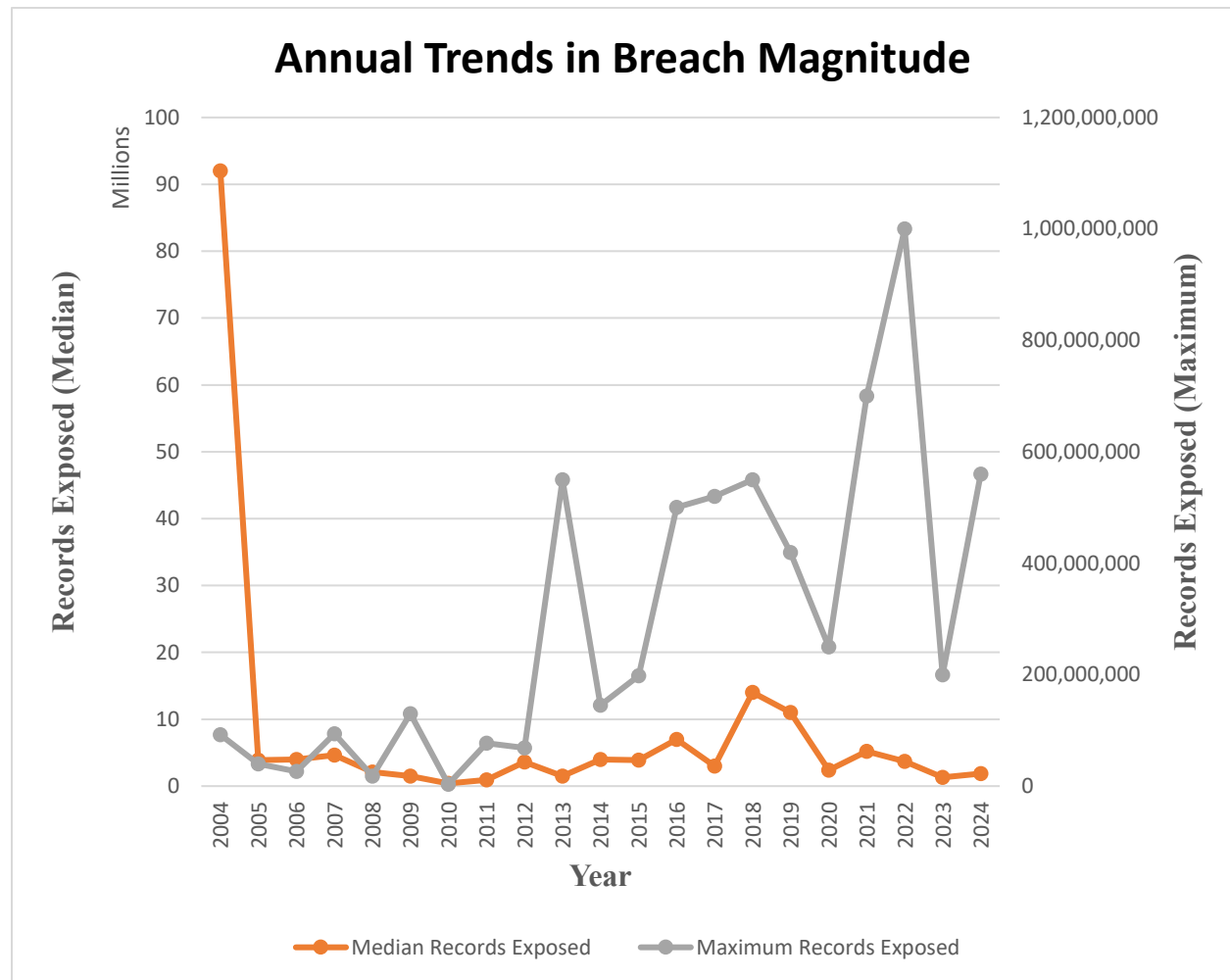


Figure 2. Annual Trends in Breach Magnitude. Median and Maximum Records Exposed Per Breach by Year, Illustrating Stable Typical Breach Sizes Alongside Intermittent Mega Breaches.

