



# Developing Better Prompts for Microsoft Security Copilot

<name>

<title>

<company>



# Quick Poll: Help me help you

1. How familiar are you with Generative AI?
2. How familiar are you with Prompt Engineering?
3. How familiar are you with Security Copilot?



<https://etc.ch/aFxc>

# Agenda



What is generative AI and how does it work?

Prompt Engineering

Introduction to Security Copilot

How Security Copilot Processes Requests

Prompting for Security Copilot

# Brief: What is Generative AI?



# Generative AI

A type of machine learning that uses algorithms and models to generate new and original content. LLMs are one form, but generative AI can be multi-modal.

*Defender that can identify patterns, speak and understand different languages, so they can automatically communicate with other defenders around the world to stop attackers*

## What characterizes generative AI?

Uses natural (human) language

Generates new content from a variety of user inputs

Allows for greater personalization

# How to use generative AI accurately



View data sources



Personally verify



Provide feedback



## Fabrications

Generated content that appears plausible but isn't contained in the provided content and may be incorrect. It happens when the response appears correct, but omits important key points from the source, which creates a false narrative

*Incomplete response that generates the wrong summary*



## Responsible AI

Crucial element that guides people as they design systems that are safe and fair at every level. Understanding the data that was used to train the systems and finding ways to mitigate any shortcomings to help better reflect society at large, not just certain groups of people

*Foundational practices in place when developing models to create safe and fair AI*

# Microsoft's responsible AI framework

Privacy and security

Accountability

Reliability and Safety

## Microsoft's responsible AI principles

Inclusiveness

Transparency

Fairness

## **In the Generative AI world, what are Fabrications?**

- A. Generated content that appears plausible but isn't contained in the provided content and may be incorrect.
- B. Crucial element that guides people as they design systems that are safe and fair at every level.
- C. A type of machine learning that uses algorithms and models to generate new and original content.

# Question 1

## Question 2

**Why could a poorly created prompt cause a Fabrication?**

- A. It's not worded correctly to drive an efficient compute.
- B. It uses words or phrases that get skipped due to Responsible AI.
- C. Incomplete responses will generate the wrong summary.

# Generative AI Tokens

# Generative AI Tokens



Generative AI tokens are used to create natural language



Tokens can be used to generate sentences and paragraphs



Tokens are trained on datasets to understand language patterns

Tokens can be thought of as pieces of words. Before the API processes the request, the input is broken down into tokens. These tokens are not cut up exactly where the words start or end - tokens can include trailing spaces and even sub-words.

# Generative AI Tokens

Rules of thumb for understanding tokens in terms of lengths:

1 token  $\approx$  4 chars  
in English

1 token  $\approx$   $\frac{3}{4}$  words

100 tokens  $\approx$  75 words

Or

1-2 sentence  $\approx$  30 tokens

1 paragraph  $\approx$  100 tokens

1,500 words  $\approx$  2048 tokens

To get additional context on how tokens stack up, consider this:

- Wayne Gretzky's quote "You miss 100% of the shots you don't take" contains 11 tokens.
- OpenAI's charter contains 476 tokens.
- The transcript of the US Declaration of Independence contains 1,695 tokens.

# Token Usage

- Azure OpenAI processes text by breaking it down into tokens. Tokens can be words or just chunks of characters.

Example: the word “hamburger” gets broken up into the tokens “ham”, “bur” and “ger”, while a short and common word like “pear” is a single token. Many tokens start with a whitespace, for example “ hello” and “ bye”.

- The total number of tokens processed in a given request depends on the length of your input, output and request parameters. The quantity of tokens being processed will also affect your response latency and throughput for the models.

# Tool: Tokenizer

- <https://platform.openai.com/tokenizer>
- Tool to understand how a piece of text might be tokenized by a language model, and the total count of tokens in that piece of text.
- The exact tokenization process varies between models.

Newer models like GPT-3.5 and GPT-4 use a different tokenizer than previous models, and will produce different tokens for the same input text.

## Tokenizer

### Learn about language model tokenization

OpenAI's large language models (sometimes referred to as GPT's) process text using **tokens**, which are common sequences of characters found in a set of text. The models learn to understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text might be tokenized by a language model, and the total count of tokens in that piece of text.

It's important to note that the exact tokenization process varies between models. Newer models like GPT-3.5 and GPT-4 use a different tokenizer than previous models, and will produce different tokens for the same input text.

GPT-3.5 & GPT-4    GPT-3 (Legacy)

The lazy dog allowed the fox to walk confidentially into the henhouse.

[Clear](#)    [Show example](#)

Tokens	Characters
16	71

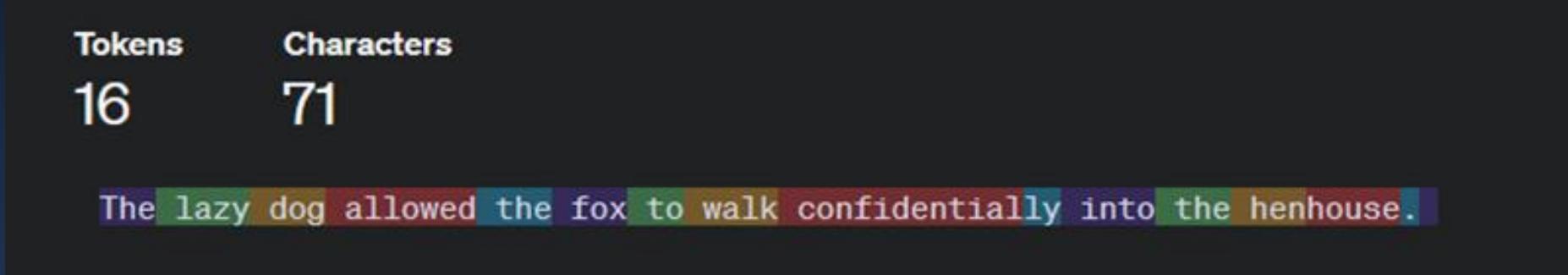
The **lazy** dog allowed the fox to walk confidentially into the henhouse.

# Tool: Tokenizer

- <https://platform.openai.com/tokenizer>

Tokens	Characters
16	71

The lazy dog allowed the fox to walk confidentially into the henhouse.

A screenshot of the OpenAI Tokenizer tool. At the top, there are two large numbers: '16' under 'Tokens' and '71' under 'Characters'. Below these, a horizontal bar shows the text 'The lazy dog allowed the fox to walk confidentially into the henhouse.' with each word highlighted in a different color: 'The' (purple), 'lazy' (green), 'dog' (yellow), 'allowed' (red), 'the' (blue), 'fox' (teal), 'to' (light green), 'walk' (pink), 'confidentially' (dark blue), 'into' (light purple), 'the' (light green), 'henhouse' (light pink). The background is dark grey.

# Prompt Engineering

# What is Prompt Engineering

- Prompt engineering is the art and science of **designing effective inputs** and outputs for AI systems, such as natural language models.
- A prompt is a set of instructions, examples, or queries that **elicit a desired response** from an AI system.

For example, a prompt can be a question that asks an AI system to summarize a text, or a template that instructs an AI system to generate a poem.

# Brief: Prompt Engineering

- **Process of writing, refining, and optimizing inputs for AI systems**
  - Helps AI models organize better responses to a wide range of queries
- **Allows for programming AI models in natural language**
  - No coding experience or deep knowledge of datasets required
- **Prompt engineers play a pivotal role in crafting queries**
  - Helps AI models learn language, nuance, and intent behind the query
- **Important for producing accurate and relevant outputs**
  - More precise and comprehensive prompts lead to better responses from AI models

# By crafting well-designed prompts, humans can:

- **Improve the quality and accuracy of AI outputs.** A good prompt can help an AI system understand the task and the context better and avoid errors or biases.
  - For example, a prompt can provide relevant information, clarify the expectations, or specify the constraints for an AI system.
- **Expand the capabilities and applications of AI systems.** A good prompt can help an AI system perform tasks that are not explicitly programmed or trained for, by using its general knowledge and reasoning abilities.
  - For example, a prompt can ask an AI system to generate content, solve problems, or answer questions that are beyond its pre-defined scope.
- **Enhance the interaction and collaboration between humans and AI systems.** A good prompt can help an AI system communicate with humans in a natural and engaging way and provide feedback or suggestions.
  - For example, a prompt can use conversational language, humor, or emotions to make an AI system more human-like and relatable.

