Beyond CVEs: Mastering the Landscape with Vulnerability-Lookup

FIRSTCON25 - 37th ANNUAL FIRST CONFERENCE

https://www.vulnerability-lookup.org

Alexandre Dulaunoy - alexandre.dulaunoy@circl.lu

June 25, 2025

CIRCL https://www.circl.lu

TLP:CLEAR

Origin of the project

Who is behind Vulnerability-Lookup?

Vulnerability-Lookup[1] is an Open Source project led by **CIRCL**.
It is co-funded by **CIRCL** and the **European Union**[2].
Used by many organisations including CSIRTs and ENISA (EUVD).
A reference implementation to **GCVE** standards.

vulnerability
-lookup

[1] https://www.vulnerability-lookup.org
[2] https://github.com/ngsoti

Origin

- `cve-search`[3] is an open-source tool initially developed in late 2012, focusing on maintaining a **local** CVE database.
- `cve-search` is widely used as an **internal** tool.
- The design and scalability of cve-search are limited. Our operational public instance at https://cve.circl.lu has reached a hard limit of 20,000 queries per second.
- Vulnerability sources have **diversified**, and the **NVD CVE is no longer the sole source** of vulnerability information.

[3] `https://github.com/cve-search/cve-search`

Initial Challenges

- **Volume of data:** Handling a substantial dataset and heavy network traffic, currently over 1,360,500 security advisories and more than 90,000 sightings[4].
- **Flexibility:** Balancing ongoing development with legacy issues while designing a future-proof architecture. It's complex and yes, sometimes chaotic[5].
- **Robustness:** Validating data even when external entities don't comply with their own JSON schemas. It's not always pretty.
- **Fast lookup:** Rapidly correlating identifiers across **diverse sources**, including unpublished advisories.

[4] The first sighting on Exploit-DB dates back 26 years.

[5] We enjoy challenges, especially when they lead to practical solutions.

TLP:CLEAR

Ongoing Challenges and Development

- **CPE fragmentation:**[6] Tackling the fragmentation of CPEs (e.g., `cpe:/a:oracle:java` vs. `cpe:/a:sun:java`) by introducing *Organizations* as unified containers.
- **CVD process:** Building an open-source tool that fully supports the Coordinated Vulnerability Disclosure (CVD) process.[7]
- **Vulnerability numbering:** Enabling a new distributed approach through the Global CVE Allocation System.[8]
- **Scoring vulnerabilities:** Aggregating a large volume of observations from diverse advisory types to improve vulnerability scoring.

[6] Well, another mess to clean up!

[7] Aligned with NIS 2 and the Cyber Resilience Act.

[8] `https://gcve.eu`

Current Sources in Vulnerability-Lookup

- **CISA Known Exploited Vulnerability** (HTTP)
- **NIST NVD CVE** (API 2.0)
- **CVEProject - cvelist** (Git submodule)
- **Fraunhofer FKIE** (Git submodule)
- **Cloud Security Alliance - GSD** (Git submodule)
- **GitHub Advisory DB** (Git submodule)
- **PySec Advisory DB** (Git submodule)
- **CSAF 2.0** (HTTP CSAF)

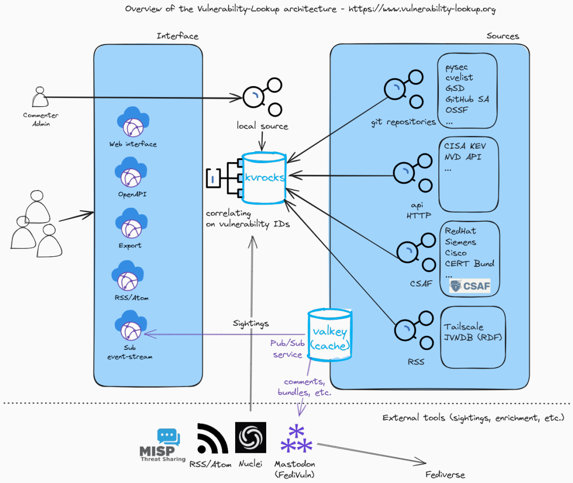
  CERT-Bund, Cisco, Siemens, Red Hat, Microsoft, NCSC-NL, CISA, etc.
- **VARIoT** (API)
- **Japan - JVN DB** (HTTP)
- **Tailscale** (RSS)
- **GCVE.eu all GNA sources**
- **CWE, CAPEC, MITRE EMB3D or KEV**
- **Growing...**

Open Data Initiative: Regular JSON dumps published[9].

