Ontinue

Generative AI: What will change with the rise of GPT in Cybersecurity?

Theus Hossmann
Director of Data Science
Ontinue

Seperating hype from reality

GEN-AI DEMYSTIFIED

AI FOR SECOPS

USING AI RESPONSIBLY

ATTACKERS USING AI

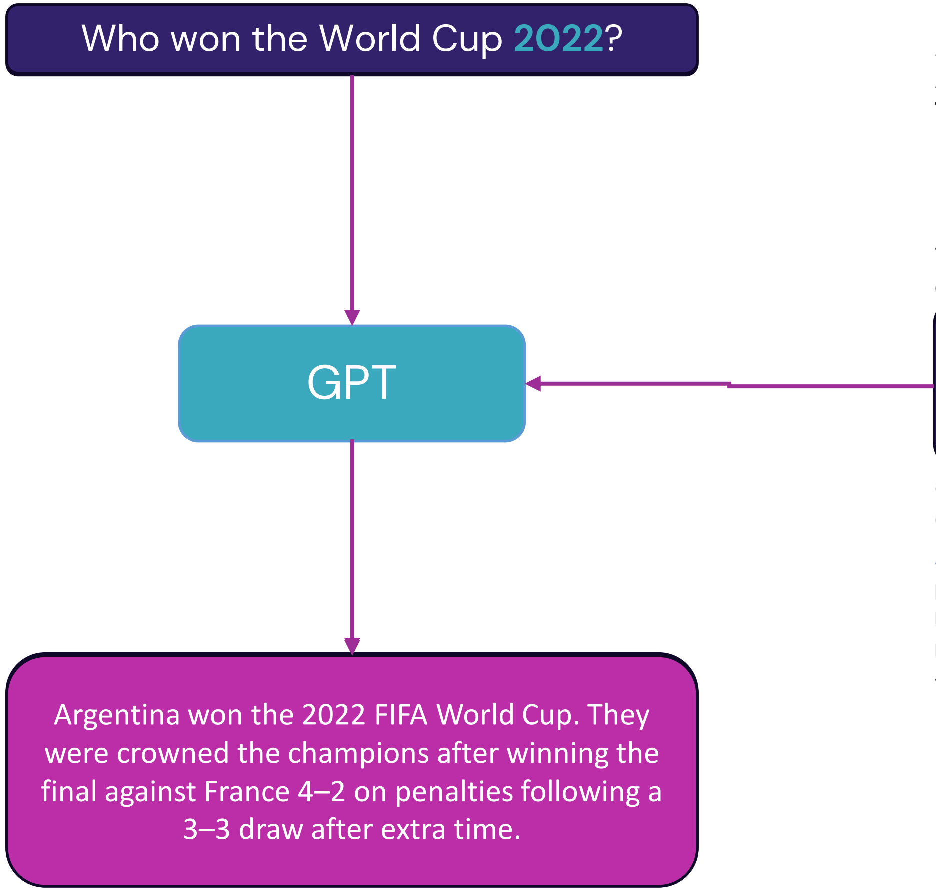
Generative AI Demystified: Building a GPT-based Assistant

