



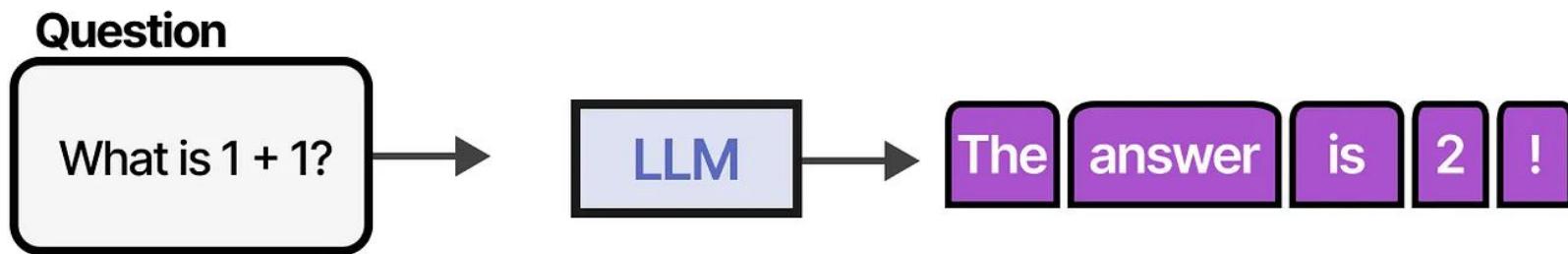
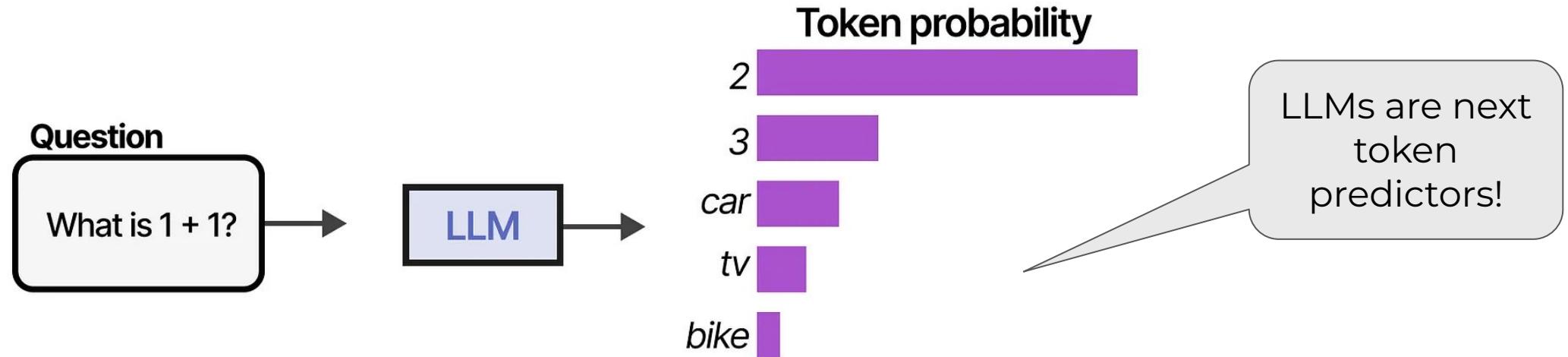
# Building Web Security Agents

**eCrime 2025**

11/03/2025  
Mohamed Nabeel



# LLMs to Agentic Systems



\* Source: An Illustrated Guide to LLM Agents

## Conversation

Question 1

What is  $1 + 1$ ?

LLM

The answer is 2!

Question 2

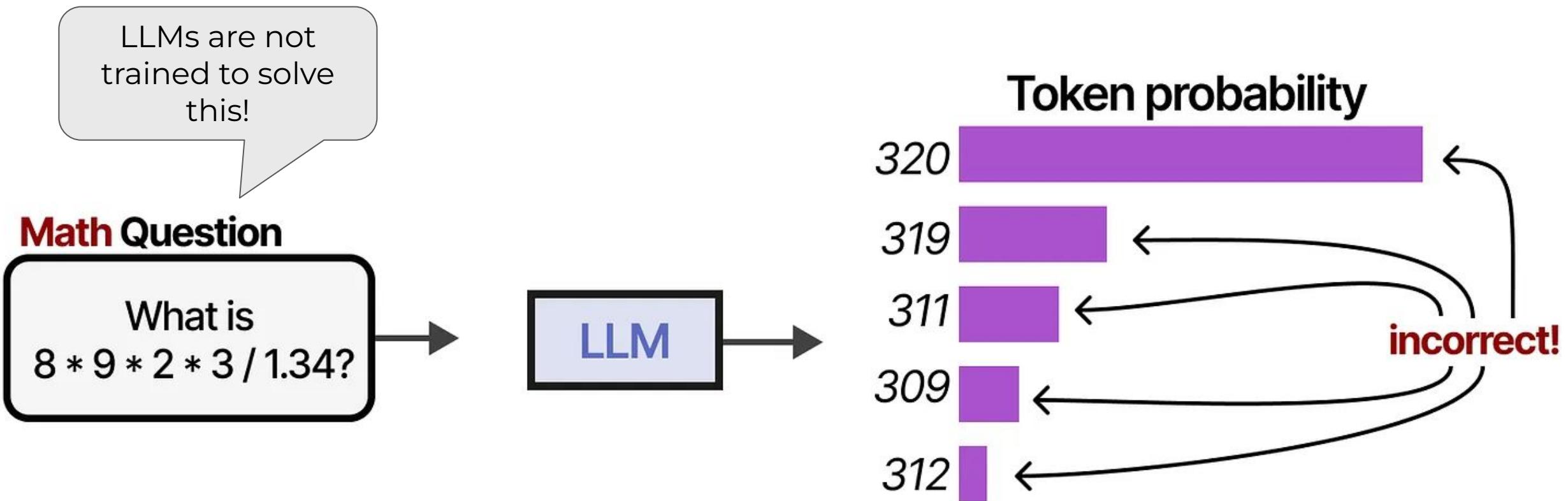
What was my  
**previous question?**

LLM

I have **no knowledge** about  
any previous question.

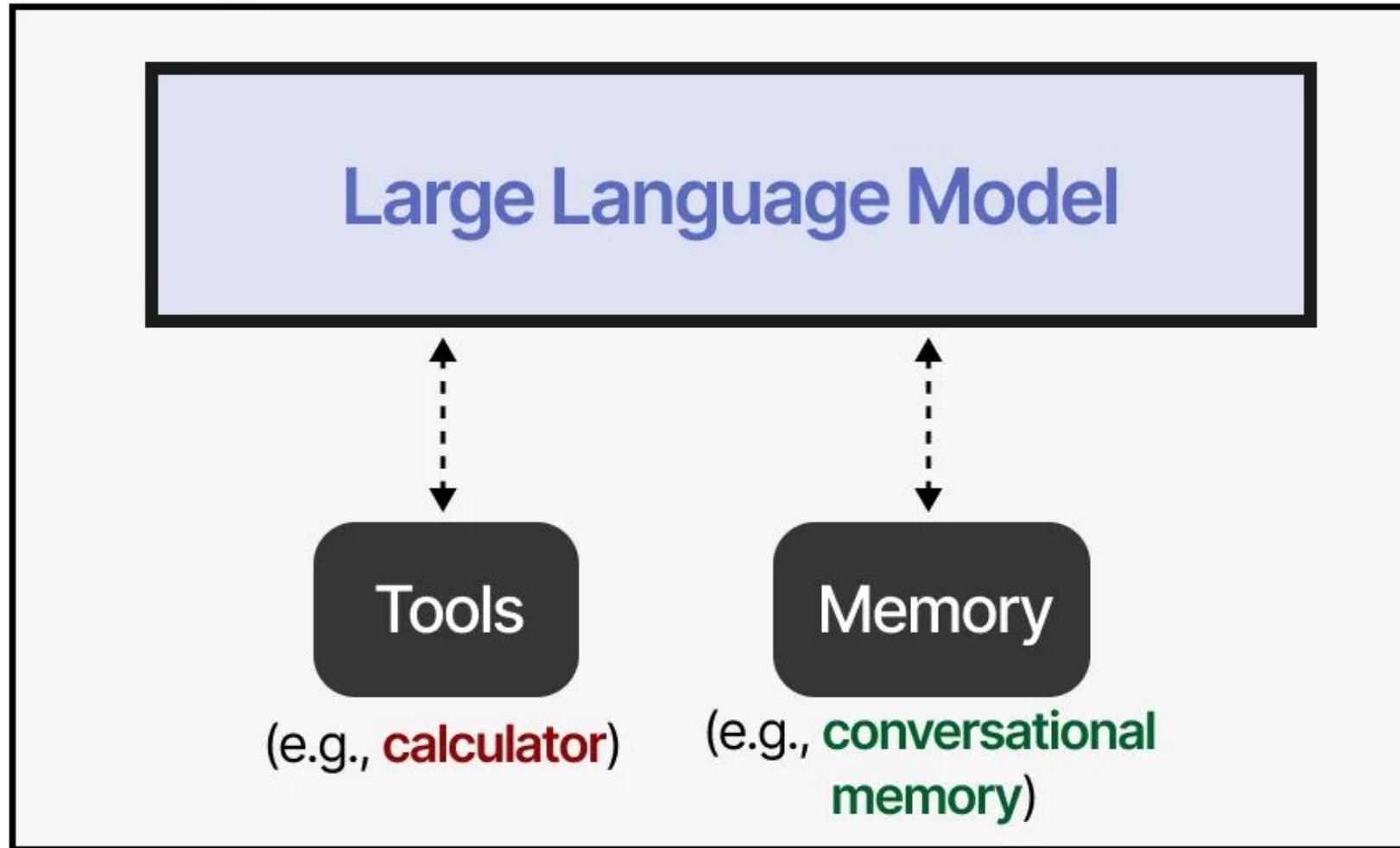
LLMs have no  
memory!

\* Source: An Illustrated Guide to LLM Agents



\* Source: An Illustrated Guide to LLM Agents

# The Augmented LLM



\* Source: An Illustrated Guide to LLM Agents

## Math Question

What is  
 $8 * 9 * 2 * 3 / 1.34?$

Tool use!

## The “Augmented” Large Language Model

322.389

Tools  
(calculator)

Memory  
(conversational)

identifies a **math question**  
and chooses the  
corresponding **tool** to use

\* Source: An Illustrated Guide to LLM Agents

# Why Agents?

Task	GPT-3.5		ChatGPT		GPT-4	
	Base	+SELF-REFINE	Base	+SELF-REFINE	Base	+SELF-REFINE
Sentiment Reversal	8.8	<b>30.4</b> ( $\uparrow$ 21.6)	11.4	<b>43.2</b> ( $\uparrow$ 31.8)	3.8	<b>36.2</b> ( $\uparrow$ 32.4)
Dialogue Response	36.4	<b>63.6</b> ( $\uparrow$ 27.2)	40.1	<b>59.9</b> ( $\uparrow$ 19.8)	25.4	<b>74.6</b> ( $\uparrow$ 49.2)
Code Optimization	14.8	<b>23.0</b> ( $\uparrow$ 8.2)	23.9	<b>27.5</b> ( $\uparrow$ 3.6)	27.3	<b>36.0</b> ( $\uparrow$ 8.7)
Code Readability	37.4	<b>51.3</b> ( $\uparrow$ 13.9)	27.7	<b>63.1</b> ( $\uparrow$ 35.4)	27.4	<b>56.2</b> ( $\uparrow$ 28.8)
Math Reasoning	<b>64.1</b>	<b>64.1</b> (0)	74.8	<b>75.0</b> ( $\uparrow$ 0.2)	92.9	<b>93.1</b> ( $\uparrow$ 0.2)
Acronym Generation	41.6	<b>56.4</b> ( $\uparrow$ 14.8)	27.2	<b>37.2</b> ( $\uparrow$ 10.0)	30.4	<b>56.0</b> ( $\uparrow$ 25.6)
Constrained Generation	28.0	<b>37.0</b> ( $\uparrow$ 9.0)	44.0	<b>67.0</b> ( $\uparrow$ 23.0)	15.0	<b>45.0</b> ( $\uparrow$ 30.0)

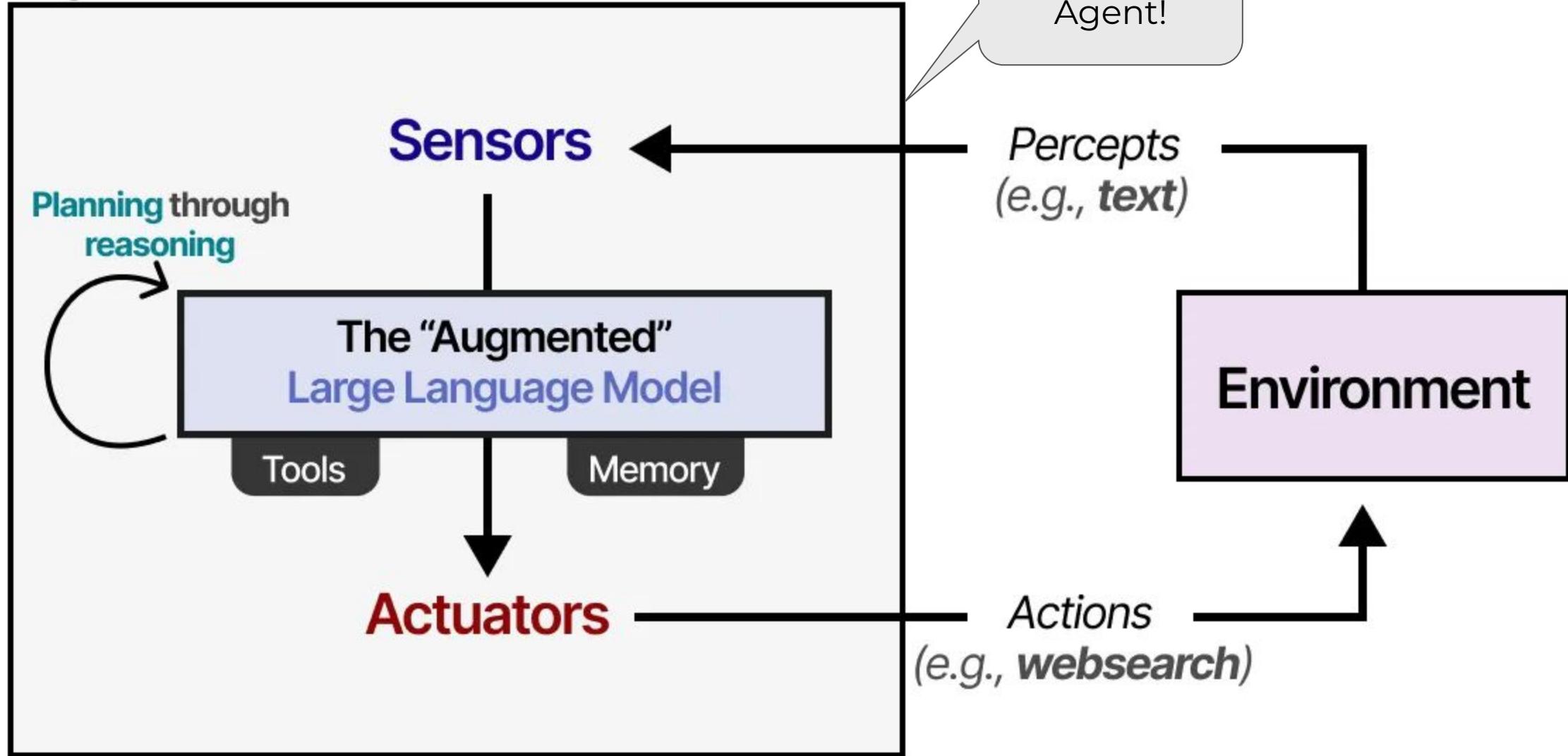
Table 1: SELF-REFINE results on various tasks using GPT-3.5, ChatGPT, and GPT-4 as base models. SELF-REFINE consistently improves LLM. Metrics used for these tasks are defined in Sect.