Sectoral Differences in Breach Impact

Breach magnitude differs notably across organizational sectors. As illustrated in Figure 3, certain sectors consistently exhibit higher median and average breach sizes, reflecting differences in data aggregation practices, user base size, and system centralization. In several sectors, average breach size substantially exceeds the median, indicating the influence of occasional high-impact “mega breach” events. While some sectors exhibit smaller typical breach sizes, extreme incidents remain present across organizational contexts.

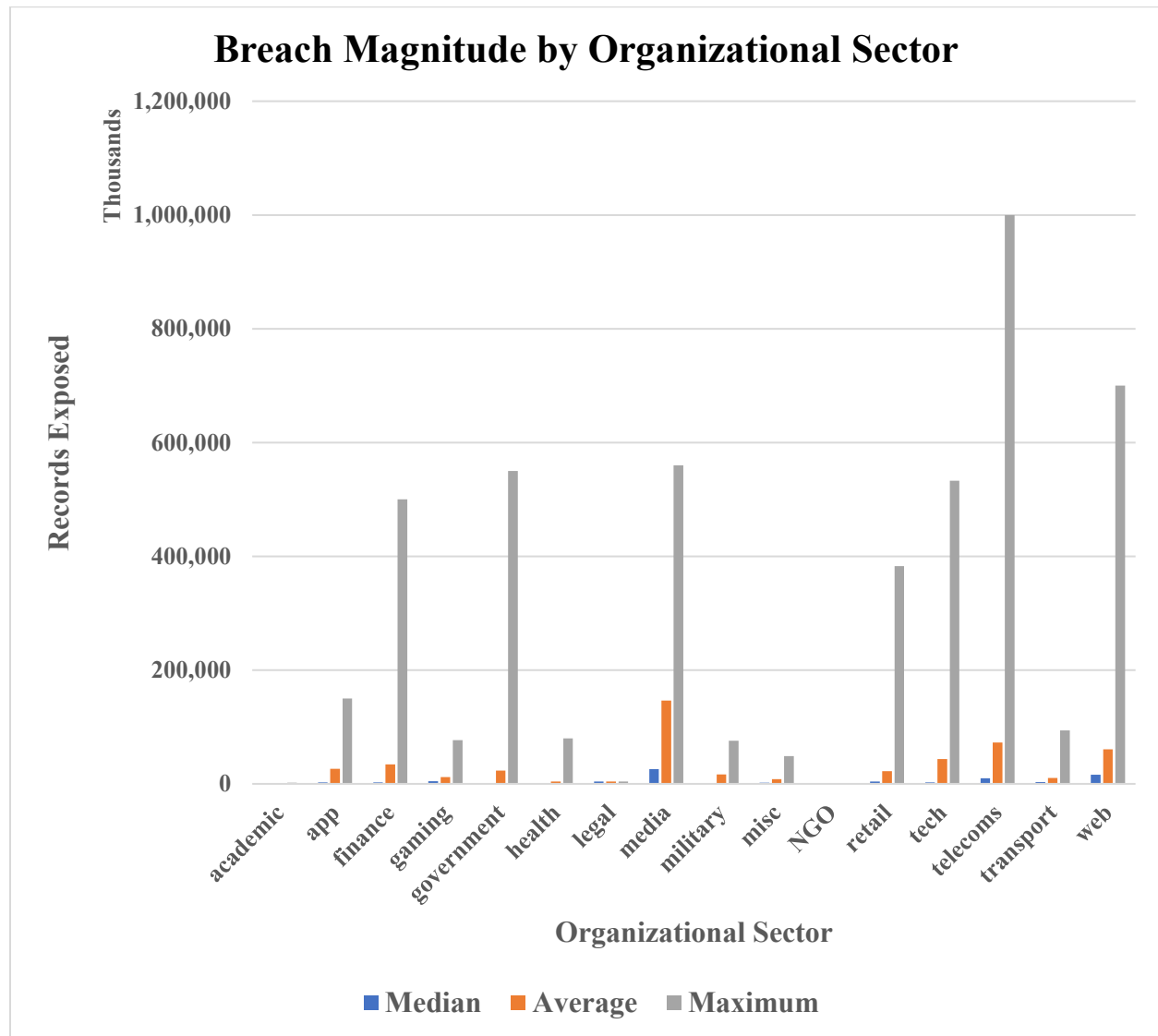


Figure 3. Breach Magnitude by Organizational Sector. Median, Average, and Maximum Records Exposed Per Breach Across Organizational Sectors, Illustrating Differences in Typical Breach Impact and the Influence of High-Impact Incidents.

