



[HTTPS://CONFERENCE.HITB.ORG/HITBSECCONF2024BKK](https://conference.hitb.org/hitbseccconf2024bkk)

## Secret Scanning in Open Source at Scale (in-depth)

- Hassan Khan Yusufzai
- Danish Tariq

Directors, Laburity



Main Track

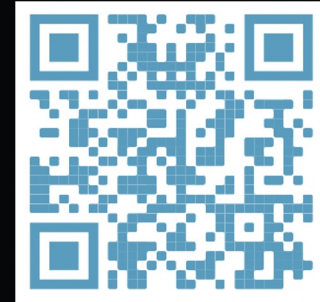
29 AUG

#HITB2024BKK

`$ /usr/local/bin/whoami`

## Hassan Khan Yusufzai

- Director, Laburity - Security Researcher
- Blackhat MEA 2022 & 2023, ThreatCon 2023, DevSecCon 2023, EOCON 2023 Security Analyst Summit 2023.
- Author of Rails Security Guide
- Helped and got acknowledged by companies like include Google, Microsoft, Dell, Intel, Magento and other 200+
- Author of multiple CVE's
- Travelling around the world and hack
- Love to perform at scale research  $\backslash(^o^)/$
- OSCP



#HITB2024BKK

## Disclaimer

This talk/presentation is for educational purposes and is intended to spread awareness in a way that you could be vigilant. Nothing here is done or presented to be used in any malicious/illegal/unethical way.

This talk is all about education and awareness! We're like the friendly neighborhood watch for the internet, here to share knowledge, not malicious stuff.

#HITB2024BKK



## Open Source Software

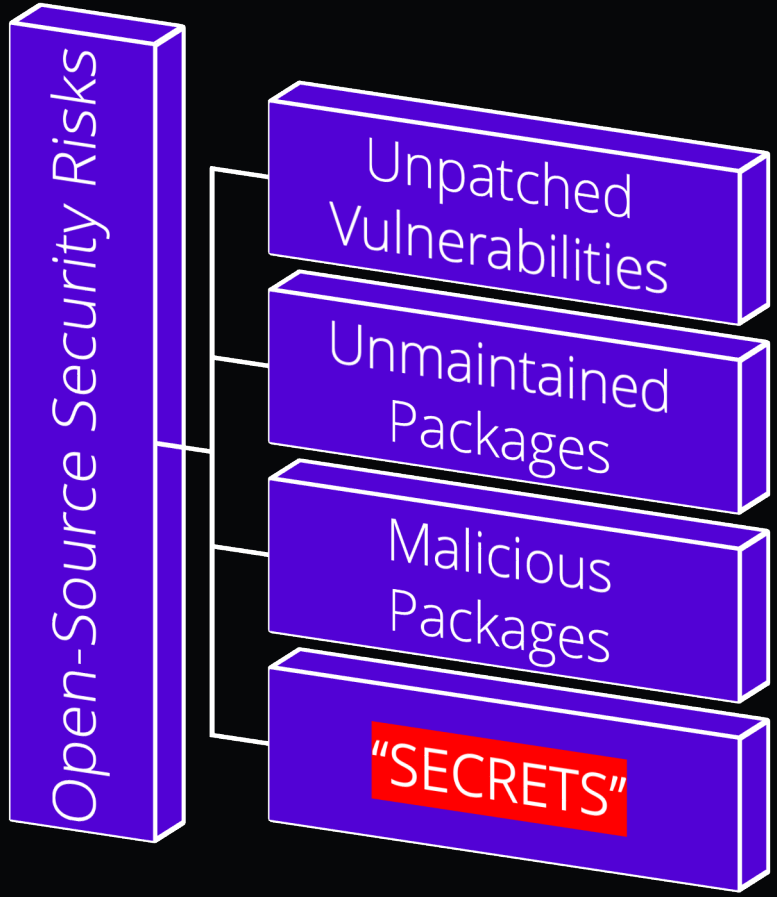
---

A software that is distributed with its source code, making it available for use, inspection, modification, enhancement, and possibly the redistribution.



