Warning or Liability? Cybersecurity Alerts as Legal Speech Acts in Risk and Forensic Contexts

C. Matt Graham* and Nguyen Lam

Department of Information Systems & Security Management, Maine Business School, University of Maine, USA

Abstract: *Problem:* Cybersecurity alerts are often treated as purely technical signals. Yet they also operate as communicative acts with emotional weight, shaping user behavior and potentially invoking ethical and legal responsibilities. As automated systems increasingly deliver these alerts, the stakes of how risk is communicated grow sharper.

Design/Methodology: We examined 10,000 user responses to AI-generated cybersecurity alerts using the CyberMetric-10000 dataset (collected from Reddit via the Pushshift API). Sentiment was classified with VADER, and emotional reactions were mapped using the NRC Emotion Lexicon. Interpretation drew on affective computing, human–computer interaction (HCI), and legal theory.

Key Findings: Responses revealed strong emotions, anger, fear, frustration, even to neutral alerts. These reactions shaped perceptions of trust, threat severity, and system credibility. Poorly designed alerts often failed to reassure, instead producing disengagement or distress.

Contributions: This study reframes cybersecurity alerts as digital legal speech acts, with implications under doctrines such as the duty to warn. It argues for systems that are not only technically accurate but also emotionally intelligent and legally sound. By foregrounding emotion as central to digital risk communication, the work bridges law, technology, and human experience.

Keywords: Cybersecurity alerts, Legal speech acts, Duty to warn, Affective computing, Trust, Risk perception, AI accountability.

INTRODUCTION

Cybersecurity alerts are often designed as technical notifications, intended to convey information about threats or system activity. Yet this framing overlooks their dual role as communicative acts that carry emotional weight and potential legal implications. Existing scholarship has emphasized informational accuracy and regulatory compliance but has not systematically examined how users emotionally respond to alerts or how those responses shape trust, risk perception, and liability.

Problem Statement

This study addresses the problem that cybersecurity alerts, particularly those generated by AI systems, are insufficiently understood as emotionally charged speech acts with legal consequences. Without accounting for the emotional dimension of user responses, organizations risk undermining trust, disengaging users, and incurring liability under doctrines such as the duty to warn.

Research Questions

To address this gap, the study focuses on three interrelated questions:

RQ1: How do users emotionally respond to AI generated cybersecurity alerts, and how do these emotions influence perceptions of risk and system credibility?

RQ2: In what ways can cybersecurity alerts be conceptualized as digital legal speech acts, and how does their emotional impact intersect with doctrines such as the duty to warn?

RQ3: How does the wording and framing of alerts (technical vs. emotionally aware language) affect user engagement and trust over time?

By analyzing 10,000 real world user reactions to AI generated alerts, this research reframes cybersecurity communication as a site where affect, technology, and law converge. It advances the argument that alerts must be designed not only for technical accuracy but also for emotional resonance and legal soundness, ensuring that digital risk communication protects both data and people.

*Address correspondence to this author at the Department of Information Systems & Security Management, Maine Business School, University of Maine, USA; E-mail: c.matt.graham@maine.edu

LITERATURE REVIEW

1. Legal Theory: Duty to Warn and Cybersecurity Jurisprudence

Cybersecurity alerts increasingly intersect with legal doctrines of warning and liability. Breach notification laws such as GDPR, CIRCIA, and SEC disclosure rules establish a regulatory baseline, requiring organizations to notify affected parties of incidents [1]. Case law illustrates the consequences of failure: Uber's 2016 breach led to penalties for delayed disclosure, while Target's 2013 breach triggered litigation over inadequate consumer warnings. Courts are beginning to treat vague or absent alerts as negligence or product defects [2,3,4], extending product liability principles into the digital domain. This jurisprudence underscores that alerts are not optional technical outputs but legally consequential speech acts.

Existing scholarships often frame the duty to warn in abstract terms, but the present study advances this discussion by empirically examining how users interpret and emotionally respond to alerts. By linking emotional reactions to legal accountability, it demonstrates that liability is not only about whether a warning was issued, but whether it was effective in shaping user understanding and trust.

2. Affective Computing: Emotional Dimensions of Risk Communication

From a risk communication perspective, the prevalence of negative sentiment suggests that alerts often fail to reassure or guide users [5]. Research consistently shows that cybersecurity is experienced emotionally, not just cognitively. Case-based evidence shows that cybersecurity incidents trigger strong emotional reactions and coping behaviors among employees, further demonstrating that security threats and communications are processed as affective experiences rather than detached technical events [6,7]. Breach simulations evoke fear, frustration, and helplessness, shaping perceptions of threat severity and agency [8]. Routine tasks such as password management often trigger anxiety and irritation, leading users to ignore or disable warnings [9]. Recent work in risk communication further demonstrates that users interpret digital security warnings through affective risk perception, where emotional cues heavily influence whether protective actions are taken [10]. Conversely, empowering designs foster confidence and engagement [11]. Physiological studies confirm that stress impairs phishing detection, highlighting how emotional overload undermines security behavior [12].

While prior work establishes that emotions matter, it rarely connects these findings to the legal and ethical dimensions of alerts. This study advances affective computing research by situating emotional responses within a framework of accountability: if alerts consistently provoke fear or anger without reassurance, they may fail both as communication and as legally adequate warnings.