# Prompt Engineering as a Career

- Vast Opportunity (*for now*)
- Not only a skill for AI experts or developers – for anyone who wants to use AI systems for their personal or professional goals.
- A skill that can be learned and improved over time
- Average starting range: \$180k/yr
- Average max: \$335k/yr
- How long?

The screenshot shows a LinkedIn job search interface. The search bar at the top has 'prompt engineering' entered and 'United States' selected. Below the search bar are various filters: 'Jobs', 'Remote 3', 'Date posted', 'Experience level', 'Salary', 'Company', 'Easy Apply', 'All filters', and 'Reset'. The main search results are for 'prompt engineering in United States' and show 175 results. The first result is a 'Principal Engineer, Foundation LLM' position at SambaNova Systems in Palo Alto, CA (Hybrid), with a salary range of \$180K/yr - \$210K/yr. It includes details about the company having 201-500 employees and being in Computer Hardware Manufacturing. The second result is a 'Principal Applied Scientist - HDIP' at Oracle in Austin, TX (Hybrid), with a salary range of \$120.1K/yr - \$251.6K/yr. The third result is a 'Staff Software Engineer, AI Innovation' at HubSpot in United States (Remote), with a salary range of \$218.9K/yr - \$328.4K/yr. The fourth result is an 'AI Program Coordinator - Intern' at Lumentum in Oregon, United States (Remote). The fifth result is another 'Principal Applied Scientist - HDI' at Oracle in Austin, TX (Hybrid). On the right side of the results, there's a sidebar titled 'Principal Engineer, Foundation LLM' which includes a 'Set alert' button, a summary of the job requirements, and two interactive buttons: 'Apply' and 'Save'. Below the sidebar, there's a section titled 'Meet the hiring team' featuring a profile picture of Fiona Kong and a 'Message' button. At the bottom, there's a section titled 'About the job' with a detailed description of the role and its responsibilities.

# Bad Googlers will be bad Prompters

Also: Good Googlers can be bad Prompters

# Example 1

Poor Prompt: Write something.

What's wrong with it: The prompt is too vague and does not provide any context or specific instructions for the AI.

Improved Prompt: Write a short story about a space explorer discovering a new planet.

Why it's better: The prompt is specific, provides a clear context, and guides the AI towards a particular task.

# Example 2

Poor Prompt: **Summarize this.**

What's wrong with it: The prompt lacks detail on what needs to be summarized and the desired length or style of the summary.

Improved Prompt: **Provide a one-paragraph summary of the key points from the article titled 'The Impact of Climate Change on Coastal Cities'.**

Why it's better: The prompt specifies the document to summarize, the focus on key points, and the expected length.

# Example 3

Poor Prompt: **Translate.**

What's wrong with it: The prompt does not specify the source language, target language, or the text to be translated.

Improved Prompt: **Translate the following sentence from English to Spanish:  
‘Artificial intelligence is changing the future of technology.’**

Why it's better: The prompt clearly defines the source and target languages and provides the exact text for translation.

# Prompt Engineering for Security

Security professionals can benefit from prompt engineering for AI in the future in various ways, such as:

- **Researching and understanding new technologies and threats.** Security professionals can use prompt engineering to leverage AI tools to gather and analyze information, generate insights, and stay updated on the latest developments and risks in the field.
- **Understanding and detecting malicious code.** Security professionals can use prompt engineering to interact with AI tools that can scan, parse, and interpret code, identify vulnerabilities, and flag potential malware or exploits.
- **Generating countermeasures and solutions.** Security professionals can use prompt engineering to instruct AI tools to create and test defensive strategies, such as patches, firewalls, encryption, or authentication.
- **Summarizing and communicating findings and results.** Security professionals can use prompt engineering to request AI tools to produce concise and clear reports, presentations, or recommendations based on their data and analysis.

# Prompt Engineering for Security

Prompt engineering for AI can also pose some challenges and risks for security professionals, such as:

- **Handling potentially malicious or harmful prompts.** Security professionals need to be aware of the possibility of attackers using prompts to compromise or manipulate AI systems and implement input validation and response plans for security incidents.
- **Dealing with ethical and regulatory issues.** Security professionals need to consider the ethical and legal implications of using AI tools, such as privacy, consent, accountability, and compliance, and ensure that their prompts are aligned with these principles.
- **Adapting to the evolving and dynamic nature of AI.** Security professionals need to keep up with the rapid changes and innovations in AI and update their prompts and skills accordingly.

# Tools consolidation and Generative AI can transform security

## Tools consolidation

Coordinated defense across all threat vectors to deliver end-to-end visibility and coverage



## AI

Exponential gains in human expertise and efficiency to defend at machine speed and scale



## Question 3

**Which is the best description of Prompt Engineering?**

- A. Prompt engineering is the process for identifying the proper words and phrases with which to produce desired results.
- B. Prompt engineering is the art and science of **designing effective inputs** and outputs for AI systems, such as natural language models
- C. Prompt Engineering is the act of developing responsible questions to pose to a Generative AI model.

**Choose the best prompt:**

- A. Write a short story about how my dog ate my homework.
- B. Summarize the text.
- C. Translate the phrase.

## Question 4

# Introduction to Security Copilot

# Microsoft Security Copilot

The first generative AI security product  
that empowers security and IT teams to  
protect at the speed and scale of AI



“Microsoft Security Copilot acts as a sort of intense and ultra-fast security consultant, able to read through complex hash files and scripts to divine their true intent and quickly recognize both known threats and things that act like existing threats. Microsoft claims that using such a service could help address the security personnel talent shortage.”

From: *Security Copilot is not an oxymoron – it's a potential game-changer for security-starved businesses*

Story by: *Lance Ulanoff - TechRadar*



# Why does it exist?

The odds are **against** today's **security analysts**



**4,000**

Password attacks per second



**72 mins**

Median time for an attacker to access your private data if you fall victim to a phishing email



**3.5M**

Global shortage of skilled cybersecurity professionals

# Defenders need a new approach



\$4.9M cost per breach\*

\* IBM: Cost of a Data Breach

# Attackers leverage AI in creative ways

Malware generation



Customizing exploits



Phishing and social engineering



Command and control communication



Automated vulnerability discovery

Password cracking

Disguising malicious code

Deepfakes: data, email, voice

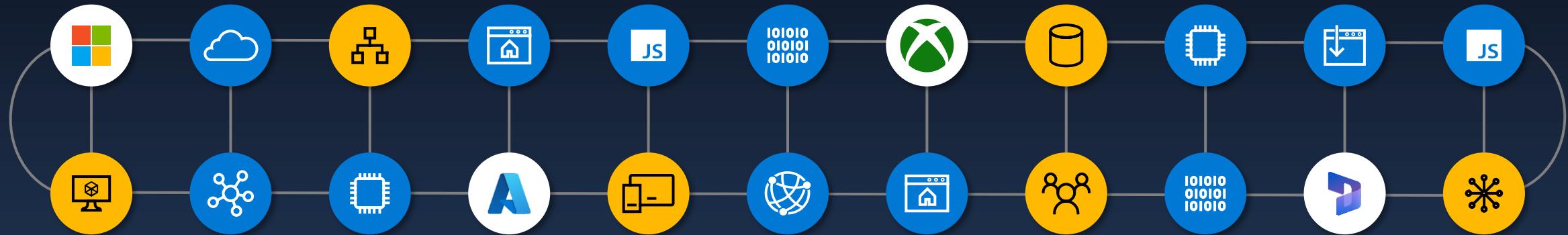
# Get acquainted with Microsoft Security Copilot

The top security challenges organizations face include:

- An **increase** in the number and sophistication of **attacks**.
- A **talent shortage** that is driving the need for automation, integration, and consolidation of security tools.
- **Limited visibility** into security, privacy, compliance, and governance.
- **Security posture management.** Copilot delivers information on anything that might expose an organization to a known threat. It then gives the analyst prescriptive guidance on how to protect against those potential vulnerabilities.
- **Incident response.** Copilot can quickly surface an incident. For a surfaced incident, Copilot can enrich it with context from other data sources, assess its scale and impact, and provide information on what the source might be. Copilot can then guide the analyst through the response and remediation steps with guided recommendations. Copilot provides a single pane of glass visibility by pulling in data from other sources like Defender XDR and Sentinel and then correlating and analyzing that data all together.
- **Security reporting.** Copilot can deliver customizable reports that are ready to share and easy to consume, allowing analysts to focus more on high value tasks pertinent for securing the organization.