#HITB2024BKK

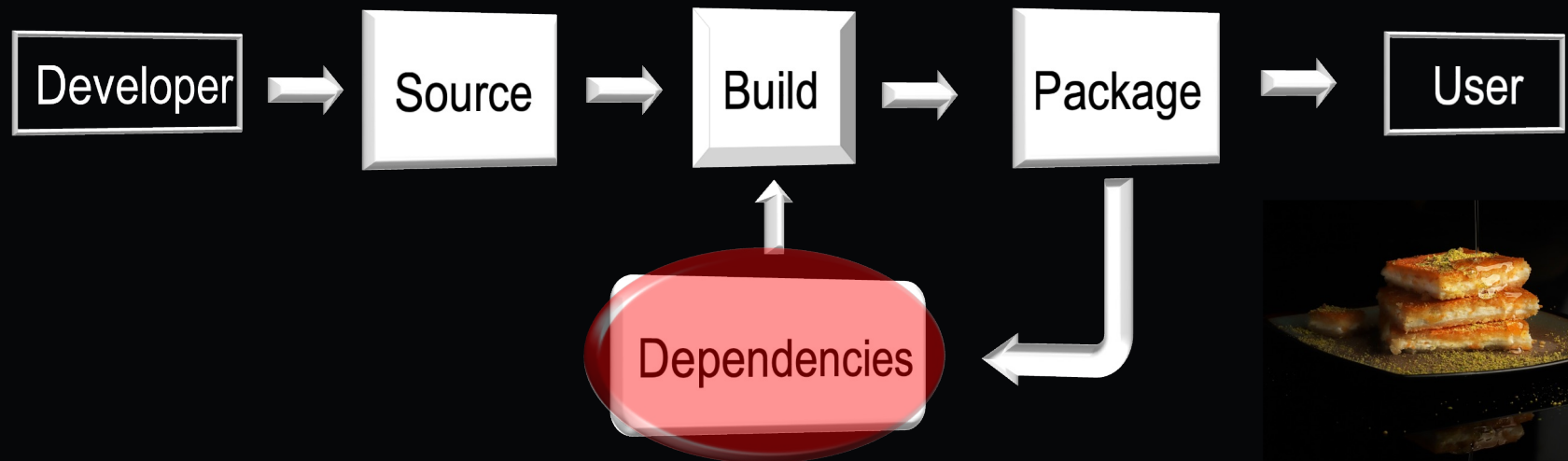




#HITB2024BKK

## Software Supply Chain

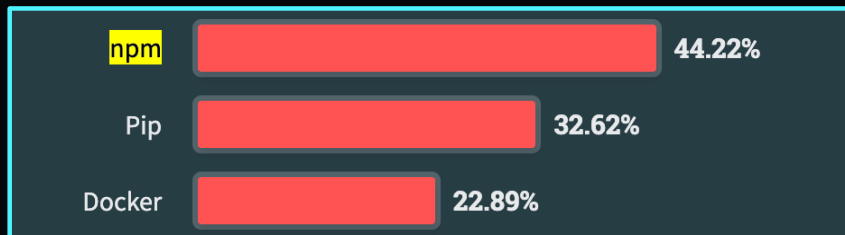
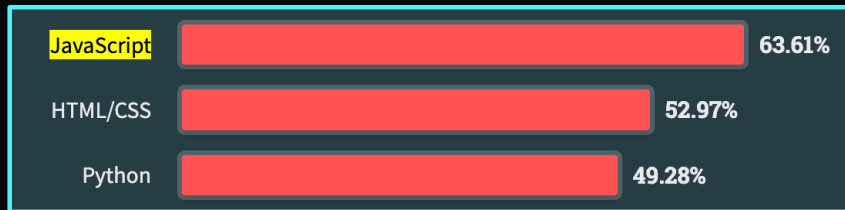
Software supply chain is anything that goes into the codebase until it made it to production whether we talk about the dependencies, binaries, or other components.



#HITB2024BKK

# JS Packages

2023 continues JavaScript's streak as its eleventh year in a row as the most commonly-used programming language.



- 2023 Developer Survey – Stack OverFlow

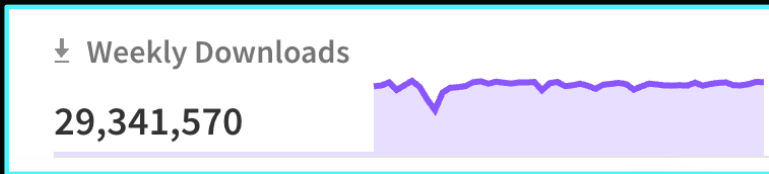
#HITB2024BKK

# Example of a package

**express** 

4.18.2 • Public • Published a year ago

 [Readme](#)  [Code](#) Beta  31 Dependencies  74,270 Dependents  270 Versions



- <https://www.npmjs.com/package/express>

#HITB2024BKK



# Our last year's research— Account Takeover in NPM Packages

The Register®

## How to find NPM dependencies vulnerable to account hijacking

Security engineer outlines self-help strategy for keeping software supply chain safe

 [Thomas Claburn](#)