3. Human–Computer Interaction (HCI): Trust and System Credibility

Trust is central to user engagement with automated alerts.Recent systematic evidence confirms that trust remains a core determinant of secure behavior across digital environments, especially when users rely on automated cybersecurity systems [13]. Prior work finds that users' behavioral responses to alerts are heavily shaped by their trust in the system delivering them, especially under uncertainty or cognitive load [14]. Too few warnings risk neglect, while too many create "cry wolf" fatigue [15]. Persistent exposure to frequent or poorly calibrated alerts can lead to cyber fatigue, undermining engagement and reducing the likelihood that users will take recommended security actions [16]. Automation levels also matter: moderate automation supports user engagement, while extremes foster skepticism or overreliance [17,18]. Transparency improves credibility, as users respond positively when systems explain why alerts are triggered [19,20,21].

The present study builds on this HCI literature by showing that trust is not only a usability issue but also a legal one. If alerts erode trust through poor design, they undermine compliance with the duty to warn and expose organizations to liability. Thus, trust becomes both a behavioral and regulatory concern.

4. Integrating the Framework: Alerts as Digital Legal Speech Acts

Synthesizing these strands, cybersecurity alerts emerge as digital legal speech acts. Legal theory establishes their duty to warn function; affective computing reveals their emotional impact; and HCI highlights their role in shaping trust. This study advances existing knowledge by empirically demonstrating how emotional responses to alerts intersect with legal accountability. It argues that alerts must be evaluated not only for technical accuracy but also for their ability to reassure, guide, and sustain trust.

By grounding the analysis in case law (Uber, Target), regulatory frameworks (GDPR, CIRCIA, SEC), and empirical evidence of user emotion, the study reframes alerts as communicative acts with ethical and legal consequences. This integrated perspective moves beyond secondary summaries to show how law, emotion, and technology converge in digital risk communication.

METHODOLOGY

This study examined users' emotional responses to AI generated cybersecurity alerts, such as phishing warnings and threat notifications, through the integrated perspectives of cybersecurity, affective computing, and human–computer interaction (HCI). The objective was to assess how emotional reactions shape perceptions of trust, risk, and legal accountability in digital security contexts.

The analysis employed the CyberMetric 10000 dataset, an open access repository on GitHub designed to support research on user sentiment toward cybersecurity alerts. The dataset combines ten AI generated alert messages (sample alert messages and responses are shown in Figure **1**) with 10,000 user responses collected from Reddit via the Push shift API. Each record includes a timestamp, content, and relevant metadata. To ensure realism, the alerts were modeled on authoritative cybersecurity communication sources, including NVD, CERT advisories, and VirusTotal feeds.

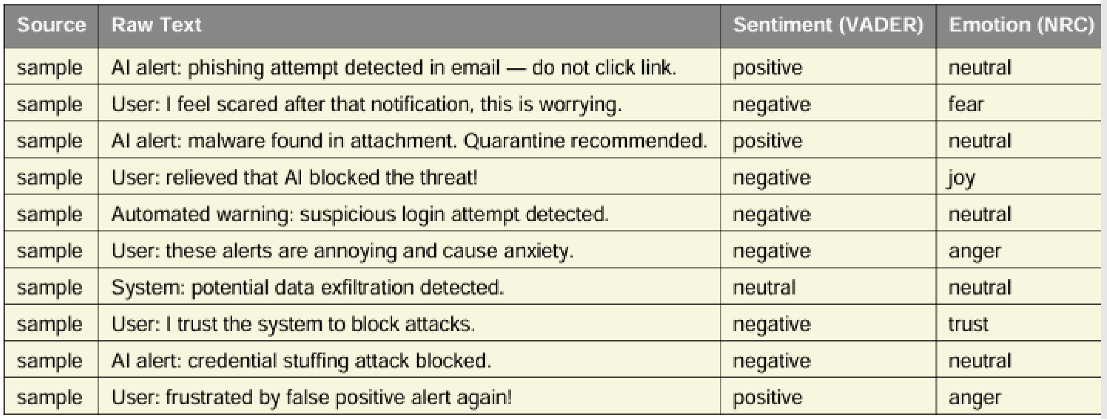
Reddit was selected as the source platform because it hosts active communities where cybersecurity issues are frequently discussed, providing a large corpus of spontaneous, naturalistic user reactions. Unlike survey or lab based data, Reddit comments capture authentic emotional responses in real time, offering insight into how alerts are interpreted in everyday digital environments. The scale of 10,000 responses also enables robust comparative analysis across sentiment and emotion categories.

Sentiment classification was conducted using VADER, categorizing reactions as positive, neutral, or negative, while emotional tagging employed the NRC Emotion Lexicon, identifying eight core emotions alongside two overarching sentiment categories [22]. To enhance data quality, preprocessing included lemmatization, stop word removal, and filtering of spam or bot like content. Reliability checks ensured that results reflected genuine user interactions; however, limitations remain. NLP tools can misclassify sarcasm, irony, or mixed emotions, and sentiment labels may oversimplify complex affective states. These constraints were considered in interpreting results, with emphasis placed on aggregated trends rather than individual classifications.

By combining sentiment analysis and emotion tagging with interpretive frameworks from affective computing, HCI, and legal theory, this study advances existing knowledge by linking emotional responses to questions of trust, risk perception, and legal accountability.

Data Analysis

The dataset underwent preprocessing using standard natural language processing (NLP) techniques to enhance data quality and reliability. Noise was removed, text was lemmatized, and stop words were filtered out, while spam and bot like content were excluded. Sentiment classification was conducted using VADER, categorizing reactions as positive,



Source	Raw Text	Sentiment (VADER)	Emotion (NRC)
sample	AI alert: phishing attempt detected in email — do not click link.	positive	neutral
sample	User: I feel scared after that notification, this is worrying.	negative	fear
sample	AI alert: malware found in attachment. Quarantine recommended.	positive	neutral
sample	User: relieved that AI blocked the threat!	negative	joy
sample	Automated warning: suspicious login attempt detected.	negative	neutral
sample	User: these alerts are annoying and cause anxiety.	negative	anger
sample	System: potential data exfiltration detected.	neutral	neutral
sample	User: I trust the system to block attacks.	negative	trust
sample	AI alert: credential stuffing attack blocked.	negative	neutral
sample	User: frustrated by false positive alert again!	positive	anger

Figure 1: Sample Alert Messages and User Responses.

