International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

Artificial Intelligence-Driven Risk Prioritization in Automated Web Vulnerability Assessment

Kulyk Anton

CEO and Founder, Cyber Trust Innovations LLC, Tampa, FL, USA

Abstract: ***The article examines the transition from static models of web vulnerability assessment to intelligent, context-dependent risk prioritization enabled by artificial intelligence. The relevance of the study is driven by the rapid growth in the number of vulnerabilities, the increasing complexity of web application architectures, and the mounting overload experienced by security analysts, for whom traditional scales, primarily CVSS, no longer provide an adequate distinction between formally critical and genuinely exploitable threats. The aim of the article is to provide a theoretical substantiation and conceptualization of approaches to improving the efficiency of automated web vulnerability assessment by integrating probabilistic models, machine learning, and contextual analysis. The scientific novelty of the work lies in the interdisciplinary synthesis of DAST methods, NLP, predictive models such as EPSS, multimodal ML ensembles, and the practical case of the VULNWatch platform within a unified analytical ecosystem. The principal conclusions demonstrate that AI-driven prioritization enables more accurate forecasting of exploitation risk, substantially reduces the share of false positives, decreases the burden on SOC teams, and shifts vulnerability management from a reactive mode to a predictive one, provided that it is mandatorily accompanied by explainable AI mechanisms and regulatory oversight. The article will be useful for researchers, cybersecurity specialists, SOC analysts, DevSecOps engineers, and developers of vulnerability management platforms.***

Keywords: artificial intelligence, risk prioritization, web vulnerabilities, cybersecurity, VULN Watch

1. Introduction

The contemporary cybersecurity landscape is characterized by an increase in the number of detected software vulnerabilities and by the complexity of attack vectors, which are increasingly generated by attackers themselves through automated and intelligent systems [1]. In the coming years, the global network will face more than 1 million documented common vulnerabilities and exposures, an increase relative to indicators observed at the beginning of the current decade [2]. Under conditions of such rapid scaling of the digital attack surface, traditional methods of vulnerability assessment and management, relying predominantly on static rules, manual triage, and isolated scanners, are demonstrating their conceptual and operational insufficiency [3].

Specialists in security operations centers are confronted with a critical level of information overload known as alert fatigue [4]. On average, large corporate networks and distributed web applications generate thousands of vulnerability alerts daily [5]. At the same time, a full manual investigation of a single incident or the validation of a complex web vulnerability, such as a multi-stage SQL injection or a blind cross-site request forgery, requires hours of work by highly qualified engineers [6]. This makes it physically impossible to process the entire volume of incoming threats, forcing organizations to make decisions under conditions of extreme time and context scarcity.

The principal problem with the existing paradigm is the inefficiency of widely accepted static assessment systems, such as the Common Vulnerability Scoring System. Historically, CVSS has provided a standardized industry metric for determining the baseline technical severity of an identified vulnerability. Despite this, the approach fundamentally ignores the dynamic context of the deployment environment, the actual probability of exploitation in the wild, and the business criticality of the specific targeted asset [7]. As a result, organizations squander substantial computational and human resources on the urgent remediation of vulnerabilities with high baseline CVSS scores. At the same time, such vulnerabilities may have an almost zero probability of real-world exploitation, while concealed yet critical infrastructure breaches go unnoticed.

In this connection, the need to implement artificial intelligence systems and machine learning algorithms for automating web vulnerability assessment and prioritization is critically urgent. Artificial intelligence opens the door to a qualitative transition from reactive defense to predictive security. This enables systems to continuously learn from historical data, analyze software dependency graphs, extract semantic context from threat reports, and dynamically rank risks.

The aim of the study is to develop theoretically grounded approaches to improving the efficiency of web vulnerability assessment and risk prioritization in automated systems using artificial intelligence technologies.