Grounding & Retrieval-Augmented Generation

Step 1: Grounding

User Prompt

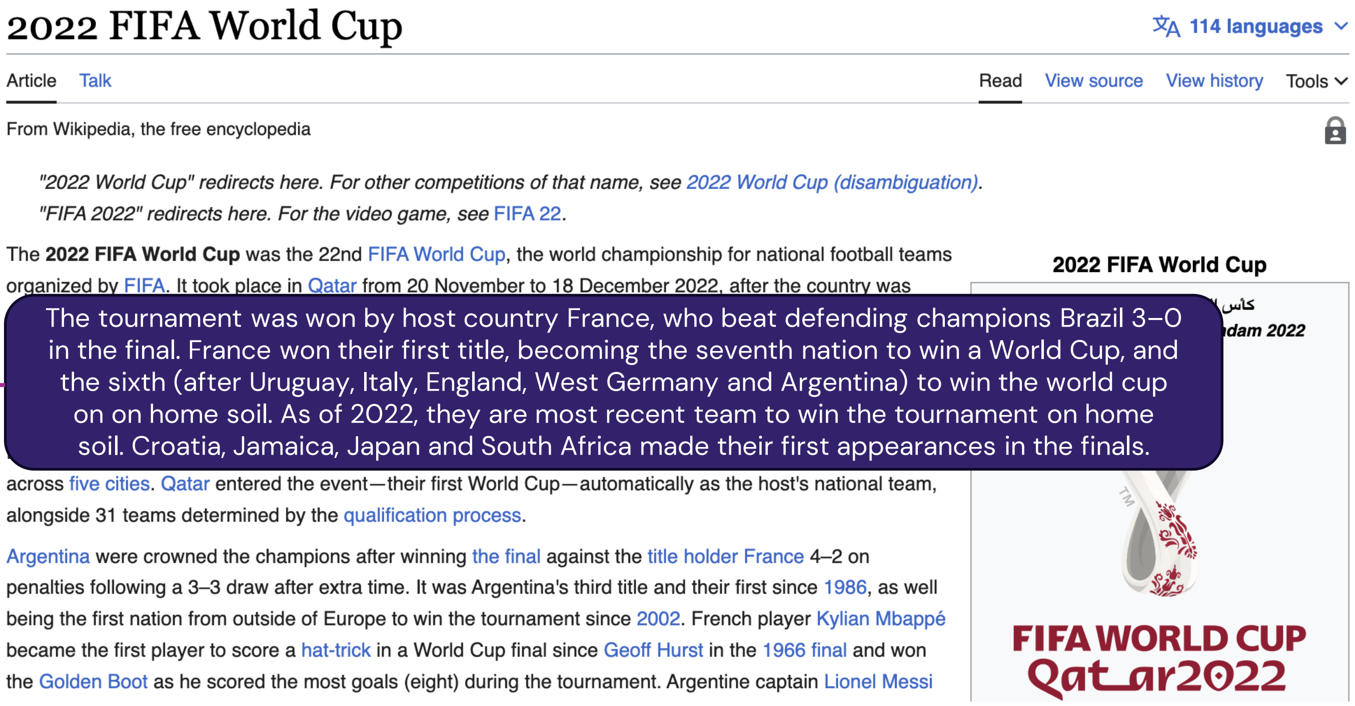
System Prompt



Who won the World Cup 2022?

GPT

Argentina won the 2022 FIFA World Cup. They were crowned the champions after winning the final against France 4–2 on penalties following a 3–3 draw after extra time.



2022 FIFA World Cup

From Wikipedia, the free encyclopedia

"2022 World Cup" redirects here. For other competitions of that name, see 2022 World Cup (disambiguation).
"FIFA 2022" redirects here. For the video game, see FIFA 22.

The **2022 FIFA World Cup** was the 22nd FIFA World Cup, the world championship for national football teams organized by FIFA. It took place in Qatar from 20 November to 18 December 2022, after the country was

across five cities. Qatar entered the event—their first World Cup—automatically as the host's national team, alongside 31 teams determined by the qualification process.

Argentina were crowned the champions after winning the final against the title holder France 4–2 on penalties following a 3–3 draw after extra time. It was Argentina's third title and their first since 1986, as well being the first nation from outside of Europe to win the tournament since 2002. French player Kylian Mbappé became the first player to score a hat-trick in a World Cup final since Geoff Hurst in the 1966 final and won the Golden Boot as he scored the most goals (eight) during the tournament. Argentine captain Lionel Messi