Breach Method and Impact

Breach incidents were categorized according to reported method descriptors, including externally driven attacks (e.g., hacking), insider activity, lost devices, poor security practices, and third-party exposure. Table 3 and Figure 4 illustrate notable variation in breach impact across method categories. Hacked incidents exhibit the highest median breach size, followed closely by poor security practices and third-party exposure, suggesting comparatively greater typical organizational impact. Although poor security practices demonstrate the highest average breach magnitude, the difference between average and median values indicates the influence of occasional large-scale incidents. By contrast, lost device incidents exhibit the lowest median breach size, suggesting comparatively smaller typical breach severity. These findings indicate that breach method is associated with meaningful differences in organizational exposure outcomes.

Table 3. Breach Impact Statistics by Method Category. Descriptive Breach Impact Statistics by Method Category, Including Incident Frequency, Average Exposure, Maximum Breach Magnitude, and Median Breach Size.

Method	Count	Records Exposed (Millions)		
		Average	Maximum	Median
hacked	338	35.3	1000	3.95
inside job	20	18.0	200	1.90
lost device	49	4.4	76	1.12
poor security	69	62.4	550	3.30
via a third party	12	37.4	340	3.18

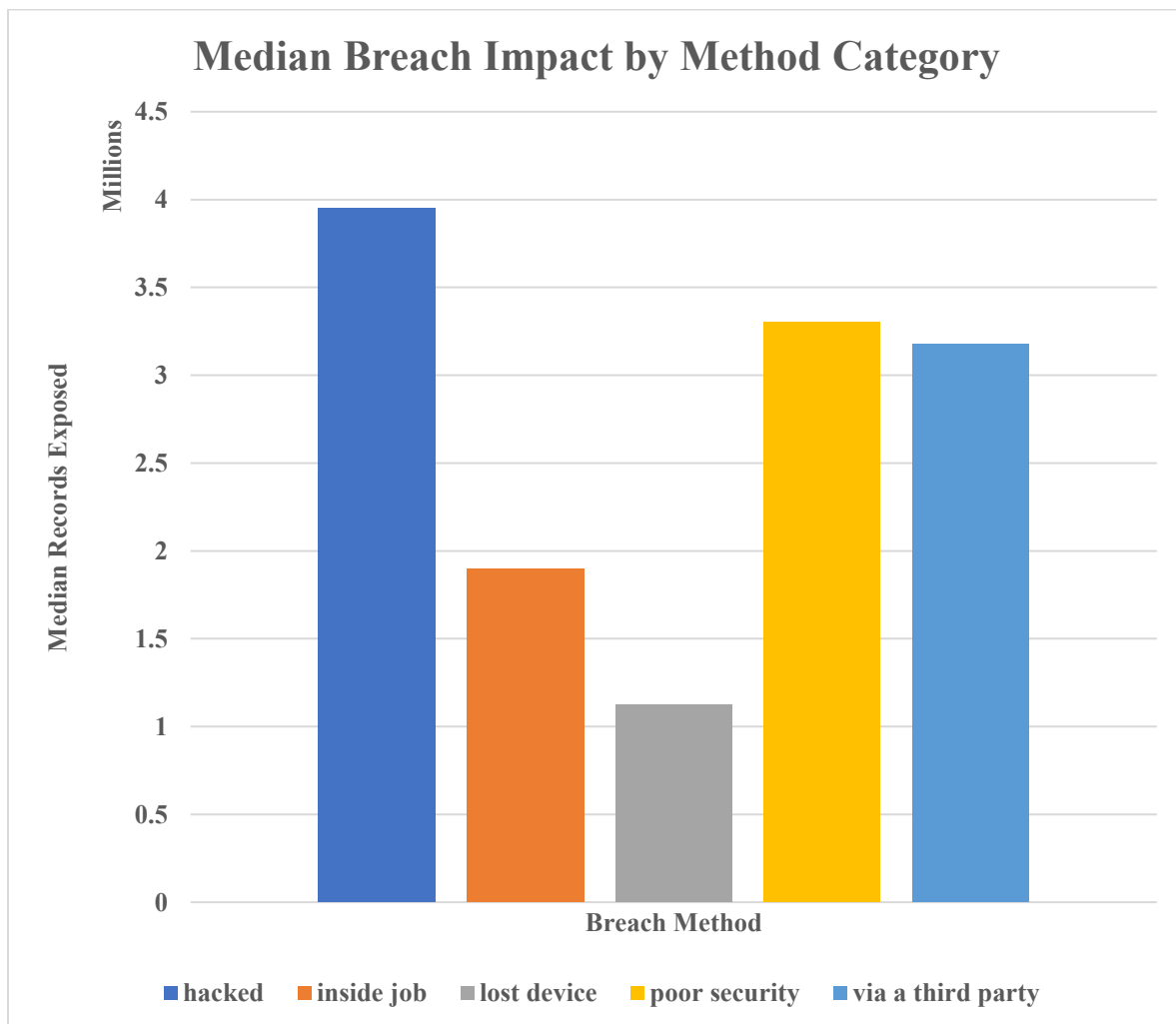


Figure 4. Median Breach Impact by Method Category. Median Records Exposed Per Breach Across Method Categories, Illustrating Variation in Typical Breach Impact Associated with Different Breach Mechanisms.

Concentration of Breach Impact Across Organizational Sectors

As illustrated in Figure 5, the impact of breaches is highly concentrated across organizational sectors. When sectors are ranked by the total number of records exposed, a relatively small number of sectors account for

most compromised records. The Web sector alone accounts for approximately 44% of the total records exposed, while the top four sectors collectively account for approximately 75% of all exposed records. This concentration suggests that cybercriminals frequently target sectors that aggregate large volumes of sensitive information, making them attractive targets for financially motivated attacks. The findings underscore the importance of prioritizing cybersecurity investments and resilience strategies in sectors that manage large, centralized repositories of sensitive data.