To achieve this aim, the following objectives were formulated:

- Identify the key problems with traditional approaches to assessing and prioritizing web vulnerabilities amid growing cyber threats and the increasing complexity of web application architectures.
- Analyze the current state and development trends of web vulnerability assessment systems: the transition from static metrics to probabilistic and multimodal models based on artificial intelligence.
- Analyze the architecture of contemporary automated AI platforms for threat aggregation, using the author's project, VULNWatch, as an example.
- Determine the factors that have driven these changes, along with the associated risks and limitations of introducing AI into vulnerability management processes.
- Substantiate possible ways of improving web vulnerability assessment and risk prioritization through

the use of AI models, explainable AI methods, and regulatory governance mechanisms.

The scientific novelty of this work lies in the comprehensive interdisciplinary synthesis of theoretical concepts of multimodal machine learning with real market data and proprietary web traffic metrics of information security platforms. The study examines in detail the convergence of dynamic testing tools, natural language processing algorithms, and global databases of artificial intelligence incidents within a unified analytical ecosystem.

2. Literature Survey

Contemporary studies on vulnerability prioritization proceed from the premise that formal severity assessment, in itself, weakly reflects an organization's operational risk. A review of exploitability assessment studies demonstrates that the literature is gradually shifting from manual, deterministic schemes toward probabilistic models [8]. These models use vulnerability description features, information about already observed exploitation, and additional contextual parameters. The same direction is evident in the line of work on exploitation forecasting, where remediation priority is linked to the probability of an attack emerging in a real environment [9]. For automated web vulnerability assessment, this is especially important, since the volume of findings after scanning is usually high, while the value of each finding for decision-making varies sharply depending on external context, the maturity of the exploit, and the criticality of the affected component. From this research trajectory, the central conclusion of the present article follows. Intelligent prioritization should operate as a layer above automated detection results and translate technical findings into a ranked list of actions suitable for practical risk management.

The literature on automated web application testing shows that the field of detection has long moved beyond simple pattern matching. A systematic mapping of studies on web application security testing records a diversity of methods. These combine static analysis, dynamic testing, hybrid procedures, and tooling chains with varying depth of coverage [10]. More recent studies focus on complex defect classes for which superficial scanning yields limited results. Thus, works on the automatic detection of access control violations show that the real problem lies in reconstructing the semantics of permissions, state logic, and hidden transitions between requests [11]. This means that the quality of the initial assessment is determined by the method's ability to extract semantic dependencies from application behavior. Therefore, in the web environment, artificial intelligence proves useful above all where there is a need to connect disparate features, eliminate noise, reduce the proportion of false positives, and isolate genuinely dangerous scenarios from a vast array of technical signals.

The most promising direction over recent years is context-dependent ranking, in which data about a discovered vulnerability are enriched with external threat intelligence, temporal threat features, and characteristics of the specific infrastructure. Statistical models under partial information indicate that risk ranking must account for uncertainty as an independent factor; otherwise, the order of vulnerability remediation becomes fragile and poorly transferable to a production environment [12]. In parallel, an infrastructural foundation for such models is taking shape. Extended datasets are emerging that combine multiple sources of information about vulnerabilities, exploits, and defensive measures. This enables training systems to perceive a vulnerability as a node within a dense network of relationships. Practice-oriented studies confirm the practical value of this approach and show that integrating temporal threat information, expert input, and asset characteristics yields a more meaningful remediation sequence [8]. Taken together, these studies lead to the logic of the present article. The most mature formulation of the problem consists of combining automated web assessment with an intelligent prioritization module that takes into account the probability of exploitation, asset context, incompleteness of knowledge, and the structure of incoming data.

3. Problem Definition

The problem addressed in this study is that existing approaches to web vulnerability assessment cope poorly with the realities of the modern digital environment, where the number of detected defects is growing rapidly, web application architectures are becoming increasingly distributed, and attacks are evolving faster than formal severity scales can be updated. Traditional ranking systems rely predominantly on fixed indicators of technical risk and therefore poorly reflect the likelihood of practical exploitation of a vulnerability, the significance of the affected resource, the characteristics of the network environment, and the consequences for a specific organization. Because of this, automated protection environments accumulate a large volume of signals, making it difficult to identify truly critical threats in a timely manner, leading to analyst overload, increased false positives, and inefficient resource allocation. Under such conditions, a scientific and applied task arises: to develop an approach that, on the basis of artificial intelligence, makes it possible to combine technical, behavioral, and contextual features, calculate risk more flexibly and accurately, and then form a substantiated sequence for eliminating web vulnerabilities in a mode aligned with continuous development and deployment.