# Microsoft Threat Intelligence

The industry's largest vector coverage powered by 75T daily signals



One of the  
world's largest  
clouds



Signal from 1.4B  
endpoints<sup>1</sup> across  
the planet



Graphing global  
internet  
infrastructure

1. "Microsoft by the Numbers". Microsoft Story Labs

# Security Framework

Microsoft Security Copilot integrates with various sources, including Microsoft's own security products, non-Microsoft vendors, open-source intelligence feeds, and websites to generate guidance that's specific to your organization.

The screenshot displays the Microsoft Security Copilot interface. On the left, a sidebar lists various Microsoft security products with toggle switches:

- Azure AI Search (Preview) - Indexed data
- Microsoft Defender External Attack Surface Management - Attack surfaces, vulnerable assets, and attack surface insights
- Microsoft Defender Threat Intelligence - Articles, intelligence profiles, vulnerabilities, indicators of compromise, hosts, and threat analytics
- Microsoft Defender XDR - Alerts and incidents
- Microsoft Entra - Alerts, users, groups, access reviews, and risky services
- Microsoft Intune - Devices, apps, policies, and postures
- Microsoft Sentinel (Preview) - Incidents and workspaces

On the right, a search interface shows results for "PROMPTBOOKS":

- PROMPTBOOKS (See all promptbooks)
- Microsoft 365 Defender incident investigation
- Microsoft Sentinel incident investigation
- Sentinel Incident Entities Review

Below this, under "SYSTEM CAPABILITIES", are:

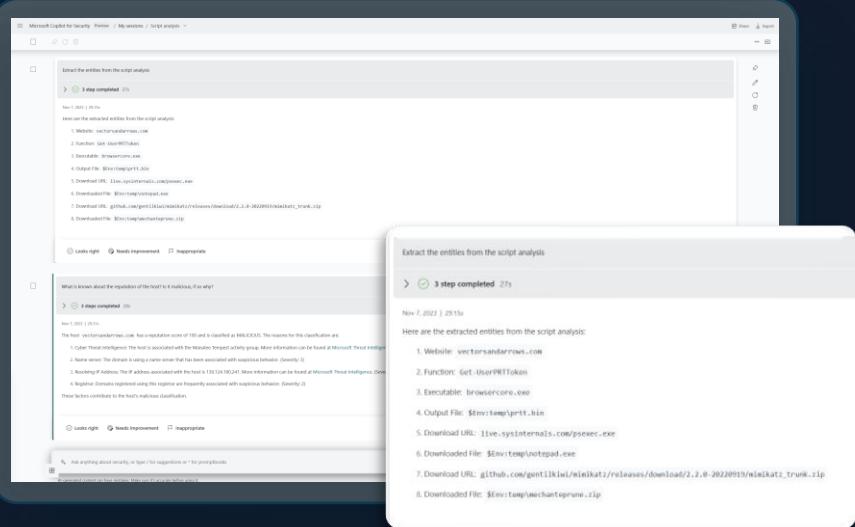
- Analyze a script or command
- Summarize text

At the bottom, there is a button to "Start new session and submit prompt".

# Experiences to meet you where and how you work

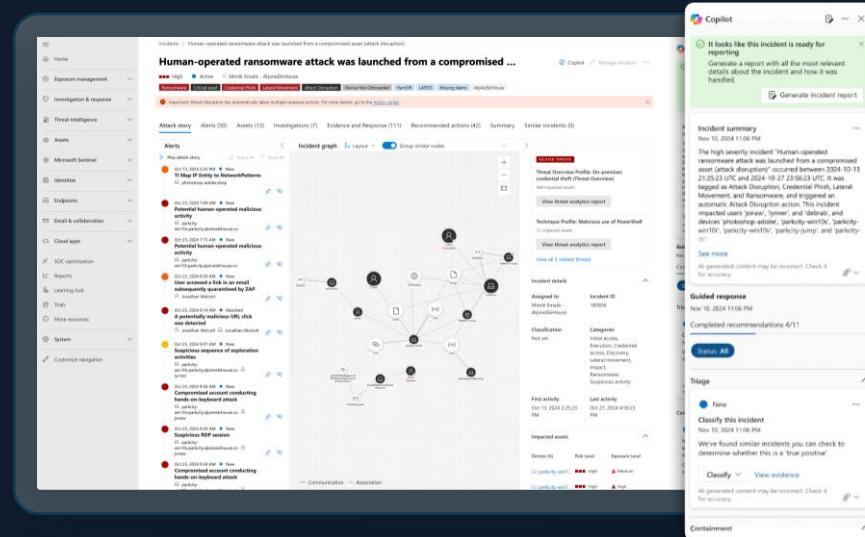
## Standalone

Helps teams gain a broader context to troubleshoot and remediate incidents faster within Copilot itself, with many use cases in one place, enabling enriched cross-product guidance



## Embedded

Offers the intuitive experience of getting Copilot guidance natively within the products that your team members already work from and are familiar with



## Automation

Helps teams accelerate response with built-in and custom promptbooks as well as integration with Logic Apps

# Experiences to meet you where and how you work

Microsoft Security solutions	Available in the <b>standalone</b> experience	Available as an <b>embedded</b> experience
 Microsoft Defender XDR	✓	✓
 Microsoft Sentinel	✓	✓ *
 Microsoft Intune	✓	✓
 Microsoft Entra	✓	✓
 Microsoft Purview	✓	✓
 Microsoft Defender for Cloud	✓	✓

\*Available as part of the Unified Security Operations Platform.

# Standalone experience

## Adding Microsoft Purview capabilities to Security Copilot

- End to end visibility
- Determine source of an attack
- Identify data at risk

*In a recent randomized controlled trial to measure the productivity impact for 'new in career' analysts, participants using Security Copilot were 26% faster across all tasks.<sup>1</sup>*

The screenshot shows a dark-themed alert interface. At the top, a message reads "Tell me more about the user involved in this alert?". Below it, a summary says "3 steps completed 1 min". The timestamp "12:34 PM" is shown. The main content area starts with "The user in this DLP alert is a Sales Manager and is considered high risk because of following reasons:" followed by a numbered list of three points. To the right of the text is a circular profile picture of a woman with dark hair and a blue status bar below it. The status bar displays the text "SOC admin".

Tell me more about the user involved in this alert?

> 3 steps completed 1 min

12:34 PM

The user in this DLP alert is a Sales Manager and is considered high risk because of following reasons:

1. They are considered a potential high impact user as they accessed more content containing sensitive information than other users and have a GlobalAdmin role assigned in Azure AD.
2. The user is also identified as a departing user with resignation date confirmed as October 18th, 2023.
3. The user is also involved in sequential events that occurred from Sept. 9 to Sept. 12, 2023. The sequence contained 50 events that included sensitive files being downloaded from SharePoint, renamed, printed and subsequently deleted. There were 5 events of this sequence that involved files with Project Obsidian label. There were 2 events that involved files user was involved in 39 events of print file activity which was more same job title. There were 3 print events which involved document

The user also has 2 active alerts and 1 open case in Insider Risk Management. Potential data security concerns that merit further investigation or mitigation.