[9] https://vulnerability.circl.lu/dumps/

Design and Implementation

Vulnerability-Lookup High-Level Architecture


Overview of the Vulnerability-Lookup architecture - https://www.vulnerability-lookup.org

Extended API

```
$ curl -s https://vulnerability.circl.lu/api/vulnerability/last/csaf_redhat/10 | jq .[2].document.title
"Red Hat Security Advisory: Red Hat Ceph Storage 6.1 security and bug fix update"

$ curl -s https://vulnerability.circl.lu/api/vulnerability/last/csaf_redhat/10 | jq .[2].vulnerabilities[0].cve
"CVE-2021-4231"
```

- **Documented API** (OpenAPI): https://vulnerability.circl.lu/api
- Pagination and filtering by source
- CPE search by vendor and product name
- **Many endpoints available via RSS and Atom**[10]

[10] https://www.vulnerability-lookup.org/documentation/feeds.html

Empowering the Community

Crowd-Sourced Threat Intelligence

- **Bundles:** Group similar vulnerabilities and aggregate sightings for easier tracking.
- **Comments:** Additional context such as PoCs, remediations, related insights.
- **Tags:** Use the MISP Vulnerability Taxonomy to annotate comments[11]. Example:

```
vulnerability:information=remediation
```

- **Sightings:** Report real-world observations of vulnerabilities, including metadata like timestamps and sources.

```
{
  "uuid": "f9ec8b2c-2ceb-4c05-b052-264b51c6a3ee", "vulnerability_lookup_origin": "1a89b78e-f703-45f3-bb86-59eb712668bd",
  "author": "9f56dd64-161d-43a6-b9c3-555944290a09", "creation_timestamp": "2025-04-17T19:14:32.000000Z",
  "vulnerability": "CVE-2025-32433",
  "type": "exploited",
  "source": "https://gist.github.com/numanturle/b7333fb02a4ee3618995babc9b62c507"
}
```

[11] https://www.misp-project.org/taxonomies.html#_vulnerability_3

Types of Sightings

Type	Description	Negative/Opposite
`seen`	The vulnerability was mentioned, discussed, or observed by the user.	-
`confirmed`	The vulnerability has been verified by an analyst.	X
`exploited`	The vulnerability was actively exploited and observed by the user reporting the sighting.	X
`patched`	The vulnerability was successfully mitigated or patched by the user reporting the sighting.	X

Table 1: Types of vulnerability sightings

TLP:CLEAR

Automated Sightings: Tools and Sources

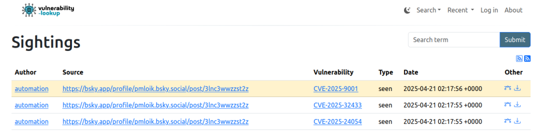
Automatically gathering crowd-sourced intelligence without requiring direct user contributions to our platform.

- **Social Platforms:** Fediverse, Bluesky
- **Threat Intelligence Tools:** MISP, Nuclei
- **Content Feeds:** RSS/Atom, curated web pages, GitHub Gist
- **Specialized Projects:** ShadowSight, ExploitDBSighting
- **Community Contributions:** Passive signals and indirect data enrichment

Scoring Vulnerabilities

Sightings Detection Rate and Types of Sightings

- A high rate of sightings (type *seen*) often correlates with high or critical severity vulnerabilities[12].
- Early sightings of type *exploited* (e.g., proof-of-concept code) or *confirmed* (e.g., detection templates for tools like Nuclei) can signal emerging threats.
- Sightings can sometimes be detected **before any official advisory is published**.



