

EVOLUTION OF SOFTWARE SECURITY VULNERABILITIES: A LARGE-SCALE EMPIRICAL ANALYSIS OF CVE SEVERITY, ATTACK VECTORS, EXPLOITABILITY, AND VULNERABILITY TRENDS

Asim Ismail^{*1}, Aqeel Hassan², Shah Zaib Sajid³, Nazia Azim⁴, Hameed Hussain⁵

¹Girne American University

²Southern cross university

³Institute: Ebryx (pvt) Ltd

⁴Lecturer department of computer science Abdul Wali Khan University Mardan

⁵Department of Computer Science, University of Buner, 19290, Khyber Pakhtunkhwa, Pakistan

¹asimismail7@gmail.com, ²aqeelhassan451@gmail.com, ³shahzaib.sajid@ebryx.com,

⁴n.azim@awkum.edu.pk, ⁵dr.hameed@ubuner.edu.pk

DOI: <https://doi.org/10.5281/zenodo.22828737>

Keywords

CVE; CVSS; NVD; CISA KEV; CWE; software vulnerabilities; vulnerability management; exploitability; cybersecurity; empirical analysis.

Article History

Received: 11 June 2026

Accepted: 02 September 2026

Published: 18 September 2026

Copyright @Author

Corresponding Author: *

Asim Ismail

Abstract

Software vulnerability disclosure has become a high-volume global security process, yet raw CVE counts, technical severity, and real-world exploitation are often conflated. This study analyzes 366,873 unique published CVE records spanning 1999–1 September 2026 and performs a focused longitudinal analysis for the complete years 2016–2025. The snapshot contains 195,907 records with CVSS scores, 182,710 with CWE mappings, and 1,687 CISA Known Exploited Vulnerabilities (KEV). Among scored records, Medium severity accounted for 46.85%, High for 38.21%, Critical for 10.51%, Low for 4.38%, and 108 records (0.06%) had a base score of 0.0 and were classified as None. Severity distributions changed across 2016–2025 ($\chi^2(36)=1296.4$, $p<0.001$; Cramér's $V=0.048$); a CVSS v3.x-only sensitivity analysis of 124,082 records yielded a similarly small association ($\chi^2(36)=1046.0$, $p<0.001$; $V=0.046$). Metadata completeness increased sharply over time, with CVSS coverage rising from 1.50% in 2016 to 93.86% in 2025. KEV-listed vulnerabilities had higher CVSS scores than non-KEV-listed records (median 8.8 vs 6.9). In the adjusted logistic model ($n=195,784$; 1,686 KEV-listed), each one-point increase in CVSS score was associated with 1.99-fold higher odds of KEV listing (95% CI 1.90–2.08; $p<0.001$). Publication-to-KEV catalog-addition lag shortened markedly across publication eras, while normalized CWE trends showed persistent web weaknesses and recent growth in authorization-related categories. These findings support vulnerability prioritization that combines technical severity with exploitation evidence, attack conditions, asset exposure, and data-quality context.

1. Introduction

Software security increasingly depends on the ability to interpret an expanding stream of publicly disclosed vulnerabilities. The Common Vulnerabilities and Exposures (CVE) program supplies standardized identifiers, while the National Vulnerability Database (NVD) and vulnerability-scoring standards add technical metadata that support triage, remediation, and research [1-4]. As software ecosystems have expanded across cloud services, open-source dependencies, embedded systems, mobile platforms, and enterprise infrastructure, vulnerability-management teams face a scale problem: they must decide which disclosed flaws deserve the fastest attention.

A central difficulty is that three different concepts are often treated as interchangeable. CVE publication volume measures the output of the disclosure ecosystem; CVSS describes technical severity and exploit conditions; and observed exploitation indicates that attackers have actually used a vulnerability in the wild. NIST explicitly cautions that CVSS is a severity measure rather than a complete measure of organizational risk [2]. Earlier empirical research likewise found that high CVSS scores alone are insufficient for efficient remediation prioritization and that exploitation evidence provides additional value [11-18].

This study examines how the software-vulnerability landscape has changed over time using record-level CVE data. It quantifies changes in CVE publication volume, CVSS severity, and attack-vector characteristics; evaluates the association between severity and CISA KEV status; measures the interval between CVE publication and KEV addition; and identifies the most frequently reported CWE weakness classes. By considering these measures together, the analysis distinguishes vulnerability disclosure volume from technical severity, exploitability conditions, and evidence of exploitation in the wild. The study contributes a single reproducible record-level framework that jointly evaluates disclosure growth, CVSS severity and version effects, attack conditions, metadata completeness, KEV listing, publication-to-catalog lag, and

normalized CWE composition. This combination allows observed temporal changes to be separated more carefully from changes in scoring-system adoption and database enrichment.

2. Background and Related Literature

2.1 CVE, NVD, and vulnerability measurement

The CVE system provides stable identifiers for disclosed vulnerabilities and is now supported by a distributed network of CVE Numbering Authorities. The NVD enriches CVE records with standardized information, including CVSS data and product/configuration metadata [1-3]. These databases have enabled longitudinal security research for two decades, including early large-scale vulnerability-lifecycle studies and empirical examinations of disclosure, patching, zero-day exploitation, and vulnerability ageing [17-20].

2.2 CVSS: severity rather than complete risk

CVSS provides a standardized description of vulnerability severity through numerical scores and vector components. CVSS v3.x and v4.0 classify scores from 9.0-10.0 as Critical, 7.0-8.9 as High, 4.0-6.9 as Medium, and 0.1-3.9 as Low [3,4]. The vector describes exploitation conditions such as Attack Vector, Attack Complexity, Privileges Required, and User Interaction. CVSS v4.0 also introduced additional distinctions, including Active and Passive forms of user interaction [3]. Recent user studies have shown that CVSS scoring can be inconsistent across assessors, reinforcing the need to interpret scores carefully [25,26].

2.3 Exploitation evidence and prioritization

The CISA KEV catalog provides a different signal: inclusion requires evidence that a vulnerability has been exploited in the wild and is intended to support remediation prioritization [5,6]. This distinction aligns with a substantial research literature showing that only a subset of disclosed vulnerabilities is exploited and that exploit evidence can improve prioritization beyond static severity thresholds [11-16]. EPSS and SSVC are prominent examples of efforts to supplement severity with threat or stakeholder-specific context

[12,14,15]. The present study does not train predictive models; instead, KEV is treated as an observed exploitation label for descriptive and inferential comparison.

2.4 Data quality and current operational constraints

Public vulnerability repositories change as records are added, revised, and enriched. Prior research has documented missing or inconsistent NVD information and difficulties in interpreting some CVSS ratings [24]. NIST also revised its NVD enrichment operations in 2026 in response to record growth and processing constraints [10]. For this reason, missing CVSS and CWE fields are reported as part of the dataset rather than filled through imputation.

3. Data and Methods

3.1 Dataset

The analysis uses three CSV shards derived from the public CVE corpus and frozen on 1 September 2026: `cve_2022_and_before.csv`, `cve_2023_to_2025.csv`, and `cve_2026_to_now.csv`. After concatenation, the dataset contained 366,873 rows and exactly 366,873 unique CVE identifiers, so no duplicate CVE IDs required removal. Publication dates ranged from 29 September 1999 to 1 September 2026. The columns used in this study included CVE ID, publication/update dates, CVSS version, base score, vector, CWE identifiers, CISA KEV indicator and date, SSVC-related fields, vendor, product, and CPE fields.

Table 1. Record availability in the 1 September 2026 CVE snapshot.