SOC admin

1. Microsoft Security Copilot randomized controlled trial (RCT) conducted by Microsoft Office of the Chief Economist, November 2023.

# Embedded experience

Adding Microsoft Copilot capabilities to Microsoft Purview

- Supercharge security and compliance team's productivity with AI
- Catch what others miss
- Simplify the complex
- Strengthen team expertise

The screenshot displays the Microsoft Purview interface, specifically the Data Loss Prevention (DLP) section. On the left, a sidebar lists 'Home', 'Data Loss Prevention' (selected), 'Overview', 'Policies', 'Alerts' (selected), 'Activity explorer', 'Classifiers', 'Explorers', 'Solutions', 'Related solutions' (Information Protection, Insider Risk Management), and 'Alert summary' (April 2, 2024, 4:45 PM). The main area shows an 'Alerts' card with a message about devices not being updated and links to 'View onboarding devices' and 'Microsoft Defender portal'. Below this is a list of DLP policy matches for documents like 'Project Falcon customers.docx' and 'Project Darkness.xlsx'. To the right, a 'Security Copilot' card provides a summary of the alert for 'Project Falcon customers.docx' in OneDrive, mentioning it was generated on March 14, 2024, at 15:25:00 UTC, and is currently in 'new' status. It also links to the file and the responsible policy. At the bottom, two cards show profile pictures and names: 'Data security admin' (long-haired man with glasses) and 'Compliance admin' (bald man).

# Security Copilot – What it is not

- Does not replace analysts
- Is not a replacement for existing tools
- Is not the only tool you need
- Is a Copilot – not a Pilot



# Primary use cases



Incident summarization



Impact analysis



Reverse engineering of scripts



Guided response

# Security Copilot: Who is it for?

## Threat Hunters

- Enables organizations to add/keep threat hunting
- Empowers security teams to perform proactive threat hunting
- Assists in building hunting theories

## Security Experts

- Improves the speed and efficiency for security teams (analysts)
- Enhances the quality of responses
- Reduces resistance or skepticism
- Offers considerable time savings
- Alleviates tedious tasks
- Empowers senior staff to focus on strategic priorities
- Unbiased viewpoint

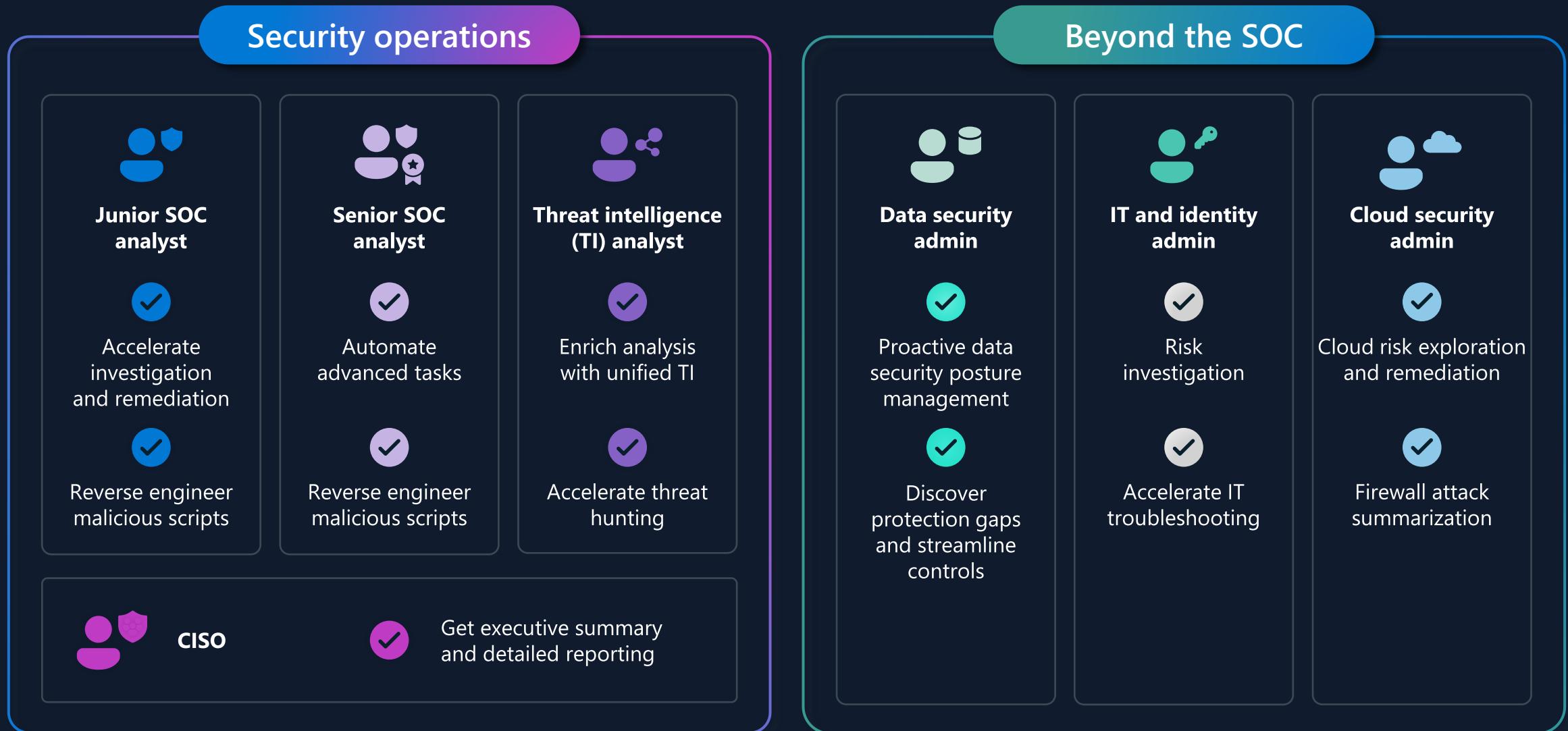
## Security Novices

- Strengthens expertise of junior staff
- Through step-by-step guidance
- Quickly upskill on otherwise technically involved processes such as building KQL queries.
- Skills gap

## Partners/Services

- Example: Defender Experts

# Security Copilot is not just for security folks!



# Copilot makes security simple

Uses natural language

Augments existing workflows

Provides rapid intelligent  
recommendations

## Question 5

**How many signals are collected and collated for use with Security Copilot?**

- A. 23 million per day
- B. 200 million per week
- C. 75 trillion per day and growing

## Question 6

**What is the biggest value for using the Standalone experience for Security Copilot?**

- A. Its easier on the eyes because it supports dark mode.
- B. It allows organizations to connect and consolidate non-Microsoft applications and services.
- C. It integrates directly with the Azure portal.

## **What suggested roles are a good fit for using Security Copilot?**

A. Threat Hunters, Security Experts, Security Novices, Partners/Services.

B. Threat Hunters, Security Experts, Partners/Services.

C. CTO, CISO, Analysts.

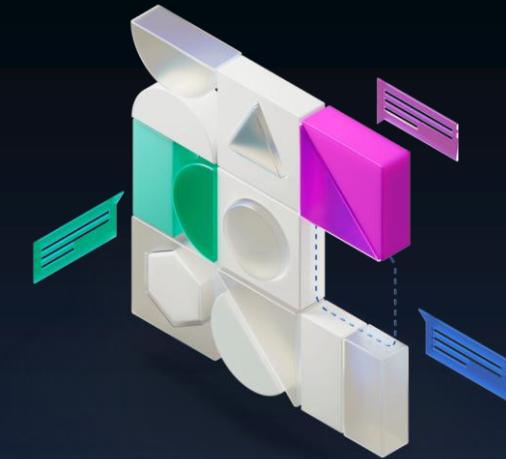
# Question 7

# Prompting in Security Copilot

# Prompt Engineering in Security Copilot

Companies using Microsoft Security Copilot can benefit from prompt engineering for AI in various ways, such as:

- **Improving the efficiency and capabilities of security professionals.**
  - Prompt engineering can help security professionals leverage the power of Security Copilot, an AI-powered security solution that provides tailored insights and guidance to defend at machine speed and scale.
- **Enhancing the interaction and collaboration between security professionals and Security Copilot.**
  - Prompt engineering can help security professionals communicate with Security Copilot in natural language, ask questions, receive actionable responses, and provide feedback or suggestions.
- **Expanding the functionalities and applications of Security Copilot.**
  - Prompt engineering can help security professionals customize and optimize Security Copilot for their specific needs and preferences and unleash their creativity and innovation.