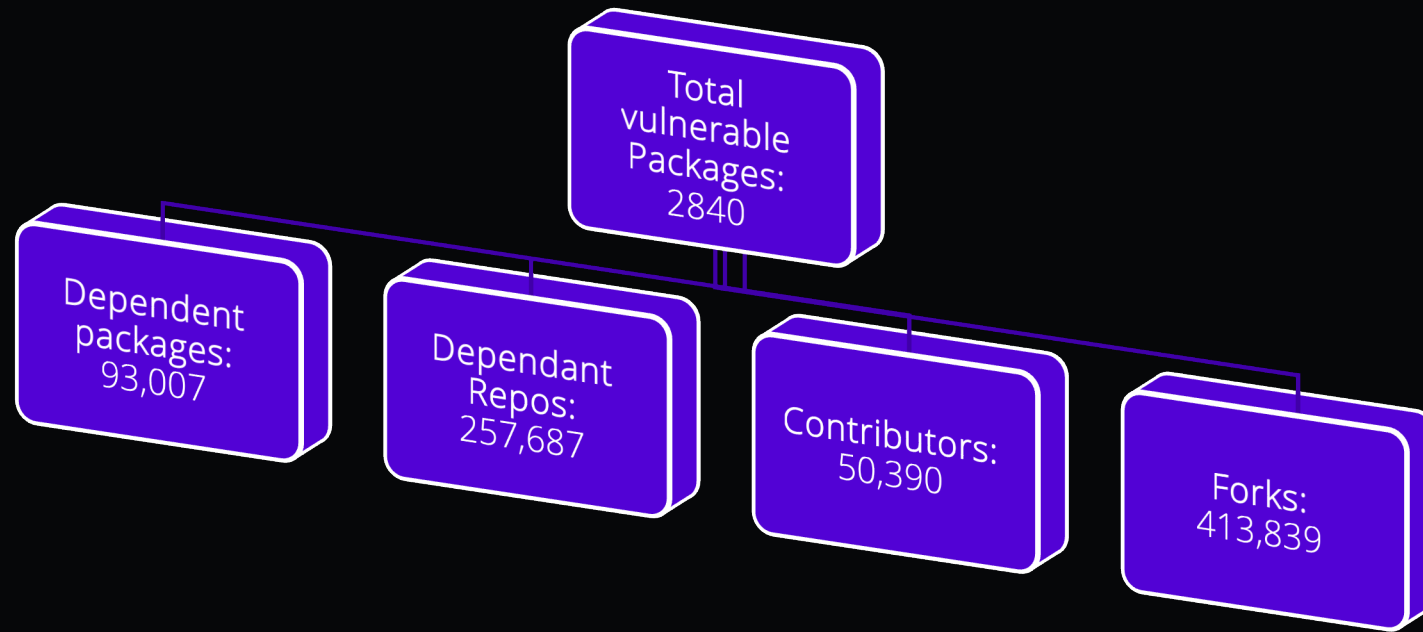
Mon 23 May 2022 // 07:58 UTC

Following the recent disclosure of a technique for hijacking certain NPM packages, security engineer [Danish Tariq](#) has proposed a defensive strategy for those looking to assess whether their web apps include dependencies tied to subvertable email domains.

Credits: The Register

#HITB2024BKK

# Our last year's research— Account Takeover in NPM Packages



#HITB2024BKK

## Secrets

---

Secrets exist as artifacts in code that an app uses to connect to an external service, account, or application. Secrets used by developers include API keys; encryption keys; OAuth tokens; certificates; Webhooks and passwords etc.

Why do they exist?

Ease! Speed! To prevent reinventing the wheel! A lot of good reasons!

What if the secrets gets leaked?

### The World's Largest Breach. Root cause: Secrets in Code

- BluBracket, 2022

On July 3rd, Changpeng Zhao, the highly regarded CEO of cryptocurrency exchange **Binance**, posted a tweet under his widely known moniker CZ, alerting the world to a massive data breach that in part read, "our threat intelligence detected 1 billion resident records for sell in the dark web, including name, address, national id, mobile, police and medical records from one asian country". [[https://twitter.com/cz\\_binance/status/1543700689611792386](https://twitter.com/cz_binance/status/1543700689611792386)]

#HITB2024BKK

## Major incidents related to Secrets mishandling

“Microsoft confirmed the consumer signing key used to breach email accounts in May was stolen from the software giant's own network.” - TechTarget

Our investigation found that a consumer signing system crash in April of 2021 resulted in a snapshot of the crashed process (“crash dump”). The crash dumps, which redact sensitive information, should not include the signing key. In this case, a race condition allowed the key to be present in the crash dump (this issue has been corrected). The key material’s presence in the crash dump was not detected by our systems (this issue has been corrected).

We found that this crash dump, believed at the time not to contain key material, was subsequently moved from the isolated production network into our debugging environment on the internet connected corporate network. This is consistent with our standard debugging processes. Our credential scanning methods did not detect its presence (this issue has been corrected).

- Microsoft - Results of Major Technical Investigations for Storm-0558 Key Acquisition - 2023

#HITB2024BKK

## Major incidents related to Secrets mishandling

---

“Apology and notification regarding the possibility of leakage of customer email addresses” - **Toyota** (Auto-Translated)

September 15, 2022. It has been discovered that the released source code contains an access key to the data server, which can be used to access email addresses and customer management numbers stored on the data server. On the same day, we immediately made the source code private on GitHub, and on September 17th we took measures such as changing the data server access key, and no secondary damage has been confirmed.

- Toyota (Auto-Translated) - <https://global.toyota/jp/newsroom/corporate/38095972.html>

#HITB2024BKK

## Major incidents related to Secrets mishandling

---

“On December 29, 2022, we were notified of suspicious activity on our GitHub account. Upon investigation, we discovered that a limited number of Slack employee tokens were stolen and misused to gain access to our externally hosted GitHub repository. Our investigation also revealed that the threat actor downloaded private code repositories on December 27. No downloaded repositories contained customer data, means to access customer data, or Slack’s primary codebase.”

- Slack.

#HITB2024BKK

## Major incidents related to Secrets mishandling

---

Sourcegraph experienced a security incident on August 30, 2023 where a malicious actor used a leaked admin access token in our public Sourcegraph instance at Sourcegraph.com. The malicious external user used their privileges to increase API rate limits for a small number of users.