4. Methodology

To ensure a comprehensive, objective, and representative analysis of the problems associated with the application of artificial intelligence in automated web vulnerability assessment, this study employed an integrated methodological apparatus. The principal scientific toolkit consisted of literature review methods, comparative analysis of conceptual models of risk classification, content analysis of cybersecurity system technical documentation, and case-study methodology with the application of quantitative performance metrics. Structuring the source base enabled the analysis to be divided into theoretical, practical, and regulatory-market categories.

The first and most extensive group of sources, which formed the theoretical and algorithmic basis of the study, comprised current peer-reviewed scientific publications for 2020-2026 indexed in the international databases Scopus and Web of

Science, as well as in the electronic libraries IEEE Xplore, ACM Digital Library, SpringerLink, and ArXiv. The search was carried out using the key lexemes machine learning vulnerability prioritization, DAST deep learning, CVSS EPSS artificial intelligence, and web vulnerability assessment frameworks. From this group, fundamental concepts were extracted regarding the application of ensemble machine learning methods, natural language processing, and graph neural networks to risk classification. This group of sources ensured scientific rigor in describing the mathematical principles underlying AI prioritization.

The second group consisted of sources that provided the practical and empirical basis, grounded in case studies and the content analysis of documentation for real information security systems. The main focus of this part of the study was the technical specifications and web traffic analytics data for the VULNWatch platform. The analysis of this authorial project enabled examination of the real architecture of vulnerability aggregation using AI to consolidate outputs from a broad pool of specialized tools, including OWASP Zed Attack Proxy, Burp Suite, Metasploit, Nmap, sqlmap, WPScan, Decodo proxy, and others. The integration of the platform's empirical web traffic data enabled validation of demand for such solutions in the international market of specialists.

The third group of sources encompasses the regulatory-legal, managerial, and market bases of the study. Since the deployment of autonomous AI algorithms is associated with colossal risks of black-box behavior, hallucinations, and adversarial attacks, the methodology included an analysis of trusted AI structures. The analysis of risks associated with the exploitation of artificial intelligence models is based on studying the global AI incident database, which indexes real cases of compromise, such as large-scale distillation attacks against LLMs. All public sources were carefully filtered to exclude unreliable blogs and news press releases, thereby ensuring an exclusively professional focus of the methodology.

5. Results and Discussion

The introduction of machine learning and artificial intelligence models into the web vulnerability assessment ecosystem radically transforms the very nature of the work performed by security analysts, shifting it from the plane of endless reactive elimination of symptoms to the plane of strategic, probabilistic risk forecasting.

5.1 Evolution of risk assessment systems

The shift in risk-prioritization methodology is associated with the gradual yet inevitable abandonment of the exclusive use of baseline CVSS metrics in favor of stochastic algorithms that account for the continuously changing environment and the real threat of exploitation by attackers. The nature of the baseline CVSS assessment is static and deterministic: it is fixed at the time a vulnerability is published in the National Vulnerability Database and is revised extremely rarely [13]. However, from the standpoint of corporate risk management, the presence of a critically rated vulnerability in an isolated internal subnet protected by zero-trust policies represents a completely different, orders-of-magnitude lower level of real risk than a vulnerability located on a publicly accessible load-balancer web server for which working exploit code has already been published on the darknet.

The solution to this problem was the creation and active implementation of predictive systems, the flagship of which is the Exploit Prediction Scoring System. Developed by a group of experts [14], EPSS differs fundamentally from deterministic CVSS formulas. EPSS uses logistic regression and gradient-boosted decision trees to process vast amounts of global cyber threat intelligence data. The model is trained on sixteen predictors, forecasting the probability that a specific software vulnerability will be exploited in the wild within the next 30 days. The integration of EPSS with machine learning models enables analysis of dynamic features such as the age of the vulnerability, mentions on social media, discussions on darknet forums, port-scanning activity, and the availability of modules for automated platforms such as Metasploit.