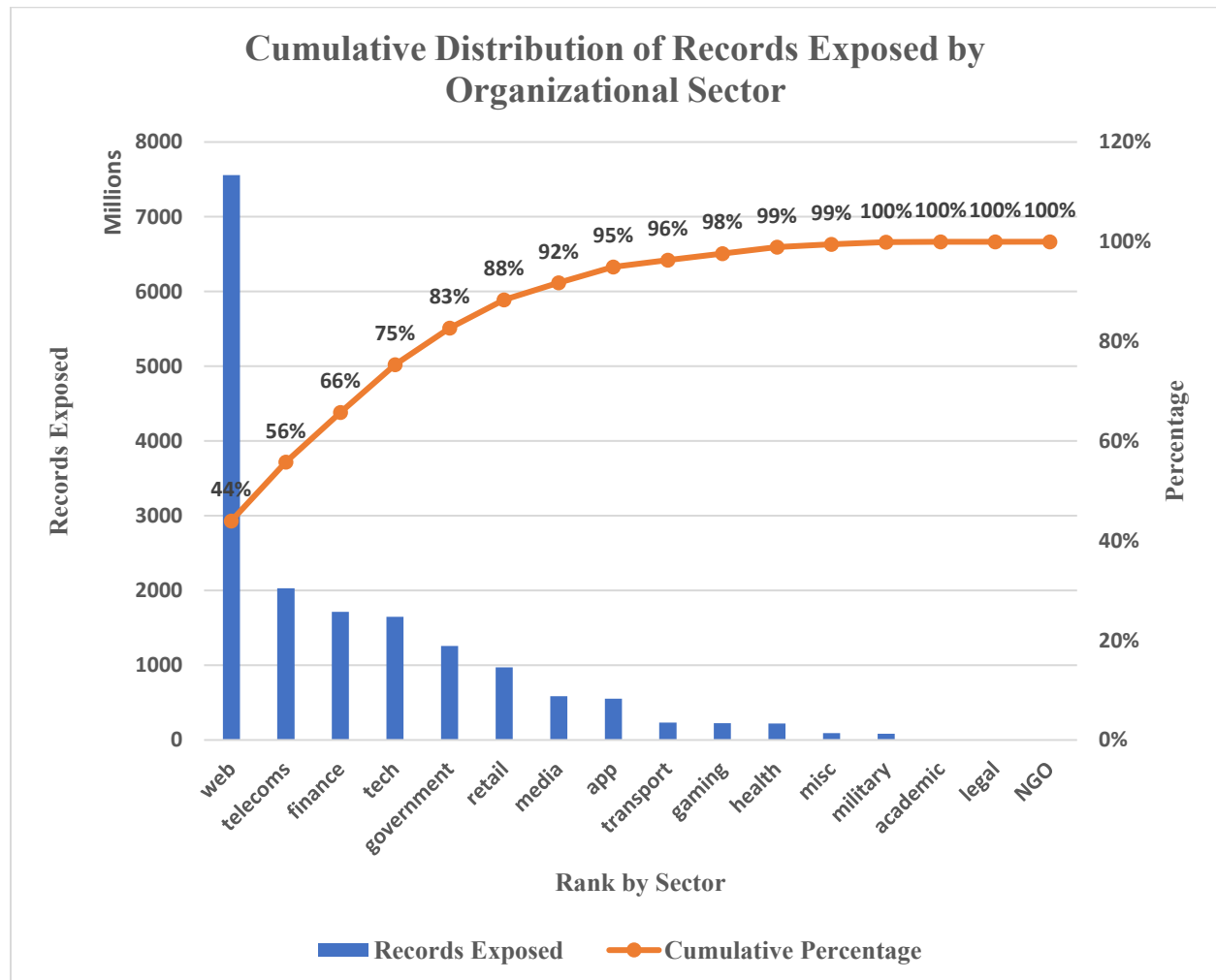


Figure 5. Cumulative Distribution of Records Exposed by Organizational Sector. Organizational Sectors are Ranked According to the Total Number of Records Exposed. The Cumulative Percentage Shows that a Relatively Small Number of Sectors Account for Most Records Exposed, Highlighting the Concentration of Breach Impact Across Sectors.

Discussion

Interpreting Temporal Persistence of Large Breaches

The findings indicate that organizational data breaches remain a persistent cybersecurity challenge despite continued advances in cybersecurity technologies, regulatory oversight, and organizational investment in security controls. As demonstrated in the temporal analysis, breach magnitude does not exhibit a sustained downward trend. Instead, large-scale "mega breaches" continue to occur intermittently throughout the

observation period, suggesting that technological progress alone has not eliminated the underlying structural conditions that enable high-impact cyber incidents. Several factors may contribute to this persistence. Organizations continue to consolidate ever-increasing volumes of personally identifiable information (PII), financial records, authentication credentials, healthcare information, and other high-value digital assets within centralized enterprise systems and cloud environments (Malomo et al., 2025). As these repositories expand, they become increasingly attractive targets for financially motivated cybercriminals seeking maximum economic return from a single successful compromise. Consequently, while cybersecurity investments may reduce the likelihood of certain attack vectors, they do not necessarily eliminate the potential for low-frequency, high-impact breach events.

These findings suggest that organizational cybersecurity effectiveness should be evaluated not only by reductions in breach frequency but also by an organization's ability to prevent or withstand catastrophic breach events that continue to characterize the modern cyber threat landscape.

Sectoral Risk Profiles and Organizational Data Concentration

The observed variation in breach magnitude across organizational sectors appears closely associated with differences in organizational scale, data concentration, and digital service centralization. As illustrated in Figure 5, a relatively small number of sectors account for the majority of publicly reported records exposed, indicating that breach impact is highly concentrated rather than evenly distributed across industries. Sectors such as Web services, Telecommunications, Finance, and Technology frequently maintain extensive customer databases, identity repositories, payment information, cloud-hosted applications, and digital platforms supporting millions of users. Consequently, successful compromises within these environments may expose substantially larger volumes of sensitive information than breaches affecting organizations with smaller or more decentralized data assets.