Dataset characteristic	Records	Share of all CVEs
All published CVEs	366,873	100.00%
CVSS score available	195,907	53.40%
CWE mapping available	182,710	49.80%
CISA KEV	1,687	0.46%

3.2 Analytical period and variables

All 366,873 records were used for overall descriptive statistics. Temporal hypothesis tests were restricted to complete calendar years 2016–2025 to avoid treating the partial 2026 year as comparable with full years. CVSS severity bands were assigned using version-specific thresholds: CVSS v2 scores of 7.0–10.0 were High, whereas v3.x/v4.0 scores of 9.0–10.0 were Critical and 7.0–8.9 High. Scores of exactly 0.0 were retained as the explicit CVSS qualitative category None; this accounts for 108 scored records and resolves the fifth-category degrees of freedom in the severity tests. Attack Vector, Attack Complexity, Privileges Required, and User Interaction were parsed directly from CVSS vector strings. CWE fields were split into individual identifiers so a CVE with multiple mappings could contribute to multiple weakness categories.

3.3 Statistical analysis

Categorical distributions were summarized as counts and percentages. Pearson chi-square tests assessed whether severity and attack-vector distributions varied across 2016–2025, with Cramér's V used to quantify effect size. CVSS scores for KEV-listed and non-KEV-listed records were compared using a two-sided Mann–Whitney U test. A severity-by-KEV chi-square test assessed association between qualitative severity and KEV listing. To evaluate measurement-system sensitivity, the severity trend analysis was repeated using CVSS v3.x records only. Year-specific CVSS, CWE, and vector completeness was calculated using all published CVEs in each year as the denominator. A multivariable logistic regression modeled KEV listing using CVSS base score, Attack Vector, Attack Complexity, Privileges Required, harmonized User Interaction, and publication year; adjusted odds ratios (aORs) and 95% confidence intervals are reported. Publication-to-KEV lag was calculated as catalog

addition date minus CVE publication date and stratified by publication era. Negative lags were retained as timestamp/catalog chronology artifacts and are reported transparently. Finally, longitudinal CWE patterns were normalized by the number of CWE-mapped CVEs in each publication year, avoiding interpretation of raw growth as compositional change.

3.4 Reproducibility

The analysis was implemented in Python using pandas, NumPy, SciPy, statsmodels, and Matplotlib. No machine-learning model was trained. The complete processing workflow used the three frozen CSV shards, deterministic severity classification, CVSS-vector parsing, inferential tests, adjusted logistic regression, lag stratification, and normalized CWE aggregation. The exact

analysis script is supplied as Supplementary File S1 (CVE_analysis_reproducibility.py), with derived result tables supplied in Supplementary Data S2; file names and transformation rules are documented in the script so the reported analyses can be regenerated from the same snapshot.

4. Results

4.1 Growth in published CVEs

Within the principal longitudinal window, annual publication volume increased from 6,457 CVEs in 2016 to 48,153 in 2025, reaching 7.46 times its 2016 level (a 645.7% increase). As shown in Figure 1, growth accelerated after 2022: the dataset contains 25,000 CVEs for 2022, 28,847 for 2023, 39,925 for 2024, and 48,153 for 2025. Because 2026 is incomplete at the 1 September snapshot, it is not used as a full-year comparator.

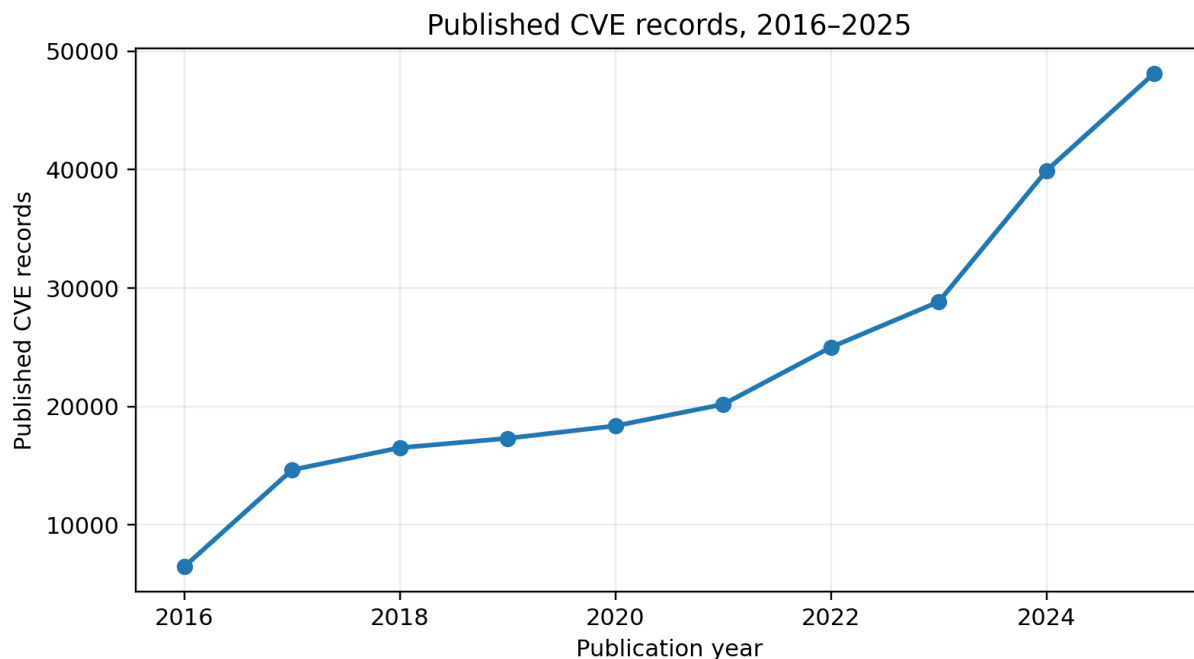


Figure 1. Published CVE records by year, 2016–2025, calculated from the merged dataset.

Table 2. Annual CVE counts used in the longitudinal analysis.

Year	Published CVEs
2016	6,457
2017	14,642
2018	16,510
2019	17,308
2020	18,363

2021	20,177
2022	25,000
2023	28,847
2024	39,925
2025	48,153

4.2 CVSS severity distribution

Table 3. Overall CVSS severity distribution (n = 195,907 scored records; None denotes CVSS base score 0.0).

Severity	Records	Share of scored CVEs
Critical	20,592	10.51%
High	74,854	38.21%
Medium	91,781	46.85%
Low	8,572	4.38%
None (score 0.0)	108	0.06%

Medium severity was the largest category (46.85%), followed by High (38.21%), Critical (10.51%), Low (4.38%), and None (0.06%). Across 2016–2025, the five-category severity distribution changed significantly ($\chi^2(36)=1296.4$, $p<0.001$), but Cramér's $V=0.048$ indicates a small association between year and severity. Figure 2

visualizes this distribution over time and shows the conspicuous early fluctuation in Critical severity. Because scoring coverage and CVSS-version composition changed strongly over the same period, this pattern is interpreted alongside the sensitivity and completeness analyses below rather than as a direct change in underlying software risk.