In addition to global systems such as EPSS, studies demonstrate the exceptionally high effectiveness of custom, local ML frameworks for automatically converting and adapting assessments to the specific needs of a particular enterprise. An important academic achievement is the study [15], which successfully resolved the problem of the shortage of up-to-date CVSS v3.x data for older vulnerabilities. The researchers applied machine learning algorithms to automatically transform obsolete CVSS v2.0 metrics into more accurate, contextually relevant CVSS v3.x assessments. The use of an ensemble of methods, including the k-nearest neighbors algorithm with cosine similarity and Euclidean distance, a naïve Bayes classifier, probabilistic neural networks, and support vector machines, made it possible to minimize the human factor. The application of natural language processing using the NLTK library and the optimization of feature vectors via principal component analysis enabled the system to extract semantics from CVE textual descriptions with high accuracy. Similar approaches are actively used in adjacent areas, including IoT network security auditing and threat detection [16].

Even more sophisticated multimodal architectures that directly analyze program code and configuration vectors combine semantic representations from graph neural networks, for example, GraphCodeBERT, with traditional statistical features. One such study demonstrates the use of ordinal logistic regression and class-weighted gradient boosting for predicting risk severity. The trained multimodal model achieved high recall accuracy for critical vulnerabilities. Remarkably, the architecture was intentionally programmed to maintain a conservative security profile: the algorithm overestimated risk in half of ambiguous situations, thereby preventing catastrophic false-negative omissions while reducing the overall need for manual verification [17].

To aid visual conceptualization and the generalization of the described fundamental shifts in prioritization methodologies, Table 1 presents a comparative analysis of traditional and AI-driven systems.

Table 1: Comparative analysis of paradigms and methodologies for prioritizing web vulnerabilities

Evaluation characteristic	Traditional approach, CVSS Base	Predictive global approach, EPSS	Local multimodal AI triage, ML-Ensembles
Nature of the calculated metric	Static, strictly deterministic	Probabilistic, 0–100%, dynamic	Context-dependent, adaptive, semantic
Update frequency and triggers	Extremely rare, only during manual CVE review	Daily probability recalculation	Real time, as logs/code are updated
Consideration of local topology and environment	Completely absent, global metric	Absent, reflects global activity	Very high, considers network graphs, CI/CD, logs
Exploit analysis methodology	Documented proof of concept, PoC	Logistic regression based on CTI feeds and darknet data	Correlation of EPSS with internal enterprise telemetry
Impact on SOC efficiency	Negative, excess of critical alerts	Positive, focus on relevant threats	Maximum, reduction of routine triage by up to 95%

5.2 Analysis of the VULNWatch analytics platform

The conceptual theory of AI-oriented risk assessment finds its direct practical embodiment in the development of next-generation aggregating cloud platforms. A good example of such synergy is the VULNWatch project. An analysis of the technical documentation and operating principles of this platform clearly demonstrates that modern web scanners are evolving. They function as high-level orchestrators that unite heterogeneous data from numerous disparate sources in order to form a unified, weighted threat picture.

A distinctive feature of the VULNWatch architecture is the rejection of reliance on a single scanning engine. The platform performs intelligent aggregation, simultaneously integrating the outputs of advanced, highly specialized tools recognized as industry standards. The tool stack managed by the platform includes dynamic testing solutions, such as DAST tools like OWASP Zed Attack Proxy and Burp Suite Professional, which intercept and modify HTTP traffic to identify architectural vulnerabilities at runtime. Network topology and attack surface mapping are provided through Nmap integration. For the targeted identification of critical database vulnerabilities, such as various forms of blind SQL injection, automated SQLmap engines are used. Specific attack vectors against content management systems, covering the core, plugins, and themes, are scanned by means of WPScan.