Prompt Method <sup>a</sup>	HotpotQA (EM)	Fever (Acc)
Standard	28.7	57.1
CoT (Wei et al., 2022)	29.4	56.3
CoT-SC (Wang et al., 2022a)	33.4	60.4
Act	25.7	58.9
ReAct	27.4	60.9
CoT-SC → ReAct	34.2	<b>64.6</b>
ReAct → CoT-SC	<b>35.1</b>	62.0

Table 2: Pass@1 results on HumanEvalFix [32]. Except for SWE-agent, we use scores as reported in Yu et al. [65].

Model	Python	JS	Java
CodeLLaMa-instruct-13B	29.2	19.5	32.3
GPT-4	47.0	48.2	50.0
DeepseekCoder-CodeAlpaca-6.7B	49.4	51.8	45.1
WaveCoder-DS-6.7B	57.9	52.4	57.3
SWE-agent w/ GPT-4 Turbo	<b>87.7</b>	<b>89.7</b>	<b>87.9</b>

# Agent



\* Source: An Illustrated Guide to LLM Agents

## Question

I have **10** apples. I gave **2** apples away. I ate **1**. How many do I have?

Let's think step-by-step.

Start **reasoning** behavior  
(typically Chain-of-Thought)

CoT  
Reasoning

Large Language Model

You have **10** apples

You gave **2** away and have **8** left

You ate **1** and have **7** left

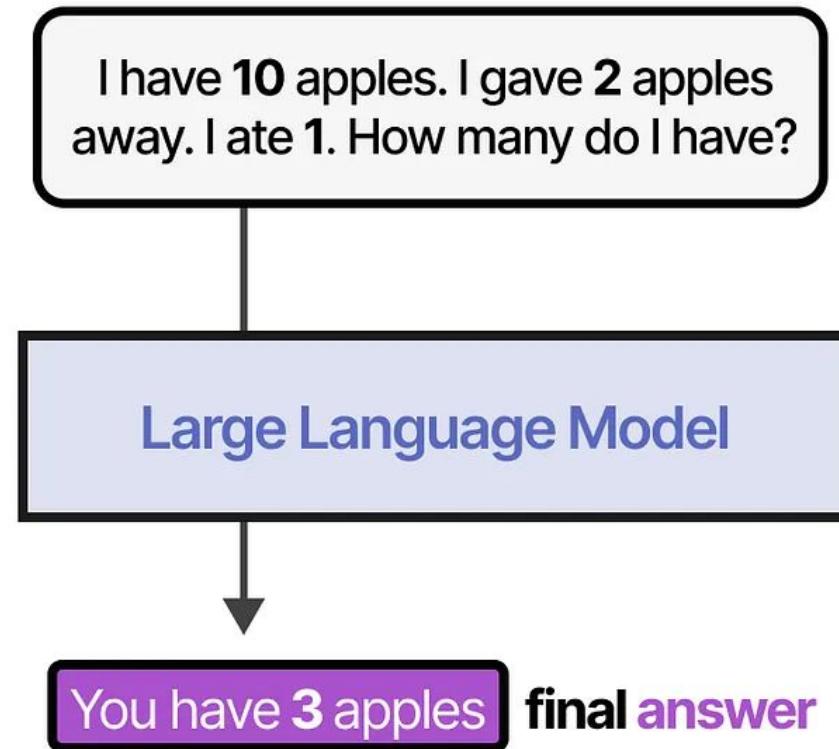
reason steps

You have **7** apples

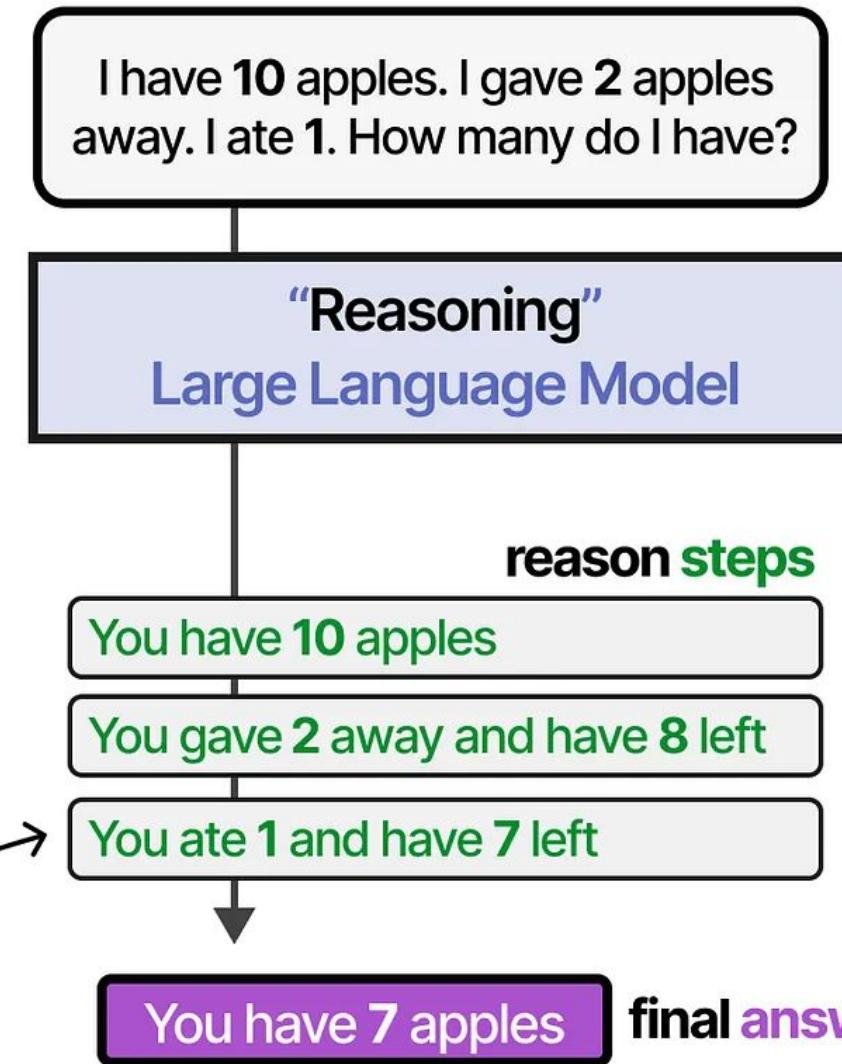
final answer

\* Source: An Illustrated Guide to LLM Agents

## “Regular” LLMs



## “Reasoning” LLMs

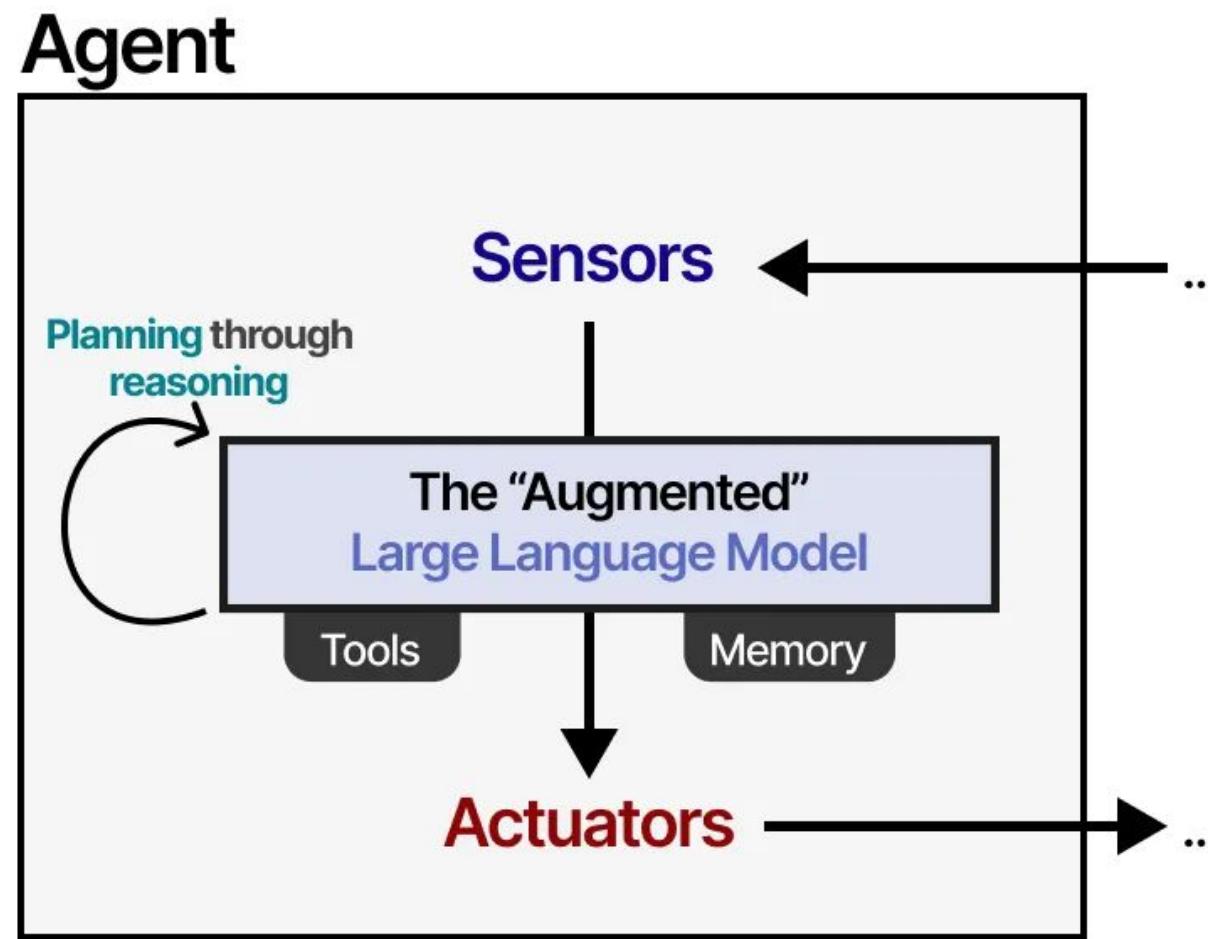


“reason” before  
answering

\* Source: An Illustrated Guide to LLM Agents

Agent  
Planning

- Step 1 **Search** the web
- Step 2 **Summarize** results
- Step 3 **Infer** best product
- ⋮
- Step *n* ....



\* Source: An Illustrated Guide to LLM Agents

Agent = LLM +  
Tools + Memory  
+ Planning

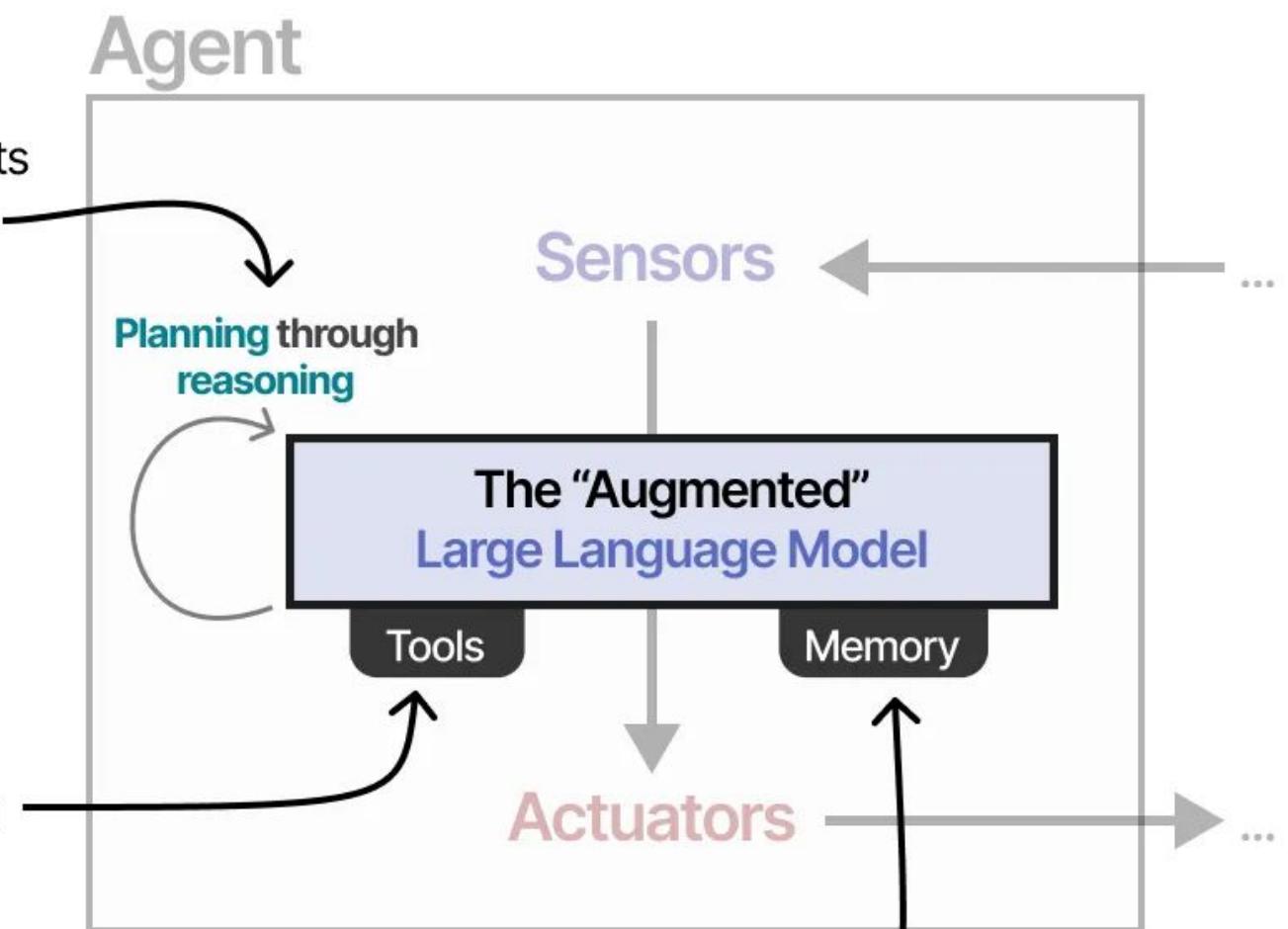
### Planning

- Chain-of-Thoughts
- Reflection
- Subgoal decomposition

### Tools

- Search
- API
- Function calling

# Agent



### Memory

- Short-term memory
- Long-term memory
- Retrieval

\* Source: An Illustrated Guide to LLM Agents

# **Can LLMs help with cyber threat intelligence?**

Query:

"Can you summarize the recent event of bulletproof host Prospero moving to Kaspersky Lab?"