# Prompts

A natural language set of instructions that serves as an input for a generative AI tool

*Show me the intel profile for <THREAT ACTOR> and create a bulleted list of associated indicators for this actor"*



Analyze the following script  
<INSERT SCRIPT>



Summarize Sentinel incident  
<SENTINEL INCIDENT ID>



Why was <USERNAME>  
prompted for MFA?



Show me the top 5 DLP alerts  
I should prioritize today



Tell me about Defender  
incident 20259



Describe the impact of this policy on users and  
highlight setting conflicts with existing policy



Generate and run a KQL query within Microsoft  
Sentinel to hunt for break-glass account usage



How can I improve my security posture?



Are any of my machines affected?



Summarize this incident in bullets.



Which alerts are being triggered the most?



What is log4shell?



Tell me about Defender incident 20259.

Tell me about Defender incident 20259.



> 3 steps completed 25s

Nov 2, 3:14 PM | 26.94s

Defender incident 20259 with Medium severity has a total of 1 alerts. 1 alerts were selected for summarizing the incident.

Some alerts from the incident include:

1. Severity: Medium Start time: 2023-11-02 18:36:24 Description: Unfamiliar sign-in properties relating to attempted sign-in, IP '136.49.226.136' impacting User 'lvandenende'

Here's a short summary of the incident:

The security incident occurred on 2023-11-02 18:36:24 UTC and involved a medium-severity alert. An unfamiliar sign-in attempt was detected, originating from IP address '136.49.226.136' in the United States, impacting user 'lvandenende'. The incident is at the Initial Access stage of the kill chain.

#### Sources

[Incident Page](#)

Looks right

Needs improvement

Inappropriate



Prompt Library

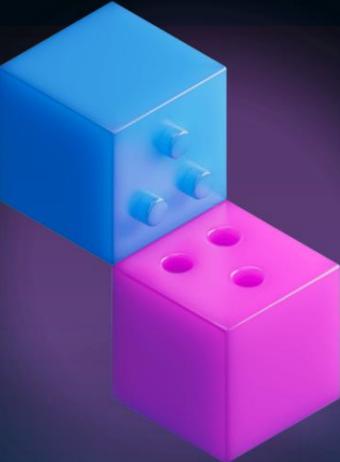




## Skill

A predefined function Copilot uses to solve part of a problem

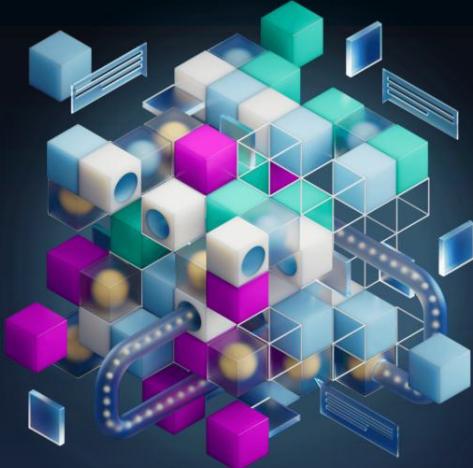
*Leverage the expertise of security analysts in areas such as threat hunting, incident response and vulnerability management*



## Plugin

A collection of skills for a particular resource that provide AI capabilities and intelligent insights into existing apps or platforms

*ServiceNow plugin available to automate initial triage efforts*



## Orchestrator

An AI architecture that evaluates a prompt, determines relevant plugins, retrieves data, ranks it for relevance, and adds grounding data to the prompt with guardrails to stay below token limits before presenting to the LLM to generate an informed response

*The 'engine' that develops a plan, gathers context and determines which skills to use to give to the LLM to generate a response*

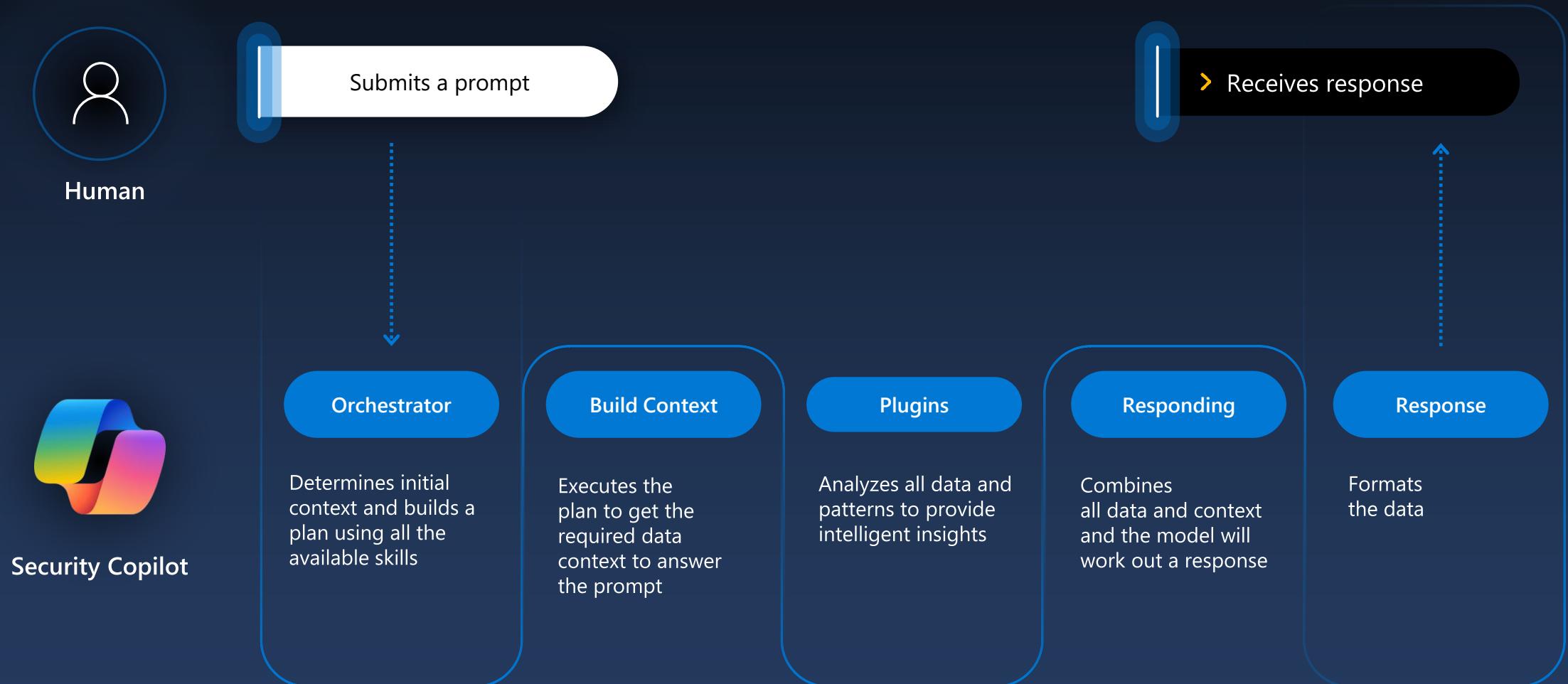
## Question 8

**Can Security Copilot do math calculations?**

- A. No. It's a Copilot that is hyper-focused on Security.
- B. Yes. But every session counts against the previsioned SCUs.