However, running multiple scanners would inevitably lead to data duplication and excessive information overload. It is precisely here that the VULNWatch AI Module, a proprietary AI module, plays a key role. This component correlates complex threat patterns by analyzing raw scanner logs and comparing them with a massive volume of global telemetry. To this end, the AI module accesses government databases, the OWASP Top Ten rankings, and open-source intelligence in real time. The analytical loop incorporates data from the GitHub Advisory Database for tracking supply-chain vulnerabilities, Certificate Transparency log search engines for detecting compromised subdomains, sandboxes for behavioral analysis of web pages, and threat intelligence feeds tracking malware distribution nodes. As an advanced function for detecting concealed redirects and masking techniques, the platform routes scanning traffic through custom proxies and nodes of The Onion Router anonymous network, simulating the real behavior of sophisticated attackers.

The final result of the AI orchestrator's operation is a list of discovered bugs and a prioritized alert pipeline with the automatic generation of individualized remediation recommendations. The effectiveness and high market demand for such a comprehensive approach are objectively confirmed by the platform's web traffic metrics shown in Figure 1.

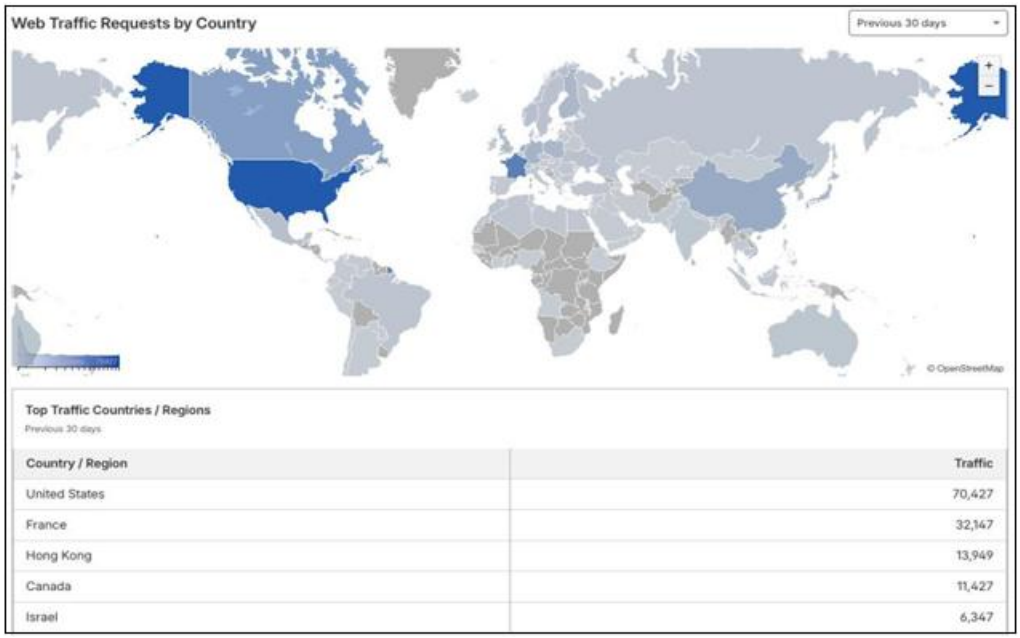

Figure 1: Geographic distribution of unique web queries to the VULNWatch AI analytics platform over a 30-day period

An analysis of the telemetry of the domain vulnwatch. tech over a typical thirty-day period demonstrates that the resource attracts cybersecurity specialists from the United States. Over the specified period, the platform recorded 12.22 thousand unique visitors. The traffic is characterized by pronounced volatility: daily activity ranges from a minimum of 344 to a peak of 1080 visitors, which correlates with periods of mass publication of new exploits or updates to CVE databases.

To systematize the operating principles of the described AI aggregator, reflecting data flows from raw logs to final instructions, a conceptual architecture diagram is provided below.