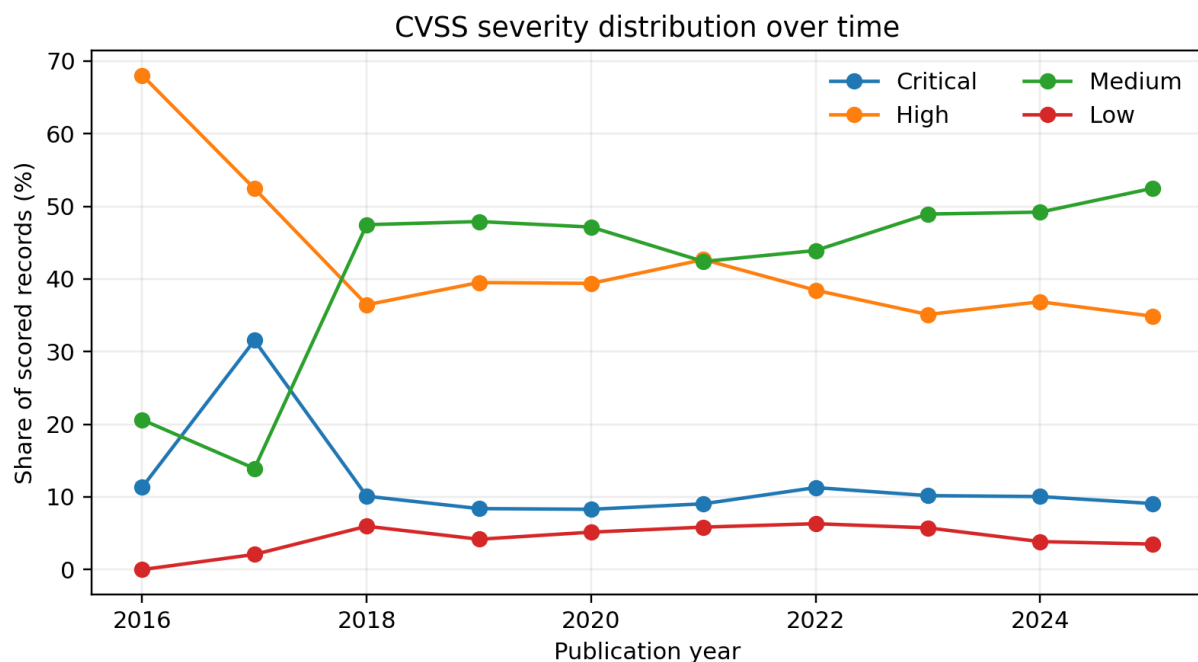


Figure 2. CVSS severity distribution among scored records, 2016–2025.

4.2.1 CVSS-version sensitivity analysis

Restricting the longitudinal analysis to 124,082 CVSS v3.x-scored records produced a very similar small year-severity association ($\chi^2(36)=1046.0$, $p<0.001$; Cramér's $V=0.046$). As Figure 3 shows, the broad post-2018 pattern remains after excluding CVSS v2 and v4.0, although 2016–2017 estimates are based on sparse scored subsets and

should be interpreted cautiously. In the v3.x-only series, Medium severity increased from 43.91% in 2022 to 49.08% in 2025, while High remained between 35% and 38% during 2023–2025. This sensitivity analysis indicates that version mixing does not explain the entire temporal pattern, but scoring coverage remains an important source of selection.

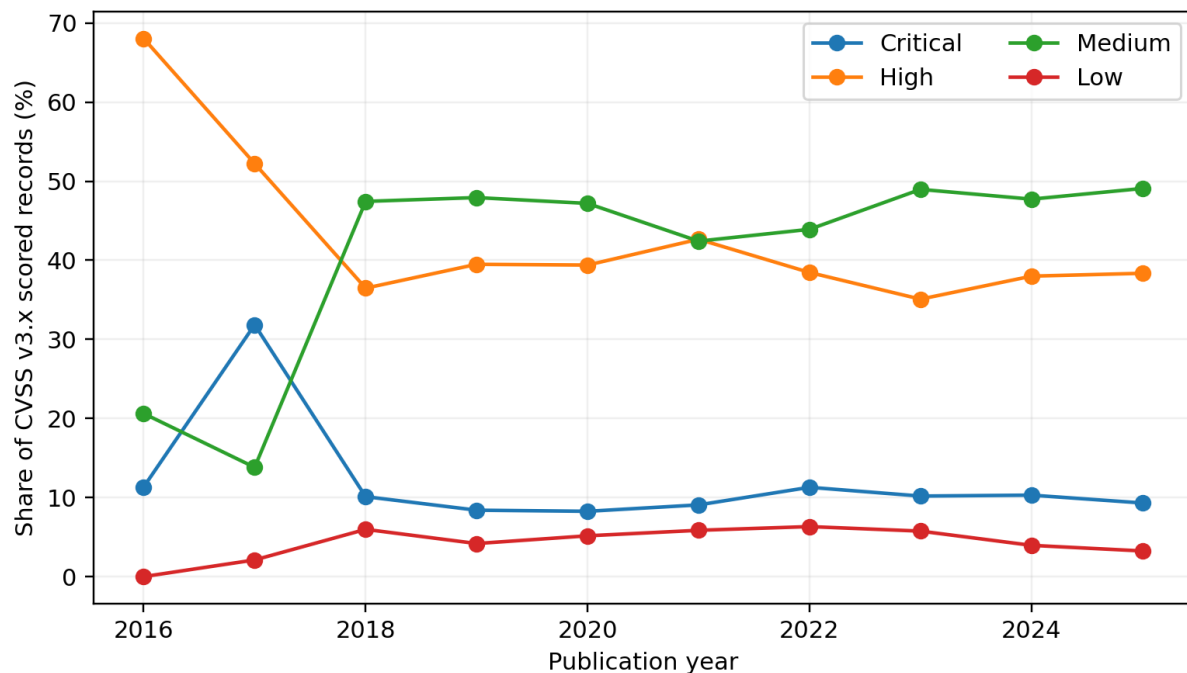


Figure 3. CVSS v3.x-only severity distribution, 2016–2025 (sensitivity analysis).

4.2.2 Year-by-year metadata completeness

Metadata completeness changed substantially across the longitudinal window. CVSS score/vector coverage rose from 1.50% of published CVEs in 2016 to 59.06% in 2022, 80.26% in 2023, 94.46% in 2024, and 93.86% in

2025. CWE coverage similarly increased from 1.01% in 2016 to 88.29% in 2025. Figure 4 makes this enrichment gradient explicit. Accordingly, early-year severity and vector percentages describe a small enriched subset of published CVEs rather than the full annual corpus.

Table 4. Year-by-year metadata completeness, 2016–2025.

Year	Published CVEs	CVSS available	CWE available	Vector available
2016	6,457	1.50%	1.01%	1.50%
2017	14,642	2.94%	6.38%	2.94%
2018	16,510	10.55%	15.36%	10.55%
2019	17,308	15.21%	13.78%	15.21%
2020	18,363	33.07%	20.37%	33.07%
2021	20,177	41.53%	35.71%	41.53%

2022	25,000	59.06%	56.47%	59.06%
2023	28,847	80.26%	75.45%	80.26%
2024	39,925	94.46%	86.56%	94.46%
2025	48,153	93.86%	88.29%	93.86%