These findings should not be interpreted as evidence that particular sectors are inherently less secure than others. Rather, they suggest that organizational characteristics—including the volume of information managed, the degree of system centralization, and the concentration of digital assets—may significantly influence the magnitude of breach consequences once a compromise occurs.

Breach Method as an Informative Organizational Risk Indicator

The comparative analysis demonstrates that breach method provides a useful empirical indicator for understanding organizational exposure patterns. Although breach method does not establish causality, different incident categories are associated with distinct breach-magnitude profiles. Incidents attributed to poor security practices exhibit particularly high average breach magnitudes, suggesting that weaknesses such as inadequate security governance, delayed patch management, weak access controls, configuration errors, and insufficient security monitoring may permit attackers to compromise large volumes of organizational data. By contrast, lost-device incidents generally exhibit smaller median breach sizes, reflecting the comparatively localized nature of these events.

The substantial variation observed among breach-method categories reinforces the importance of maintaining foundational cybersecurity controls in addition to deploying advanced defensive technologies. Effective cybersecurity resilience depends not only on detecting sophisticated attacks but also on reducing organizational exposure arising from preventable security weaknesses.

Implications for Cybersecurity Risk Assessment and Organizational Decision-Making

The findings demonstrate that publicly documented breach data can provide meaningful empirical evidence for cybersecurity benchmarking and strategic risk assessment. Even relatively limited incident descriptors—including organizational sector, breach method, and records exposed—reveal systematic

patterns that can inform organizational cybersecurity priorities. The concentration of breach impact among a relatively small number of organizational sectors suggests that cybersecurity investments should be informed not only by the probability of attack but also by the potential magnitude of organizational exposure. Organizations responsible for managing highly centralized repositories of sensitive information should prioritize resilience strategies that reduce the likelihood and consequences of catastrophic breach events.

The results also carry implications for higher education institutions. The Academic sector represented within the dataset includes educational organizations that increasingly depend on centralized Learning Management Systems (LMS), student information systems, cloud collaboration platforms, and digital administrative services. As universities continue expanding these interconnected environments, cybersecurity resilience should extend beyond preventive controls to include redundancy planning, operational continuity, backup access mechanisms, and disaster recovery capabilities. These safeguards become particularly important during high-impact academic periods such as course registration and final examinations, where service disruptions may significantly affect institutional operations even in the absence of large-scale data loss.

Finally, the study illustrates the value of openly verifiable breach datasets for cybersecurity research. Transparent datasets facilitate replication, comparative benchmarking, and cumulative scientific investigation, enabling future researchers to extend empirical analyses using predictive modeling, machine learning, and theory-driven approaches without relying exclusively on proprietary cybersecurity datasets.

Limitations

This study is subject to several limitations that should be considered when interpreting the findings. First, the analysis relies on publicly reported organizational data-breach incidents, which may be affected by reporting bias, incomplete disclosure, variations in reporting standards, and the availability of publicly accessible information. Consequently, the compiled dataset should not be interpreted as a complete census of all organizational data breaches occurring during the observation period. Second, the reported number of records exposed serves as the operational measure of breach magnitude; however, breach size does not necessarily correspond to organizational impact or the sensitivity of compromised information. For example, the exposure of one million email addresses may present substantially different organizational, legal, financial, and societal consequences than the exposure of one million Social Security numbers, payment card records, healthcare records, authentication credentials, or classified information. Accordingly, breach magnitude should be interpreted as an indicator of exposure rather than a direct measure of organizational harm. Third, organizational sector and breach-method classifications were derived from publicly available incident descriptions and therefore depend on the completeness and accuracy of external reporting. While these descriptors provide valuable comparative insights, they should not be interpreted as causal determinants of breach outcomes.

Finally, the descriptive nature of the analysis does not permit causal inference regarding the relationships among breach characteristics, organizational attributes, and breach magnitude. The findings instead provide an empirical characterization of observable breach patterns that can inform future analytical and theory-driven investigations.