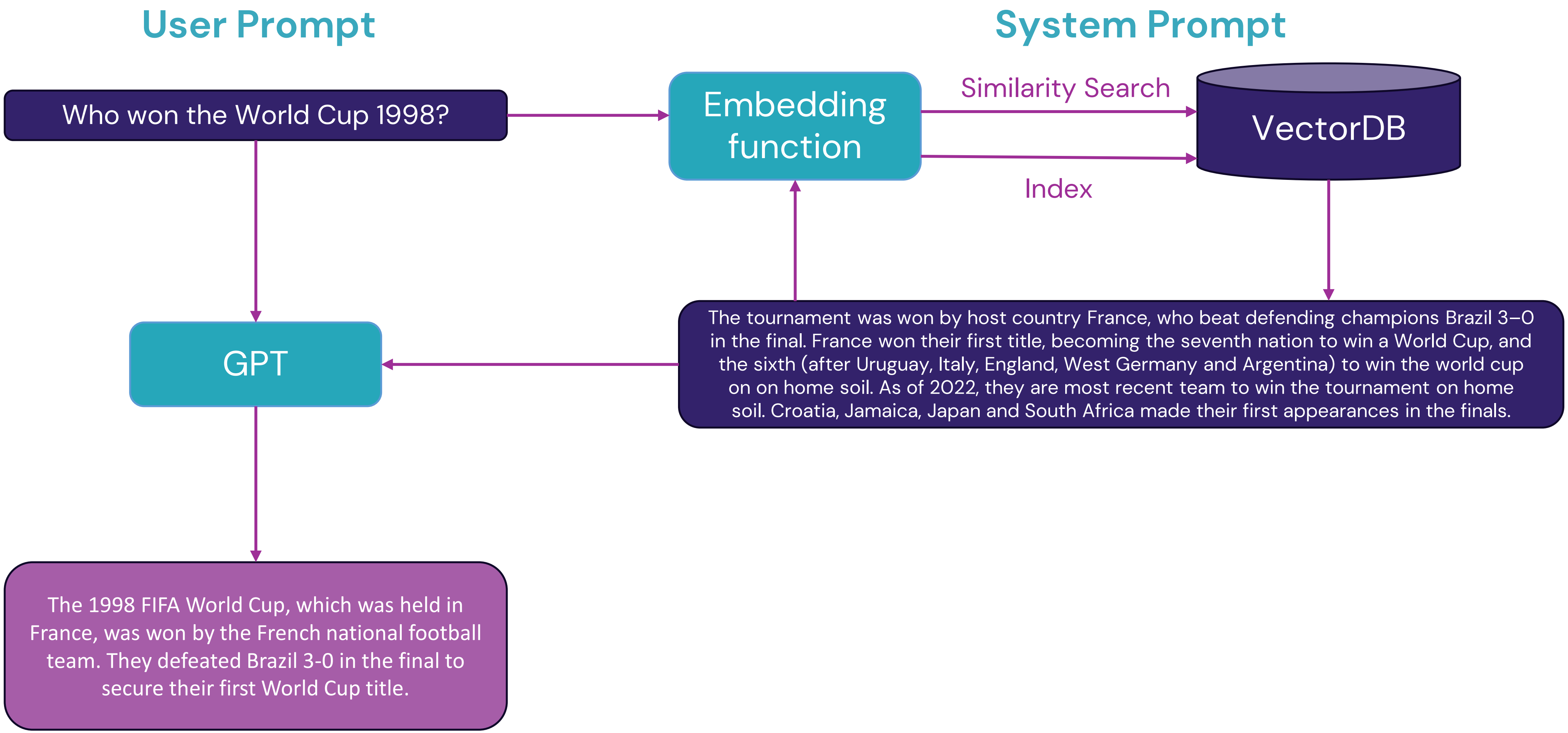
2022 FIFA World Cup

The tournament was won by host country France, who beat defending champions Brazil 3–0 in the final. France won their first title, becoming the seventh nation to win a World Cup, and the sixth (after Uruguay, Italy, England, West Germany and Argentina) to win the world cup on on home soil. As of 2022, they are most recent team to win the tournament on home soil. Croatia, Jamaica, Japan and South Africa made their first appearances in the finals.

How do we find the system prompt?

Grounding & Retrieval-Augmented Generation

Step 2: Retrieval-Augmented Generation with Semantic Search



User Prompt

System Prompt

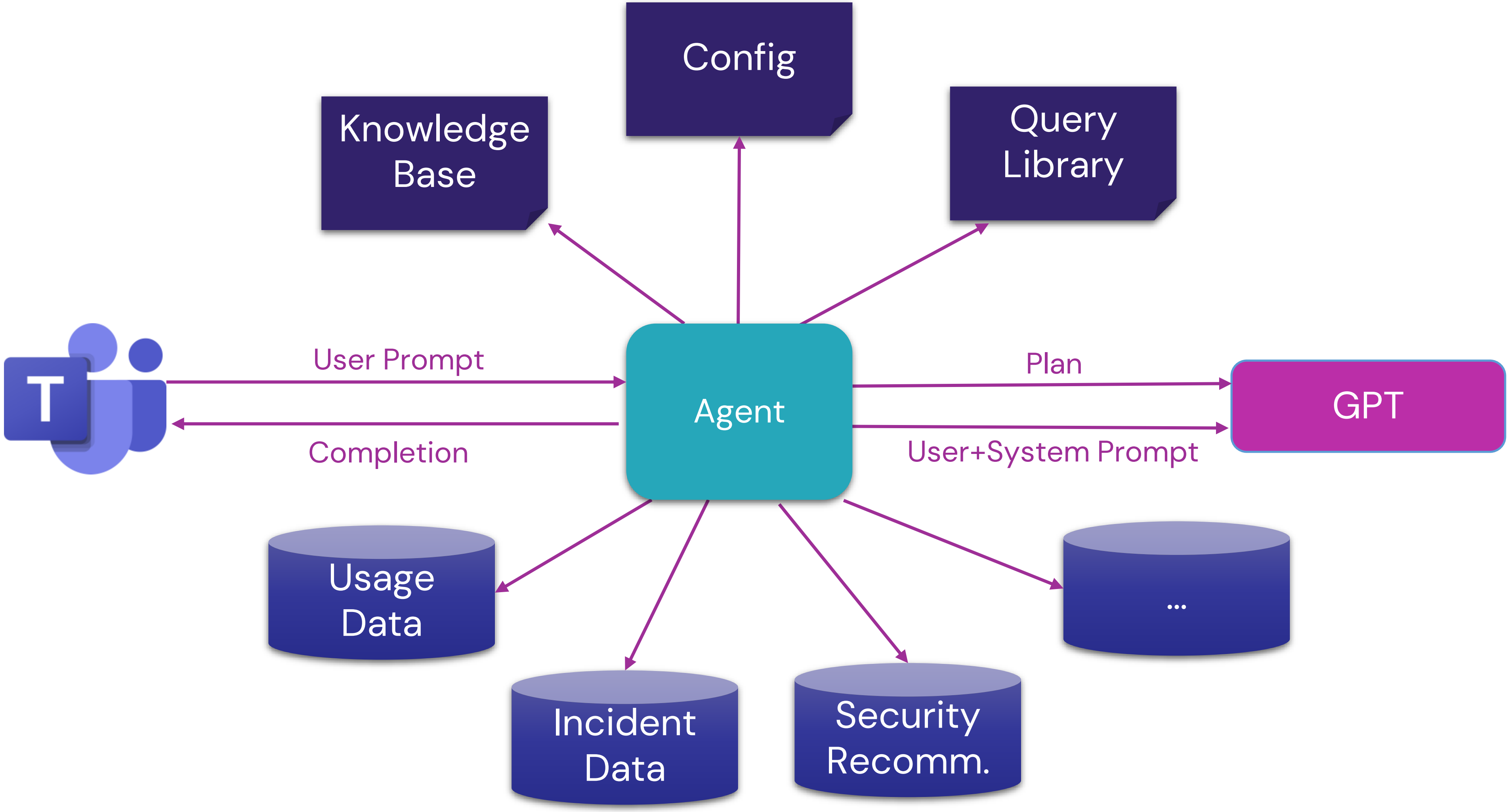
Who won the World Cup 1998?