Response (Gemma3 4b) :

Incorrect context

```
('{\n'
  "summary": "Prospero, the popular and highly-regarded cybersecurity analyst '\n    "and host of the 'Ask Prospero' YouTube channel and podcast, has joined "\n    "Kaspersky Lab as a Senior Threat Research Analyst. Here's a breakdown of the "\n    'key details:\n\n* **The Move:** Prospero officially announced his '\n    'departure from his independent ventures and joined Kaspersky Lab in late '\n    "August 2023.\n\n* **Why Kaspersky?** While he hasn't revealed all the "\n    'specifications, Prospero cited a desire to work for a company with a strong focus '\n    'on cybersecurity research and a commitment to protecting users. He '
```

Knowledge cutoff

Query:

"Can you summarize the recent event of bulletproof host Prospero moving to Kaspersky Lab?"

Response (Llama3.2 3b) :

Incorrect context

```
AIMessage(content="I couldn't find any information on a person named Prospero being the host of the Bulletproof podcast, or anyone associated with Kaspersky Lab, being 'associated with it. It's possible that you may be thinking of a different 'event or individual.\n'\n'Bulletproof is a popular self-improvement and personal development podcast 'hosted by Eric Thomas (also known as ET). I couldn't find any information 'about Prospero being the host of this podcast.\n'
```

Query:

"Can you summarize the recent event of bulletproof host Prospero moving to Kaspersky Lab?"

Response (Gemma3 4b + Search Tool) :

Correct context

('One of the most notorious providers of abuse-friendly "bulletproof" web hosting for cybercriminals, Prospero 000 (AS200593), has begun routing its operations through networks operated by Kaspersky Lab, as reported by KrebsOnSecurity. AS209030, owned by Kaspersky Lab, is providing connectivity to Prospero. The reason for this transit remains unclear.')

Up-to-date Information

Query:

"Can you summarize the recent event of bulletproof host Prospero moving to Kaspersky Lab?"

Refined response

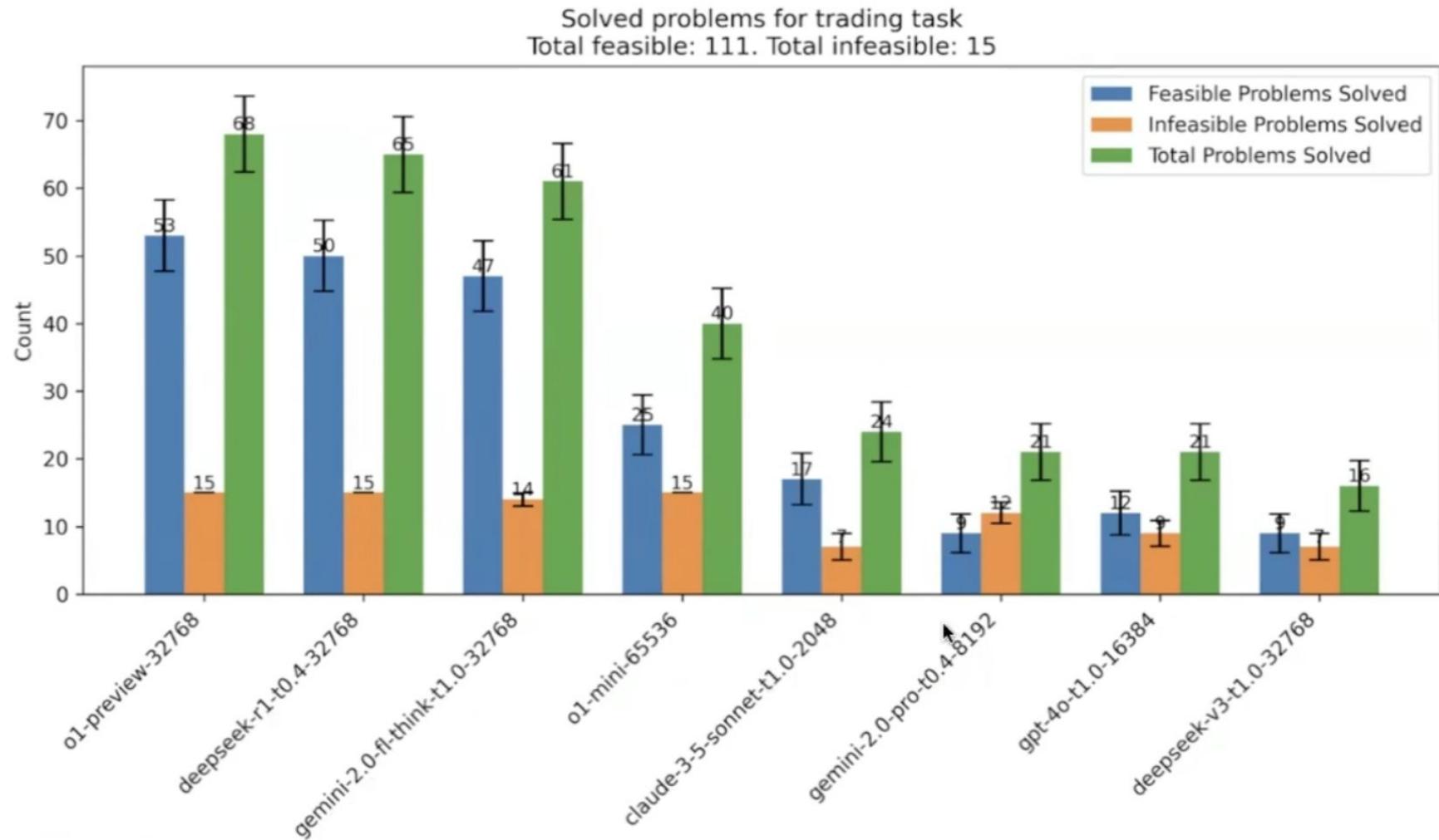
Response (Gemma3 4b + Search Tool + Reflection) :

```
(## Summary\n'
'\n'
'One of the most notorious providers of abuse-friendly "bulletproof" web '
'hosting for cybercriminals, Prospero OOO (AS200593), has begun routing its '
'operations through networks operated by Kaspersky Lab. AS209030, owned by '
'Kaspersky Lab, is providing connectivity to Prospero. The Kaspersky Security '
'Network (KSN), a complex distributed infrastructure developed by Kaspersky '
'Lab, is designed to intelligently process cybersecurity-related data streams '
'from millions of voluntary participants worldwide. KSN enables Kaspersky '
' ### Sources:\n'
'* Notorious Malware, Spam Host "Prospero" Moves to Kaspersky Lab : '
'https://krebsonsecurity.com/2025/02/notorious-malware-spam-host-prospero-moves-to-kaspersky-lab/\n'
'* Notorious Malware, Spam Host "Prospero" Moves to Kaspersky Lab – Krebs ... '
': '
https://techznews.com/2025/03/02/notorious-malware-spam-host-prospero-moves-to-kaspersky-lab-krebs-on-security/\n'
'* Notorious Malware Spam Host "Prospero" Moves to Kaspersky Lab February '
```

Up-to-date Information with  
sources

# Model vs. Task Complexity vs. Success Rate

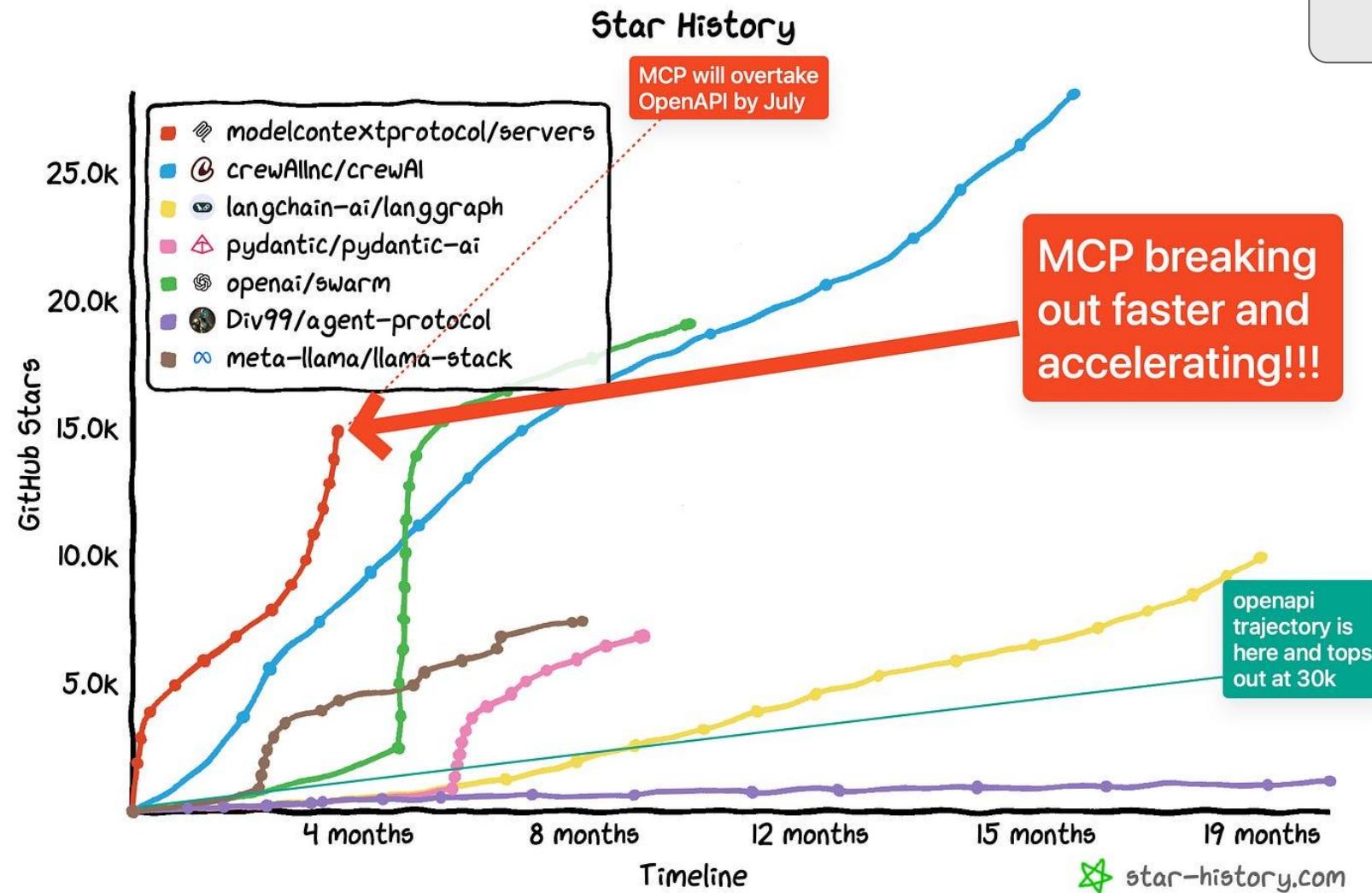
Strive for simplicity!