Sightings

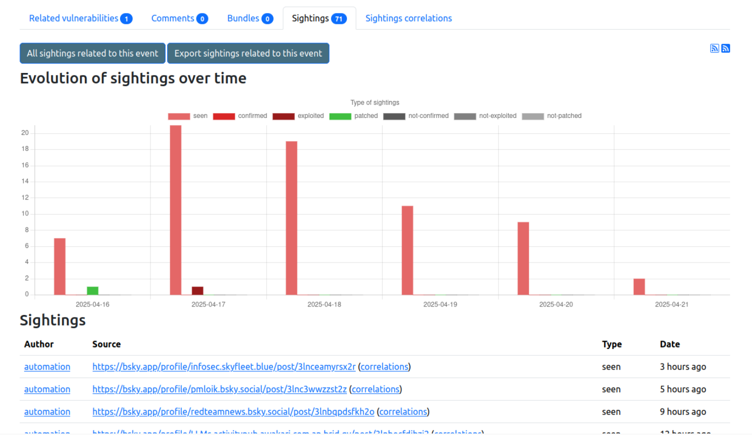
Author	Source	Vulnerability	Type	Date	Other
automation	https://bsky.app/profile/pmloik.bsky.social/post/3lnc3wwzzst2z	CVE-2025-9001	seen	2025-04-21 02:17:56 +0000	
automation	https://bsky.app/profile/pmloik.bsky.social/post/3lnc3wwzzst2z	CVE-2025-32433	seen	2025-04-21 02:17:55 +0000	
automation	https://bsky.app/profile/pmloik.bsky.social/post/3lnc3wwzzst2z	CVE-2025-24054	seen	2025-04-21 02:17:55 +0000	

- Continuous exploitation patterns are frequently observed through sources like The Shadowserver Foundation or MISP.

[12] Don't underestimate the hype surrounding some vulnerabilities.

TLP:CLEAR

Early PoC (erlang / otp)



Related vulnerabilities 1
Comments 0
Bundles 0
Sightings 71
Sightings correlations
All sightings related to this event
Export sightings related to this event

Evolution of sightings over time

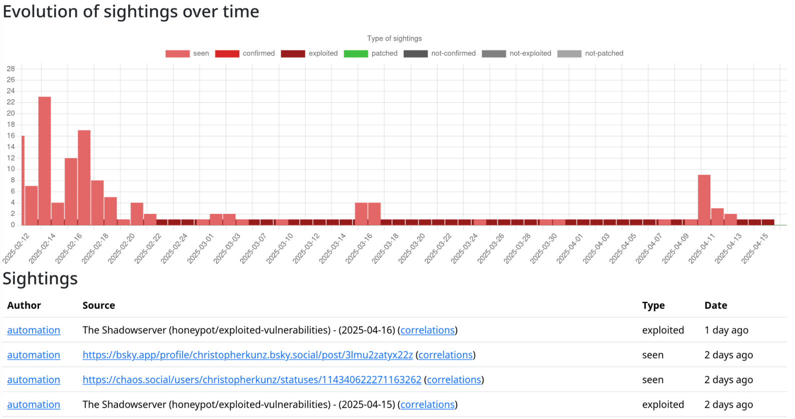
Sightings

Author	Source	Type	Date
automation	https://bsky.app/profile/infosec.skyfleet.blue/post/3lnceamyrsx2r (correlations)	seen	3 hours ago
automation	https://bsky.app/profile/pmloik.bsky.social/post/3lnc3wwzzst2z (correlations)	seen	5 hours ago
automation	https://bsky.app/profile/redteamnews.bsky.social/post/3lnbqpdsfkh2o (correlations)	seen	9 hours ago
automation	https://bsky.app/profile/LLMs.activitypub.awakari.com.ap.brid.gy/post/3lnbcfdibzi2 (correlations)	seen	12 hours ago

https://vulnerability.circl.lu/vuln/CVE-2025-32433#sightings

TLP:CLEAR

Continuous Exploitations (Palo Alto Networks / Cloud NGFW)