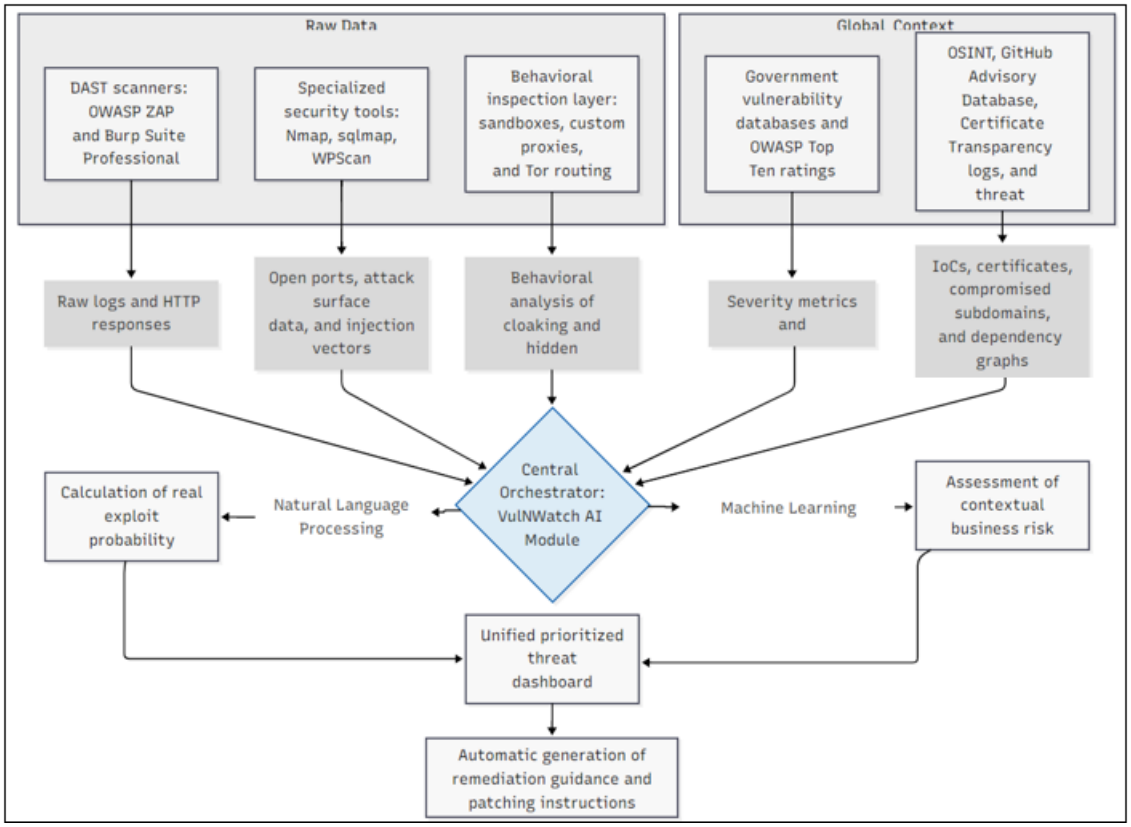

Figure 2: Logical architecture for intelligent scanning, aggregation, and prioritization of web vulnerabilities

5.3 Barriers to implementation and risks

Despite the evident operational and financial advantages, the large-scale implementation of autonomous AI-driven vulnerability prioritization is associated with several fundamental technological and ethical barriers. A critical limitation of the overwhelming majority of contemporary deep learning architectures is the problem of interpretability, widely known in the academic environment as the black box phenomenon. In the high-risk environment of corporate cybersecurity, it is categorically insufficient for a security operations center analyst to receive from an algorithm merely a final abstract estimate of attack probability, for example, Risk 89%. The specialist must understand the internal logic and causal relationships underlying the algorithm's decision in order to provide a legitimate justification for blocking a business-critical service or for urgently deploying a patch outside standard CI/CD procedures.

To overcome this conceptual barrier, advanced assessment frameworks, such as dynamic cybersecurity risk management, are actively integrating the paradigm of explainable artificial intelligence. Within such systems, the use of deep learning is supplemented by mathematical interpretation methods, among which SHAP and LIME are recognized as the most effective. In particular, the use of a vector of Shapley values allows researchers to quantitatively assess the contribution of each individual CVSS parameter or network telemetry feature to the final predicted vulnerability score. The algorithm may clearly indicate that it was precisely the factor network attack vector in combination with partial integrity violation that became the decisive trigger for elevating the alert status. The LIME toolkit, in turn, forms local linear approximations of a complex model around a specific incident, explaining in natural language why a given alert has been classified in precisely this way. The implementation of XAI components is critically important for increasing operational trust on the part of engineers and for ensuring compliance with strict regulatory audit requirements obliging corporations to transparently document all automated IT risk management procedures.