\*Ref: By Chip Huan

# Google ADK

# Agentic Frameworks

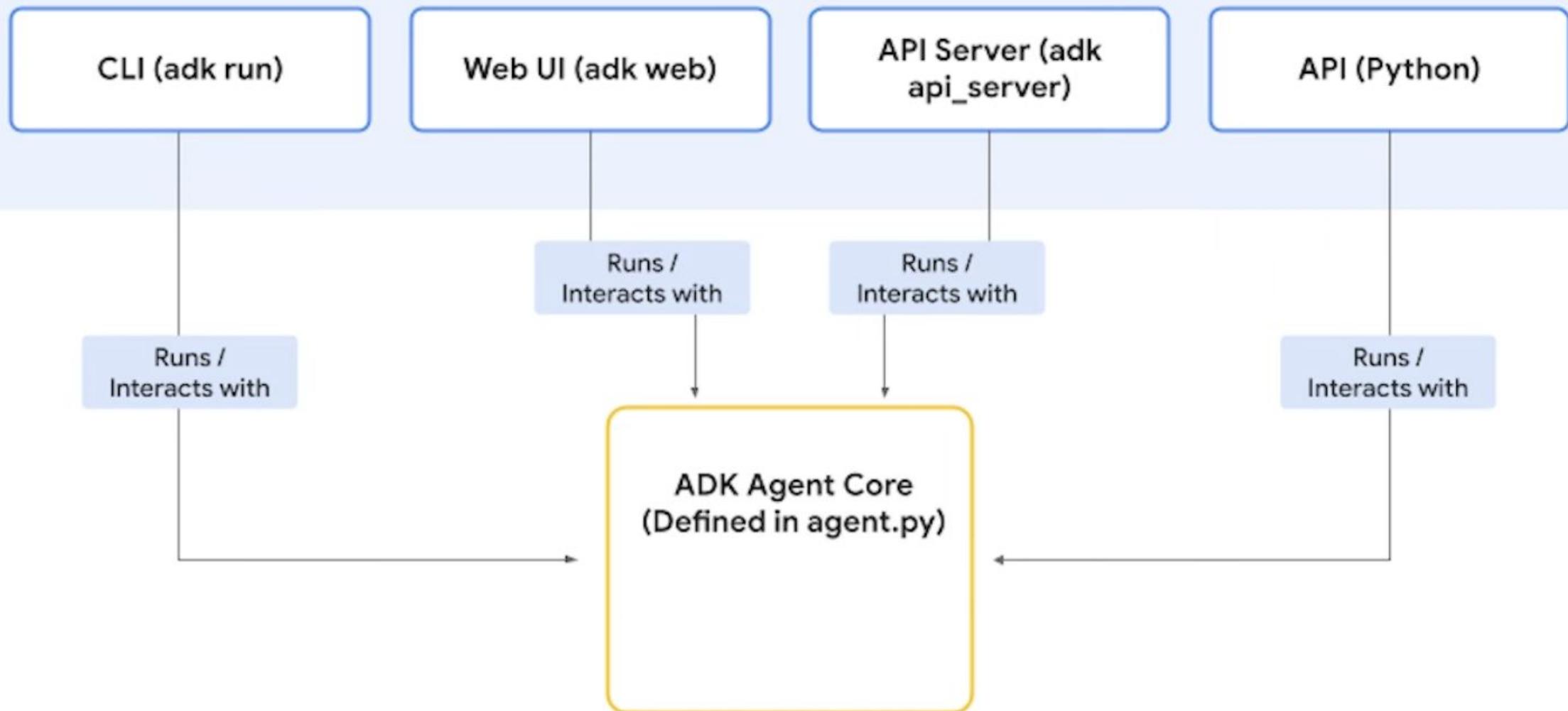


Google ADK

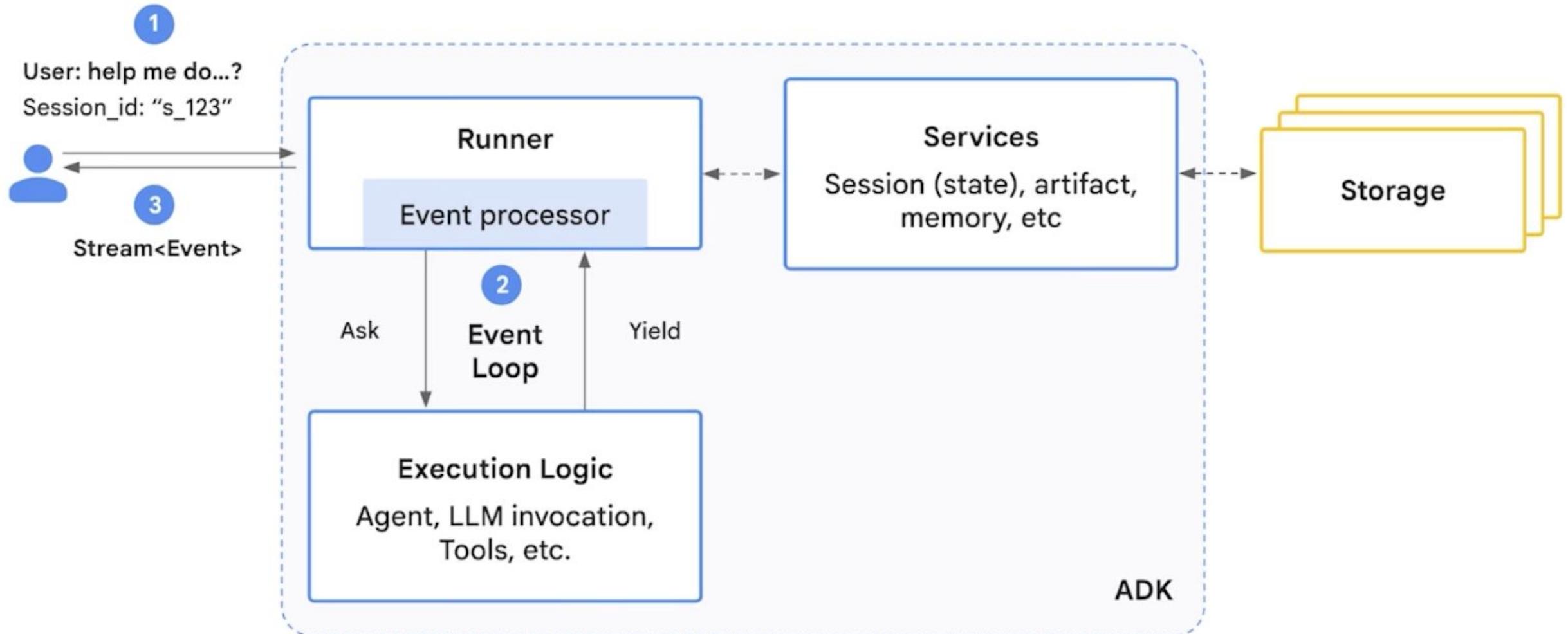
OpenAI Agents SDK

Slightly outdated chart - done before ADK and SDK.

## Interaction Interfaces



\* Source: Google official documentation



\* Source: Google official documentation

# Building a Web Security Agent

**Is paypal-verify.com benign?**

**firewallexpert.st**

create time	03/18/2025 14:04 PT
cr_id	b2d5b3e4-5ea8-4e62-8346-6ca5cb138910
type	FN (unknown -> malware)
pandb reason	not in our db as malicious, not in c2
submitter	rick.major@usu.edu
raw url	http://firewallexpert.st

[CUSTOMER] [src\_IP]:172.22.136.1 https://hybrid-analysis.com/sample/039893862988d86493976f69e42d7137fcc686bdb2536a5847707ca7dfa06893/

```

function showVerifyWindow() {
    verifyWindow.style.display = "block";
    verifyWindow.style.visibility = "visible";
    verifyWindow.style.opacity = "1";
    verifyWindow.style.top = checkboxWindow.offsetTop - 80 + "px";
    verifyWindow.style.left = checkboxWindow.offsetLeft + 54 + "px";
    if (verifyWindow.offsetTop < 5) {
        verifyWindow.style.top = "5px";
    }
    if (verifyWindow.offsetWidth + verifyWindow.offsetHeight > window.innerWidth - 10) {
        verifyWindow.style.left = checkboxWindow.offsetLeft - 8 + "px";
    } else {
        verifyWindowArrow.style.top = checkboxWindow.offsetTop + 24 + "px";
        verifyWindowArrow.style.left = checkboxWindow.offsetLeft + 45 + "px";
        verifyWindowArrow.style.visibility = "visible";
        verifyWindowArrow.style.opacity = "1";
    }
}

function closeVerifyWindow() {
    verifyWindow.style.display = "none";
    verifyWindow.style.visibility = "hidden";
    verifyWindow.style.opacity = "0";
    verifyWindowArrow.style.visibility = "hidden";
    verifyWindowArrow.style.opacity = "0";
    showCaptchaCheckboxes();
    hideCaptchaLoading();
}

```

Time Categories accept\_rate: 64% rank: 79%

Devices	Queries
26.71	32.71

VT (score: 1) Engine

Scanning Result

1/94 security vendor flagged this domain as malicious

Community Score top 1M

Join our Community and enjoy additional community insights and crowdsourced detections, plus an API key to automate checks.

Do you want to automate checks?

Trustwave	Phishing	Abusix	Clean
Acronis	Clean	ADMINUSLabs	Clean
AI Labs (MONITORAPP)	Clean	AlienVault	Clean
alphaMountain.ai	Clean	Anti-AVL	Clean
benkow.cc	Clean	Blueliv	Clean

(py39) (base) mohammednabe@M-L5727YXX3F graph\_builder % dig +short fire 104.21.14.146 172.67.159.176

mohammednabe@M-L5727YXX3F graph\_builder % dig firewallxp 104.21.14.146 <>> firewallexpert.st :+cmd

opcode: QUERY, status: NOERROR, id: 65020 ra: 1, QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 1

TION: 0, flags:; udp: 1224

ON: st. IN A

: 292 IN A 172.67.159.176

: 292 IN A 104.21.14.146

0 msec

2.10#53(10.55.2.10) 19 08:52:45 PDT 2025

: 78

is firewallexpert.st legit

All Videos Short videos Forums Images News Shopping More Tools

ScanAdviser.com https://www.scanadviser.com/check-website/bopaji... bopaji.firewallexpert.st Reviews | scam or legit check

3 days ago — bopaji.firewallexpert.st is probably not a scam but legit. We analysed the website and came with a review of bopaji.firewallexpert.st of 61%.

ScamAdviser.com https://www.scamadviser.com/check-website/curora... curora.firewallexpert.st Reviews | scam or legit check

We think curora.firewallexpert.st is legit as we found few indicators which might point to a scam. curora.firewallexpert.st review resulted in a trust score of ...

Scam Detector https://www.scam-detector.com/validator/firewall... firewallexpert.st Reviews: Is this site a scam or legit?

Is firewallexpert.st legit? Well, it's suspicious, as it received an overall low trust score based on our chart. We put to work 53 powerful factors to ...

People also ask :

Is Fame St legit?

Feedback

Are you human? Verify that you are a human  I'm not a robot CAPTCHA Privacy Terms

Home Explore Notifications Messages Grok Bookmarks Communities Premium Verified Orgs Profile More Post

Top Latest People Media Lists

RS. @redrabbytes - 36m Little update - New #Masjesu relay IP found. IP: 77[.]190[.]153[.]218 ( AS214943 ) Downloader: hxpx://77[.]190[.]153[.]218/bin/spim Potential #C2: 77[.]190[.]153[.]218:1944

cc @500mk500 @banthisguy9349

DNS-ST.BAHNHOF.NET NS1.BAHNHOF.NET 195 SOUTHEAST-2.DNS-AU.WEST-2.DNS-US.ST 26

Whois: whois.nic.st

status: ACTIVE

remarks: Registration inform

created: 1997-11-07

changed: 2022-11-15

source: IANA

# whois.nic.st

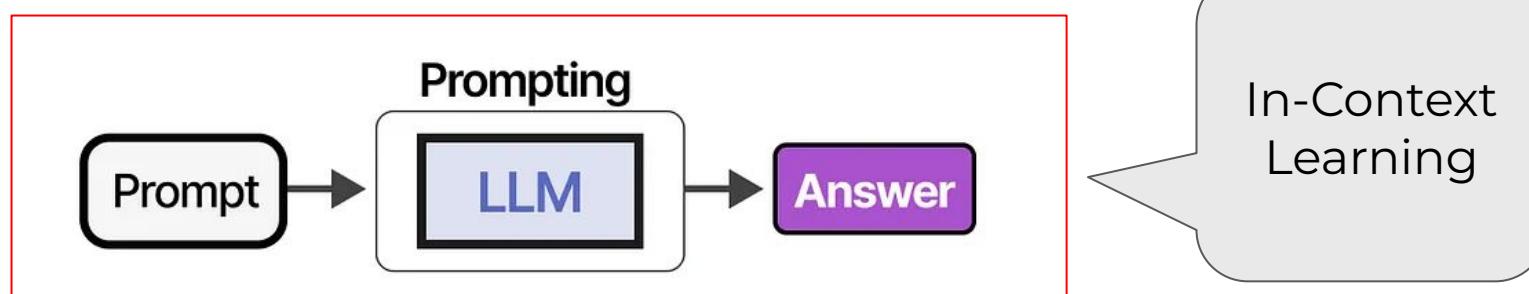
Whois Server Version 3.3.2

Domain Name: firewallexpert.st Registrar: IncogNET Name Server: blair.ns.cloudflare.com Name Server: sonny.ns.cloudflare.com Status: ok Updated Date: 2024-12-23 Creation Date: 2024-12-23 Expiration Date: 2025-12-23

Nabeel Yousuf @nabeelyx WatchingRac @RacWatchin8872 - 4h #Tycoon2FA#Phishing

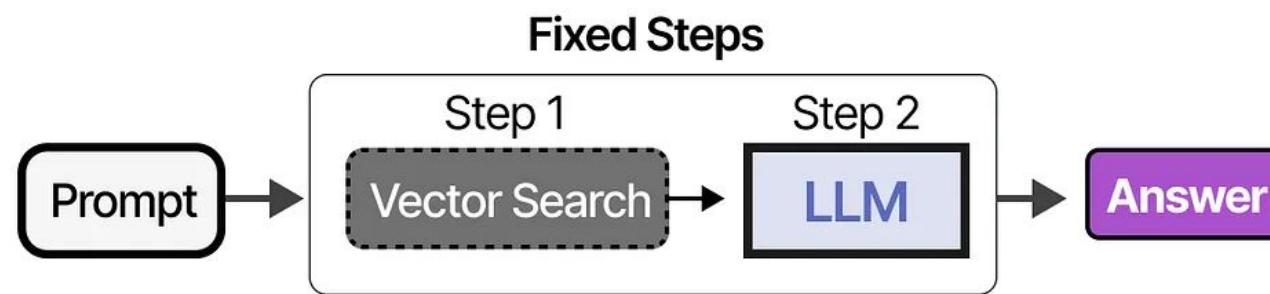
Page 1 of 1 43

# Basic Agents

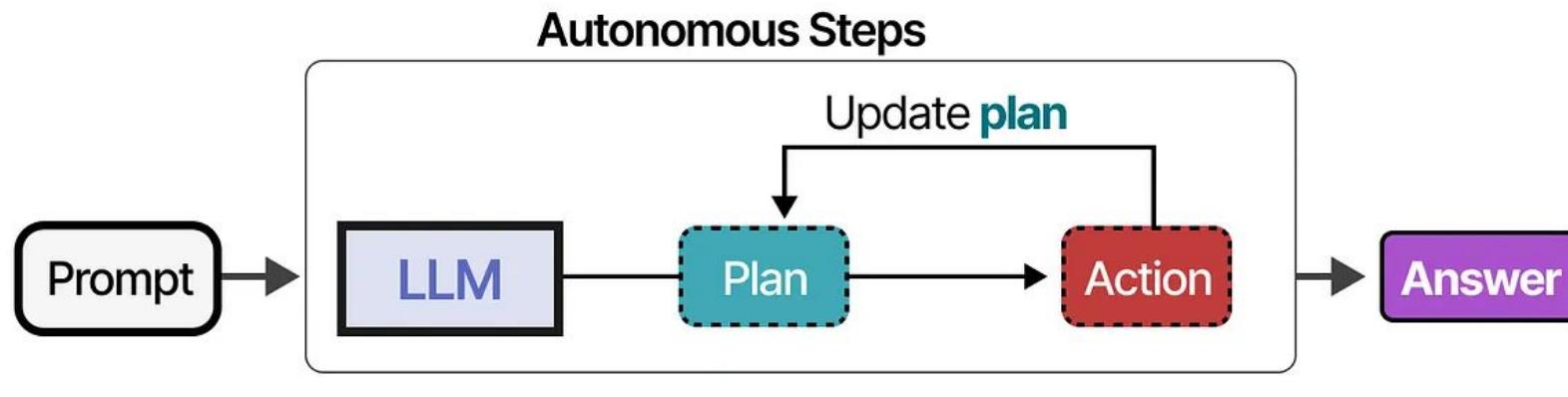


Less

Control  
(Autonomy)



More



Typically referred to as an **LLM Agent**,  
because of its **autonomous** capabilities.

\* Source: An Illustrated Guide to LLM Agents

**<https://github.com/nabeelxy/web-security-agents>**

## **Activity 1.0 Building a Basic Agent with ADK**



I cannot give a definitive "benign" or "malicious" verdict on paypal-login.com without further investigation. Here's why, and what factors I'd consider:

#### Why a direct answer is difficult:

- **Domain Similarity and Typosquatting:** The domain is extremely close to the real PayPal domain (paypal.com). This is a common tactic called "typosquatting" or "domain squatting" used by attackers. They rely on users mistyping the legitimate address.
- **Phishing Potential:** Even if the site isn't actively malicious right now, it could easily be used for phishing. A convincing fake login page could steal user credentials.
- **Dynamic Nature of Websites:** A website can be compromised. A domain that was once safe might later become malicious.

#### Factors I would investigate to make a determination:

##### 1. WHOIS Information:

- Who owns the domain? Is the registration information private? Is the registrant associated with PayPal in any way? A private or unrelated registration would be suspicious.
- How old is the domain? A newly registered domain, especially one mimicking a well-known brand, is a red flag.

##### 2. SSL Certificate:

- Does the site use HTTPS (SSL/TLS)? If not, that's a major problem. Even if it does, who issued the certificate? Is it a reputable Certificate Authority (like Let's Encrypt, DigiCert, etc.)? A self-signed certificate or one from an unknown CA is suspicious.

##### 3. Website Content:

- Does the page look exactly like the real PayPal login page? Even subtle differences can indicate a fake.
- What happens when you submit the form? Does it redirect to a suspicious URL? Does it throw an error? Does it redirect to the real PayPal site (after potentially stealing your credentials)?
- Are there any obvious signs of malware or suspicious scripts on the page?

##### 4. Reputation and Blacklists:

- Is the domain on any known blacklists (like Google Safe Browsing, Spamhaus, etc.)?
- What do other security vendors and researchers say about the domain? A quick search might reveal if it's been reported for phishing or other malicious activity.
- Use online tools like VirusTotal or URLScan.io to analyze the URL.