# How Microsoft Security Copilot processes prompt requests

# Operated with simple natural language queries



# How Microsoft Security Copilot processes prompt requests

- The process starts when a **user submits a prompt in the prompt bar**. Once the user submits their prompt, it's **sent to the Copilot backend referred to as the orchestrator**. The orchestrator is Copilot's system for composing capabilities (skills) together to answer a user's prompt.
- **Copilot bundles the user prompt and a full list of Copilot capabilities** for the enabled sources (plugins and knowledge bases) and then **sends it to Azure OpenAI** with the request to make a plan for fulfilling the user's request
- Azure OpenAI, **runs advanced LLMs to match the prompt with the available capabilities** (skills) and creates a plan (set of steps) for fulfilling the user's request. That plan is sent back to the orchestrator.
  - If no capability is matched, the response to the user's prompt is generated using the general knowledge LLM.
  - The **general knowledge LLM** is good at general knowledge and problem solving but **isn't uniquely focused on security** so there's a greater chance that the response provided isn't accurate. For this reason, it's beneficial to enable plugins and connect to knowledge bases to bring a collection of resource specific capabilities to Copilot.

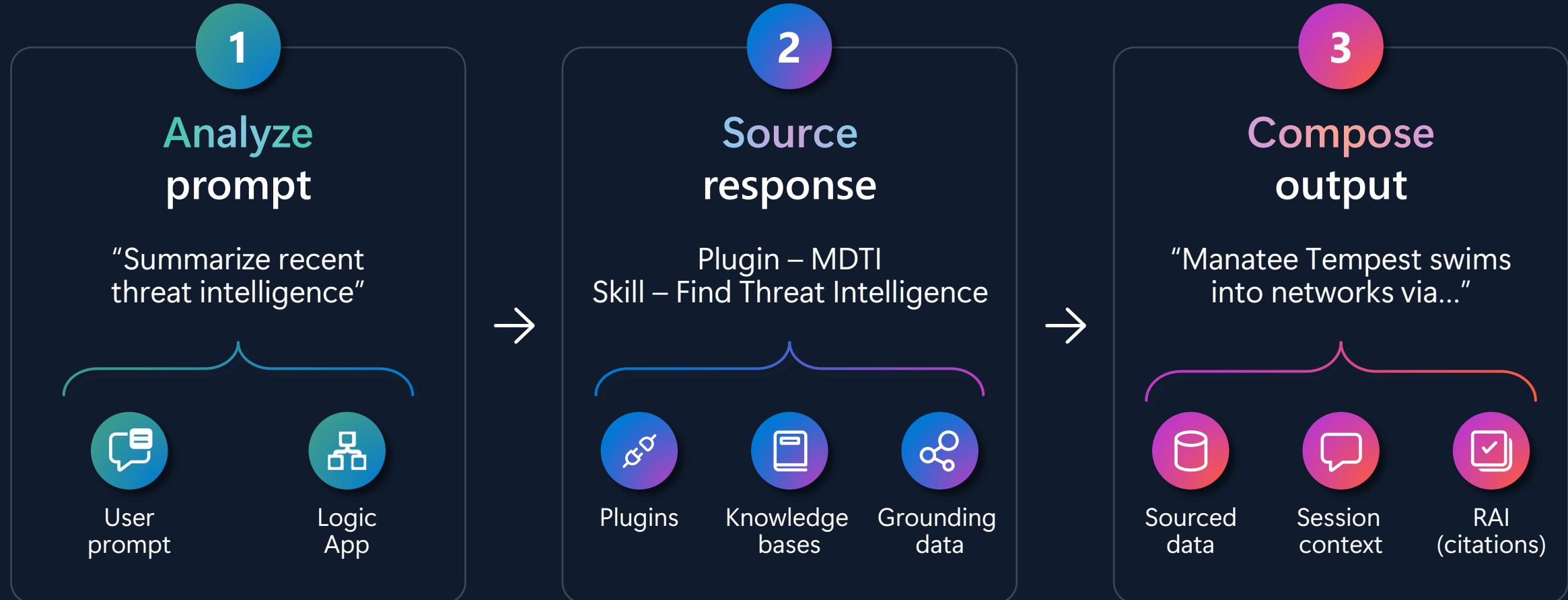
# How Microsoft Security Copilot processes prompt requests

- Copilot's **orchestrator executes the plan** by running the code for the selected plugins/capabilities and calling the appropriate application programming interfaces (API) to gather information and to take action.
- The **first- and third-party integration partner apps gather information and execute actions** based on the API call and sends the response back to Copilot.
- The orchestrator receives the response from the API calls, but it's still not ready to be sent to the user. **Copilot iterates on the process to ensure the best response**. Before a final response can be sent to the user, the **orchestrator bundles that response with the original prompt**, and sends it back to Azure OpenAI.
- Azure OpenAI uses the power of its advanced LLM to **compose a response using language that makes sense to a human being**.

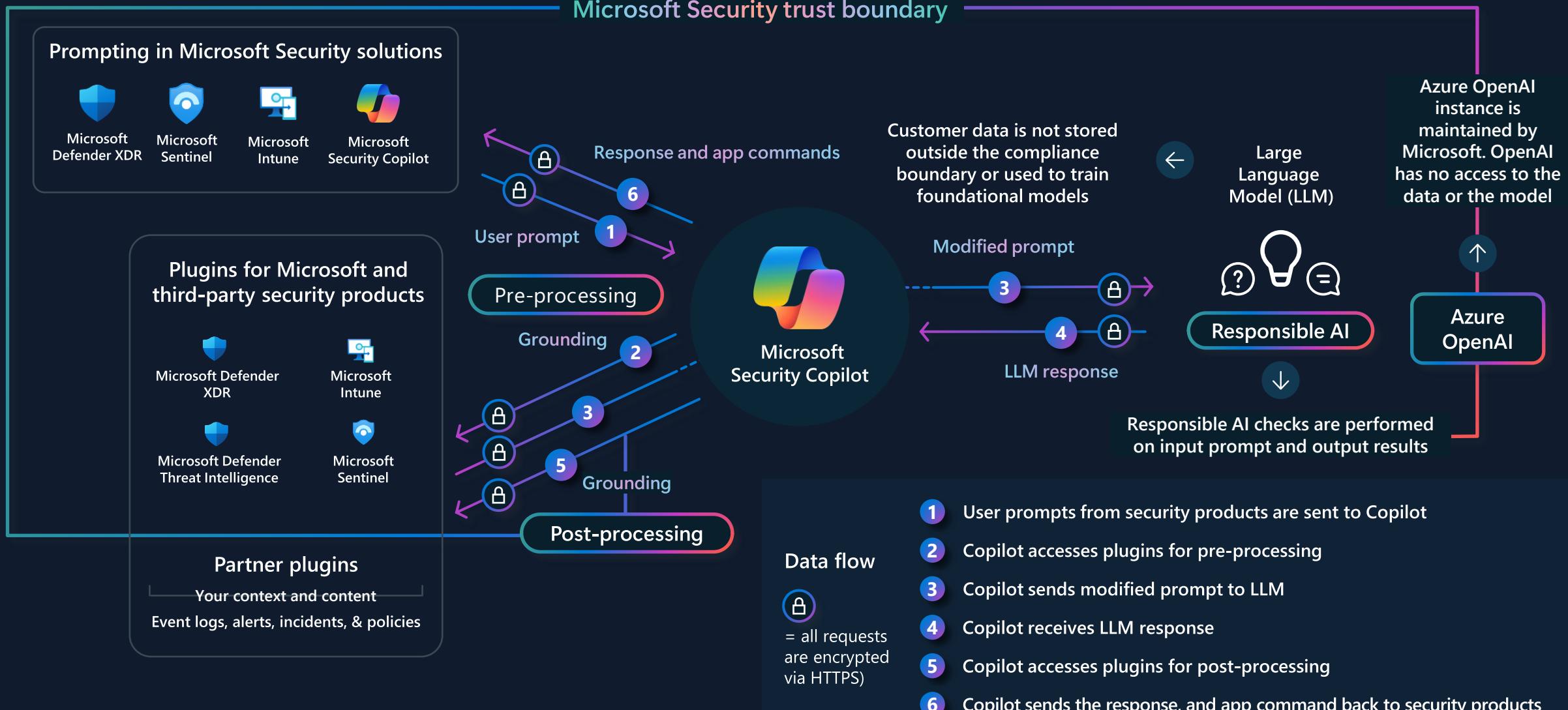
# How Microsoft Security Copilot processes prompt requests