Evolution of sightings over time

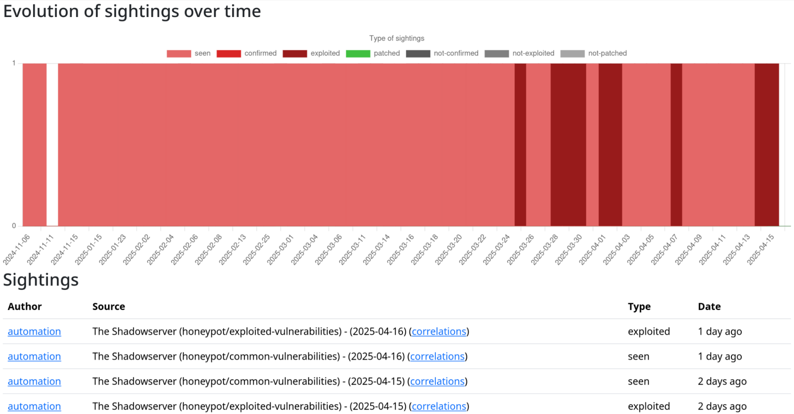
Sightings

Author	Source	Type	Date
automation	The Shadowserver (honeypot/exploited-vulnerabilities) - (2025-04-16) (correlations)	exploited	1 day ago
automation	https://bsky.app/profile/christopherkunz.bsky.social/post/3lmu2zatyx22z (correlations)	seen	2 days ago
automation	https://chaos.social/users/christopherkunz/statuses/114340622271163262 (correlations)	seen	2 days ago
automation	The Shadowserver (honeypot/exploited-vulnerabilities) - (2025-04-15) (correlations)	exploited	2 days ago

https://vulnerability.circl.lu/vuln/CVE-2025-0108#sightings

TLP:CLEAR

Continuous Exploitations (D-Link / DNS-320)



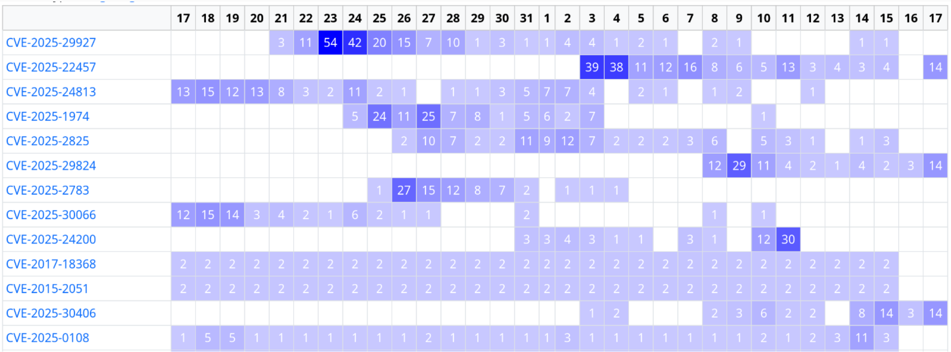
Evolution of sightings over time

Sightings

Author	Source	Type	Date
automation	The Shadowserver (honeypot/exploited-vulnerabilities) - (2025-04-16) (correlations)	exploited	1 day ago
automation	The Shadowserver (honeypot/common-vulnerabilities) - (2025-04-16) (correlations)	seen	1 day ago
automation	The Shadowserver (honeypot/common-vulnerabilities) - (2025-04-15) (correlations)	seen	2 days ago
automation	The Shadowserver (honeypot/exploited-vulnerabilities) - (2025-04-15) (correlations)	exploited	2 days ago

https://vulnerability.circl.lu/vuln/CVE-2024-10914#sightings

Last Month’s Most Sighted Vulnerabilities



	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
CVE-2025-29927					3	11	54	42	20	15	7	10	1	3	1	1	4	4	1	2	1		2	1					1	1		
CVE-2025-22457																		39	38	11	12	16	8	6	5	13	3	4	3	4		14
CVE-2025-24813	13	15	12	13	8	3	2	11	2	1		1	1	3	5	7	7	4		2	1		1	2			1					
CVE-2025-1974								5	24	11	25	7	8	1	5	6	2	7							1							
CVE-2025-2825										2	10	7	2	2	11	9	12	7	2	2	2	3	6		5	3	1		1	3		
CVE-2025-29824																							12	29	11	4	2	1	4	2	3	14
CVE-2025-2783									1	27	15	12	8	7	2		1	1	1													
CVE-2025-30066	12	15	14	3	4	2	1	6	2	1	1				2								1		1							
CVE-2025-24200															3	3	4	3	1	1		3	1		12	30						
CVE-2017-18368	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		
CVE-2015-2051	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		
CVE-2025-30406																		1	2				2	3	6	2	2		8	14	3	14
CVE-2025-0108	1	5	5	1	1	1	1	1	1	1	2	1	1	1	1	1	3	1	1	1	1	1	1	1	2	1	2	3	11	3		