- Sourcegraph, 2023

#HITB2024BKK

## Major incidents related to Secrets mishandling

As part of these scans, the hacker says they found a PowerShell script containing admin credentials for the company's Thycotic privileged access management (PAM) platform, which was used to access the login secrets for the company's other internal services.

*"ok so basically uber had a network share \\[redacted]pts. the share contained some powershell scripts.*

*one of the powershell scripts contained the username and password for a admin user in Thycotic (PAM) Using this i was able to extract secrets for all services, DA, DUO, Onelogin, AWS, Gsuite"*

The New York Times reports that the attacker claimed to have accessed Uber databases and source code as part of the attack.

To be clear, this information is from the threat actors and has not been verified by Uber, which has not responded to our requests for more information.

Credits: BleepingComputer, 2022

#HITB2024BKK



# New Supply Chain Attack Exploits Abandoned S3 Buckets to Distribute Malicious Binaries

"Malicious binaries steal the user IDs, passwords, local machine environment variables, and local host name, and then exfiltrates the stolen data to the hijacked bucket," [Checkmarx](#) researcher Guy Nachshon said.

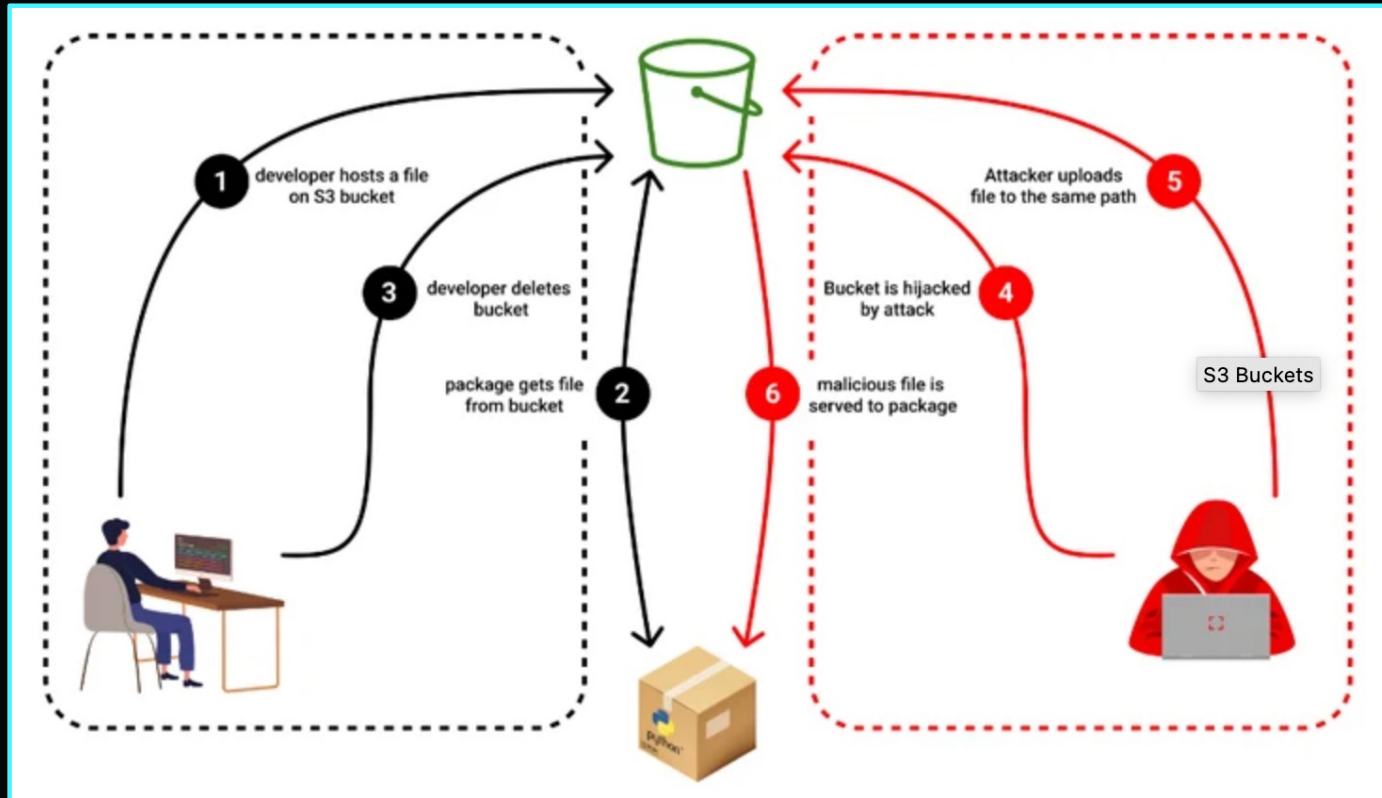
The attack was first observed in the case of an npm package called [bignum](#), which, until version 0.13.0, relied on an Amazon S3 bucket to download pre-built binary versions of an addon named node-pre-gyp during installation.

"These binaries were published on a now-expired S3 bucket which has since been claimed by a malicious third party which is now serving binaries containing malware that exfiltrates data from the user's computer," according to a [GitHub advisory](#) published on May 24, 2023.