##### 5. Network Traffic Analysis:

- Examine the HTTP requests made by the page. Are they going to legitimate PayPal servers, or to suspicious third-party sites?

#### Recommendation:

- **Assume it is malicious.** Do not enter any credentials or personal information on paypal-login.com.
- **Always double-check the URL** before entering sensitive information online, especially for financial sites.
- **Access PayPal directly** by typing "paypal.com" into your browser's address bar, or using the official PayPal app.
- **Report the suspicious domain** to PayPal and any relevant security organizations.

In conclusion, due to the high potential for phishing and typosquatting, treat paypal-login.com as highly suspicious and avoid interacting with it. A thorough investigation would be required to determine its true nature.

LLM along was  
not able to figure  
out the answer

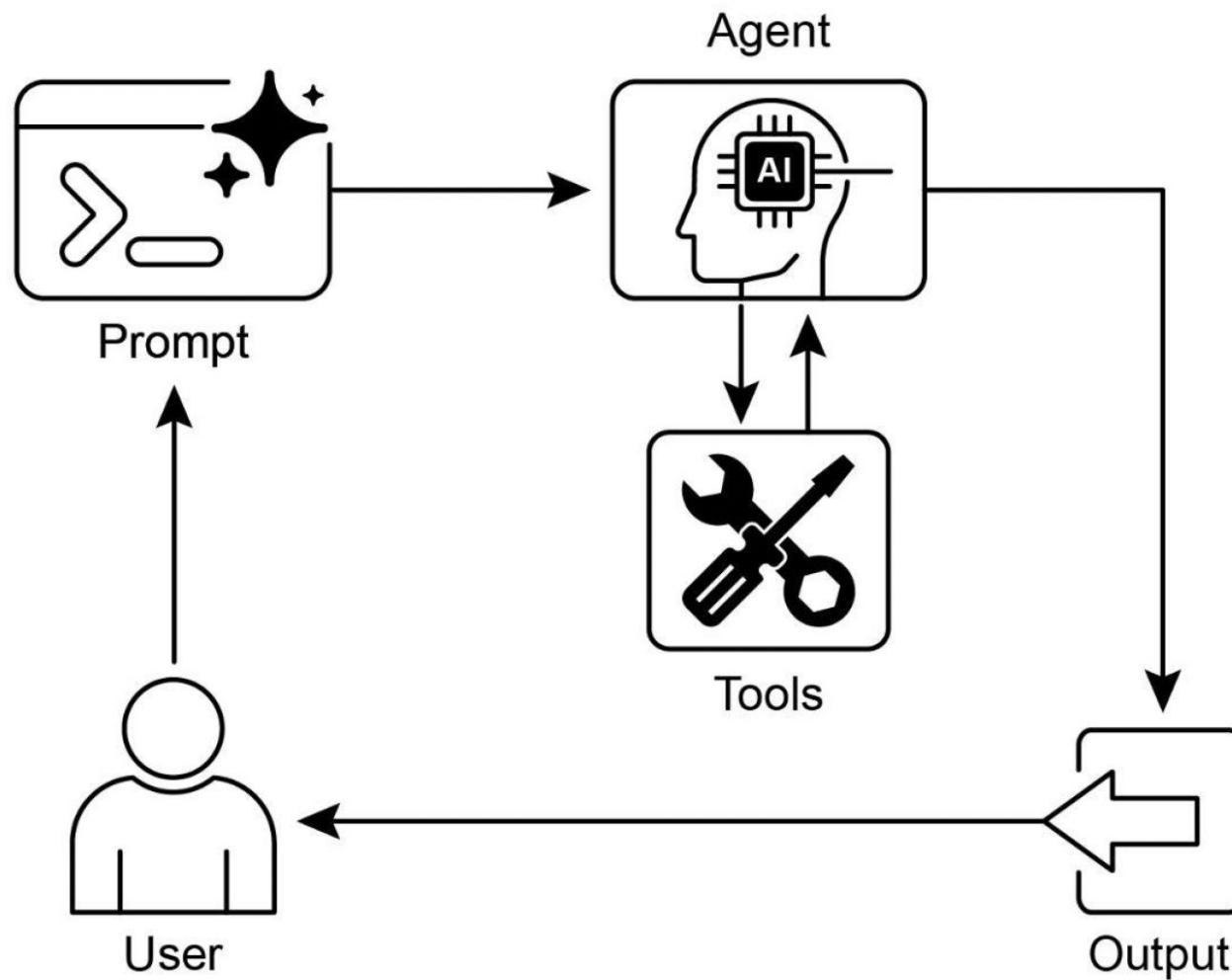


## **Activity 1.1 Building a Basic Agent with LangChain**

## **Activity 1.2 Building a Basic Agent with Ollama/ADK**

## **Activity 1.3 Building an Agent without adk web**

# Tool Calling Agents



\* Source: Agentic Design Patterns

## **Activity 2.0 Calling a Built-In Tool (google-search)**

google\_search  
tool makes an  
incorrect  
assertion

is paypal-login.com benign?



It is highly likely that "paypal-login.com" is **malicious** and used for phishing. Here's why:

- **Phishing Tactics:** Scammers often use spoofed emails that appear to be from PayPal to trick users into entering their passwords on fake websites. These emails often warn that your account is about to be suspended and ask you to log in to resolve the issue.
- **Generic Greetings:** Phishing emails commonly use impersonal greetings like "Dear User" instead of your name. Real PayPal emails will address you by your first and last name or business name.
- **Suspicious Links:** Phishing emails contain links that lead to fake websites designed to steal your login credentials. Always check the links in an email before clicking on them.
- **Sense of Urgency:** Phishing emails create a false sense of urgency, warning you to update your account immediately.
- **Requests for Personal Information:** PayPal will never ask for sensitive information like your password, bank account, or credit card number via email.
- **Invoice Scams:** Cybercriminals use legitimate-looking payment requests to hijack accounts. These scams utilize authentic-looking email addresses and URLs, making them difficult to detect.



#### How to protect yourself:

- **Check for Red Flags:** Be wary of generic greetings, suspicious links, and requests for personal information.
- **Log in Directly:** Always access your PayPal account by typing the URL into your browser or using the official app instead of clicking on links in emails.
- **Enable Two-Factor Authentication (2FA):** 2FA adds an extra layer of security, making it harder for attackers to access your account even if your credentials are compromised.
- **Report Suspicious Activity:** Forward any suspicious emails or websites to [phishing@paypal.com](mailto:phishing@paypal.com).
- **Monitor Your Account:** Regularly check your account statements and transaction history for any discrepancies.



paypal-login.com phishing

paypal-login.com security analysis

paypal-login.com scam

## **Activity 2.1 Calling a Custom Tool (whois tool)**

whois\_tool  
invocation  
helped to get the  
correct answer.

The screenshot shows a mobile application interface with a dark background. At the top right is a blue rounded rectangular button with white text that reads "is paypal-login.com benign?". To the right of this button is a small white circular icon containing a black user profile icon. In the center of the screen is a large grey rounded rectangular box containing text. On the left side of this box are three green circular icons, each with a white computer monitor and a small number (1, 2, or 3) inside. To the right of each icon is a white rounded rectangular button with a black outline. The first button contains the text "⚡ get\_domain\_whois\_v2". The second button contains the text "✓ get\_domain\_whois\_v2". The third button contains the text "The domain paypal-login.com is likely benign. The registrar is MarkMonitor, and the organization is PayPal Inc. Both of these facts suggest that the domain is brand protected." The overall layout is clean and modern, typical of a mobile security app's user interface.

is paypal-login.com benign?

1 ⚡ get\_domain\_whois\_v2

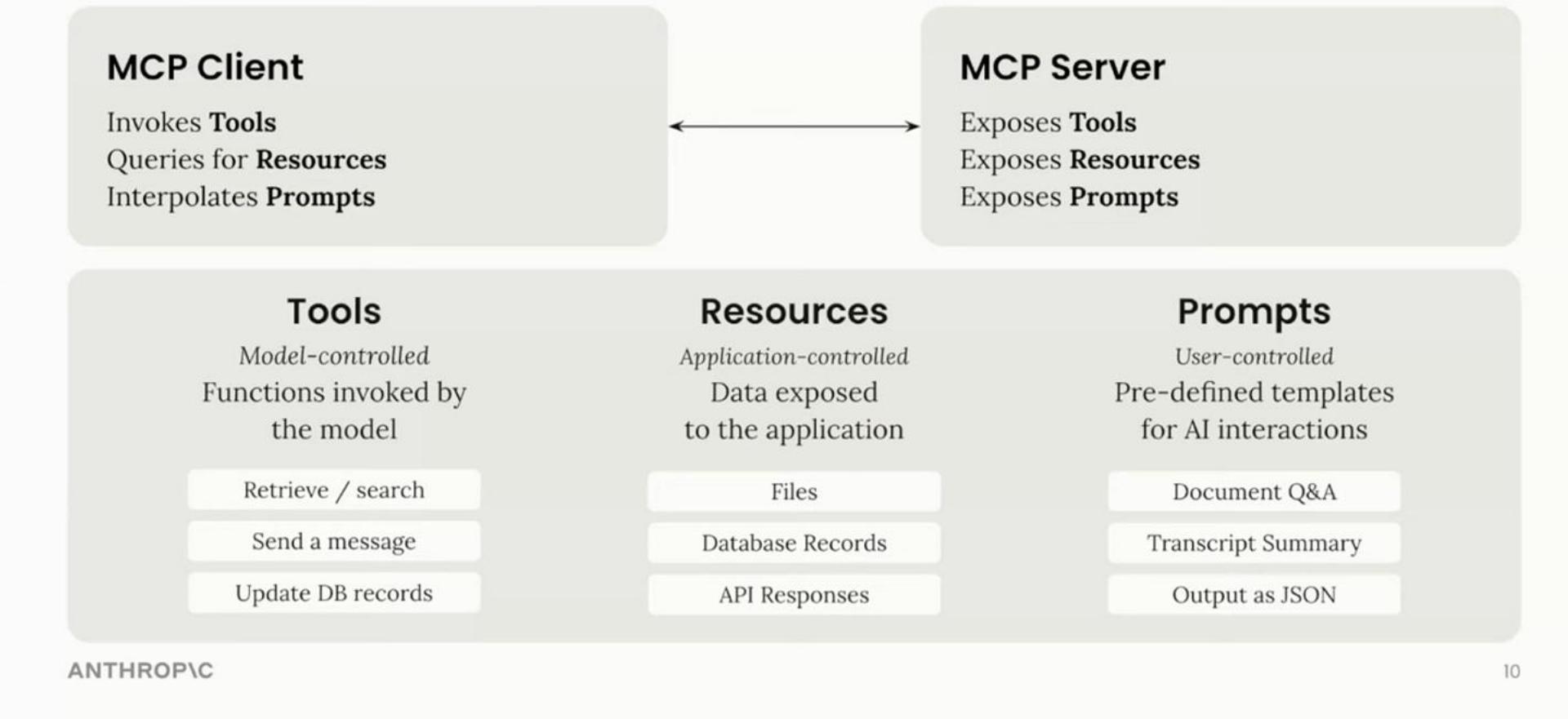
2 ✓ get\_domain\_whois\_v2

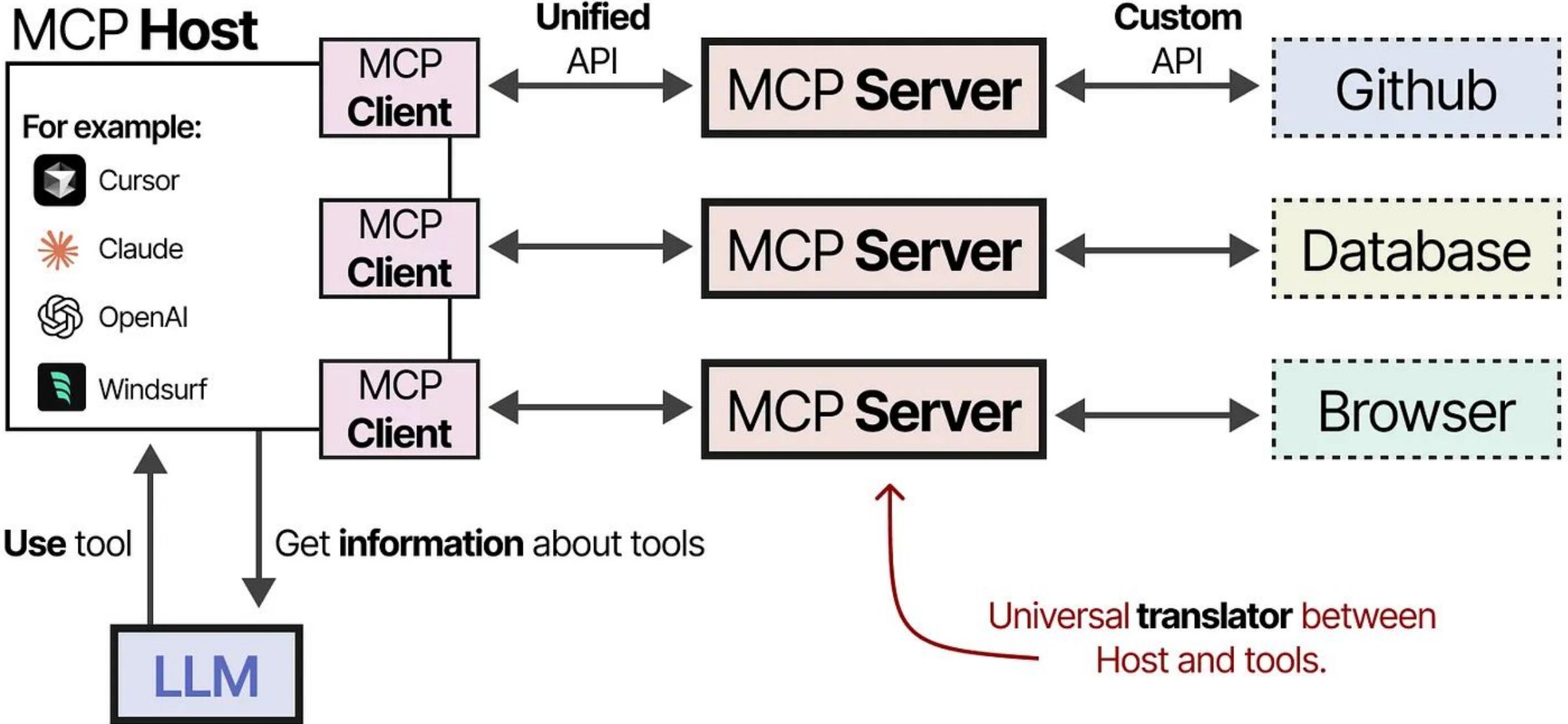
3 The domain paypal-login.com is likely benign. The registrar is MarkMonitor, and the organization is PayPal Inc. Both of these facts suggest that the domain is brand protected.