Future Research Directions

The descriptive findings presented in this study provide several opportunities for future cybersecurity research.

- Predictive breach analytics: Future studies may employ statistical modeling, machine learning, or artificial intelligence techniques to predict breach magnitude or organizational exposure using explanatory variables such as organizational sector, breach method, temporal characteristics, and additional organizational attributes.
- Sector-specific investigations: The observed differences across organizational sectors suggest opportunities for more detailed sector-level studies examining why certain industries consistently experience larger breach impacts and identifying organizational, technical, and regulatory factors that contribute to these differences.
- Data sensitivity and organizational impact: Future research should extend beyond breach magnitude by incorporating measures of data sensitivity, including distinctions among personally identifiable information (PII), financial records, healthcare data, authentication credentials, and other categories of compromised information to better estimate organizational harm.
- Cyber resilience in higher education: As educational institutions become increasingly dependent on centralized Learning Management Systems (LMS), cloud services, and digital academic platforms, future work should investigate strategies for strengthening operational resilience, redundancy planning, disaster recovery, and business continuity during cybersecurity incidents affecting higher education.
- Longitudinal cybersecurity risk modeling: Future studies may integrate additional organizational variables, threat intelligence, ransomware characteristics, or economic indicators to develop more comprehensive models of organizational cybersecurity risk and improve evidence-based decision-making.

Collectively, these research opportunities illustrate how descriptive analyses of publicly documented breach incidents can serve as a foundation for more advanced predictive, explanatory, and decision-oriented cybersecurity research.

Conclusion

This study presents a reproducible empirical analysis of organizational data breaches using a multi-year dataset of 488 publicly documented and independently verifiable incidents spanning diverse organizational sectors. The findings demonstrate persistent variability in breach magnitude across time, organizational sectors, and breach methods while also revealing that organizational breach impact is highly concentrated among a relatively small number of sectors and high-impact incidents. Beyond characterizing breach patterns, the study demonstrates that openly available incident data can provide meaningful empirical insights for comparative cybersecurity risk assessment, organizational benchmarking, and strategic prioritization. Although organizational sector and breach method should not be interpreted as causal determinants of breach magnitude, they provide informative descriptors that reveal important differences in organizational exposure and cyber-risk profiles.

The findings also carry practical implications for organizations responsible for managing large centralized repositories of sensitive information. The concentration of breach impact observed in sectors such as Web services, Telecommunications, Finance, and Technology suggests that cybersecurity investments should increasingly emphasize resilience against low-frequency, high-impact breach events rather than focusing exclusively on incident frequency. Similarly, the growing dependence of higher education institutions on centralized digital learning environments highlights the importance of operational continuity, redundancy planning, and resilient cybersecurity architectures capable of sustaining critical academic services during cyber incidents. By leveraging a transparent, reproducible, and publicly verifiable dataset, this study contributes empirical evidence that complements existing research relying on proprietary breach repositories. Ultimately, the study demonstrates how transparent, reproducible analyses of publicly

documented cyber incidents can strengthen evidence-based cybersecurity research, organizational decision-making, and strategic cyber-risk governance.