The second, no less significant barrier is the cyber vulnerability of the artificial intelligence systems themselves. Machine learning used to protect infrastructure may itself become a vector of sophisticated attack. Models are susceptible to adversarial attacks, manipulation of the training set, and theft of the algorithm's intellectual property. Contemporary neural networks regularly become targets of sophisticated distillation attacks [17]. As a textbook example, incidents of industrial scale may be considered in which attackers or competing laboratories, such as DeepSeek and

MiniMax, used so-called hydra clusters. These are accumulations of tens of thousands of fraudulent accounts passed through anonymous proxy networks. The objective of these attacks was to carry out millions of automated queries to proprietary oracle models in order to extract their weights, code-generation logic, and reasoning capabilities. The knowledge thereby obtained is used by attackers to train their own shadow algorithms devoid of built-in ethical constraints and protective filters. The dependence of corporate web scanners on external LLM-API providers poses a direct threat: distorted or fabricated chatbot responses may lead to false classification of vulnerability criticality, thereby opening backdoors in enterprise infrastructure.

To minimize such complex risks, the industry is purposefully moving toward strict standardization and the formation of rigorous governance structures. A leading role in this process is played by non-profit research initiatives such as Project Cerebellum, functioning as an analytical center and community of practitioners focused on developing trusted artificial intelligence metrics. The platform offers corporations the TAIMScore framework, a specialized maturity assessment and risk management system based on gamified threat modeling [18].

The success and economic efficiency of integrating artificial intelligence into web vulnerability assessment depend on the quality of the orchestration process. Disparate static and dynamic analysis tools, when used outside a unified environment, will continue to generate uncontrollable volumes of informational noise. The true breakthrough value of AI for the cybersecurity industry lies in its unique ability to serve as a universal cross-platform aggregator that cross-validates raw logs through the prism of global intelligence and probabilistic matrices, filtering out false signals while accounting for the topology and business context of a specific enterprise.

6. Conclusion

The conducted study of concepts, metrics, and practical case studies confirms that the fundamental transition to web vulnerability assessment and automated risk prioritization based on artificial intelligence architectures constitutes an incremental improvement of existing tools. This is a mandatory strategic condition for the survival of corporate IT infrastructure in the realities of exponential hyper-scaling of digital threats. All goals and objectives formulated at the outset of the work were fully achieved. In the course of the study, the conceptual imperative of replacing obsolete static CVSS metrics with predictive probabilistic models and multimodal machine learning ensembles was analyzed in detail and scientifically substantiated. These innovative approaches demonstrated their ability to mathematically and precisely predict the probability of defect exploitation in the wild, thereby reducing the estimated time required for urgent vulnerability remediation and minimizing the risk of system compromise by more than an order of magnitude.

The practical and economic significance of the presented work is grounded in the in-depth analysis of real corporate implementations. The metrics obtained during the study of the architecture of the aggregating VULNWatch platform clearly and quantitatively prove the effectiveness hypothesis. The application of convolutional neural networks, natural language processing technologies, and vector similarity analysis enables security analysts to reduce monotonous manual labor almost completely. Simultaneously, detection quality is improved through reduced false positives and fewer critically dangerous false negatives.

Nevertheless, the study emphasizes that accelerated AI integration is associated with serious technological challenges that require an exceptional interdisciplinary approach. The problem of semantic interpretability of complex predictive models makes the implementation of explainable AI architectures based on SHAP and LIME algorithms obligatory in order to preserve audit control and human oversight over decision-making. Uncontrolled, blind trust in language models exposes organizations to fundamentally new risks of adversarial attacks and unpredictable algorithmic hallucinations. These conditions the acute and urgent necessity of integrating strict regulatory governance mechanisms.

From a strategic perspective, current AI-driven prioritization models will serve as the technological foundation for the development of fully autonomous immune cyber systems. Within such ecosystems, the continuous process of detection, probabilistic severity assessment, automated resilience testing, and deployment of adaptive micropatches will occur in real time, without direct intervention by engineers, ensuring an unprecedented level of operational fault tolerance and data protection for contemporary digital society.