## **Activity 2.2 Calling a Custom Tool (whois tool/LangChain)**

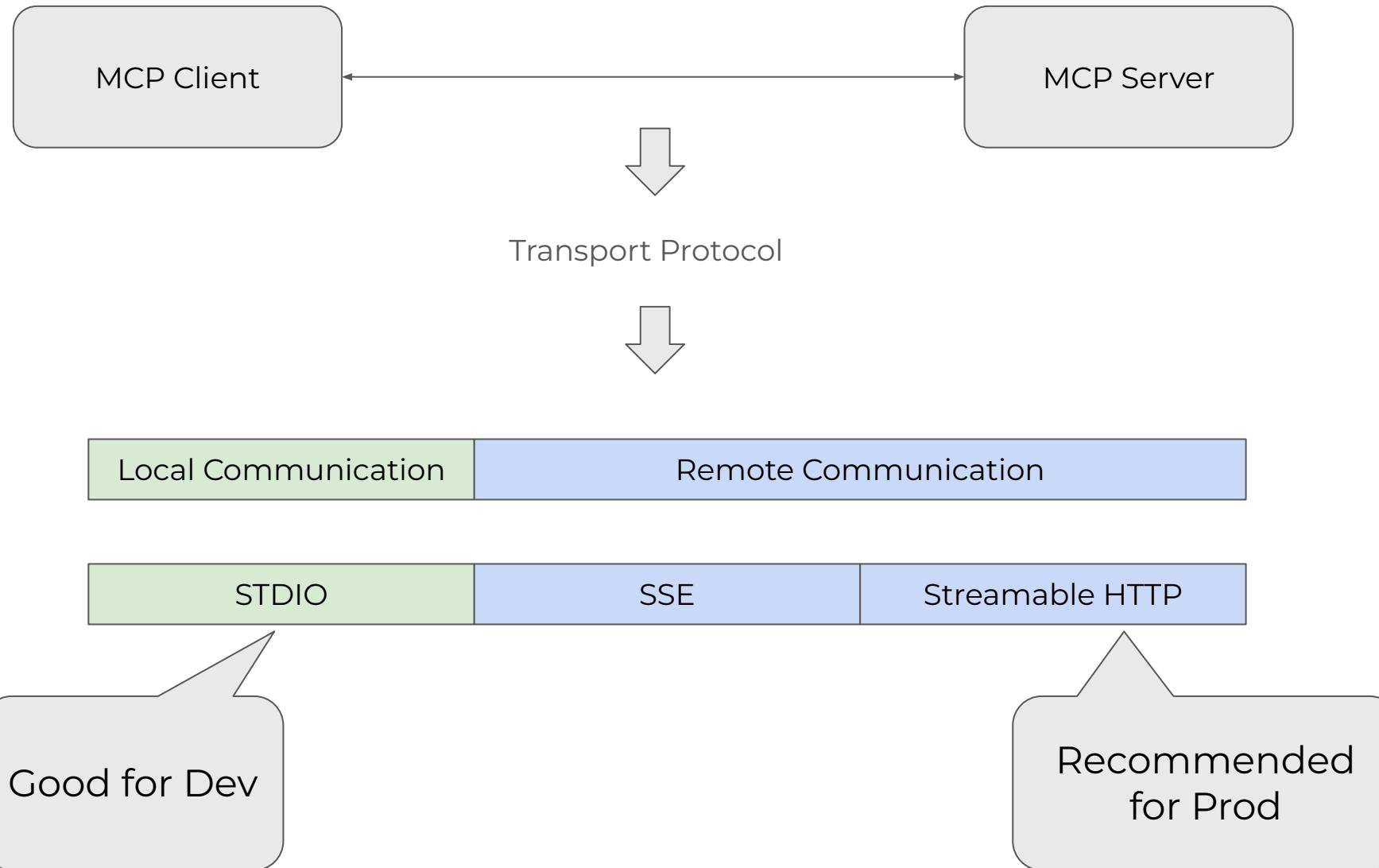
# MCP Tools (Model Context Protocol)

# MCP Deep-Dive





\* Source: An Illustrated Guide to LLM Agents



## **Activity 3.0 MCP with stdio**

## **Activity 3.0 Creating MCP with http**

The screenshot shows the Palo Alto Networks Agent Development Kit interface. On the left, there's a sidebar titled "Agent Development Kit" with a dropdown menu set to "103\_0\_mcp\_studio". Below it are tabs for "Trace", "Events", "State", "Artifacts", "Sessions", and "Eval". Under "Trace", there's a section titled "Invocations" containing a query: "is paypal-login.com benign?". This query has an "Invocation ID: e-735c351c-3f96-4386-b768-2a82aaa2d332" and a list of steps: "invocation" (10741.24ms), "invoke\_agent agent\_01..." (10740.55ms), "call\_llm" (7844.57ms), "execute\_tool get..." (5550.03ms), and another "call\_llm" (2393.63ms). On the right, the main pane shows a session with SESSION ID "3c7854da-ab1b-4990-8302-3a9d66951a1b". It displays two API invocations: "get\_domain\_whois\_mcp" (one successful, one failed) and a summary explaining why the domain is likely benign based on registrar and organization details.

SESSION ID 3c7854da-ab1b-4990-8302-3a9d66951a1b

Token Streaming | + New Session |

is paypal-login.com benign?

is paypal-login.com benign?

Invocation ID:  
e-735c351c-3f96-4386-b768-2a82aaa2d332

|→ invocation 10741.24ms

|→ invoke\_agent agent\_01... 10740.55ms

call\_llm 7844.57ms

|→ execute\_tool get... 5550.03ms

call\_llm 2393.63ms

⚡ get\_domain\_whois\_mcp

✓ get\_domain\_whois\_mcp

The domain 'paypal-login.com' is likely benign. Here's why:

- Registrar: The registrar is MarkMonitor, which is known for protecting brands.
- Organization: The registrant organization is "PayPal Inc.", a reputable company.

## **Activity 3.0 Creating MCP with http**

The screenshot shows the Palo Alto Networks Agent Development Kit (ADK) interface. On the left, a sidebar displays a session named "103\_0\_mcp\_http". The main area is titled "SESSION ID c760086c-1bc4-4160-a6bc-508255c74f97". The top right features a "Token Streaming" toggle, a "+ New Session" button, and a user profile icon.

The central part of the interface is titled "Invocations" and contains a query: "is paypal-login.com benign?". Below this, an "Invocation ID" is shown: "e-5eb5b9dd-467a-403e-8b9d-59c4ad38d70c". A detailed call stack is displayed:

- Invocation → invocation (4800.49ms)
- Invocation → invoke\_agent agent\_01... (4798.94ms)
- call\_llm (3181.77ms)
  - Invocation → execute\_tool get\_... (542.1ms)
  - call\_llm (1540.78ms)

To the right of the call stack, a summary states: "The registrar of the domain paypal-login.com is MarkMonitor, Inc. and the organization is PayPal Inc. Therefore, the domain is likely to be benign."

## **Activity 3.1 Building VT MCP Server**

 Agent Development Kit

SESSION ID 566db3a9-d99b-4ddd-b204-82a52010fee0

Token Streaming | + New Session | 

103\_1\_vt\_mcp

Trace Events State Artifacts Sessions Eval

Invocations

Investigate paypal-verify.com

Invocation ID: e-538bbe8-a855-478e-a62c-f7c4ed5886de

|→ invocation 6897.96ms  
|→ invoke\_agent agent\_03... 6896.80ms  
call\_llm 4483.11ms  
|→ execute\_tool vt\_q... 3175.16ms  
call\_llm 1935.41ms

What scanners flagged it?

Invocation ID: e-ac12b633-02d8-4c93-8589-deb548a0d6a3

|→ invocation 2390.05ms  
|→ invoke\_agent agent\_03... 2388.82ms  
call\_llm 1962.97ms

 vt\_query  
 vt\_query

Based on the VirusTotal report, 10 out of 98 scanners flagged paypal-verify.com as malicious, specifically as a phishing site. Several scanners like Forcepoint ThreatSeeker, Fortinet, Kaspersky, and Sophos have identified it as "phishing". Therefore, it is highly likely that paypal-verify.com is a phishing website and should be treated with caution.

What scanners flagged it?

Based on the VirusTotal report, the following scanners flagged paypal-verify.com as malicious:

- Criminal IP
- ESET
- Forcepoint ThreatSeeker
- Fortinet
- Kaspersky
- Lionic
- Seclookup
- Sophos
- Webroot
- alphaMountain.ai

## **Activity 3.2 Building UrlScan MCP Server**

**Agent Development Kit**

SESSION ID e6e5188f-e94a-4759-bb48-e263221febcd

Token Streaming | + New Session |

Analyze the screenshot of [inslagarm.com](#) to decide if it is malicious?

Invocations

Analyze the screenshot of [inslagarm.com](#) to decide if it is malicious?

Invocation ID: e-14d787cc-a17d-4bb6-912f-501ec86a6c9d

- |→ invocation 48332.69ms
- |→ invoke\_agent agent\_03... 48330.81ms
- call\_llm 39068.07ms
- |→ execute\_tool urls... 36862.52ms
- call\_llm 7967.67ms

Can you also analyze the crawling report for this?

Invocation ID: e-aa623e51-3437-461a-80cf-77f7b7619561

- |→ invocation 36066.15ms
- |→ invoke\_agent agent\_03... 36064.23ms
- call\_llm 21678.79ms
- |→ execute\_tool urls... 17618.87ms
- call\_llm 13072.46ms

Based on the screenshot analysis, **inslagarm.com is malicious**. It is a phishing site designed to steal your Instagram and Facebook credentials.

Here are the key findings from the analysis:

- **Convincing Replica:** The page is a high-fidelity clone of the Instagram login page, making it difficult to spot as a fake at first glance.
- **Future Copyright Date:** The copyright year is listed as "© 2025". Legitimate websites use the current year, and a future date is a strong indicator of a fraudulent site.
- **Language Inconsistency:** The page mixes English and French, which is a sign of a poorly constructed and untrustworthy website.
- **Credential Theft:** The ultimate goal of this page is to trick you into entering your login credentials, which will then be stolen by the attackers.

Due to these factors, it is highly recommended to avoid this website and never enter any personal information.

Can you also analyze the crawling report for this?

urlscan\_screenshot\_analysis

✓ urlscan\_screenshot\_analysis

Of course, here is an analysis of the crawling report for [inslagarm.com](#):

The crawling report further confirms that **inslagarm.com is a malicious phishing website**. Here are the key technical details from the report that lead to this conclusion:

**Page and Server Information:**

- **Page Title:** The page title is "Instagram", which is a deceptive practice as the domain is not the official [instagram.com](#).
- **IP Address:** The website is hosted at the IP address 209.200.246.226.
- **Hosting Location:** The server is located in Toronto, Canada.
- **ASN:** The Autonomous System Number (ASN) is AS215311, which belongs to "REGXA-CLOUD Regxa Company for Information Technology Ltd, IQ" registered in Iraq. This is highly suspicious for a website masquerading as Instagram.

**Security Verdicts:**

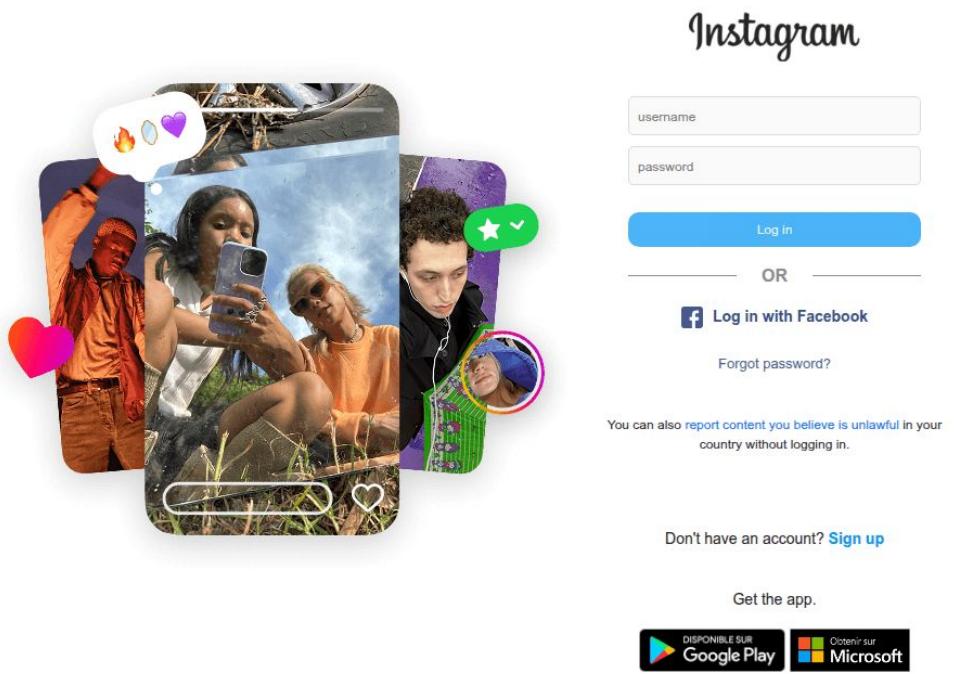
- **Malicious Verdict:** The UrlScan engine has flagged the website as **malicious**.