Reference

- Abdelraouf, M., & Hussainey, K. (2025). Systematic review of cyberrisk disclosure practices: insights and implications. *The Journal of Risk Finance*.
- Akey, P., Lewellen, S., Liskovich, I., & Schiller, C. (2023). Hacking corporate reputations. Rotman School of Management Working Paper, (3143740).
- Amani, F., Magnan, M., & Moldovan, R. (2025). Cybersecurity risks and incidents disclosure: a literature review. *Accounting Perspectives*, 24(3), 605-667.
- Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber risk and cybersecurity: a systematic review of data availability. *The Geneva papers on risk and insurance. Issues and practice*, 47(3), 698.
- Day, K., & Booker, Q. (2024). Market Reactions to Cybersecurity Incidents: A Case Study Approach. *Issues in Information Systems*, 25(4).
- Foerderer, J., & Schuetz, S. W. (2022). Data breach announcements and stock market reactions: a matter of timing?. *Management Science*, 68(10), 7298-7322.
- Ford, A., Al-Nemrat, A., Ghorashi, S., & Davidson, J. (2021, May). The impact of data breach announcements on company value in European markets. In *WEIS 2021: The 20th Annual Workshop on the Economics of Information Security* (pp. 1-8).
- Garg, M., Wang, T., & Wilkin, C. L. (2025). Impact of reporting information security breaches, accounting quality, and the opportunistic disclosure of good news and bad news. *International Journal of Accounting Information Systems*, 56, 100729.
- Liu, C., & Babar, M. A. (2024). Corporate cybersecurity risk and data breaches: A systematic review of empirical research. *Australian Journal of Management*, 03128962241293658.
- Malomo, O. O., Gao, S., Adekoya, A. A., Donald, A. M., Andrews Jr, T., Allagan, J. D., ... & Eyob, E. (2025). Toward Human-Centric Cybersecurity: Case Study Insights and the HuCRAF Model. *International Journal of Cybersecurity and Risk Assessment (IJCRA)*, 1(1), 1-61.

- Malomo, O., Gao, S., Adekoya, A. A., Eyob, E., & Gao, W. (2022, December). Cyber-criminals metrics engine for attack determination on web technologies platforms. In 2022 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1053-1058). IEEE.
- Molitor, D., Raghupathi, W., Saharia, A., & Raghupathi, V. (2023). Exploring Key Issues in Cybersecurity Data Breaches: Analyzing Data Breach Litigation with ML-Based Text Analytics. *Information*, 14(11), 600.
- Muktadir-Al-Mukit, D., & Ali, M. H. (2025). The dynamics of stock market responses following the cyber-attacks news: Evidence from event study. *Information Systems Frontiers*, 1-18.
- Newman, N. F., Trautman, L. J., & Elzweig, B. (2025). The SEC Proposed Cybersecurity Infrastructure Rules and New Disclosure Requirements. *UMKC L. Rev.*, 94, 51.
- Peng, J., Zhang, H., Mao, J., & Xu, S. (2023). Repeated data breaches and firm value. *Economics Letters*, 224, 111001.
- Rabitti, G., Khorrami Chokami, A., Coyle, P., & Cohen, R. D. (2025). A taxonomy of cyber risk taxonomies. *Risk Analysis*, 45(2), 376-386.
- Ran, H., & Murphy, F. Cyber Risk in Finance Research: A Literature Review. Available at SSRN 5413366.
- Rasoulilian, S., Grégoire, Y., Legoux, R., & Sénécal, S. (2023). The effects of service crises and recovery resources on market reactions: An event study analysis on data breach announcements. *Journal of Service Research*, 26(1), 44-63.
- Rosati, P., & Lynn, T. (2021). A dataset for accounting, finance and economics research on US data breaches. *Data in Brief*, 35, 106924.
- Schatz, D., & Bashroush, R. (2016). The impact of repeated data breach events on organisations' market value. *Information & Computer Security*, 24(1), 73-92.
- SecurityBrief UK. (2025, June 12). UK firms face rising cyber threats despite record investment. SecurityBrief UK. <https://securitybrief.co.uk/story/uk-firms-face-rising-cyber-threats-despite-record-investment>
- Singh, A. (2025). From past to present: the evolution of data breach causes (2005–2025). *LatIA*, 3, 333-333.

Singh, A., Rana, N. P., Huang, X., & Dubey, N. (2025). The Influence of Data Breach Disclosures on R&D: Firm Strategies at Play. *Information Systems Frontiers*, 1-16.

Smith, K. T., Smith, L. M., Burger, M., & Boyle, E. S. (2023). Cyber terrorism cases and stock market valuation effects. *Information & Computer Security*, 31(4), 385-403.

Vo, H., & Pham, M. D. (2025). Beware of false prophets: Cybersecurity risk and strategic voluntary disclosure. *The British Accounting Review*, 101578.

Zhang, Y. (2024). Data breach disclosure laws and social responsibility initiatives. *The International Journal of Accounting*, 2440003.