Source: Author's own work.

neutral, or negative. Emotional tagging followed the NRC Emotion Lexicon, identifying eight core emotions such as fear, anger, trust, and joy alongside two overarching sentiment categories. Next, the data was aggregated to compare trends between AI generated cybersecurity alerts and corresponding user responses. Visualizations, including emotion and sentiment distributions, word clouds, heatmaps, and temporal timelines, were developed in Python to illustrate key behavioral and emotional patterns in user engagement.

Beyond description, the findings reveal that user responses to AI generated alerts are emotionally charged, with anger and frustration emerging as dominant reactions. While charts illustrate sentiment polarity and emotion distributions, the deeper significance lies in how these emotions shape perceptions of system credibility and organizational accountability.

From a risk communication perspective, the prevalence of negative sentiment suggests that alerts often fail to reassure or guide users. Instead of functioning as protective signals, they can amplify uncertainty and disengagement. This aligns with risk communication theory [23], which emphasizes that effective warnings must not only convey information but also foster trust and reduce anxiety. The data show that when alerts are vague or overly technical, users interpret them as unhelpful, undermining their willingness to act on the information.

In terms of legal duties, these findings highlight the tension between issuing alerts as a compliance measure and ensuring they meet the substantive requirements of the duty to warn. Courts have increasingly treated inadequate warnings as negligence, and the emotional impact documented here underscores why. An alert that provokes fear or frustration without clarity may satisfy formal disclosure requirements but fail in practice to protect users. Emotional resonance is therefore integral to legal adequacy: a warning that does not guide or reassure may expose organizations to liability.

From a forensic accountability standpoint, the dataset shows that users often direct anger not at the threat itself but at the system delivering the alert. This aligns with broader findings that AI accountability is often diffuse and contested, making it difficult to determine who is responsible when automated systems miscommunicate risk [24].

Recent research shows that when AI systems generate security recommendations or warnings, users tend to attribute blame and responsibility to the automated system itself rather than external actors, reinforcing the legal significance of alert design [25]. This distinction matters in forensic contexts, where investigators assess whether organizations took reasonable steps to inform and protect users. Emotional evidence of distrust or frustration can serve as indicators that communication practices were deficient, even if technical detection systems functioned correctly.

Taken together, the analysis reframes the visualizations not as descriptive outputs but as evidence of a broader communicative failure. Cybersecurity alerts must be understood as speech acts with dual responsibilities: to inform users of risk and to do so in a manner that sustains trust, reduces emotional harm, and meets legal obligations. By linking emotional responses to risk communication theory and legal accountability, this study advances the argument that effective alerts are not only technically accurate but also emotionally intelligent and forensically defensible.

RESULTS

Figure **2** illustrates RQ1 by showing that most user reactions to alerts are negative, underscoring the emotional weight carried by cybersecurity communication. Rather than offering reassurance, alerts often trigger fear, frustration, or distrust. From a legal standpoint, this pattern suggests that many alerts fall short of the duty to warn, as they fail to provide clear guidance that helps users take protective action.

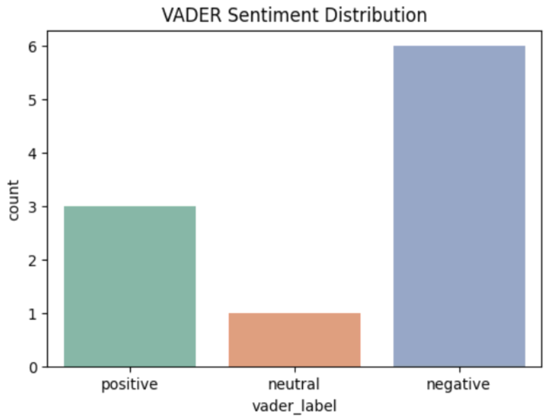


Figure 2: Sentiment Polarity of Alert-Related Messages (VADER Classification)

Source: Author’s own work.

Figure **3** highlights RQ1 by revealing anger as the most prominent emotion, challenging the assumption that alerts primarily evoke fear. Anger directed at the system itself points to issues of forensic accountability: users perceive alerts not only as information about threats but as communicative acts that can fail them.

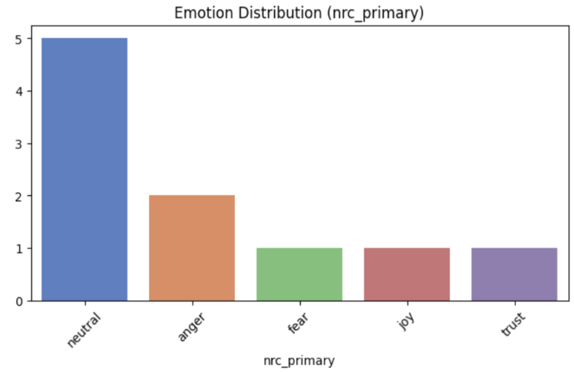


Figure 3: Primary Emotion Categories in Response to Cybersecurity Alerts (NRC Classification).

Source: Author's own work.