An unknown threat actor is said to have seized on the opportunity that the S3 bucket was once active to deliver malware when unsuspecting users downloaded the package in question.

-TheHackerNews

#HITB2024BKK



-TheHackerNews

#HITB2024BKK

"...data exposure incident on Microsoft's AI GitHub repository, including over 30,000 internal Microsoft

"People Are Pirating GPT-4 By Scraping Exposed API Keys"  
-Vice

apps leak hard-coded secrets, research shows"  
-Cybernews

"Attackers Scrape GitHub For Cloud Service Credentials, Hijack Account To Mine Virtual Currency"  
-Forbes

"Apology and notification regarding the possibility of leakage of customer email addresses"  
- Toyota (Auto-Translated)

-Codecov

Leaked secrets in GitHub account granted access to sensitive data for Equifax customers.

"..security incident on August 30, 2023 where a malicious actor used a leaked admin access token in our public Sourcegraph instance"  
- Sourcegraph

leaked via Open GitHub Repository."  
-Starbucks, Hackerone 716292

"T "Microsoft confirmed the consumer signing key used on to breach email accounts au to breach email accounts ke in May was stolen from -Bl the software giant's own network."  
- TechTarget

"Slack employees' credentials were stolen and misused to gain access to our externally hosted GitHub repository"  
-Slack

"The World's Largest Breach. Root cause: Secrets in Code"  
- BluBracket

## Another Talk another Research – SECRET SCANNING




---

- Open-Source Secret Scanning
- At scale research
- Packages were collected from the publicly available sources.
- Crafted and modified scripts + secret sauces + our high RAM servers were utilized to conduct this research.
- 33 secret scanning Signatures

#HITB2024BKK

# WordPress Plugins

Top In Open Source Usage Distribution on the Entire Internet

Technology	Websites	%
 WordPress	34,770,646	44.8
 WooCommerce Checkout	3,557,626	8.43
 Joomla!	1,109,953	1.43

-<https://trends.builtwith.com/cms/open-source/traffic/Entire-Internet>

#HITB2024BKK

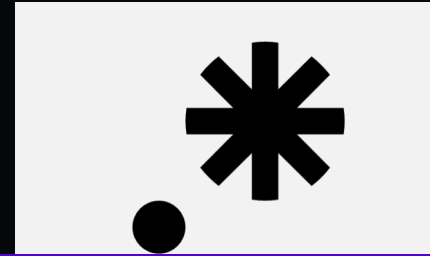
# WordPress Secret Scanning at Scale



59,705



3645



33

Active Installations: 35M

#HITB2024BKK

# WordPress Secret Scanning at Scale - breakdown

Most Common File

README.MD

Exposures: 215



#HITB2024BKK

## Readme.md: Am I a joke to you???

### Microsoft AI Researchers Accidentally Expose 38 Terabytes of Confidential Data

"The exposure came as the result of an overly permissive [SAS token](#) – an Azure feature that allows users to share data in a manner that is both hard to track and hard to revoke," Wiz [said](#) in a report. The issue was reported to Microsoft on June 22, 2023.

Specifically, the repository's [README.md](#) file instructed developers to download the models from an Azure Storage URL that accidentally also granted access to the entire storage account, thereby exposing additional private data.

Credits: TheHackerNews

#HITB2024BKK



# AWS Access Key ID in Plugins & Amazon SNS Token

AWS Key ID: 41

SNS Tokens: 4

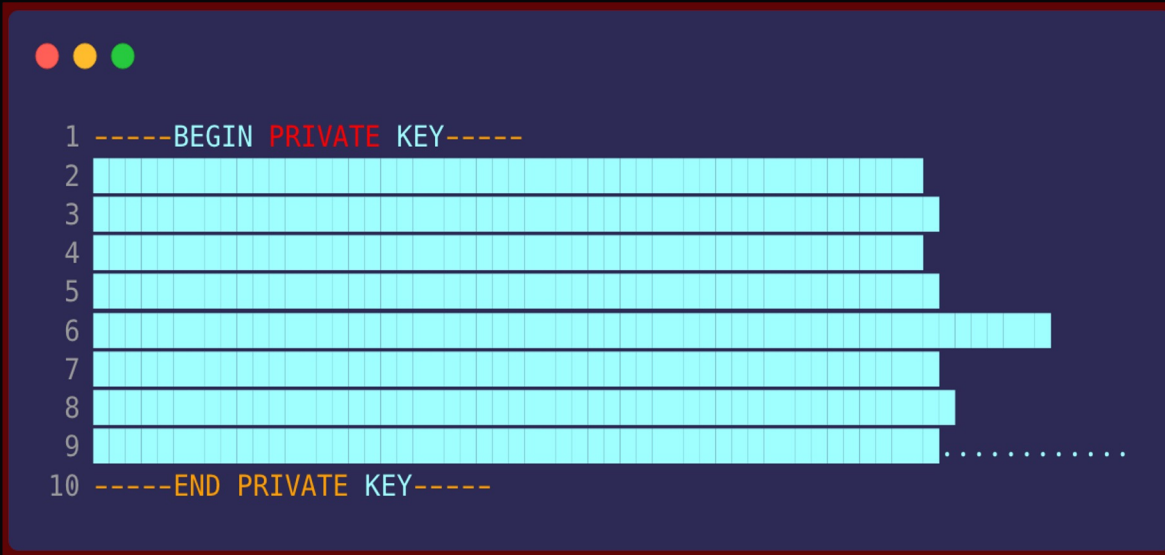
Active Installation: 6M

```
1 $api_key = $wpdb->get_var( ); // 'AKIA '
2 $secret_key = $wpdb->get_var( ); // ' '
3 $as = $wpdb->get_var( ); // ' '
4 $country = $wpdb->get_var($sql_ ); // 'com'
5 return go( [ ])
6 }
```

```
1 $options['access_key'] = trim('AKIA ');
2 $options['secret_access_key'] = trim(' ');
```