- That **response is sent back to Copilot** for review by **Microsoft's responsible AI** models that look for anything malicious, hate speech, etc.
  - If everything looks good, then the response is sent to the user.
  - If something malicious or inappropriate is detected, then an error message is generated instead of sending the information to the user.
- The **user receives the response from the Copilot.**

# Prompt processing in Security Copilot



# Data flow for Security Copilot



# The orchestrator: threat intelligence selection

1

Analyze prompt

2

Execute plugin

3

Compose output

Summarize recent threat intelligence for Manatee tempest

3 steps completed 2 minutes 15 seconds

Chose Microsoft Defender Threat Intelligence 1 minute 19 seconds

Looked up threat intelligence information 26 seconds

Prepared your response 31 seconds

Jul 29, 3:53 PM

Here's a summary of the recent threat intelligence for "Manatee Tempest":

**Highlights**

- Human-operated ransomware:** This threat is currently associated with 41 misconfigured and 0 vulnerable devices. It has an exposure score of 18, indicating low severity. The threat actor names in reports are being updated to align with the new Microsoft weather-themed naming taxonomy. Human-operated ransomware remains one of the most impactful threats to organizations. Attackers behind these campaigns exploit poor credential hygiene and legacy configurations or misconfigurations to gain easy entry and privilege escalation points in an environment.
  - Human-operated ransomware, Source: MDTI
  - Activity Profile: Human-operated ransomware, Source: Threat Analytics

**Articles**

- Cyberpunk ransomware leaves wake of tampered AVs:** This article discusses how human-operated ransomware cycles through various attack modes and entry vectors, posing a threat to industries worldwide. The success of the attack relies on multiple weaknesses in the security posture of the target enterprise. These weaknesses can be addressed through common security measures, such as maintaining credential hygiene and following the principle of least privilege.
  - Cyberpunk ransomware leaves wake of tampered AVs, Source: MDTI
- FakeUpdates: Preying on good intentions:** This article discusses how malvertising, specifically FakeUpdates, is used by cybercriminals to grow their network of implants. The impact of these attacks grows with the widespread practice of using accounts with local administrator privileges for web browsing and other day-to-day activities.
  - FakeUpdates: Preying on good intentions, Source: MDTI
  - Activity Profile: FakeUpdates: Preying on good intentions, Source: Threat Analytics

**Intel Profiles**

## Question 9

### **How does developing good prompts help Security Copilot?**

- A. It saves time when Security Copilot doesn't have to verify responsible prompts.
- B. It shortens the time it takes for the responses to be returned.
- C. It ensures the responses are better security related and not generalized.

# Effective Prompting in Security Copilot

# Elements of Effective Prompts

- Goal: Specific, security-related information needed
- Expectations: Format or target audience for the response
- **Source:** Known information, data sources, or plugins to use
- **Context:** Why the information is needed and how it will be used

# Describe the elements of an effective prompt

**Goal**  
What is the specific security-related information you need?

---

*“Give me information about incident 18718...”*



**Context**  
Why do you need it  
And how will you use the information?

---

*“...for a report that I can submit to my manager.”*



**Expectations**  
What format or audience do you want the response tailored to?

---

*“Compile the information in a list, with a short summary.”*



**Source**  
Is there a plugin, known info, or data source Security Copilot should use?

---

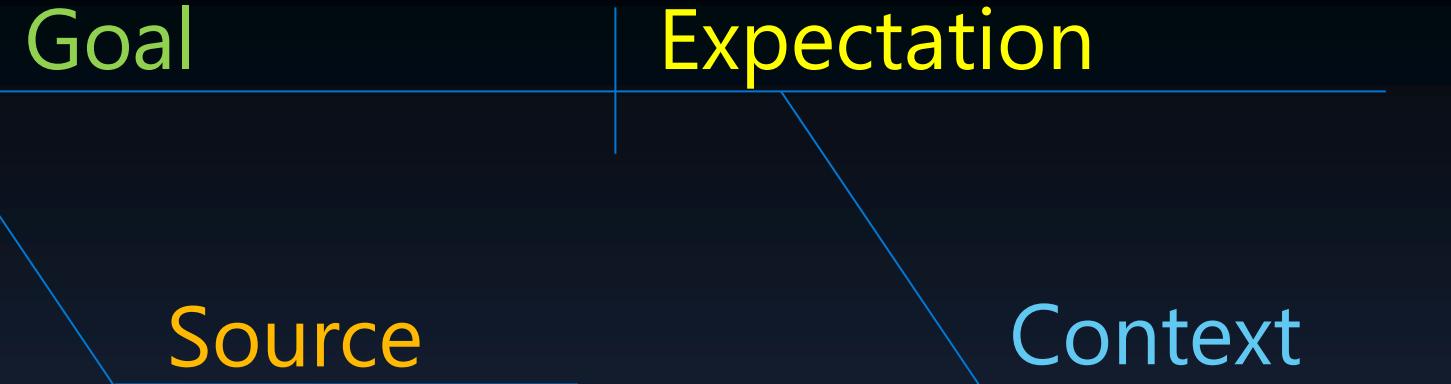
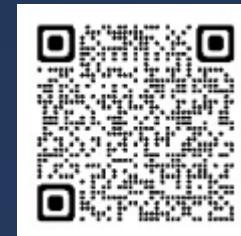
*“Look in Defender incidents.”*

# Diagram Format:

Components  
of a prompt

# Diagram example:

Identify the  
components in the  
prompt



**Prompt:** Look in Defender and tell me about the entities associated with Incident 18718 for a summarized report I can submit to my manager

Tell me about the entities in  
Incident 18718

Look in Defender

Summarization

Report to send to  
my manager

# Tips

- **What is your Goal?**- Be specific, clear, and concise as much as you can about what you want to achieve.
- **Supply Context** - Why do you need this information or how will you use it? Provide necessary context to narrow down where Copilot looks for data.
- **Set Expectations** - What format or which target audience do you want the response tailored to? Give positive instructions instead of "what not to do. Copilot is geared toward action, so telling it what you want it to do for exceptions is more productive.
- **Provide a Source**- Supply any known information, data source(s), or plugins Copilot should be used.
- **Directly address** Copilot as "You", as in, "You should ..." or "You must as this is more effective than referring to it as a model or assistant.

# Featured Prompts

- **Analyze a script or command** - Identifies script language, purpose, risks, and recommended actions
- **Summarize a security article** - Extracts main points, key takeaways, and implications for your organization
- **Generate a security query** - Converts natural language request into query language for specific data sources
- **Generate a security report** - Creates concise and informative report for specific audience using previous prompts and responses



# Top 10

1 Analyze the following script <INSERT SCRIPT>

2 If a user is listed in the incident details, show which devices they recently used and indicate if they are compliant with policies.

3 Summarize Sentinel incident <SENTINEL INCIDENT ID>.

4 Show me the top 5 DLP alerts that I should prioritize today.

5 Show me the intel profile for <THREAT ACTOR> and create a bulleted list of associated indicators for this actor.

6 Can you summarize the IOC's related to this intel profile into a list and give me direct links for Microsoft Defender Threat Intelligence portal?