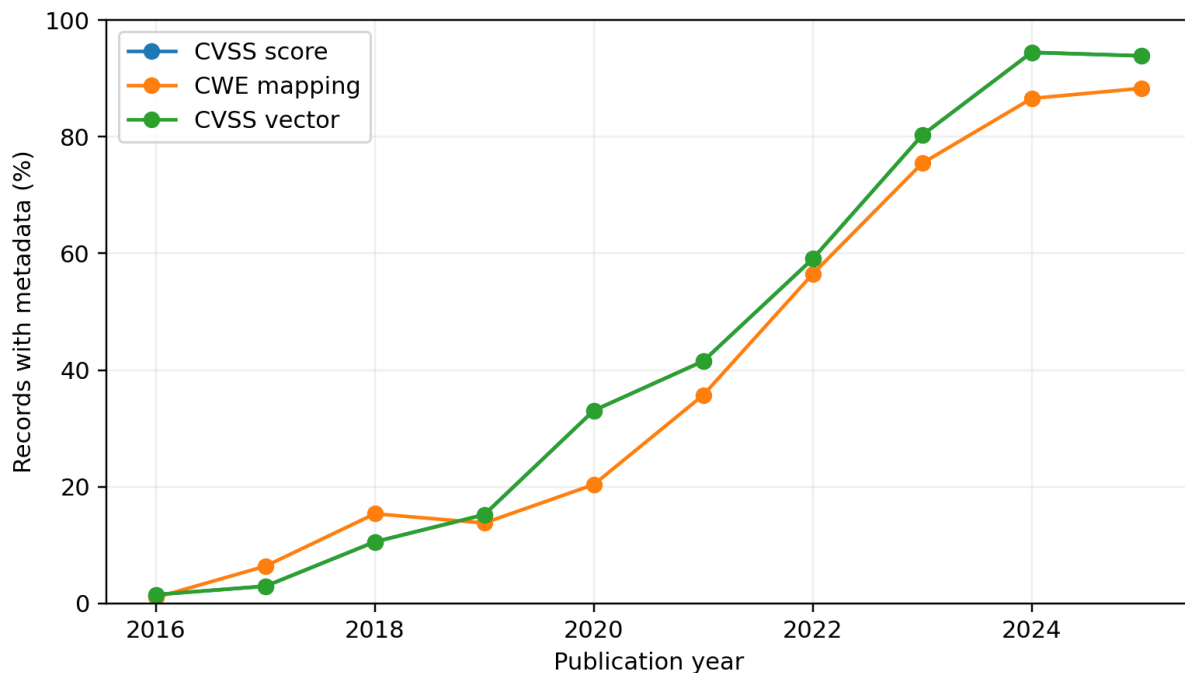


Figure 4. CVSS, CWE, and vector metadata completeness by publication year, 2016–2025.

4.3 Attack vector and exploitability conditions

Table 5. Attack Vector distribution among CVSS-vectorized records.

Attack vector	Records	Share of vectorized CVEs
Network	145,767	74.41%
Local	40,806	20.83%
Adjacent	7,569	3.86%
Physical	1,765	0.90%

As summarized in Table 5 and visualized in Figure 5, network-reachable vulnerabilities dominated the dataset: 145,767 records (74.41%) had AV:N, compared with 40,806 Local (20.83%), 7,569 Adjacent (3.86%), and 1,765 Physical (0.90%). The distribution changed significantly over 2016–2025 ($\chi^2(27)=1656.4$, $p<0.001$), again with a small effect size (Cramér's $V=0.063$). After falling to around 64% in 2020–2021, the Network share

rose to 76.66% in 2025. Across all vectorized records, 88.15% had Low Attack Complexity, 52.05% required no privileges, and 70.06% required no user interaction under the applicable CVSS vocabulary. Because CVSS v4.0 replaced the single v3.x 'Required' interaction state with Active/Passive distinctions, User Interaction is interpreted descriptively rather than as a directly stable long-term category [3].

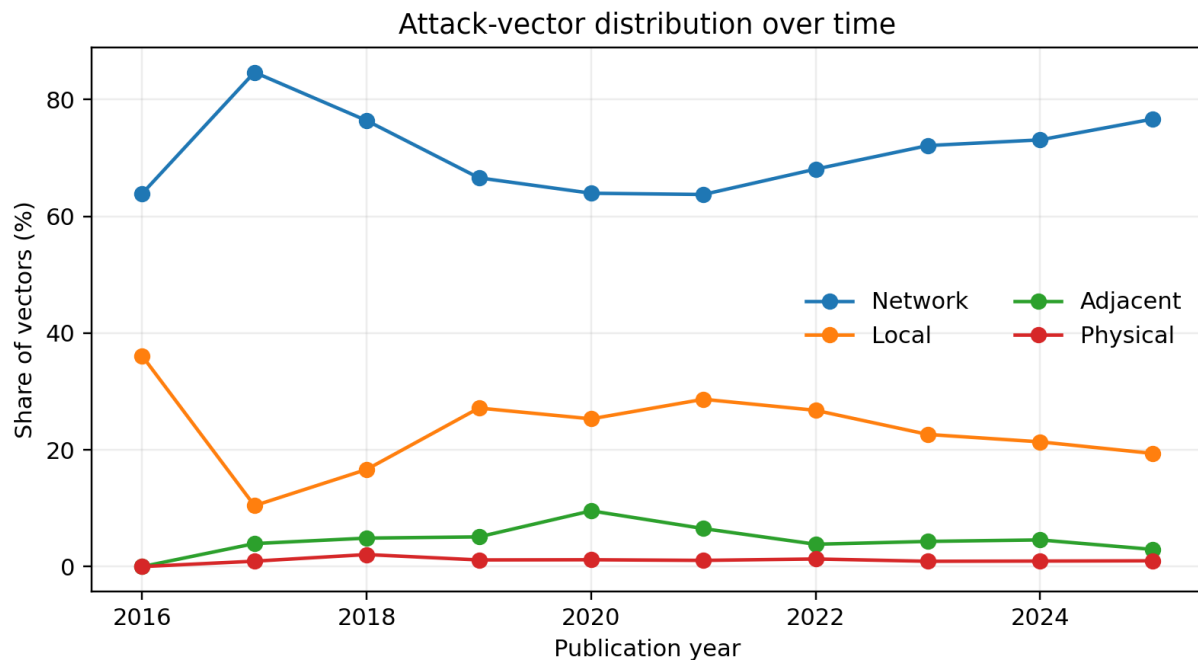


Figure 5. Attack Vector distribution among CVSS-vectorized records, 2016–2025.

4.4 Known exploitation and CVSS severity

Only 1,687 of the 366,873 CVEs (0.46%) were KEV-listed in the snapshot. KEV listing is treated as positive evidence of known exploitation, whereas absence from KEV is not interpreted as evidence of non-exploitation. KEV-listed vulnerabilities were markedly more severe: their median CVSS score was 8.8 (mean 8.36),

compared with a median of 6.9 (mean 6.79) among non-KEV-listed records. The Mann-Whitney test was highly significant ($U=250400932$, $p<0.001$), with a rank-biserial effect magnitude of approximately 0.528. Figure 6 shows the corresponding unadjusted KEV prevalence gradient across the four non-zero severity bands.

Table 6. CISA KEV prevalence by CVSS severity, including the None (0.0) category.

CVSS severity	KEV records	Scored records	KEV prevalence
Critical	585	20,592	2.841%
High	891	74,854	1.190%
Medium	204	91,781	0.222%
Low	7	8,572	0.082%
None (score 0.0)	0	108	0.000%

KEV listing was associated with the five-category severity variable ($\chi^2(4)=1541.2$, $p<0.001$). Critical vulnerabilities had the highest KEV prevalence (2.84%), followed by High (1.19%), Medium (0.22%), Low (0.08%), and None (0%). However,

211 KEV-listed entries were Medium or Low, showing that a high-severity cutoff alone can omit vulnerabilities with documented exploitation evidence.