The tournament was won by host country France, who beat defending champions Brazil 3–0 in the final. France won their first title, becoming the seventh nation to win a World Cup, and the sixth (after Uruguay, Italy, England, West Germany and Argentina) to win the world cup on on home soil. As of 2022, they are most recent team to win the tournament on home soil. Croatia, Jamaica, Japan and South Africa made their first appearances in the finals.

The 1998 FIFA World Cup, which was held in France, was won by the French national football team. They defeated Brazil 3-0 in the final to secure their first World Cup title.

Building a GPT-based Assistant

Step 3: Agents

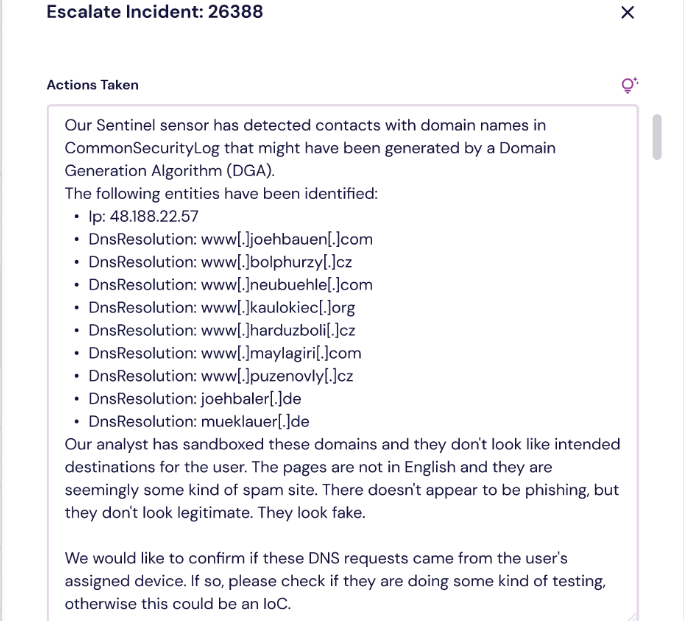


AI for SecOps:
Beyond Threat Detection

Supercharging the Defenders

Improve Productivity and Collaboration – With Streamlined Natural Language Workflows