- **CVE-2025-22457:** Ivanti / Connect Secure — Severity: 10.0 (Critical)
- **CVE-2025-29927:** Vercel / Next.js — Severity: 9.1 (Critical)

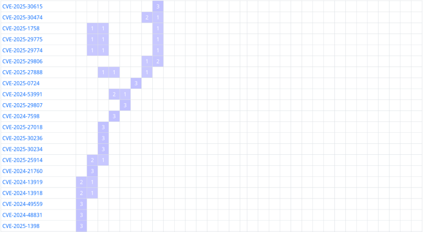
Other Examples

Vulnerability	Product	Sighting count	EPSS	Severity
CVE-2025-29927	next.js	167	89.24% (0.99521)	9.1
CVE-2025-24813	Apache Tomcat	128	93.55% (0.99827)	9.2
CVE-2024-4577	PHP	190	94.38% (0.99961)	9.8
CVE-2025-0282	Connect Secure	243	90.87% (0.99618)	9.0
CVE-2024-55591	FortiOS	126	92.79% (0.99756)	9.8
CVE-2024-10914	D-Link DNS-320	81	93.73% (0.9985)	9.2
CVE-2020-21650	Myucms	57	2.48% (0.83998)	9.1

Table 2: Top vulnerabilities from our April 2025 report, based on sightings and scoring data.

TLP:CLEAR

Least Sighted Vulnerabilities in the Last Month



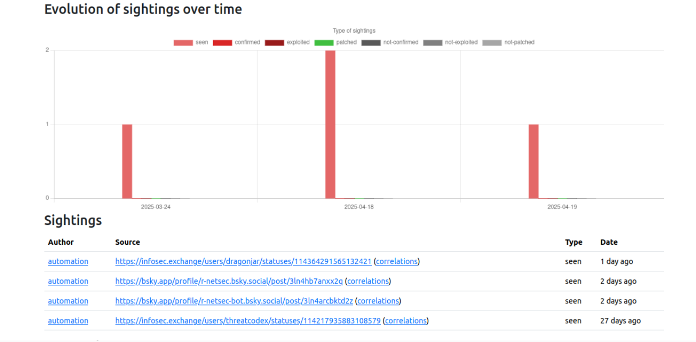
CVE-2025-30615								3
CVE-2025-30474							2	1
CVE-2025-1758		1	1					1
CVE-2025-29775		1	1					1
CVE-2025-29774		1	1					1
CVE-2025-29806							1	2
CVE-2025-27888			1	1			1	
CVE-2025-0724						3		
CVE-2024-53991				2	1			
CVE-2025-29807					3			
CVE-2024-7598				3				
CVE-2025-27018			3					
CVE-2025-30236			3					
CVE-2025-30234			3					
CVE-2025-25914		2	1					
CVE-2024-21760		3						
CVE-2024-13919	2	1						
CVE-2024-13918	2	1						
CVE-2024-49559	3							
CVE-2024-48831	3							
CVE-2025-1398	3							

- **Low-sighting outliers offer valuable intel**, even if absent from EPSS or predictive models.
- Particularly relevant in low-noise sources (e.g., MISP, private Telegram channels).
- Often rated low/medium by CVSS and have low EPSS scores.
- Trend highlights EPSS's dependence on public threat intel feeds.

TLP:CLEAR

Tracking the Exploitability of Vulnerabilities Prior to Public Disclosure

- **Google / Android:** https://vulnerability.circl.lu/vuln/CVE-2024-43093#sightings
- **Speedify VPN (macOS):** https://vulnerability.circl.lu/vuln/CVE-2025-25364#sightings
- **SourceCodester:** https://vulnerability.circl.lu/vuln/CVE-2025-3821#sightings
  - Low visibility, no EPSS score, few sightings