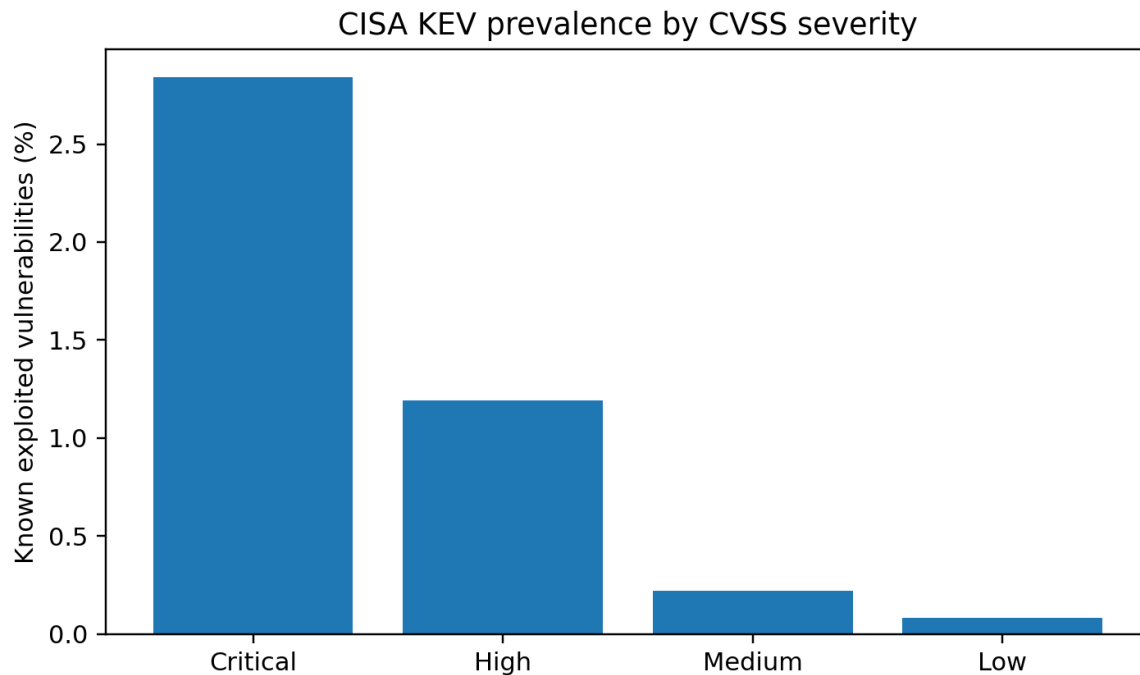


Figure 6. Proportion of scored CVEs in CISA KEV by qualitative CVSS severity.

4.4.1 Adjusted correlates of KEV listing

The multivariable logistic regression included 195,784 records with complete CVSS score and vector components, of which 1,686 were KEV-listed. After adjustment for attack conditions and publication year, each one-point increase in CVSS base score was associated with 1.99-fold higher odds of KEV listing (aOR 1.99, 95% CI 1.90–2.08, $p < 0.001$). Relative to local attack vector, network vector had lower adjusted odds (aOR 0.58, 95% CI 0.51–0.66), illustrating that base

score, vector characteristics, and publication timing are not interchangeable signals. No-user-interaction vulnerabilities also had lower adjusted odds than harmonized interaction-required vulnerabilities (aOR 0.81, 95% CI 0.72–0.92). The negative publication-year coefficient (aOR 0.73 per year, 95% CI 0.72–0.74) should not be interpreted causally: older vulnerabilities have had more time to enter KEV, and catalog coverage is time-dependent.

Table 7. Multivariable logistic regression of KEV listing (n = 195,784).

Predictor	aOR	95% CI	p
CVSS base score (per 1 point)	1.99	1.90–2.08	<0.001
Attack Vector: Network vs Local	0.58	0.51–0.66	<0.001
Attack Vector: Adjacent vs Local	0.27	0.19–0.40	<0.001
Attack Vector: Physical vs Local	0.87	0.41–1.86	0.729
Attack Complexity: Low vs High	0.87	0.72–1.04	0.114

Privileges: Low vs High	0.91	0.72-1.16	0.447
Privileges: None vs High	1.18	0.93-1.49	0.164
User Interaction: None vs Yes	0.81	0.72-0.92	0.001
Publication year (per year)	0.73	0.72-0.74	<0.001

4.5 Publication-to-KEV lag

All 1,687 KEV records had an interpretable catalog-addition date. The raw publication-to-KEV addition lag had a median of 265 days (IQR 9.5-1397.5); 59 records had negative lags, consistent with timestamp revisions or catalog chronology. Figure 7 displays the non-negative lag distribution for continuity with the primary descriptive analysis. Because the pooled distribution is strongly confounded by vulnerability age and the

later introduction of the KEV catalog, the analysis was additionally stratified by CVE publication era. Figure 8 shows a steep decline in median catalog-addition lag: 5,549.5 days for pre-2010 CVEs, 3,227 days for 2010-2015, 1,199 days for 2016-2020, 152 days for 2021-2022, 7 days for 2023-2024, and 6 days for 2025-2026. These values describe catalog addition, not time to first exploitation.

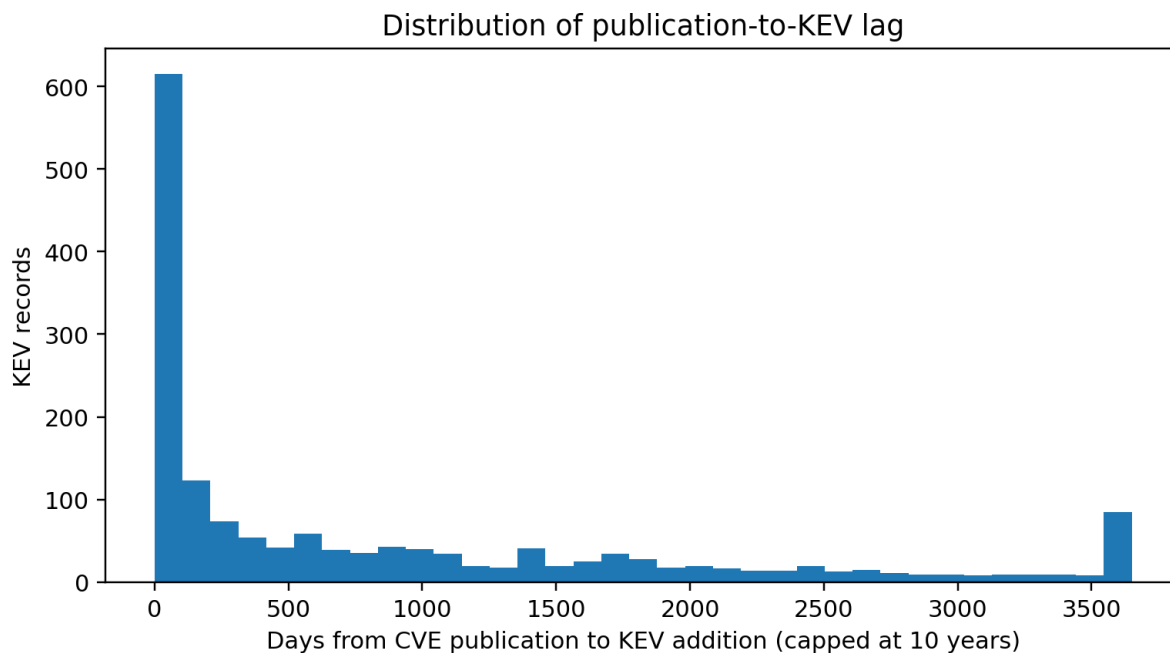


Figure 7. Distribution of non-negative CVE publication-to-KEV addition lag; values above 10 years are capped for display only.

Table 8. Publication-to-KEV catalog-addition lag by CVE publication era.