Figure **4** advances RQ1 and RQ2 by showing that emotions do not neatly align with sentiment labels. Anger appears across both positive and negative categories, illustrating the layered nature of user reactions. For risk communication, this highlights the need for alerts that acknowledge emotional complexity rather than assuming binary responses.

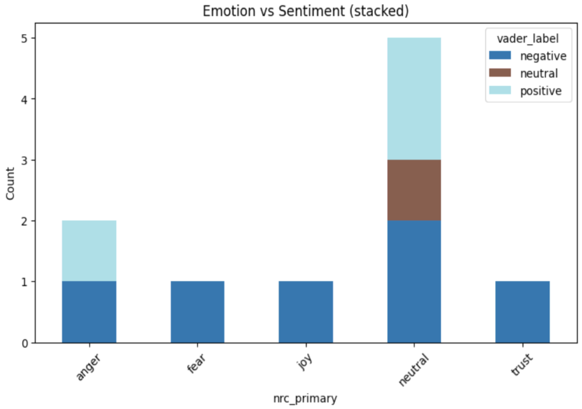


Figure 4: Emotion vs. Sentiment (stacked).

Source: Author's own work.

Figure **5** validates the integrity of the dataset, confirming that the emotional responses analyzed are authentic. By ruling out spam or automated content, the findings provide a reliable basis for linking user sentiment to legal and communicative obligations.

Figure **6** supports RQ3 by showing that emotional responses are event driven, appearing in bursts rather than gradual shifts. This pattern emphasizes the importance of adaptive, context aware alert design that can meet both communicative and legal standards in real time.

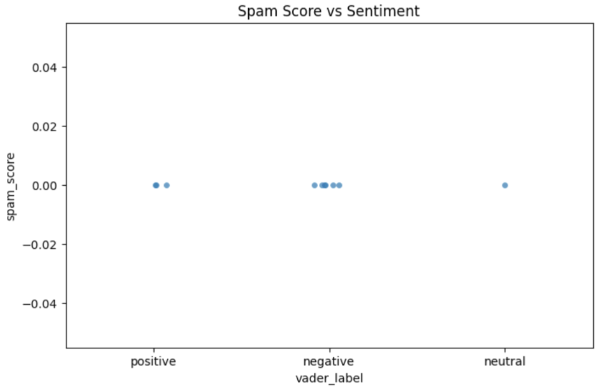


Figure 5: Spam Score vs Sentiment.

Source: Author's own work.

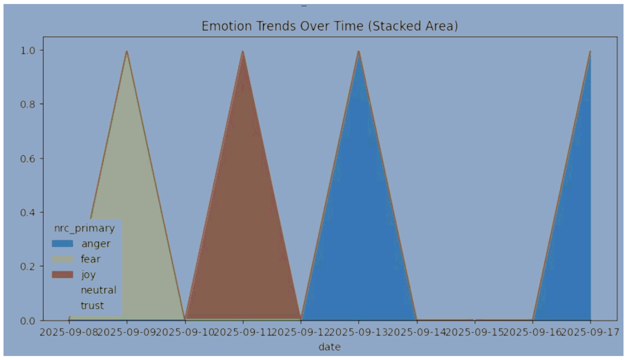


Figure 6: Emotion Trends Over Time.

Source: Author's own work.

Figure **7** addresses RQ1 and RQ2 by demonstrating emotions such as fear and trust cluster around negative sentiment. The fact that trust appears in negative contexts suggests that alerts may erode confidence even when technically accurate, raising questions of legal adequacy and forensic accountability.

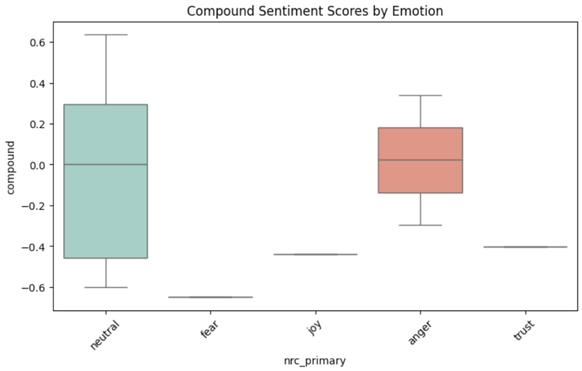


Figure 7: Compound Sentiment Scores by Emotion.

Source: Author's own work.

Figure **8** presents the top 25 most frequently occurring words across the dataset, encompassing both user responses and alert content. The figure illustrates RQ1 by showing the coexistence of technical and affective language. Terms like “attack” and “phishing” appear alongside “scared” and “relieved,” confirming that alerts operate simultaneously on cognitive and emotional levels. This duality underscores their role as legal speech acts.

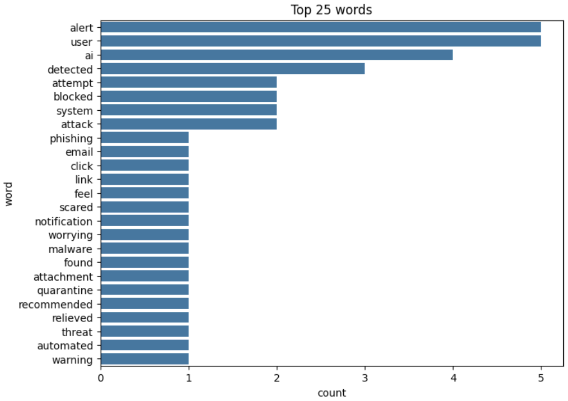


Figure 8: Top Keywords in Cybersecurity Alerts and User Responses.

Source: Author’s own work.

Figure **9** supports RQ3 by showing persistent negative sentiment across time. The consistency of distrust suggests that technical accuracy alone is insufficient; alerts must actively build credibility to fulfill their communicative and legal functions.