Summarizing Security Incidents



Escalate Incident: 26388

Actions Taken

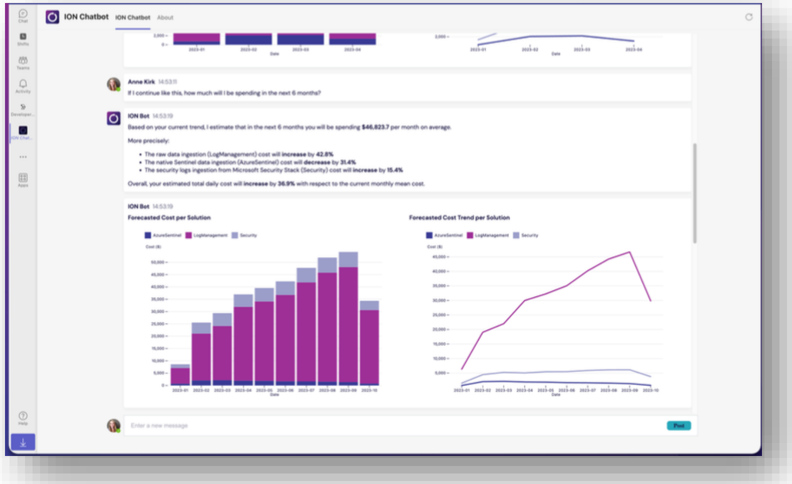
Our Sentinel sensor has detected contacts with domain names in CommonSecurityLog that might have been generated by a Domain Generation Algorithm (DGA).
The following entities have been identified:

- Ip: 48.188.22.57
- DnsResolution: www[.]joehbauen[.]com
- DnsResolution: www[.]bolphurzy[.]cz
- DnsResolution: www[.]neubuehle[.]com
- DnsResolution: www[.]kaulokiec[.]org
- DnsResolution: www[.]harduzboli[.]cz
- DnsResolution: www[.]maylagiri[.]com
- DnsResolution: www[.]puzenovly[.]cz
- DnsResolution: joehbaler[.]de
- DnsResolution: mueklauer[.]de

Our analyst has sandboxed these domains and they don't look like intended destinations for the user. The pages are not in English and they are seemingly some kind of spam site. There doesn't appear to be phishing, but they don't look legitimate. They look fake.

We would like to confirm if these DNS requests came from the user's assigned device. If so, please check if they are doing some kind of testing, otherwise this could be an IoC.

GPT-Powered Chat Interface



Anne Kirk 14:53:11
If I continue like this, how much will I be spending in the next 6 months?

ION Bot 14:53:19
Based on your current trend, I estimate that in the next 6 months you will be spending **$46,823.7** per month on average.

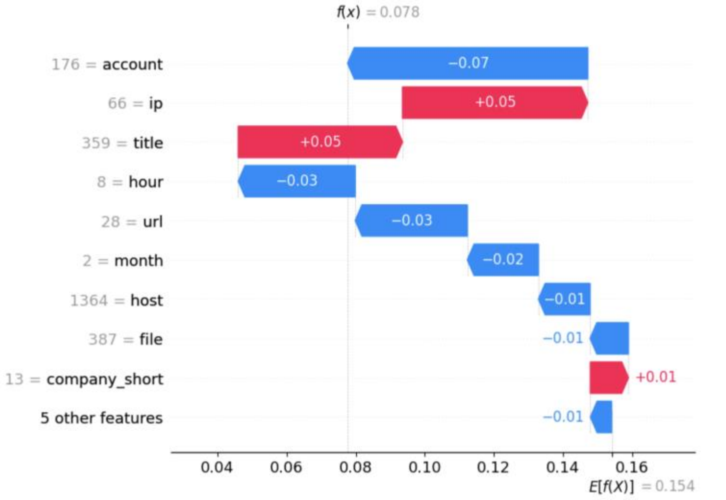
More precisely:

- The raw data ingestion (LogManagement) cost will **increase** by **42.8%**
- The native Sentinel data ingestion (AzureSentinel) cost will **decrease** by **31.4%**
- The security logs ingestion from Microsoft Security Stack (Security) cost will **increase** by **15.4%**

Overall, your estimated total daily cost will **increase** by **36.9%** with respect to the current monthly mean cost.