Publication era	KEV records	Median lag, days	IQR, days
<2010	28	5,549.5	4,746.8-6,214.3
2010-2015	168	3,227.0	2,683.8-3,824.0
2016-2020	493	1,199.0	700.0-1,730.0
2021-2022	345	152.0	33.0-307.0

2023-2024	325	7.0	0.0-105.0
2025-2026	328	6.0	0.0-34.3

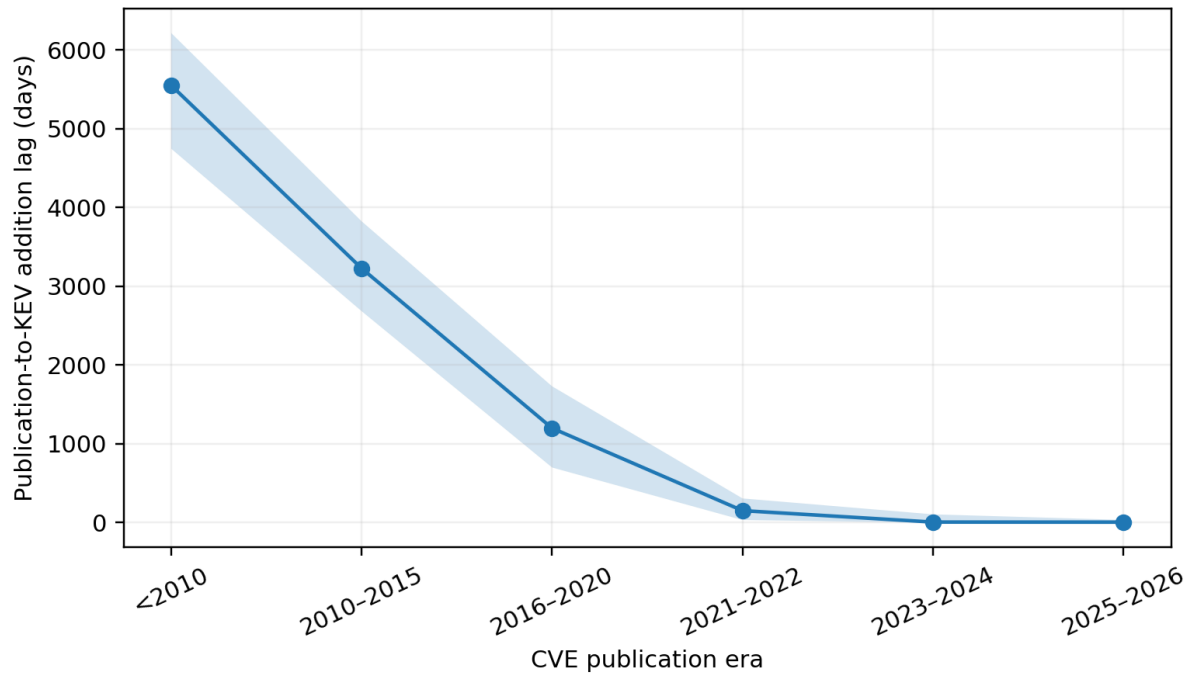


Figure 8. Median and interquartile publication-to-KEV catalog-addition lag by publication era.

4.6 CWE weakness patterns

Table 9. Twenty most frequent CWE mappings in the complete dataset. Counts are mapping occurrences; a CVE can contribute to more than one CWE category.

CWE	Weakness class	Mapped occurrences
CWE-79	Cross-site scripting	27,673
CWE-89	SQL injection	12,457
CWE-862	Missing authorization	7,865
CWE-284	Improper access control	6,377
CWE-352	Cross-site request forgery	5,947
CWE-22	Path traversal	5,497
CWE-20	Improper input validation	5,203
CWE-787	Out-of-bounds write	4,917
CWE-125	Out-of-bounds read	4,762
CWE-416	Use after free	4,662
CWE-200	Sensitive information exposure	4,553
CWE-74	Injection	4,533
CWE-78	OS command injection	4,314
CWE-94	Code injection	4,188
CWE-121	Stack-based buffer overflow	3,689
CWE-119	Memory-buffer bounds	3,151

CWE-120	Classic buffer overflow	3,071
CWE-434	Dangerous file upload	3,002
CWE-918	Server-side request forgery	2,771
CWE-400	Uncontrolled resource consumption	2,769

CWE-79 (cross-site scripting) was the most frequent mapped weakness with 27,673 occurrences, followed by CWE-89 (SQL injection; 12,457), CWE-862 (missing authorization; 7,865), CWE-284 (improper access control; 6,377), and CWE-352 (cross-site request forgery; 5,947). Figure 9 displays the 15 most frequent mappings in the complete snapshot. Because raw counts increase mechanically as the CVE corpus grows, a second analysis normalized each CWE by the

number of CWE-mapped CVEs in its publication year. Figure 10 shows that CWE-79 remained prominent, increasing from 14.35% of CWE-mapped 2022 CVEs to 18.55% in 2025; CWE-89 rose from 5.02% to 9.33%, and CWE-862 from 1.71% to 5.47%. The normalized view therefore supports a recent compositional rise in several web and authorization-related weaknesses rather than merely reflecting growth in total CVE volume.

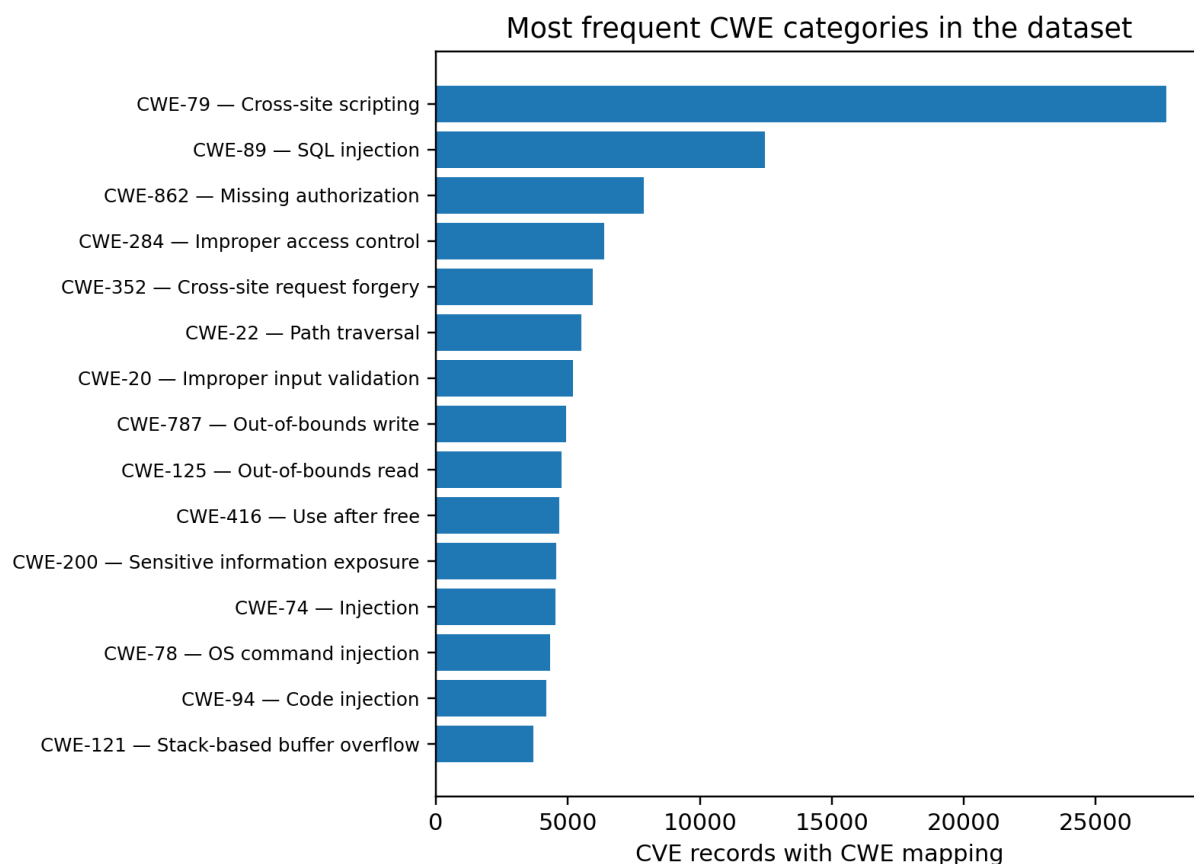


Figure 9. Fifteen most frequent CWE mappings in the 366,873-record dataset.