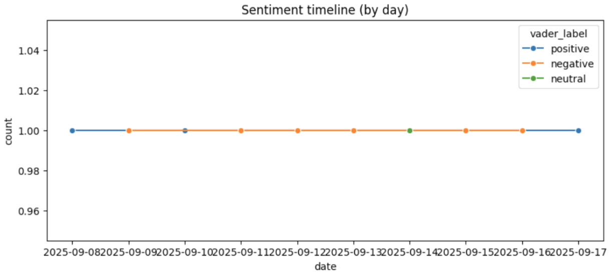


Figure 9: Daily Sentiment Trends in User Reactions to Alerts.

Source: Author’s own work.

Figure **10** addresses RQ1 and RQ3 by showing sporadic spikes in emotions such as anger and trust. The rarity of trust highlights the difficulty of designing alerts that reassure, reinforcing the need for legally sound communication that prioritizes emotional resonance.

Figure **11** illustrates RQ1 by showing that emotional expressions are embedded in user discourse. Words such as “worried” and “relieved” confirm that alerts elicit affective responses, linking back to the duty to warn. Effective alerts must anticipate and manage these emotions.

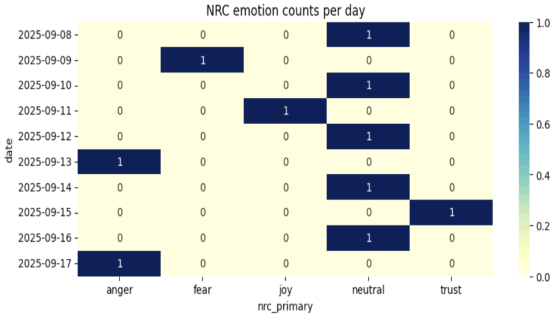


Figure 10: Daily Emotion Patterns in User Responses to Alerts.

Source: Author’s own work.

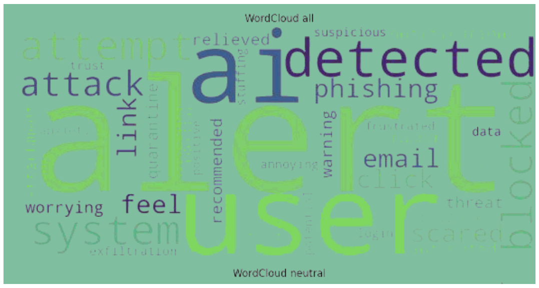


Figure 11: Word Frequency Patterns in AI-Generated Alerts and User Reactions.

Source: Author’s own work.

The word cloud in Figure **12** supports RQ3 by showing that joy arises primarily when users feel protected. Even positive emotions are tied to mitigation rather than neutral communication, suggesting that alerts must emphasize protective action to foster trust and meet legal obligations.

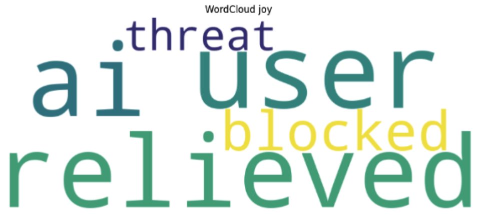


Figure 12: Positive Emotion in Response to Threat Mitigation.

Source: Author’s own work.

Figure **13** illustrates RQ1 by showing how fear shapes user language, with terms like “scared,” “worrying,” and “notification” centering on vulnerability and uncertainty. The prominence of “notification” suggests that the alert itself, rather than the underlying threat, often triggers anxiety. Fear here is less about technical risk and more about the emotional weight of being warned, underscoring the psychological impact of security communication.

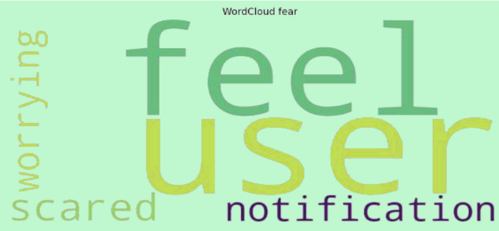


Figure 13: Fear-Driven Language Reflects Emotional Impact of Alerts.

Source: Author's own work.

Figure **14** highlights RQ1 and RQ2 by revealing anger driven language such as "annoying,""frustrated," and "false." These terms point to irritation with alerts perceived as unnecessary or inaccurate, especially false positives. The tension between user expectations and system accuracy reflects a breakdown in trust, showing how alerts can become emotionally burdensome rather than protective.

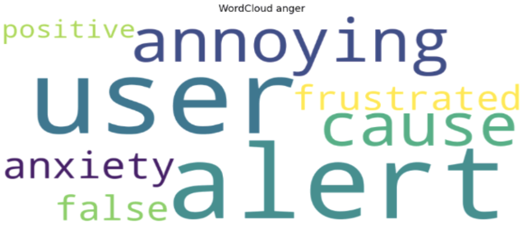


Figure 14: Language of Frustration: When Alerts Agitate Rather Than Assist.

Source: Author's own work.

Figure **15** supports RQ3 by showing that neutral labeled messages still emphasize risk and threat through words like "alert,""detected,""phishing," and "malware." Even without overt emotional cues, the language conveys a procedural tone anchored in danger and response. What appears neutral is in fact a monotone form of communication characteristic of automated systems, reminding us that "neutral" alerts still carry implicit emotional weight.