References

[1] Mohamed N. Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms. Knowledge and Information Systems. 2025 Apr 30;67:6969–7055.
[2] Wochnik J, Gräupner OS, Spranger M, Hummert C. Regarding the Exponential Growth of Security Vulnerabilities. In: Daimi K, Arabnia, Hamid R, Deligiannidis L, editors. Security and Management and Wireless Networks. Cham: Springer Nature Switzerland; 2025. p. 329–43.
[3] Malkawi M, Alhajj R. AI-Powered Vulnerability Detection and Patch Management in Cybersecurity: A Systematic Review of Techniques, Challenges, and Emerging Trends. Machine Learning and Knowledge Extraction. 2026 Jan 15;8(1):19.
[4] Tariq S, Chhetri MB, Nepal S, Paris C. Alert Fatigue in Security Operations Centres: Research Challenges and Opportunities. ACM Computing Surveys. 2025 Mar 12;57(9):1–38.
[5] Saraiva M, Mateus-Coelho N. CyberSoc Framework a Systematic Review of the State-of-Art. Procedia Computer Science. 2022; 204: 961–72.
[6] Rao S. After the Breach: Incident Response within Enterprises. arXiv. 2024.
[7] Coutinho LS, Menasche D, Miranda L, Lovat E, Kumar SG, Ramchandran A, et al. How Context Impacts Vulnerability Severity: An Analysis of Product-Specific CVSS Scores. Proceedings of the 13th Latin-American Symposium on Dependable and Secure Computing.

2024 Nov 26;17–27.
[8] Alqahtani N, Almukaynizi M. VulnScore: A deployed system for patch prioritization combining human input and temporal threat intelligence. International Journal of Information Security. 2025 Nov 27;25(1).
[9] Elder S, Rahman MR, Fringer G, Kapoor K, Williams L. A Survey on Software Vulnerability Exploitability Assessment. ACM Computing Surveys. 2024 Mar 20;56(8):1–41.
[10] Aydos M, Aldan Ç, Coşkun E, Soydan A. Security testing of web applications: A systematic mapping of the literature. Journal of King Saud University - Computer and Information Sciences. 2021 Sep;34(9).
[11] Rennhard M, Kushnir M, Favre O, Esposito D, Zahnd V. Automating the Detection of Access Control Vulnerabilities in Web Applications. SN Computer Science. 2022 Jul 15; 3: 376.
[12] Angelelli M, Arima S, Catalano C, Ciavolino E. A robust statistical framework for cyber-vulnerability prioritisation under partial information in threat intelligence. Expert systems with applications. 2024 Jun 1;255(B):124572.
[13] FIRST. CVSS v4.0 User Guide [Internet]. Forum of Incident Response and Security Teams. 2023 [cited 2026 Mar 4]. Available from: https://www.first.org/cvss/user-guide
[14] Jacobs J, Romanosky S, Suciu O, Edwards BR, Sarabi A. Enhancing Vulnerability Prioritization: Data-Driven Exploit Predictions with Community-Driven Insights. arXiv. 2023 Feb 27.
[15] Balsam A, Walkowski M, Nowak M, Oko J, Sujecki S. Automatic CVSS-Based Vulnerability Prioritization and Response with Context Information and Machine Learning. Applied Sciences. 2025;15(16):8787.
[16] Shaikhanova A, Kuznetsov O, Tokkuliyeva A, Ayapbergenov K, Olzhas S, Danir T. Security Audit of IoT Device Networks: A Reproducible Machine Learning Framework for Threat Detection and Performance Benchmarking. Sensors. 2025 Dec 11;25(24):7519.
[17] Hu H, Pang J. Stealing Machine Learning Models: Attacks and Countermeasures for Generative Adversarial Networks. Proceedings of 2021 Annual Computer Security Applications Conference. 2021 Dec 6.
[18] McGregor S. Preventing Repeated Real World AI Failures by Cataloging Incidents: The AI Incident Database. Proceedings of the AAAI Conference on Artificial Intelligence. 2021 May 18;35(17):15458–63.