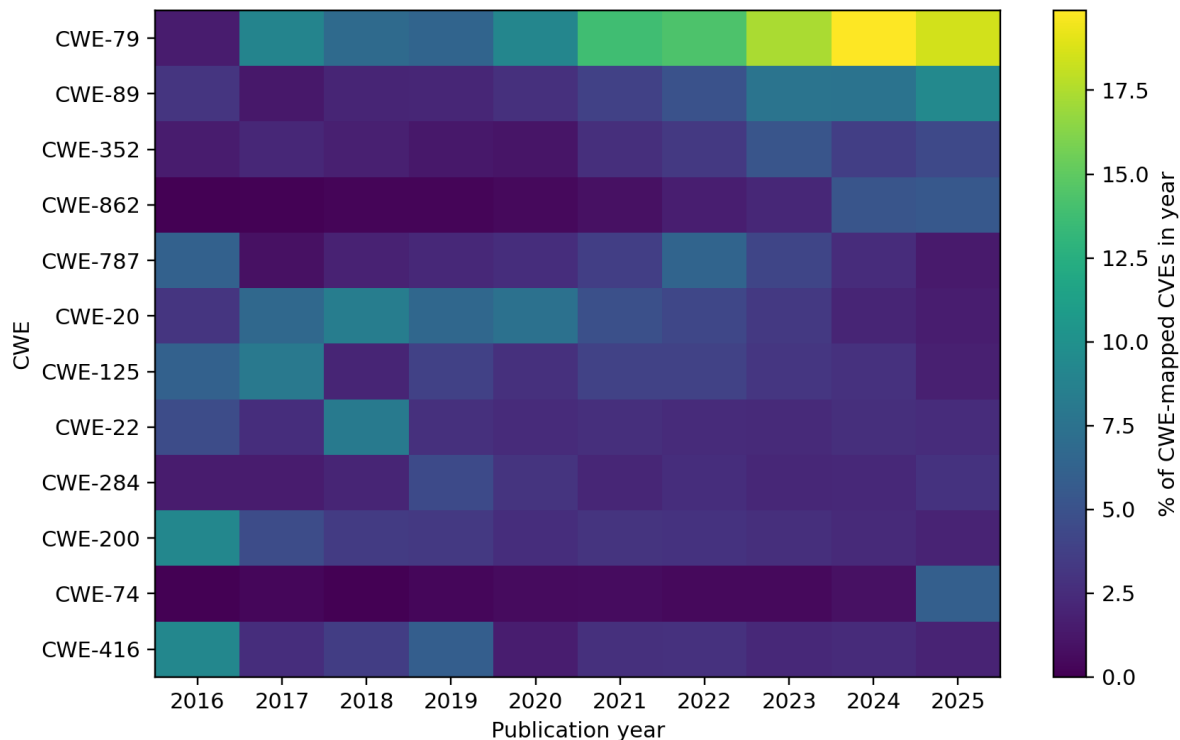


Figure 10. Normalized annual prevalence of the 12 most frequent CWE mappings, 2016–2025 (% of CWE-mapped CVEs per year).

5. Discussion

5.1 Vulnerability disclosure is expanding faster than simple risk narratives suggest

The increase from 6,457 CVEs in 2016 to 48,153 in 2025 is operationally important, but Figure 1 should not be interpreted as a direct measure of worsening software security. Earlier lifecycle research emphasized that disclosure counts depend on discovery, assignment, vendor participation, and publication processes [17–20]. The present results extend that argument by showing that metadata enrichment changed just as dramatically: Figure 4 documents CVSS coverage rising from 1.50% to 93.86% across the same decade. NIST's 2026 move toward risk-based NVD enrichment further reinforces the need to distinguish disclosure volume from metadata availability [10].

5.2 Severity is associated with exploitation, but does not replace exploitation evidence

The KEV analyses reinforce prior evidence that CVSS severity and exploitation evidence are related but non-equivalent. Allodi and Massacci [11] showed that severity alone is an inefficient remediation proxy, while EPSS explicitly models exploitation probability [12] and SSVC adds stakeholder context [14,15]. In the present dataset, KEV-listed vulnerabilities had higher CVSS scores, but 211 KEV-listed vulnerabilities were only Medium or Low. The adjusted model strengthens this result: CVSS score remained independently associated with KEV listing (aOR 1.99 per point), yet vector characteristics and publication year also contributed. This supports using CVSS as one input rather than treating it as a complete risk score, consistent with NIST/FIRST guidance and critiques of direct CVSS-based prioritization [2,3,13].

5.3 Attack conditions matter at scale

Nearly three quarters of vectorized vulnerabilities were network-reachable and almost nine in ten had low attack complexity. Figure 5 shows that these proportions also vary over time. Operationally, attack conditions matter because vulnerabilities with identical base scores can differ in reachability, required privileges, and interaction. The adjusted KEV model further shows that these components should not be interpreted in isolation: conditional associations can differ from simple prevalence patterns because score, vector components, publication year, and catalog exposure are correlated.

5.4 CWE results point to persistent classes of preventable weakness

The raw CWE ranking in Figure 9 confirms persistent concentrations in cross-site scripting, SQL injection, authorization failures, traversal, input validation, and memory-safety weaknesses. The normalized heatmap in Figure 10 adds an important longitudinal qualification: some categories, particularly CWE-862 and CWE-89, increased as a share of mapped CVEs in recent years. This strengthens the case for class-level prevention through safer frameworks, access-control design, memory-safe components, code review, and automated testing rather than relying only on post-disclosure patching [7-9].

5.5 Implications for vulnerability-management policy

The findings support a triage process that begins with the organization's own exposure: whether the affected software is present, reachable, and important to operations. Evidence of active exploitation, including KEV status, should then be considered alongside CVSS vector characteristics such as network reachability, required privileges, attack complexity, and user interaction. The CVSS base score remains useful as a standardized severity measure, but the final remediation order should also reflect asset criticality, compensating controls, patch availability, and potential business impact. This approach is consistent with the context-sensitive

reasoning used in SSVC and related prioritization frameworks [14,15].

6. Limitations

This study uses disclosed and published CVEs, not the unobservable universe of all software vulnerabilities. Counts are affected by reporting practices, CNA participation, vendor disclosure policies, and retrospective updates. Metadata missingness is substantial and strongly time-dependent: Figure 4 shows that early-year CVSS/CWE coverage is especially sparse. The sensitivity analysis reduces CVSS-version heterogeneity but cannot eliminate selection into scoring. NVD operational changes and enrichment backlogs can also alter metadata availability [10,24].

CVSS versions are not perfectly equivalent. Version-specific severity thresholds and compatible Attack Vector categories were used, and the v3.x-only sensitivity analysis produced similar small effect sizes, but metric semantics still evolve across versions. CISA KEV is authoritative evidence for vulnerabilities known to have been exploited, not a complete census of every exploited CVE; therefore non-KEV-listed records are not labeled 'unexploited.' The logistic regression is associational and is affected by differential time at risk for catalog inclusion. Finally, KEV addition date is a catalog timestamp rather than the true first-exploitation date, so Figures 7 and 9 measure publication-to-catalog addition rather than publication-to-attacker use.

7. Conclusion

Analysis of 366,873 CVE records documents rapid expansion of the public vulnerability corpus together with persistent variation in severity, attack conditions, weakness classes, metadata completeness, and exploitation evidence. Annual publication reached 7.46 times its 2016 level by 2025. The previously hidden fifth severity category was identified as 108 CVSS 0.0 records (None), resolving the denominator and degrees-of-freedom inconsistency. Severity trends remained statistically detectable but small in both the mixed-version and CVSS v3.x-only analyses, while sharply

improving metadata coverage cautions against treating early-year percentages as directly comparable to recent years.

CVSS severity and KEV listing were related but captured different information. Higher CVSS score remained associated with KEV listing after adjustment, yet vector characteristics and publication timing also mattered, and documented exploitation evidence occurred among Medium and Low vulnerabilities. Publication-to-KEV addition lag shortened strongly for recent publication eras, while normalized CWE trends identified persistent web weaknesses and rising authorization-related categories. In practice, prioritization should combine severity, attack conditions, KEV/other exploitation evidence, asset exposure, patchability, and organizational context rather than rely on a single score.