#HITB2024BKK

# Private Key Exposure



Private keys: 2539

Common file

Private.key

Active installations: 12M

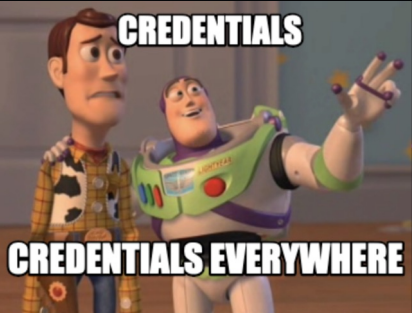
#HITB2024BKK



# Basic Authorization Credentials Check

```
1 http://[redacted]:[redacted]@x[redacted].com:8042/en/[redacted]
```

BAC exposures: 166



Active Installations: 9M

#HITB2024BKK



# Google API Keys

```
1 var config = {apiKey:"AIza [REDACTED]",authDomain:"[REDACTED].firebaseapp.com".....
```

Google API Keys: 734

Most common file

Map.php

Active Installations: 11M

#HITB2024BKK



# Github App Token and Github Personal Token

```
1 [REDACTED]_TOKEN=ghs_[REDACTED]
```

Github App & Personnel Tokens: 12

Active Installations

22,870

#HITB2024BKK



# Slack Webhook and Slack API Key

```
1 ```php
2 $webhook = 'https://hooks.slack.com/[REDACTED]';
```

Slack Exposures: 47

```
1 ```php
2 $[REDACTED]Token = 'xoxb-[REDACTED]-[REDACTED]-[REDACTED]';
3 $channel = '#[REDACTED]';
```

Active Installations: 374K



#HITB2024BKK



# Firestore Database Detect

```
1      'general' => (Object)[
2        'Firebase' => (Object)[
3          'apiKey' => '[REDACTED]',
4          'authDomain' => '[REDACTED].firebaseapp.com',
5          'databaseURL' => 'https://[REDACTED].firebaseio.com',
6          'storageBucket' => '[REDACTED].appspot.com',
7          '[REDACTED]Id' => '[REDACTED]',
8          '[REDACTED]' => '[REDACTED]',

```

Firestore Exposures: 43

Active Installations: 107K

#HITB2024BKK

# Everywhere I go I see ~~his face~~ exposed SECRETS



Code Climate Token

Exposure: 11 Active Installations: 1M

Mailchimp API Key

Exposure: 19 Active Installations: 1M

Telegram Bot

Exposure: 8 Active Installations: 2K

Twitter Secret

Exposure: 6 Active Installations: 560

Zapier Webhook

Exposure: 15 Active Installations: 552K

#HITB2024BKK



## Seekrets - OSS

---

<https://github.com/Laburity/seekrets-oss>



#HITB2024BKK

# whoami



## Danish Tariq

- Director, Laburity – Security Researcher
- Spoke @ BlackHatMEA 2022 & 2023, ThreatCon 2023, EOCON 2023, DevSecCon 2023, and AllDayDevOps 2023
- Was a moderator @ OWASP 2022 Global AppSec APAC.
- Former top-rated professional at Upwork.
- Extensive research on supply chain attacks and open-source security landscape.
- Featured in "The Register" for an initial workaround for the NPM dependency attacks.
- Helped and got acknowledged by companies like include Microsoft, Apple, Nokia, Blackberry, and Adobe to name a few.
- Traveller.

#HITB2024BKK



# JS Packages we downloaded and Scanned

2,400,447 (2.4 Million+)

#HITB2024BKK



# Total Secrets Found in Packages

129,535 – False Positives

93,863 Secrets

## Top 5 Files with such secrets

<b>readme.md</b>	7514
<b>index.js</b>	3953
<b>server.key</b>	1519
<b>index.html</b>	1117
<b>key.pem</b>	1087