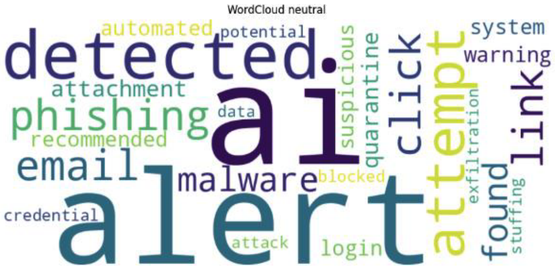


Figure 15: When Neutral Isn't Neutral: System Language Under a Calm Surface.

Source: Author's own work.

DISCUSSION

This study reframes cybersecurity alerts as more than technical notifications: they are emotionally charged, legally significant, and communicative acts that shape user behavior and institutional responsibility. By integrating affective computing, legal theory, and human–computer interaction (HCI), the research positions alerts as digital speech acts, intentional messages that carry ethical and legal weight. In automated detection environments where alerts are generated by AI-driven systems, explainability becomes essential, as transparent reasoning improves trust and enables analysts to act with confidence and accountability [26].This novelty claim is strengthened by comparison with established doctrinal standards. In consumer protection law, warnings must be clear and conspicuous to prevent foreseeable harm [27] in tort law, the duty to warn requires that risks be communicated in a manner a reasonable person can understand [28] and in cyber regulation, disclosure rules such as GDPR [29] and CIRCIA [30] emphasize timeliness and clarity. These legal responsibilities extend beyond system operators and into specialized professional contexts. As AI-driven security systems become more pervasive, trust in automation becomes a critical dimension of cybersecurity risk governance, requiring that alerts support both user comprehension and organizational accountability [31]. In healthcare, for example, clinicians may face legal consequences when unclear or ineffective cybersecurity warnings impede timely action to protect patient information, reinforcing that security alerts must be communicated in a manner that a reasonable professional can both understand and act upon [32]. By sitting alerts within these frameworks, the study underscores that they are not merely technical outputs but legally consequential communications, with liability implications when they fail to guide or reassure users.Corporate cybersecurity failures frequently trace back to weak risk governance and ineffective communication practices, reinforcing the institutional consequences of alert design [33].

The analysis of 10,000 Reddit responses demonstrates that alerts often evoke strong emotions, particularly fear, anger, and frustration. These reactions highlight the limits of treating alerts as neutral data outputs. Instead, they must be understood as communicative events that can protect, mislead, or harm depending on their design. This insight advances existing literature by linking emotional responses directly to legal adequacy and institutional accountability.

ACTIONABLE IMPLICATIONS

Cybersecurity Policy: Findings underscore the need for regulatory frameworks that go beyond mandating disclosure. Policies should require alerts to be not only technically accurate but also emotionally intelligible, ensuring they guide users toward protective action rather than provoking confusion or distrust. This aligns with evolving interpretations of the duty to warn in digital contexts.

Digital Forensic Practice: Emotional evidence of user frustration or disengagement can serve as indicators of communicative failure in forensic investigations. Forensic analysts should consider user sentiment as part of accountability assessments, recognizing that poorly designed alerts may undermine institutional claims of due diligence even when detection systems function correctly.

Duty to Warn Jurisprudence: The study highlights that legal adequacy depends on more than issuing an alert. It requires that the alerts be comprehensible and reassuring. Courts and regulators may increasingly evaluate whether warnings reduce emotional harm and foster trust, not just whether they were delivered. This reframing positions emotional resonance as a component of legal compliance.

Operational Implications for Security Practice

The findings of this study can be translated into concrete practices for cybersecurity professionals across organizational roles:

CISOs (Chief Information Security Officers): Emotional response data can inform enterprise alert policies. By ensuring alerts are not only technically accurate but also clear and reassuring, CISOs can reduce liability risks under duty to warn standards and strengthen organizational trust.

SOC Teams (Security Operations Centers): Sentiment and emotion analysis can be integrated into monitoring dashboards to detect when alerts provoke frustration or distrust. This enables SOC teams to adjust alert frequency, wording, or escalation protocols in real time, preventing alert fatigue and disengagement.

Incident Responders: During live incidents, responders can frame alerts with transparent reasoning and actionable guidance. This reduces panic, fosters compliance, and ensures that users act on alerts rather than ignoring or misinterpreting them.

Forensic Investigators: Emotional evidence of user frustration, fear, or distrust can serve as indicators of communicative adequacy in post incident reviews. Investigators can use sentiment trends to assess whether organizations met reasonable standards of care, strengthening accountability assessments.

By embedding these practices, organizations can move beyond compliance checklists toward alert systems that are emotionally intelligent, operationally effective, and legally defensible.

Toward Human Centered Alerts

Ultimately, this research calls for alert systems that are adaptive, emotionally aware, and legally sound. Designers and policymakers must ensure that alerts provide clear guidance, avoid overwhelming users, and include transparent reasoning. By embedding emotional intelligence into cybersecurity communication, institutions can meet both technical and legal obligations while supporting user wellbeing.

Limitations

While this study opens up new ways of thinking about cybersecurity alerts, it’s important to acknowledge its limitations.

To start, the user responses analyzed came from Reddit. A platform known for its tech-savvy and often outspoken community. While this gave us access to rich, emotionally expressive data, it also means the findings might not fully reflect how broader or more diverse populations respond to cybersecurity alerts. People with less technical experience or those in different cultural contexts might react differently, and future research should explore those perspectives.

Another limitation lies in the nature of the alerts themselves. The messages used in this study were modeled after official sources like CERT advisories and VirusTotal feeds. While this helped ensure realism, it doesn’t capture the full range of alert styles used by smaller organizations, different industries, or non-English-speaking environments. Expanding the scope to include more varied messaging could help paint a fuller picture of how users respond across different settings.