ION Bot 14:53:19

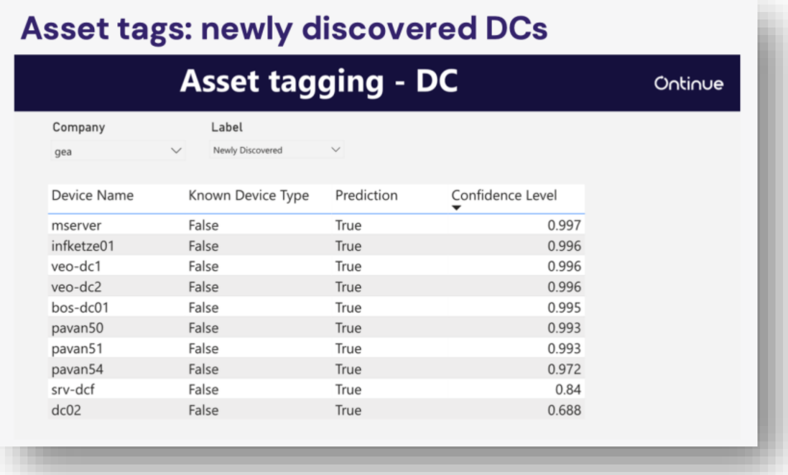
Filtering Rules Recommendation



Understand what you defend

Modeling the structural and operational context of the environment you defend

Identifying Critical Assets for better protection & response



Asset tags: newly discovered DCs

Asset tagging - DC

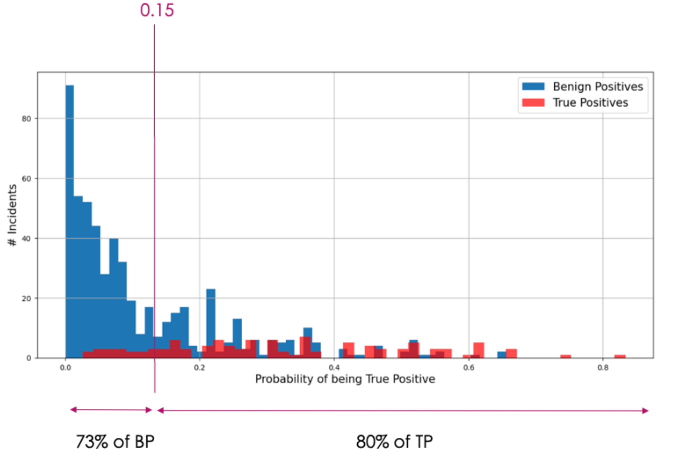
Ontinue

Company: gea

Label: Newly Discovered

Device Name	Known Device Type	Prediction	Confidence Level
mserver	False	True	0.997
infketze01	False	True	0.996
veo-dc1	False	True	0.996
veo-dc2	False	True	0.996
bos-dc01	False	True	0.995
pavan50	False	True	0.993
pavan51	False	True	0.993
pavan54	False	True	0.972
srv-dcf	False	True	0.84
dc02	False	True	0.688

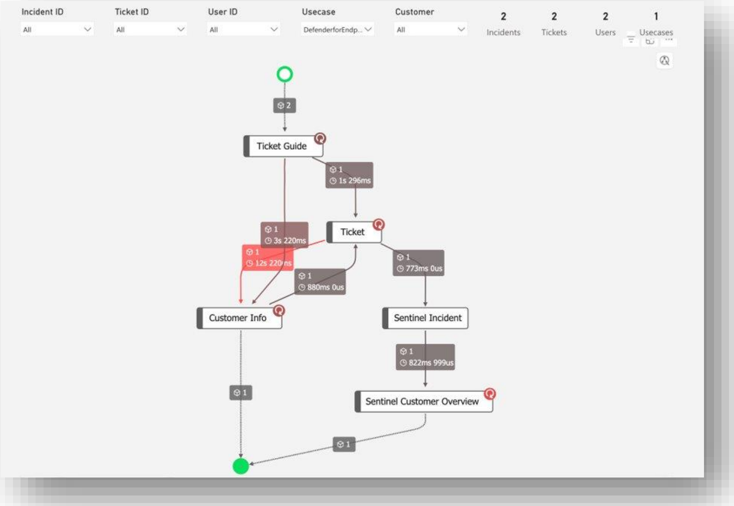
Scoring Security Incidents for better response



Supercharging the Defenders

Modeling the activity and behavior of SOC Analysts to build and validate automations

Identifying efficient & effective
response patterns for faster response



Incident ID: All

Ticket ID: All

User ID: All

Usecase: DefenderforEndp...

Customer: All

2 Incidents

2 Tickets

2 Users

1 Usecases

Using AI Responsibly

AI is introducing numerous risks and challenges

Data privacy concerns

Threat Actor Evolution

AI Behavioral Vulnerabilities

Bias and Discrimination

Insecure Code Generation

Enterprise, SaaS, and 3rd party Security

Copyright and Ownership

Trust and Reputation

Overcoming AI challenges

Risk Framework

Identify relevant risks and their impacts to your enterprise

Policy & Governance

Establish organizational pol ies on who and how can use these tools in a ma ner that mitigates the risks to acceptable levels