## Some interesting files

<b>.env</b>	506
<b>app.js</b>	427

#HITB2024BKK

# Amazon AWS Access keys, SNS, Session, and MWS

AWS Access Keys 6245

SNS Tokens 1984

Sessions Tokens 4

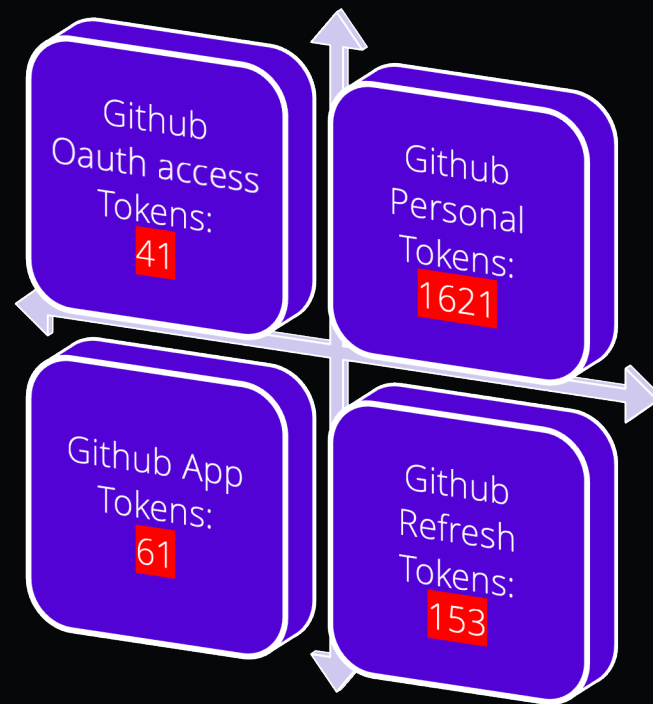
MWS Auth token 1

Common file  
index.js

```
1 storage: {  
2   type: 's3',  
3   accessKeyId: 'AKIA[REDACTED]',  
4   secretAccessKey: '[REDACTED]',  
5   bucketName: '[REDACTED]',  
6 },
```

#HITB2024BKK

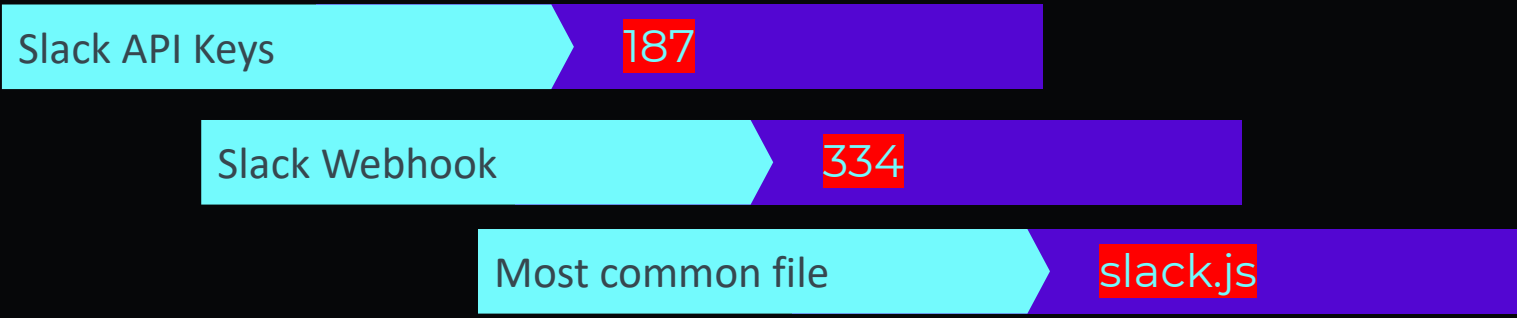
# Github Oauth access, personal, app, and refresh Tokens



#HITB2024BKK



# Slack API Keys and Webhooks in JS Packages



```
1 ({  
2   token: 'xoxb-[REDACTED]-[REDACTED]',  
3 });
```

```
1   "url": "https://hooks.slack.com/services/[REDACTED]/[REDACTED]/[REDACTED]",  
2   "payload": {  
3     "username": "[REDACTED]",  
4   },
```

#HITB2024BKK





# Basic Authorization Credentials in JS Packages

Basic Auth. Creds.

15,896

Top 4 Files with such secrets

readme.md	4042
index.js	1059
package.json	328
index.d.ts	252

```
1 List: [  
2   "http://[REDACTED]:[REDACTED]@[REDACTED].io:80/"  
3 ],
```

```
1 var mongodb = 'mongodb://[REDACTED]:[REDACTED]/[REDACTED]';  
2 [REDACTED].connect(mongodb);
```

#HITB2024BKK

# Zendesk Secret Keys in JS Packages

Zendesk Secret Keys

15

One of the common files

.env

```
1 export ZENDESK_URL=https://[REDACTED].zendesk.com
2 export ZENDESK_EMAIL=[REDACTED]
3 export ZENDESK_TOKEN=[REDACTED]
```

#HITB2024BKK

# Zapier Webhooks in JS Packages

Leaked Zapier Webhooks

54

Most common file

zapier.yml

```
1 npm publish  
2 curl -X POST -d '{"repo": "██████████"}' https://hooks.zapier.com/hooks/catch/██████/██████/
```

#HITB2024BKK







# MailChimp API Keys and Sendgrid API Keys in JS Packages

MailChimp API Keys

36

Most common file

index.js

```
1 const __[REDACTED]__MAILCHIMP_API_KEY__ = '[REDACTED]-us9';  
2 const __[REDACTED]__API_LIST_ID__ = '[REDACTED]';
```