Evolution of sightings over time

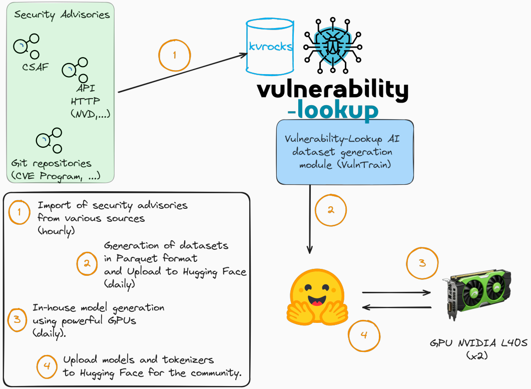
Sightings

Author	Source	Type	Date
automation	https://infosec.exchange/users/dragonjar/statuses/114364291565132421 (correlations)	seen	1 day ago
automation	https://bsky.app/profile/r-netsec.bsky.social/post/3ln4hb7anxx2q (correlations)	seen	2 days ago
automation	https://bsky.app/profile/r-netsec-bot.bsky.social/post/3ln4arcbktd2z (correlations)	seen	2 days ago
automation	https://infosec.exchange/users/threatcodex/statuses/114217935883108579 (correlations)	seen	27 days ago

TLP:CLEAR

Toward Practical AI Applications

Completing Missing Data with AI


Security Advisories
CSAF
API HTTP (NVD,...)
Git repositories (CVE Program, ...)
1
kvrocks
vulnerability-lookup
Vulnerability-Lookup AI dataset generation module (VulnTrain)
2
3
4
GPU NVIDIA L40S (x2)
1 Import of security advisories from various sources (hourly)
2 Generation of datasets in Parquet format and Upload to Hugging Face (daily)
3 In-house model generation using powerful GPUs (daily).
4 Upload models and tokenizers to Hugging Face for the community.

- Some vulnerabilities are published without an assigned CVSS score.
- To address this, we developed **VLAI Severity**[a], a model trained on the Vulnerability-Lookup dataset.
- **Predicts severity from the vulnerability description** before an official score is available.
- Available as a standalone model or via the CIRCL public instance.

[a] https://www.vulnerability-lookup.org/user-manual/ai/

TLP:CLEAR

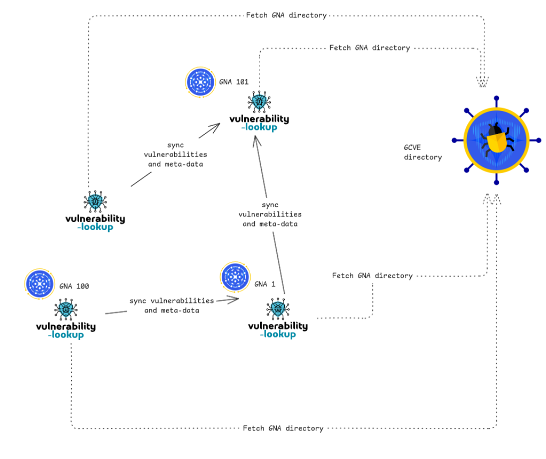
Lookup is Cool, but Publishing is Even Cooler

GCVE.eu - Role

- The primary role of GCVE[13] is to provide **globally unique identifiers** to GCVE Numbering Authorities (GNAs).
- **GNAs operate autonomously**, with full control over how they assign and manage identifiers.
- **GCVE publishes Best Current Practices (BCPs)** on directory management, Coordinated Vulnerability Disclosure (CVD), and publication protocols.
- GCVE maintains and publishes the **official directory of all GNAs**, including their publication endpoints.

[13] https://gcve.eu/

Decentralized Publication Standard



TLP:CLEAR

Closing

Future Development

- Deeper analysis of the content and context of sightings, including **source reliability assessment**.
- Full-text search capabilities across all integrated sources.
- Integration of scoring models such as Vuln4Cast[14], with testing planned on our dataset to enhance reproducibility.
- **Improved notification capabilities** for newly observed vulnerabilities via webhooks.

> The project is evolving rapidly — feedback and feature suggestions are always welcome!

[14] https://github.com/FIRSTdotorg/Vuln4Cast

References

https://www.vulnerability-lookup.org

CIRCL public instance https://vulnerability.circl.lu

Source code https://github.com/vulnerability-lookup/vulnerability-lookup

Dataset, AI Model Training, Models
https://github.com/vulnerability-lookup/VulnTrain