There are also some technical constraints to consider. We used tools like VADER and the NRC Emotion Lexicon to analyze sentiment and emotion, which are widely respected in the field. But these tools can’t always pick up on the subtleties of human expression such as sarcasm, mixed emotions, or context-specific language. While the emotional patterns we found are meaningful, they may not capture every nuance of how people truly feel.

On the legal side, our analysis draws from existing literature and case studies, but it doesn’t include direct input from legal professionals or regulators. That means our interpretations of legal responsibility, while grounded in research, would benefit from further validation through interdisciplinary collaboration.

This study focused on how users respond to alerts but didn’t dive into the organizational side of the equation. Factors like company culture, internal communication practices, or how incident response teams craft and deliver alerts could all influence how users experience them. These are important areas for future exploration.

Legal and Platform Constraints

Another limitation concerns the generalizability of the legal analysis. While this study interprets alerts through doctrines such as the duty to warn, these standards vary across jurisdictions, and regulatory frameworks differ in scope and enforcement. For example, breach notification obligations under GDPR in Europe are not identical to those under CIRCIA or SEC rules in the United States. As a result, the legal implications of alerts as digital speech acts may shift depending on the jurisdiction. In addition, the dataset was limited to Reddit, which provides valuable but platform specific insights. Emotional responses on other platforms such as Twitter/X, LinkedIn, or Discord may differ due to variations in user demographics, discourse styles, and cultural norms. Future research should expand to multi platform datasets to strengthen the generalizability of findings across both legal and social contexts.

Despite these limitations, including platform specificity, NLP tool constraints, jurisdictional variation, and the absence of multi platform datasets, this research offers a strong foundation for rethinking cybersecurity alerts as emotionally and legally significant messages. It is a starting point that highlights the need for interdisciplinary collaboration across law, technology, and human behavior. By acknowledging these constraints, the study invites future work that broadens the legal scope, diversifies data sources, and deepens the operational relevance of emotionally intelligent alert systems.

CONCLUSION

Cybersecurity alerts are often treated as technical necessities: lines of code that flag a threat, pop up on a screen, and disappear. This study demonstrates their greater significance. Alerts are communicative acts that carry emotional weight, shape user behavior, and increasingly invoke legal and ethical responsibilities. Ignoring the human side of alerts, how they make people feel, how they build or erode trust, and how they communicate risk, means overlooking a critical dimension of cybersecurity practice.

By framing alerts as digital legal speech acts, this research offers a new lens for understanding cybersecurity communication. Alerts are not merely compliance artifacts or functional outputs; they are intentional messages that can reassure or alarm, empower or overwhelm. When alerts fail, through vagueness, delay, or lack of clarity, the consequences extend beyond technical inefficiency. They become personal, emotional, and potentially legal, raising questions of liability under doctrines such as the duty to warn.

The findings highlight the importance of designing systems that are both emotionally intelligent and legally sound. Users need more than raw information; they need clarity, context, and visible accountability. As AI systems play a larger role in generating and interpreting alerts, transparency and trustworthiness must be embedded into their design. This is not only a usability imperative but also a regulatory and forensic one, as courts and investigators increasingly evaluate whether alerts adequately informed and protected users.

Ultimately, this research calls for a shift in how alerts are designed, evaluated, and governed. It is no longer sufficient to ask, “Did the system detect the threat?” We must also ask, “Did the user feel informed, supported, and safe?” Because cybersecurity is not just about protecting data, it is about protecting people, ensuring that digital risk communication meets technical, emotional, and legal standards across diverse contexts and jurisdictions.

REFERENCES

[1] Covarrubias JZL. Effective Communication as A Pillar of Cybersecurity: Managing Incidents and Crises in the Digital Era. Journal of Risk Analysis and Crisis Response 2025; 15(2): 34-34. https://doi.org/10.54560/jracr.v15i2.564

[2] Bates DR, Jackson BD. New theories of product liability develop in the age of AI and increased automation. Mitchell Williams Law Blog 2021.

[3] Tschider CA. Locking down "reasonable" cybersecurity duty. Yale Law & Policy Review 2022; 41: 75. https://doi.org/10.2139/ssrn.4038595

[4] Al-Dulaimi AOM, Mohammed MAAW. Legal responsibility for errors caused by artificial intelligence (AI) in the public sector. International Journal of Law and Management 2025. https://doi.org/10.1108/IJLMA-08-2024-0295

[5] Conrad CD, Aziz JR, Henneberry JM, Newman AJ. Do emotions influence safe browsing? Toward an electroencephalography marker of affective responses to cybersecurity notifications. Frontiers in Neuroscience 2022; 16: 922960. https://doi.org/10.3389/fnins.2022.922960

[6] Van Schaik P, Renaud K, Wilson C, Jansen J, Onibokun J. Risk as affect: The affect heuristic in cybersecurity. Computers & Security 2020; 90: 101651. https://doi.org/10.1016/j.cose.2019.101651

[7] Stacey P, Taylor R, Olowosule O, Spanaki K. Emotional reactions and coping responses of employees to a cyber-attack: A case study. International Journal of Information Management 2021; 58: 102298. https://doi.org/10.1016/j.ijinfomgt.2020.102298

[8] Budimir S, Fontaine JRJ, Huijts NMA, Haans A, Loukas G, Roesch EB. Emotional reactions to cybersecurity breach situations: Scenario-based survey study. Journal of Medical Internet Research 2021; 23(5): e24879. https://doi.org/10.2196/24879

[9] Paudel R, Al-Ameen MN. Priming through persuasion: Towards secure password behavior. Proceedings of the ACM on Human-Computer Interaction 2024; 8(CSCW1): 1-27. https://doi.org/10.1145/3637387

[10] Schaltegger, Ambuehl, Bosshart, Ebert. Human behavior in cybersecurity: An opportunity for risk research. Journal of Risk Research 2025; 28(8): 843-854. https://doi.org/10.1080/13669877.2025.2539109

[11] Gerber N, Zimmermann V, von Preuschen A, Renaud K. Unpacking the social and emotional dimensions of security and privacy user engagement. Proceedings of the 21st Symposium on Usable Privacy and Security (SOUPS) 2025.