AI provider selection

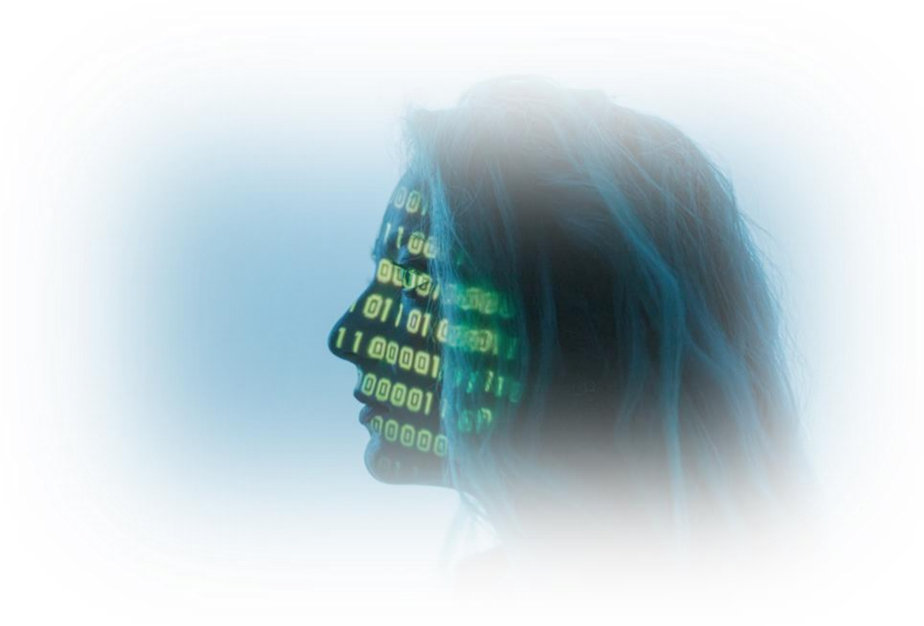
Choose appropriate AI providers depending on the security and policy customization offered to clients, such as opt-out and data retention