7 Describe the impact of this policy on users and highlight setting conflicts with existing policy.

8 Why was <USERNAME> prompted for MFA?

9 Generate and run a KQL query within Microsoft Sentinel to hunt for break-glass account usage.

10 Append comment To ServiceNow Incident.

## Question 10

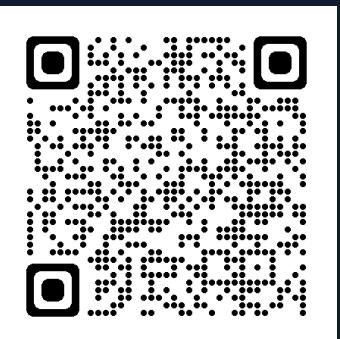
**What are the four elements of an effective prompt for Security Copilot in order?**

- A. Goal, Expectations, Source, Context
- B. Expectations, Goal, Source, Context
- C. Source, Context, Expectations, Goal

**Choose the best prompt:**

- A. Show the authentication methods setup for each user involved in that incident. Especially indicate whether they have MFA enabled.
- B. Create a report for Incident 2818.
- C. Translate Incident 2818 into Chinese.

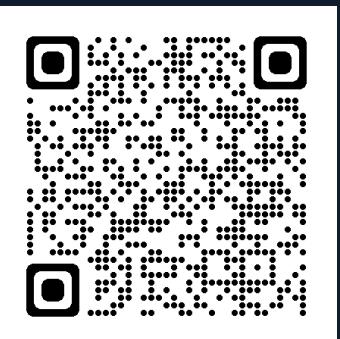
## Question 11



## Question 12

**Choose the best prompt:**

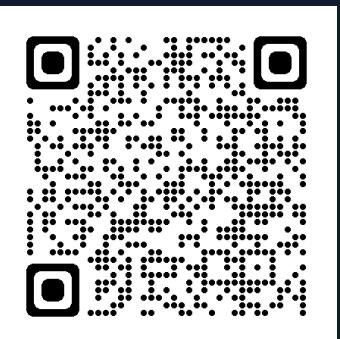
- A. Write an executive report.
- B. Write an executive report summarizing this investigation. It should be suited for a non-technical audience.
- C. Write a report in the first person.



# Question 13

**Choose the best prompt:**

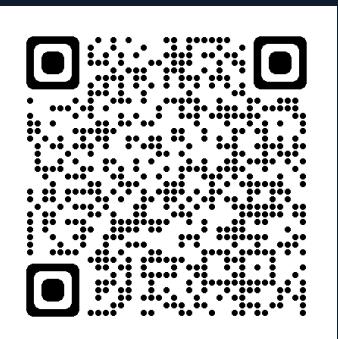
- A. What is "service-logins.com"?
- B. Get the reputations for hostname "service-logins.com" and put it in a report to show my manager.
- C. Get the reputations for hostname "service-logins.com"



## Question 14

**Choose the best prompt:**

- A. Summarize recent threat intelligence.
- B. Generate and run Defender Hunting Queries
- C. Describe the impact of this policy on users and highlight setting conflicts with existing policy.



Take it further

# Promptbooks

- **Promptbooks: A Collection of Prompts**
  - Used for specific security-related tasks
  - Examples: incident investigation, threat actor profile, suspicious script analysis, vulnerability impact assessment
  - Existing prompt books can be used as templates or examples
  - Can be modified to suit your needs

Explore with Copilot

Featured prompts Promptbooks

**Suspicious script analysis**  
Get a report analyzing the intent, intelligence, threat actors, and impacts of a suspicious script.

Microsoft Security · 7

**Threat actor profile**  
Get a report profiling a known actor with suggestions for protecting against common tools and tactics.

Microsoft Security · 5

**Microsoft Sentinel incident investigation**  
Get a report about a specific incident, along with related alerts, reputation scores, users, and devices.

Microsoft Security · 7



# Using Promptbooks

- **Incident Investigation**
  - Summarize incident, assess impact, provide remediation steps
- **Threat Actor Profile**
  - Get executive summary about specific threat actor
- **Suspicious Script Analysis**
  - Analyze and interpret command or script
- **Vulnerability Impact Assessment**
  - Assess impact of publicly disclosed vulnerability

Explore with Copilot

Featured prompts Promptbooks

**Microsoft 365 Defender incident investiga...**  
Get a report about a specific incident, with related alerts, reputation scores, users, and devices.

**Threat actor profile**  
Get a report profiling a known actor with suggestions for protecting against common tools and tactics.

**Microsoft Sentinel incident investiga...**  
Get a report about a specific incident, along with related alerts, reputation scores, users, and devices.

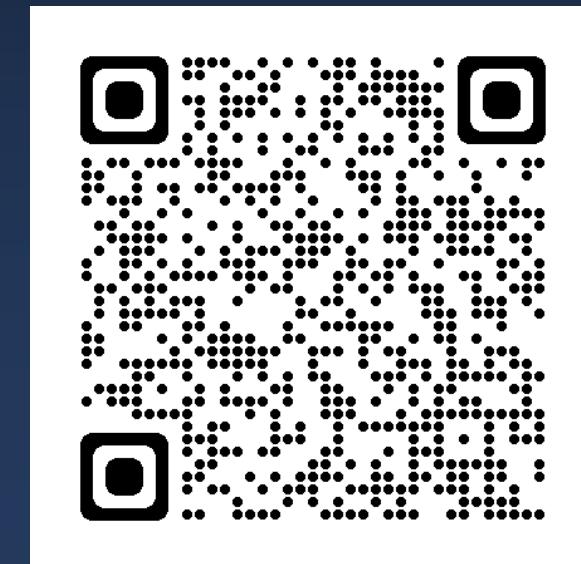
**PROMPTBOOKS**

**Suspicious script analysis**  
Get a report analyzing the intent, intelligence, threat actors, and impacts of a suspicious script.

**Microsoft 365 Defender incident investigation**  
Get a report about a specific incident, with related alerts, reputation scores, users, and devices.

**Microsoft Sentinel incident investigation**  
Get a report about a specific incident, along with related alerts, reputation scores, users, and devices.

...



# Expanding Knowledge

Knowledge base (KB) connections, a feature of Microsoft Security Copilot currently in preview, allows you to integrate your organization's knowledge base as another source of information.

- Azure AI Search plugin
- File upload
- Public Web

In the uploaded files, How Does Your Organization Know If It's Experiencing a DDoS Attack?

Manage sources

Plugins

Files Preview

Upload files, like your internal policies, so your organizational knowledge will inform Copilot's responses. When you prompt, specify a file name or 'uploaded files' so Copilot will use them. Only you will be able to see your uploaded files.

Files must be 3 MB or less, and in the format of .docx, .pdf, .txt, .md

17 out of 20 MB remaining

Upload file

Uploads

File Name	Uploaded On	Size	Action	Status
CDMDDataModelDocument_v4.1.1_508C.pdf	4/9/2024	1.91 MB	trash	On
understanding-and-responding-to-distributed-denial-of-service-attacks_508c.pdf	4/9/2024	1.03 MB	trash	On
Brief 6MDM Episodes.docx	4/9/2024	15.13 KB	trash	On

Websites

Public web Content downloads

## Question 15

**A security analyst is tasked with automating investigation flows to streamline repetitive steps in Copilot. After selecting relevant prompts from an existing session, the analyst wants to create a promptbook. What should the analyst do next to accomplish this task?**

- A. Select the 'Create promptbook' icon to open the page where they can name, tag, and further customize the promptbook.
- B. Select the 'Export prompts' button to download the selected prompts for manual execution.
- C. Choose the 'Save session' option to preserve the current state of the session for later use.

# Conclusion

- Precise and comprehensive prompts produce accurate, relevant responses
  - Improves speed and efficiency of generative AI tasks
  - Mitigates biases and reduces output errors
  - No coding experience or deep knowledge required
- Prompt engineering best practices can help security teams utilize the power of generative AI
  - Improves workflow and focus on higher-level tasks
  - Minimizes tedious work
- Learn more about Security Copilot at <https://aka.ms/SecurityCopilot>



# Stay Connected and Informed

**Community is everything!**