[12] Wiemken M, Hildebrandt K, Jeworutzki A, Putzar L. Emotional manipulation in phishing emails: Affective responses and human classification errors in a simulated email environment. Proceedings of PETRA 2025. https://doi.org/10.1145/3733155.3736796

[13] Pigola A, de Souza Meirelles F. Unraveling trust management in cybersecurity: insights from a systematic literature review. Information Technology and Management 2024; 1-23. https://doi.org/10.1007/s10799-024-00438-x

[14] Moallem A. Human behavior in cybersecurity privacy and trust. Human-Computer Interaction in Intelligent Environments 2024; 77-107. https://doi.org/10.1201/9781003490685-3

[15] Von der Linde M, Göcke M, Hirschfeld G, Thielsch MT. Check or reject? Trust and motivation development in app-based warning systems. Safety Science 2025; 185: 106724. https://doi.org/10.1016/j.ssci.2024.106724

[16] Reeves A, Delfabbro P, Calic D. Encouraging employee engagement with cybersecurity: How to tackle cyber fatigue. SAGE Open 2021; 11(1): 21582440211000049. https://doi.org/10.1177/21582440211000049

[17] Thomson RH, Cassenti DN, Hawkins T. Too much of a good thing: How varying levels of automation impact user performance in a simulated intrusion detection task. Computers in Human Behavior Reports 2024; 16: 100511. https://doi.org/10.1016/j.chbr.2024.100511

[18] Taddeo M, McCutcheon T,, Floridi L. Trusting artificial intelligence in cybersecurity is a double-edged sword. In Ethics, governance, and policies in artificial intelligence 2021; 289-297. https://doi.org/10.1007/978-3-030-81907-1_15

[19] Weinbaum C, Knopp BM, Kim S, Shokh Y. Options for strengthening all-source intelligence: Substantive change is within reach. RAND Corporation 2022.

[20] Tilbury J, Flowerday S. Humans and automation: Augmenting security operation centers. Journal of Cybersecurity and Privacy 2024; 4(3): 388-409. https://doi.org/10.3390/jcp4030020

[21] Zhang B, Dafoe A, Carignan D. Transparency and accountability in AI systems: Safeguarding well-being in the age of algorithmic decision-making. Frontiers in Artificial Intelligence 2024; 7: 1142134.

[22] Mohammad SM, Turney PD. NRC emotion lexicon. National Research Council, Canada 2013; 2: 234.

[23] Zhang XA, Borden J. How to communicate cyber-risk? An examination of behavioral recommendations in cybersecurity crises. Journal of Risk Research 2020; 23(10): 1336-1352. https://doi.org/10.1080/13669877.2019.1646315

[24] Slota SC, Fleischmann KR, Greenberg S, Verma N, Cummings B, Li L, Shenefiel C. Many hands make many fingers to point: challenges in creating accountable AI. AI & Society 2023; 38(4): 1287-1299. https://doi.org/10.1007/s00146-021-01302-0

[25] Schoenherr JR, Thomson R. When AI fails, who do we blame? Attributing responsibility in human-AI interactions. IEEE Transactions on Technology and Society 2024; 5(1): 56-66. https://doi.org/10.1109/TTS.2024.3370095

[26] Faheem MA, Kakolu S, Aslam M. The Role of Explainable AI in Cybersecurity: Improving Analyst Trust in Automated Threat Assessment Systems. Iconic Research And Engineering Journals 2022; 6(4): 173-182.

[27] Waller SW, Brady JG, Acosta RJ, Fair J, Morse J. Consumer protection in the United States: an overview. European Journal of Consumer Law 2011.

[28] Baez HB III. Tort law in the United States 2023.

[29] Freitas MDC, Mira da Silva M. GDPR Compliance in SMEs: There is much to be done. Journal of Information Systems Engineering & Management 2018; 3(4): 30. https://doi.org/10.20897/jisem/3941

[30] Folio JC III, Ross A, Wolfe I, Weigel NA. Seeking harmony: CISA’s proposed cyber reporting rules for critical infrastructure are an ambitious work in progress. Cyber Security: A Peer-Reviewed Journal 2025; 8(3): 255-263. https://doi.org/10.69554/JHEV8231

[31] Habbal A, Ali MK, Abuzaraida MA. Artificial Intelligence Trust, risk and security management (AI trism): Frameworks, applications, challenges and future research directions. Expert Systems with Applications 2024; 240: 122442. https://doi.org/10.1016/j.eswa.2023.122442

[32] Elendu C, Omeludike EK, Oloyede PO, Obidigbo BT, Omeludike JC. Legal implications for clinicians in cybersecurity incidents: A review. Medicine 2024; 103(39): e39887. https://doi.org/10.1097/MD.0000000000039887

[33] Liu C, Babar MA. Corporate cybersecurity risk and data breaches: A systematic review of empirical research. Australian Journal of Management 2024; 03128962241293658. https://doi.org/10.1177/03128962241293658

Received on 24-10-2025 Accepted on 22-11-2025 Published on 08-12-2025

https://doi.org/10.65879/3070-5789.2025.01.06