REFERENCES

- CVE Program. Common Vulnerabilities and Exposures (CVE): CVE Records and Program documentation. The MITRE Corporation/CVE Program, accessed 2 September 2026.
- National Institute of Standards and Technology. National Vulnerability Database: Vulnerability Metrics—Common Vulnerability Scoring System. NIST, accessed 2 September 2026.
- FIRST. Common Vulnerability Scoring System Version 4.0: Specification Document, Version 1.2. Forum of Incident Response and Security Teams, 2024.
- FIRST. Common Vulnerability Scoring System Version 3.1: Specification Document. Forum of Incident Response and Security Teams, 2019.
- Cybersecurity and Infrastructure Security Agency. Known Exploited Vulnerabilities Catalog. CISA, accessed 2 September 2026.
- Cybersecurity and Infrastructure Security Agency. Binding Operational Directive 22-01: Reducing the Significant Risk of Known Exploited Vulnerabilities. 2021.
- MITRE. Common Weakness Enumeration (CWE). The MITRE Corporation, accessed 2 September 2026.
- Souppaya M, Scarfone K, Dodson D. Secure Software Development Framework (SSDF) Version 1.1: Recommendations for Mitigating the Risk of Software Vulnerabilities. NIST SP 800-218. 2022. <https://doi.org/10.6028/NIST.SP.800-218>.
- Souppaya M, Scarfone K. Guide to Enterprise Patch Management Planning: Preventive Maintenance for an Organization's Technology. NIST SP 800-40 Rev. 4. 2022. <https://doi.org/10.6028/NIST.SP.800-40r4>.
- National Institute of Standards and Technology. NIST Updates NVD Operations to Address Record CVE Growth. 15 April 2026.
- Allodi L, Massacci F. Comparing Vulnerability Severity and Exploits Using Case-Control Studies. *ACM Transactions on Information and System Security*. 2014;17(1):Article 1, 1–20. <https://doi.org/10.1145/2630069>.
- Jacobs J, Romanosky S, Edwards B, Adjerid I, Roytman M. Exploit Prediction Scoring System (EPSS). *Digital Threats: Research and Practice*. 2021;2(3):Article 20, 1–17. <https://doi.org/10.1145/3436242>.
- Spring JM, Hatleback E, Householder AD, Manion A, Shick D. Time to Change the CVSS? *IEEE Security & Privacy*. 2021;19(2):74–78. <https://doi.org/10.1109/MSEC.2020.3044475>.
- Spring JM, Hatleback E, Householder AD, Manion A, Shick D. Prioritizing Vulnerability Response: A Stakeholder-Specific Vulnerability Categorization. Workshop on the Economics of Information Security (WEIS), 2020.

- Spring JM, Householder AD, Hatleback E, Manion A, Oliver M, Sarvepalli VS, Tyzenhaus L, Yarbrough CG. Prioritizing Vulnerability Response: A Stakeholder-Specific Vulnerability Categorization (Version 2.0). Software Engineering Institute, Carnegie Mellon University, 2021.
- Bullough BL, Yanchenko AK, Smith CL, Zipkin JR. Predicting Exploitation of Disclosed Software Vulnerabilities Using Open-source Data. Proceedings of the 3rd ACM International Workshop on Security and Privacy Analytics. 2017:45-53. <https://doi.org/10.1145/3041008.3041009>.
- Frei S, May M, Fiedler U, Plattner B. Large-scale Vulnerability Analysis. Proceedings of the 2006 SIGCOMM Workshop on Large-Scale Attack Defense. 2006:131-138. <https://doi.org/10.1145/1162666.1162671>.
- Bozorgi M, Saul LK, Savage S, Voelker GM. Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits. Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2010:105-114. <https://doi.org/10.1145/1835804.1835821>.
- Bilge L, Dumitras T. Before We Knew It: An Empirical Study of Zero-Day Attacks in the Real World. Proceedings of the 2012 ACM Conference on Computer and Communications Security. 2012:833-844. <https://doi.org/10.1145/2382196.2382284>.
- Ozment A, Schechter SE. Milk or Wine: Does Software Security Improve with Age? Proceedings of the 15th USENIX Security Symposium. 2006:93-104.
- Sabottke C, Suciu O, Dumitras T. Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting Real-World Exploits. Proceedings of the 24th USENIX Security Symposium. 2015:1041-1056.
- Li F, Paxson V. A Large-Scale Empirical Study of Security Patches. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017:2201-2215. <https://doi.org/10.1145/3133956.3134072>.
- Zhang S, Caragea D, Ou X. An Empirical Study on Using the National Vulnerability Database to Predict Software Vulnerabilities. In: Database and Expert Systems Applications (DEXA 2011), Lecture Notes in Computer Science, vol. 6860. Springer; 2011:217-231. https://doi.org/10.1007/978-3-642-23088-2_15.
- Wunder J, Corona A, Hammer A, Benenson Z. On NVD Users' Attitudes, Experiences, Hopes, and Hurdles. Digital Threats: Research and Practice. 2024;5(3):Article 33, 1-19. <https://doi.org/10.1145/3688806>.
- Wunder J, Kurtz A, Eichenmüller C, Gassmann F, Benenson Z. Shedding Light on CVSS Scoring Inconsistencies: A User-Centric Study on Evaluating Widespread Security Vulnerabilities. 2024 IEEE Symposium on Security and Privacy. 2024:1102-1121. <https://doi.org/10.1109/SP54263.2024.00058>.
- Zhang S, Cai M, Zhao L, de Carné de Carnavalet X, Massacci F, Zhang M. Tool-Assisted CVSS Vulnerability Scoring: A Controlled Quantitative Study of Human Assessment. Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26). ACM; 2026:Article 1219, 1-24. <https://doi.org/10.1145/3772318.3790409>.

Wunder J, Aurich J, Benenson Z. From Chaos to Consistency: The Role of CSAF in Streamlining Security Advisories. Proceedings of the 2024 European Symposium on Usable Security (EuroUSEC 2024). 2024:187–199. <https://doi.org/10.1145/3688459.3688463>.

Xu J, Zhang H, Zhou X, Yang Y, Hu J, Li X, Yang L, Liu B. Automated Localization of Affected Libraries and Versions from Vulnerability Reports. IEEE Transactions on Software Engineering. 2026;52(7):2127–2147. <https://doi.org/10.1109/TSE.2026.3688559>.

Spring JM, Hatleback E, Householder AD, Manion A, Shick D. Towards Improving CVSS. Software Engineering Institute, Carnegie Mellon University, 2018.

National Institute of Standards and Technology. National Vulnerability Database: CVE/CPE APIs and Data Feeds. NIST, accessed 2 September 2026.

CVE Program. CVE Record Format and CVE Numbering Authority documentation. CVE.org, accessed 2 September 2026.

FIRST. Common Vulnerability Scoring System Version 4.0: User Guide, Version 1.2. Forum of Incident Response and Security Teams, 2024.

Cybersecurity and Infrastructure Security Agency. Stakeholder-Specific Vulnerability Categorization (SSVC) guidance. CISA, accessed 2 September 2026.

Cybersecurity and Infrastructure Security Agency. Known Exploited Vulnerabilities Catalog: usage guidance and catalog criteria. CISA, accessed 2 September 2026.

flying-coyote. nvd-csv: CVEProject/cvelistV5 JSON converted to CSV. GitHub repository; daily refreshed tabular conversion of the authoritative CVEProject/cvelistV5 JSON corpus. Snapshot used in this study: 1 September 2026. <https://github.com/flying-coyote/nvd-csv>.