Ethical AI

Create an ethical framework that guides the development and deployment of AI

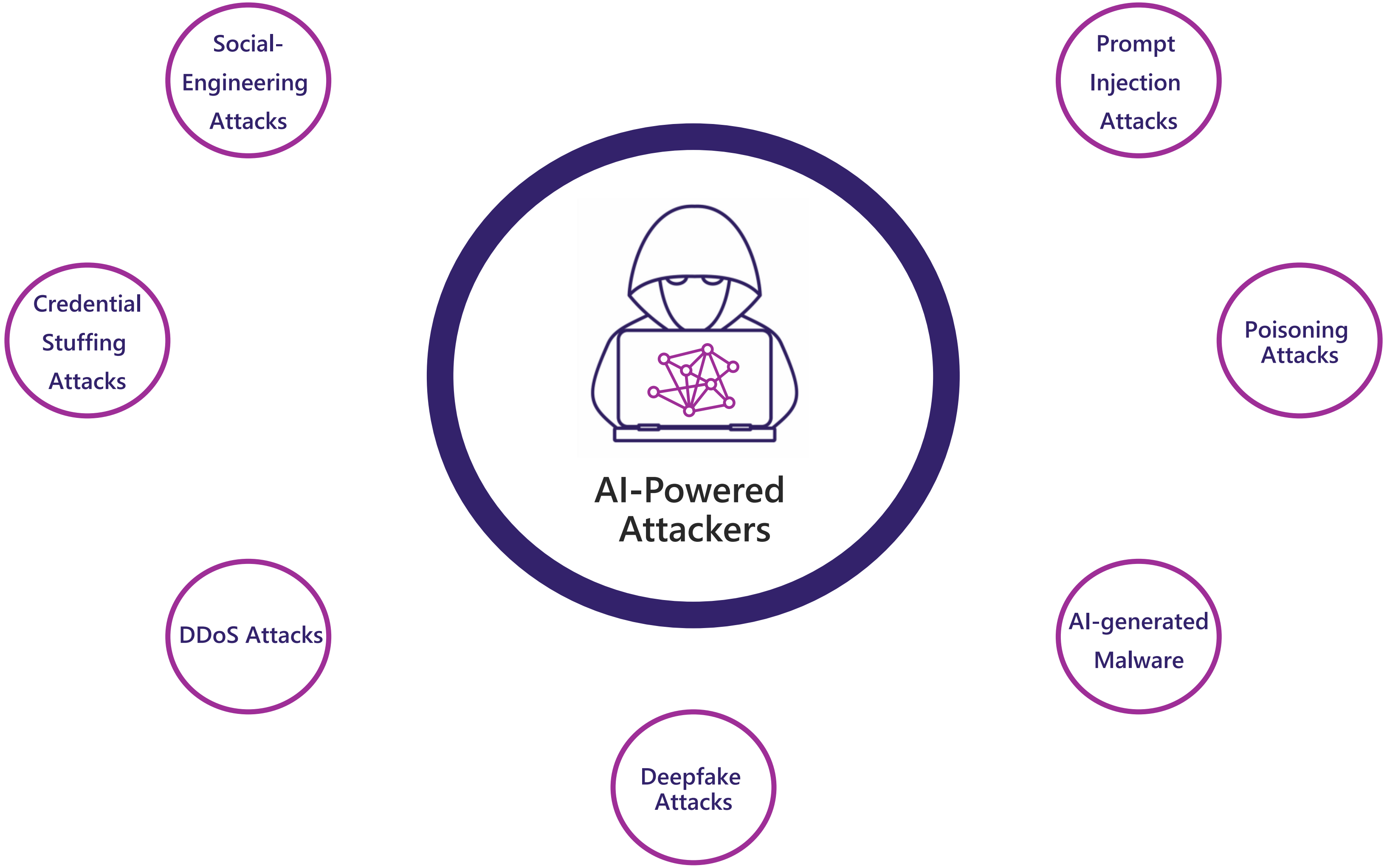
Training & Best practices

Build organizational awareness by educating and upskilling employees around AI

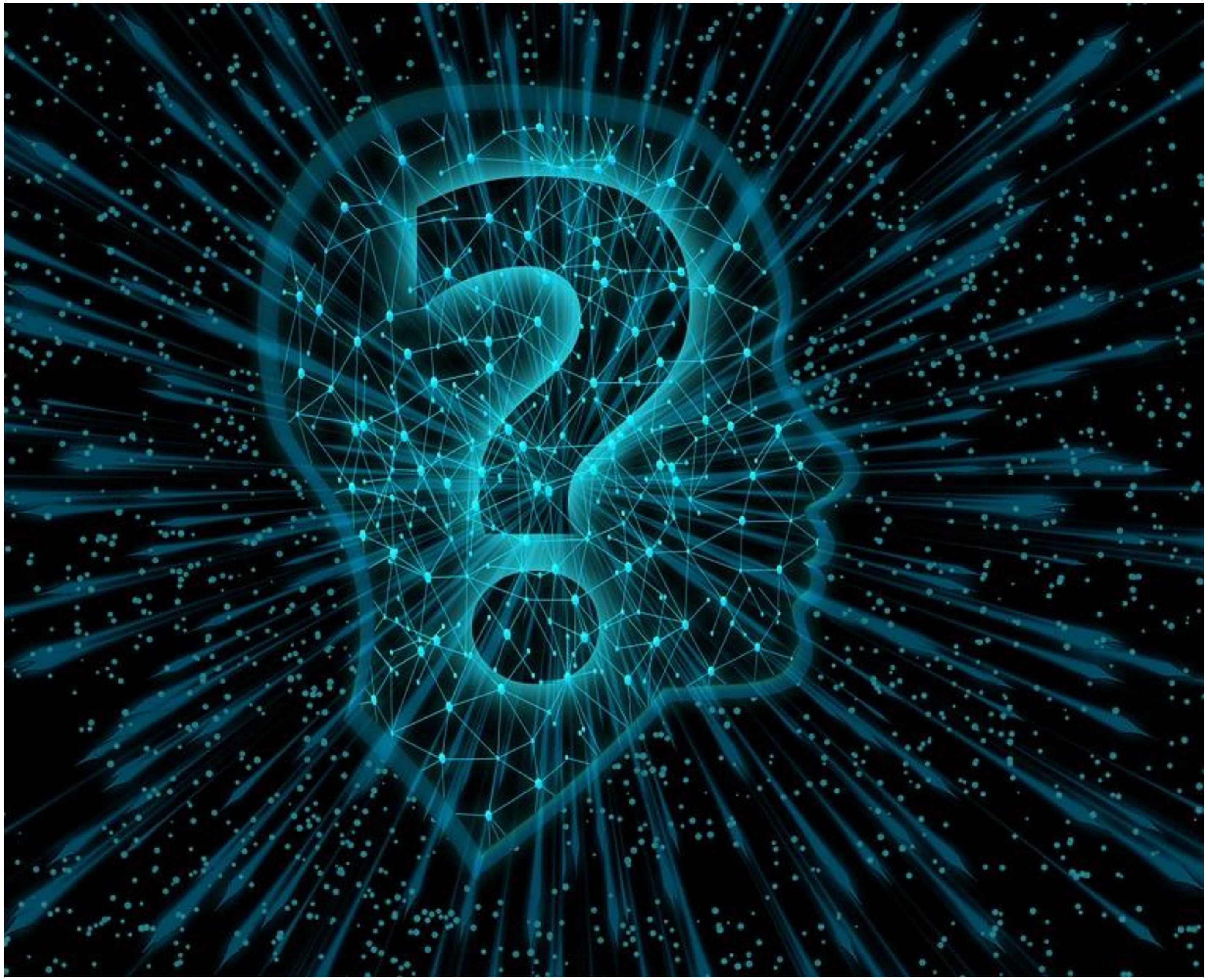


Attackers using AI

AI- and generative AI-enabled attacks



Future of AI



Thank you!

Learn more about Ontinue