Sendgrid API Keys

27

Most common file

mail.js

```
1 export const SMTP = {  
2   SENDGRID_API_KEY: "SG.[REDACTED]"  
3   sender: "[REDACTED]@[REDACTED].com"  
4 }
```

#HITB2024BKK

# Google API Key in JS Packages

Google API Keys

14,623

## Top 5 Files with such secrets

index.js	819
index.html	818
readme.md	261
google-services.json	235
config.js	216

```
1  "api_key": [  
2    {  
3      "key": "AIza[REDACTED]"  
4    }  
]
```

#HITB2024BKK



# GitLab Personal Access Token in JS Packages

Gitlab Personal access tokens

58

Most common file

index.js

```
1 method: "GET",
2 headers: {
3   "content-type": "application/json",
4   "PRIVATE-TOKEN": "glpat-[REDACTED]",
5 }
```

#HITB2024BKK

# Firestore database detect in JS Packages

Firestore DB

8,225

```
1 {
2   "project_info": {
3     "project_number": "██████████",
4     "firebase_url": "https://██████████.firebaseio.com",
5     "project_id": "██████████",
6     "storage_bucket": "██████████"
7   },
8   "api_key": [
9     {
10      "current_key": "AIza██████████"
11    }
12  ]
13 }
```

### Top 4 Files with such secrets

readme.md	1427
index.js	491
google-services.json	221
firebase.js	198

#HITB2024BKK



# Private Key Detect in JS Packages

Private Keys Exposure

43,163

Most common file

server.key

```
1 -----BEGIN RSA PRIVATE KEY-----  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27 -----END RSA PRIVATE KEY-----
```

#HITB2024BKK

## Bonus: Some more for fun

---

Heroku API Keys 8

Paypal BT Token 1

Twilio API Keys 5

Discord Secret 3

#HITB2024BKK

## Ruby Gems we downloaded and Scanned

---

Almost all scanned

Secrets: **16,830**

#HITB2024BKK

# Ruby Gems - Secrets



#HITB2024BKK

What's next ?

---

PHP & Python Packages?

#HITB2024BKK





## Significance of leaking secrets

---

“Use of stolen or compromised credentials remains the most common cause of a data breach. Stolen or compromised credentials were the primary attack vector in 19% of breaches in the 2022 study and also the top attack vector in the 2021 study, having caused 20% of breaches. Breaches caused by stolen or compromised credentials had an average cost of USD 4.50 million.”

- IBM Cost of a data breach 2022

#HITB2024BKK

# Keep secrets secret - Solution before and after leak of secrets

---

Should you only delete the file or Repo ? Would it solve your problem of leakage ?

Revoke it

Review access logs

Assess if your file or repo needs to be public

Take measures to prevent it from happening again.

#HITB2024BKK

## Keep secrets secret - Solution before and after leak of secrets

---

- **DEPLOY-REQ-3:** For code that is already in the repository and ready to be deployed, a security scanning sub-feature should be invoked to detect the presence of secrets in the code, such as keys and access tokens.

NIST SP 800-204D

#HITB2024BKK

# Keep secrets secret - Solution before and after leak of secrets

## About secret scanning

GitHub scans repositories for known types of secrets, to prevent fraudulent use of secrets that were committed accidentally.

Secret scanning alerts for partners runs automatically on public repositories and public npm packages to notify service providers about leaked secrets on GitHub.com.

Secret scanning alerts for users are available for free on all public repositories. Organizations using GitHub Enterprise Cloud with a license for GitHub Advanced Security can also enable secret scanning alerts for users on their private and internal repositories. For more information, see "[About secret scanning](#)" and "[About GitHub Advanced Security](#)."

Credits: Github

#HITB2024BKK

## Keep secrets secret - Solution before and after leak of secrets

---

- Use some good secrets management platforms.
- Secret stores like HashiCorp Vault, Azure Key Vault and AWS SSM Parameter store allow you to store your secrets securely and encrypted, with only resources that are allowed being able to decrypt and read the secret at run-time in your production environment.
- Have some automations within your CI/CD pipeline to detect those occurrences !
- Don't commit the secrets !!
- Talk to us?

#HITB2024BKK

🔄 Scan Package

☰ Scan Code File

🔒 Scan ZIP File

🗄️ Previous Scans

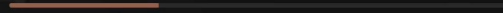
# Welcome to your dashboard!

## Completed Scans



9 ↑

## Scans In-Progress



1 ↑

## Scans Failed



0 ↑

### Past Searches

[View All Past Searches](#)

PACKAGE NAME	SCAN TIME	RESULT
redacted-package-2	2024-08-29 00:15:28	<a href="#">VIEW RESULT</a>
redacted-package	2024-08-29 00:16:50	<a href="#">VIEW RESULT</a>

**Join the waitlist**

---

<https://laburity.com/waitlist>



**#HITB2024BKK**





# ART OF MASS SCANNING

\$1,000.00

Attend In-person

#HITB2024BKK







## SOFTWARE SUPPLY CHAIN SECURITY: FROM OFFENSE TO DEFENSE

\$1,000.00

Attend In-person

#HITB2024BKK

Any questions?

#HITB2024BKK



Thank You!

#HITB2024BKK