**Deceptive Tactics:**

## Behind the scene, screenshot analysis results

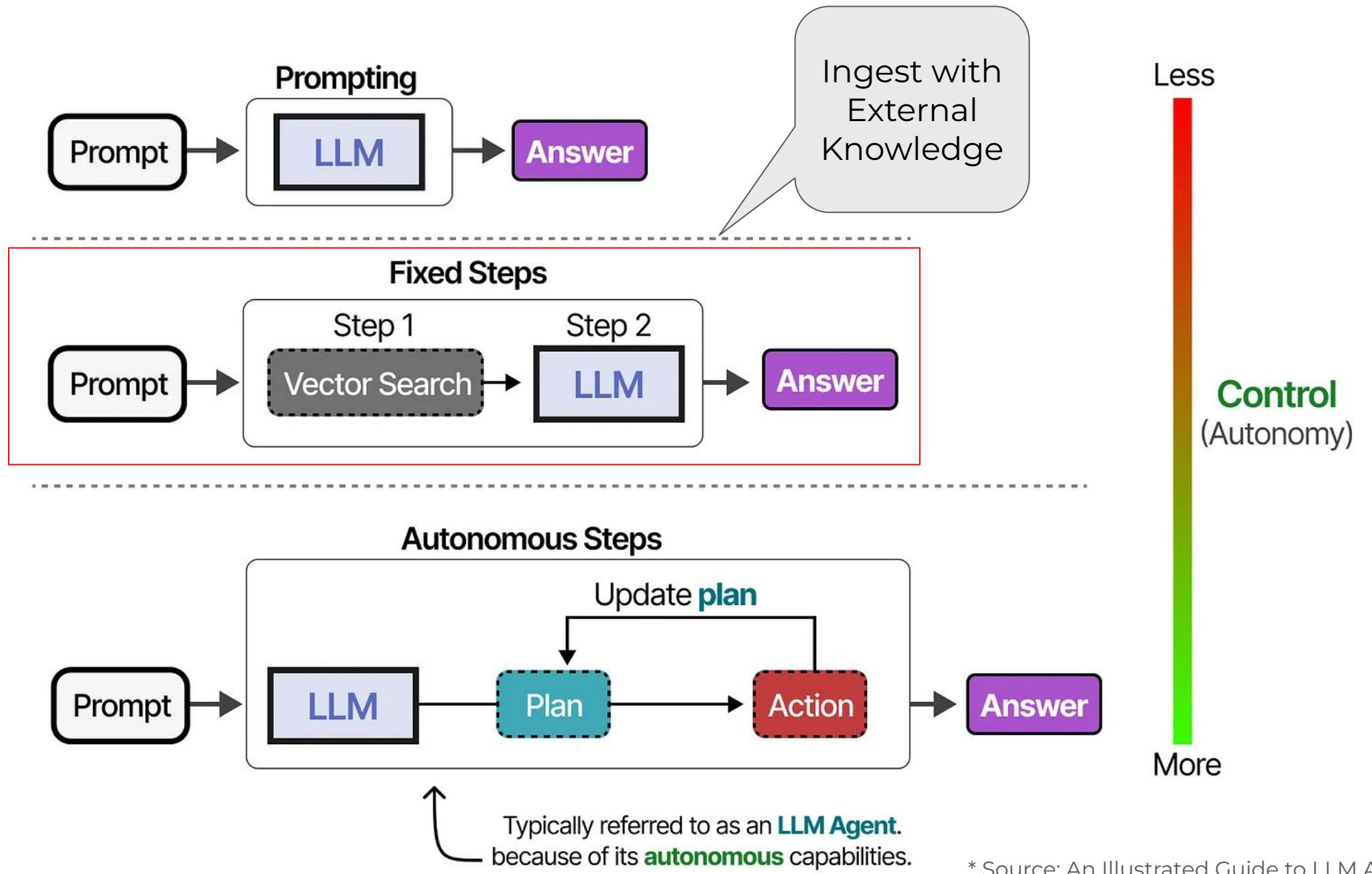
```
{  
    "brands": [  
        "Instagram",  
        "Facebook",  
        "Meta",  
        "Google Play",  
        "Microsoft"  
    ],  
    "security_risks": [  
        "***Potential Phishing Attack:** The page is designed to harvest user credentials (username and password). If this is not the official Instagram domain, any information entered will be compromised.",  
        "***Future Copyright Year:** The copyright notice shows the year 2025, which is in the future. This is a common mistake on hastily created phishing pages and a strong indicator of illegitimacy.",  
        "***Language Inconsistency:** The primary language of the page is English, but the text on the app store buttons ('DISPONIBLE SUR' and 'Obtenir sur') is French. Legitimate company websites are typically consistent in their language presentation.",  
        "***Malicious Downloads:** The 'Get the app' buttons for Google Play and Microsoft could lead to malicious software (malware, spyware) downloads instead of the official Instagram app."  
    ],  
    "lure": "The page impersonates the official Instagram login page to trick users into entering their username and password. By replicating the familiar design and branding, it leverages the user's trust in the Instagram brand to steal their account credentials.",  
    "suspicious_elements": [  
        {  
            "element": "Copyright Year",  
            "reason": "The footer displays '© 2025 Instagram from Meta'. A future copyright date is a significant red flag for a fraudulent website."  
        },  
        {  
            "element": "App Store Buttons",  
            "reason": "The buttons for Google Play and Microsoft contain French text ('DISPONIBLE SUR' / 'Obtenir sur') while the rest of the page is in English. This linguistic inconsistency is highly suspicious."  
        }  
    ],  
    "additional_observations": "The overall visual design is a very close replica of the legitimate Instagram web login page, making it highly deceptive. The absence of a browser address bar in the screenshot prevents verification of the URL, which is the most critical element for identifying a phishing site. However, the combination of the future copyright date and the language mismatch provides strong evidence that this is a phishing page."  
}
```

The screenshot analyzed

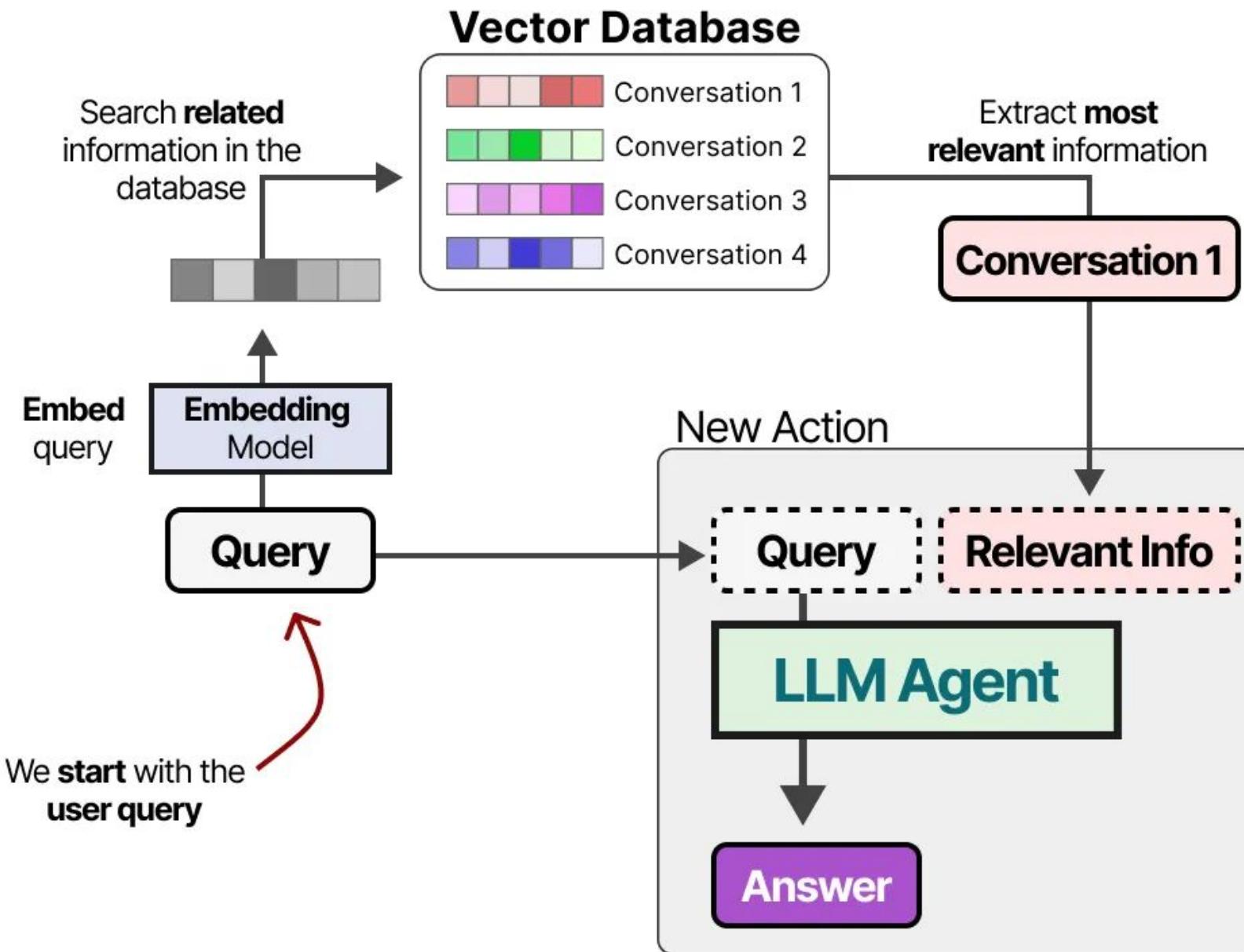


[Meta](#) [About](#) [Blog](#) [Jobs](#) [Help](#) [API](#) [Privacy](#) [Cookie Settings](#) [Terms](#) [Locations](#) [Instagram Lite](#) [Meta AI](#) [Meta AI Articles](#) [Threads](#) [Contact](#) [Uploading & Non-Users](#) [Meta Verified](#)  
English © 2025 Instagram from Meta

# RAG (Retrieval Augmented Generation)



\* Source: An Illustrated Guide to LLM Agents



\* Source: An Illustrated Guide to LLM Agents

## **Activity 4.0 Building RAG Agent to Talk to Threat Reports**



ADVERSARY

## Kimsuky | Ongoing Campaign Using Tailored Reconnaissance Toolkit



PLATFORM ▾ SOLUTIONS ▾ COMPANY ▾ RESEARCH ▾ RESOURCES ▾ BLOG

### Scattered Spider: Still Hunting for Victims in 2025

THREAT

April 8, 2025

Advertisement

**knowbe4**  
One Platform, Complete Visibility.  
Train Humans and AI

SUPERCHARGE  
PRODUCTIVITY NOW



# Krebs on Security

In-depth security news and investigation



HOME ABOUT THE AUTHOR ADVERTISING/SPEAKING

## Phishers Spoof USPS, 12 Other Nat'l' Postal Services



Hunt.io

Product ▾

Features ▾

Pricing New!

Blog

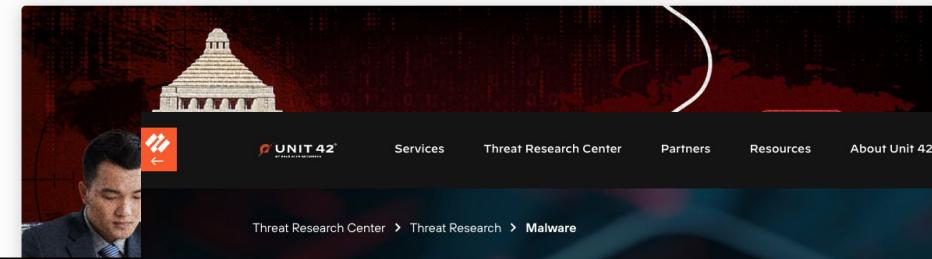
About

Login

Get Started

## Multilingual ZIP Phishing Campaigns Targeting Financial and Government Organizations Across Asia

Published on Oct 29, 2025



UNIT 42<sup>TM</sup>

of CYBERSECURITY

Services

Threat Research Center

Partners

Resources

About Unit 42



Under Attack?

English ▾

Threat Research Center > Threat Research > Malware

### MALWARE

## The Smishing Deluge: China-Based Campaign Flooding Global Text Messages

11 min read

### RELATED PRODUCTS

Advanced DNS Security

Advanced URL Filtering

 Agent Development Kit

SESSION ID 853cac3d-aa52-4e17-aa0f-c71e85809904

Token Streaming | + New Session | 

104\_rag

Trace Events State Artifacts Sessions Eval

Invocations

Can you update me on the recent developments on scattered spider campaign including IOCs, TTPs and any interesting tidbits.

⚡ retrieval\_augmented\_generation

✓ retrieval\_augmented\_generation

散布蜘蛛 (Scattered Spider) campaigns are actively using multi-lingual phishing lures. They are using domains such as gjqygs[.]cn, zxp0010w[.]vip, and jppjz1[.]cc for command and control and hosting infrastructure. They are targeting Traditional Chinese, Japanese, and English speakers with lures related to tax invoices, import/export declarations, and salary revisions. The attackers use techniques like spearphishing links (T1566.002) to deliver malicious files (T1204.002) and gather victim identity information (T1592) through open websites and domains (T1593). They also acquire infrastructure such as domains (T1583.001) and compromise web services (T1584.001) for phishing.

Can you provide a list of IOCs?

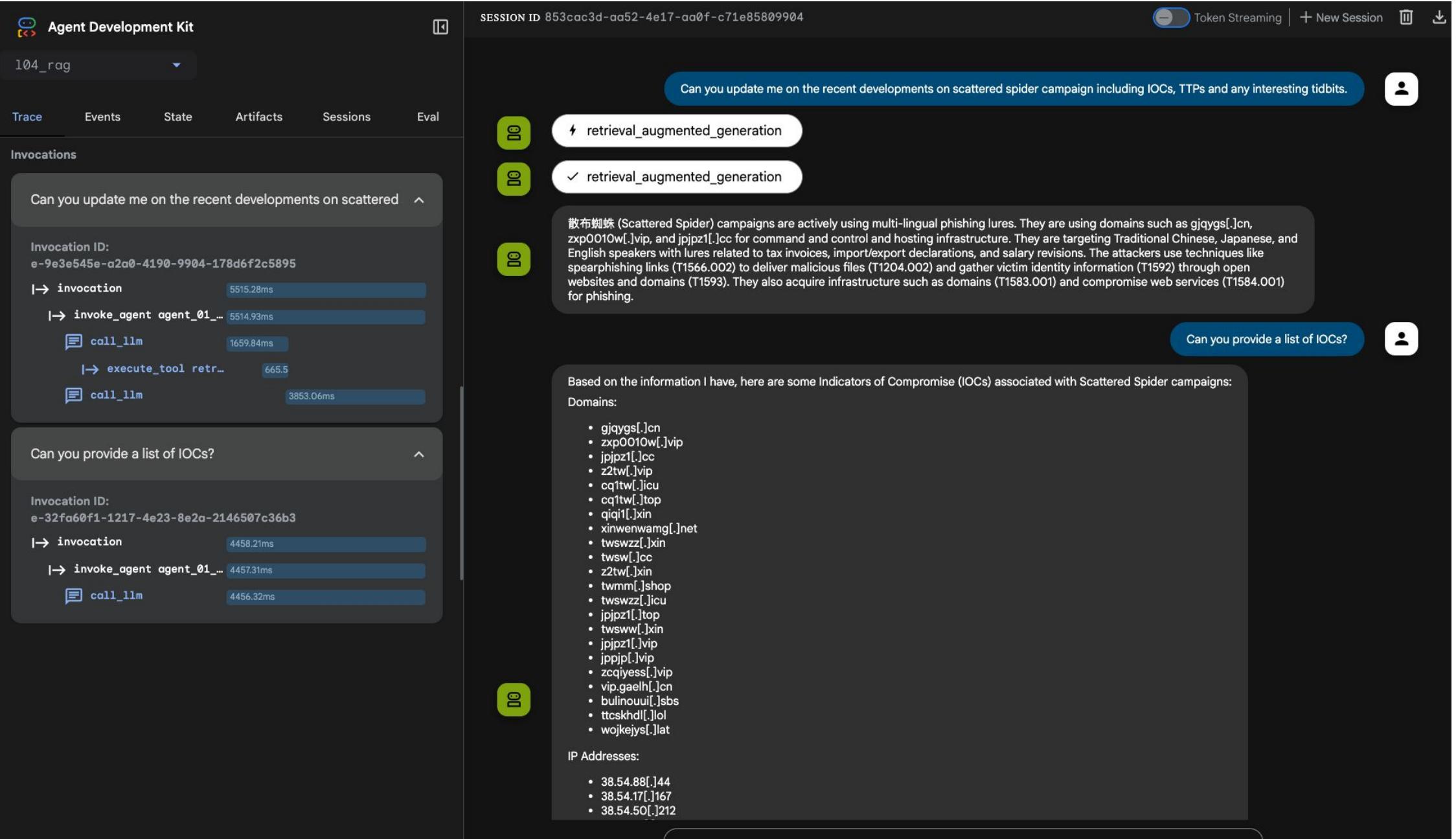
Based on the information I have, here are some Indicators of Compromise (IOCs) associated with Scattered Spider campaigns:

Domains:

- gjqygs[.]cn
- zxp0010w[.]vip
- jppjz1[.]cc
- z2twl[.]vip
- cq1tw[.]icu
- cq1tw[.]top
- qiqi1[.]xin
- xinwenwamg[.]net
- twswzz[.]xin
- twsw[.]cc
- z2twl[.]xin
- twmm[.]shop
- twswzz[.]icu
- jppjz1[.]top
- twsww[.]xin
- jppjz1[.]vip
- jppjp[.]vip
- zcqiyess[.]vip
- vip.gaelh[.]cn
- bulinouui[.]sbs
- ttcskhdl[.]lol
- wojkejys[.]lat

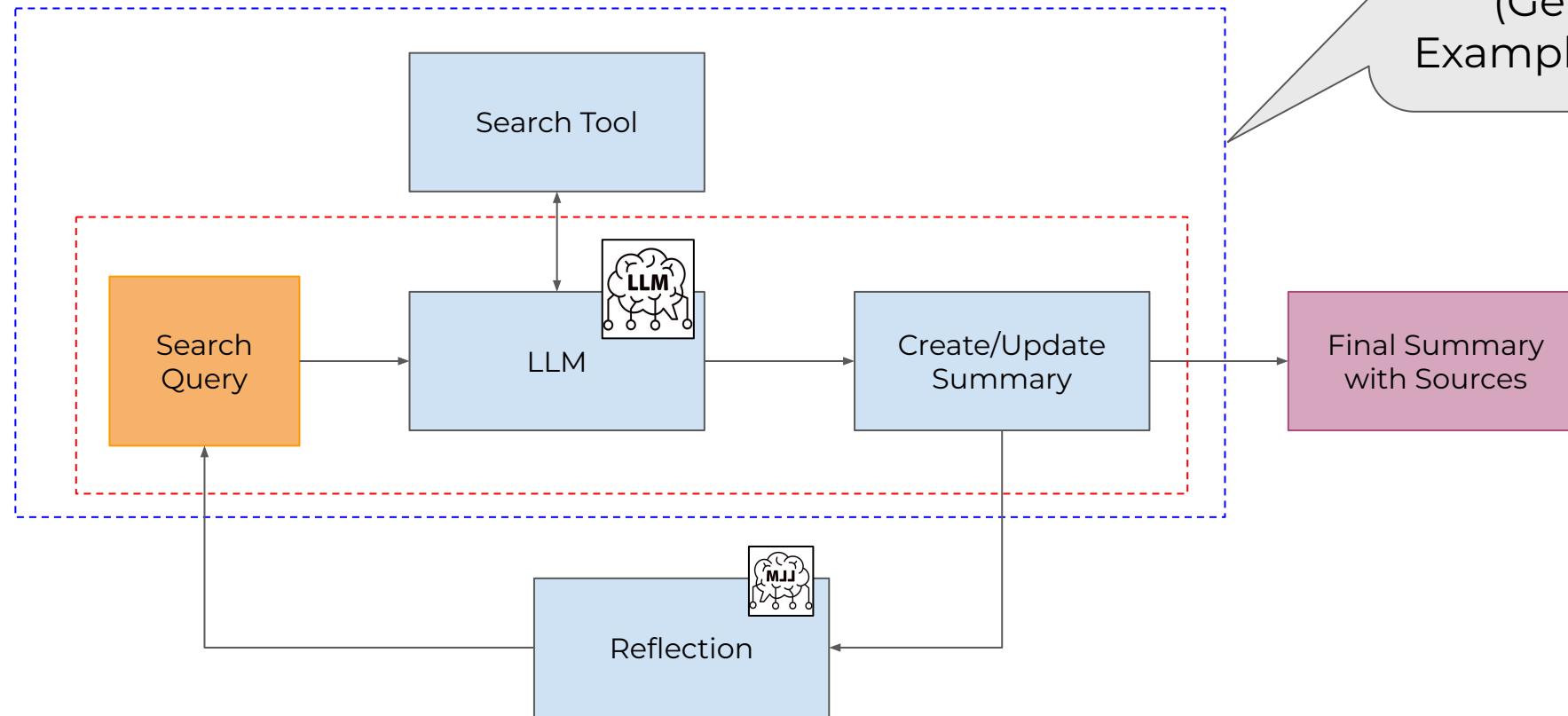
IP Addresses:

- 38.54.88[.]44
- 38.54.17[.]167
- 38.54.50[.]212

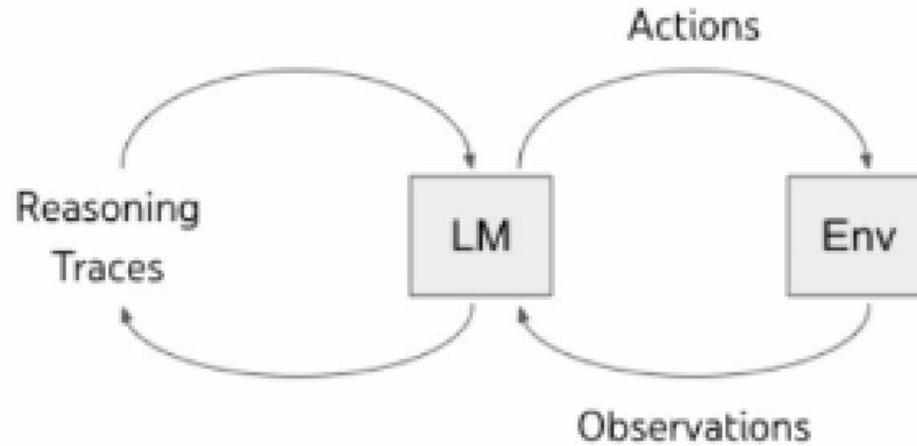


# Understanding Multi-Agents

Reflection  
Design Pattern  
(Gemma  
Example Earlier)

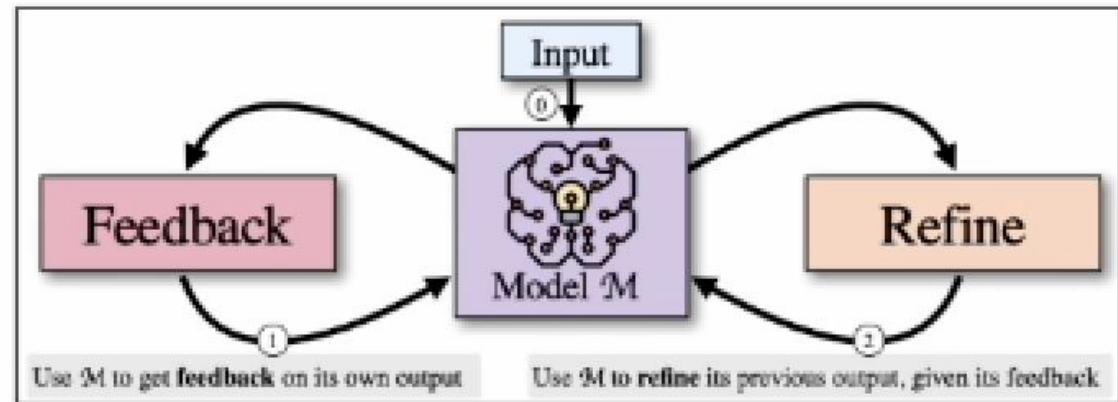


# Commonly Used Simple Agent Planning Patterns



ReAct (Reason + Act)

REACT: SYNERGIZING REASONING AND ACTING IN LANGUAGE MODELS

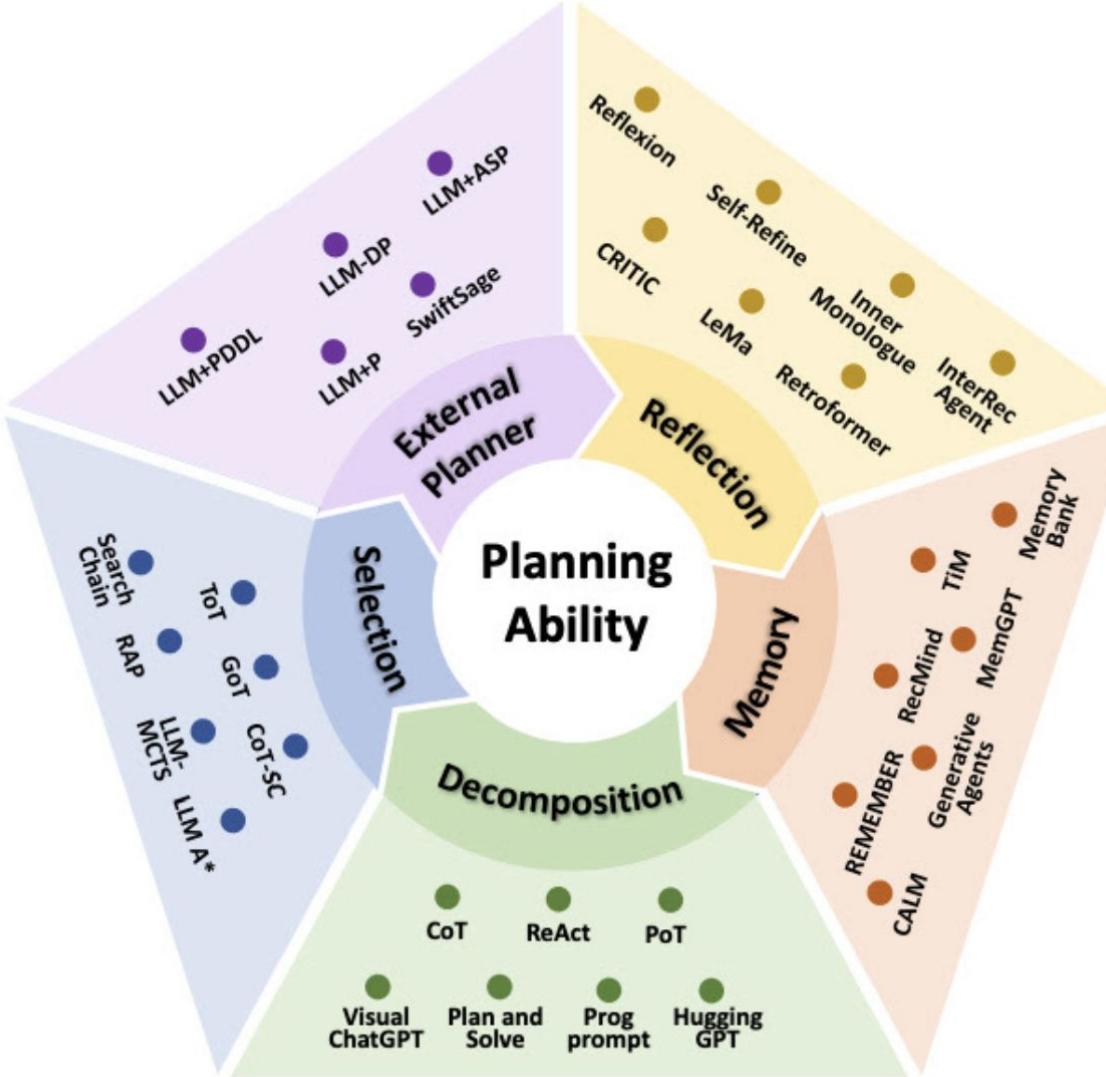


SELF-REFINE:  
Iterative Refinement with Self-Feedback

\*ReAct: Synergizing Reasoning and Acting in Language Models, <https://arxiv.org/pdf/2210.03629.pdf>, ICLR 2023

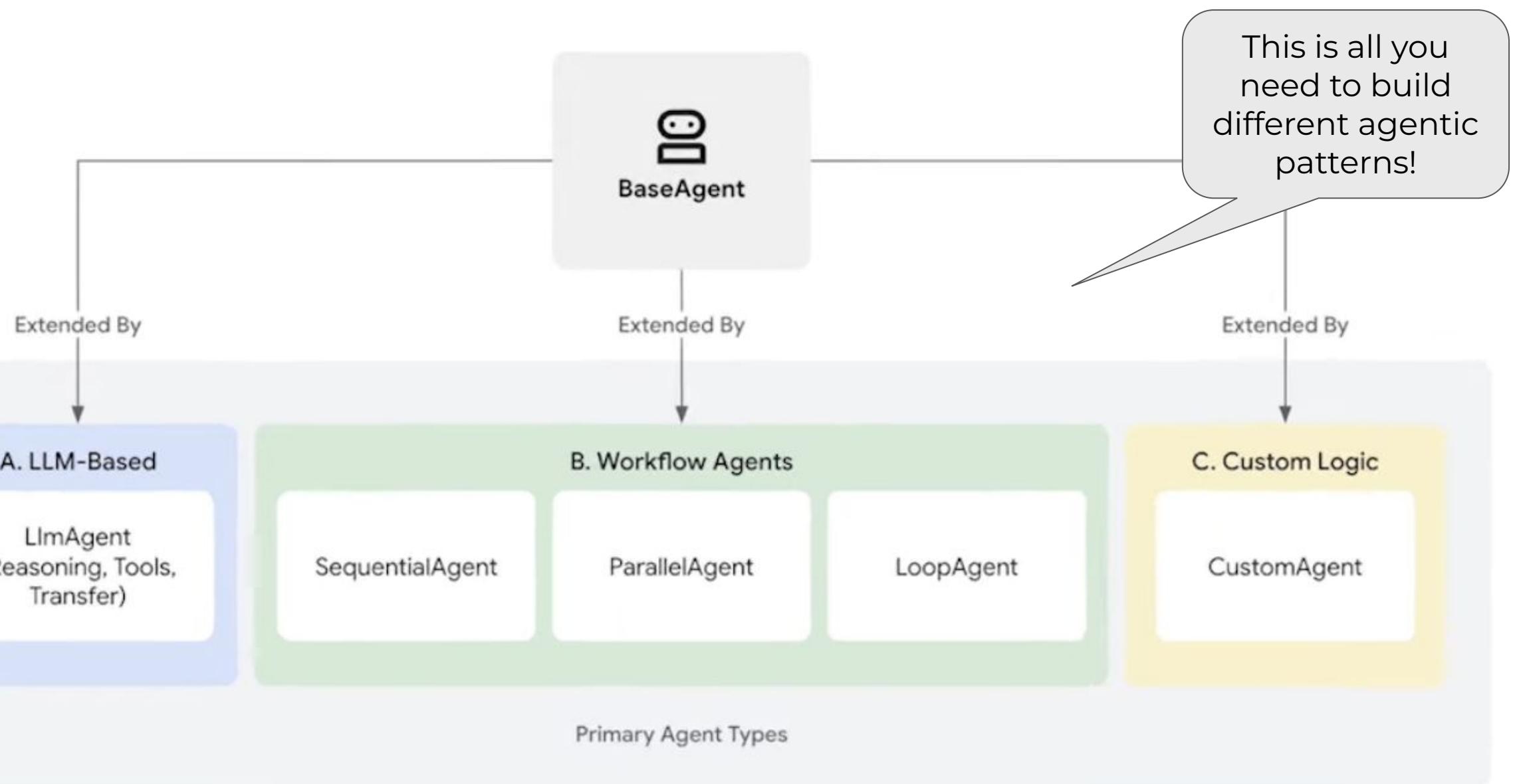
\*SELF-REFINE: Iterative Refinement with Self-Feedback, <https://arxiv.org/pdf/2303.17651.pdf>, NeurIPS Poster 2023

# Agent Planning

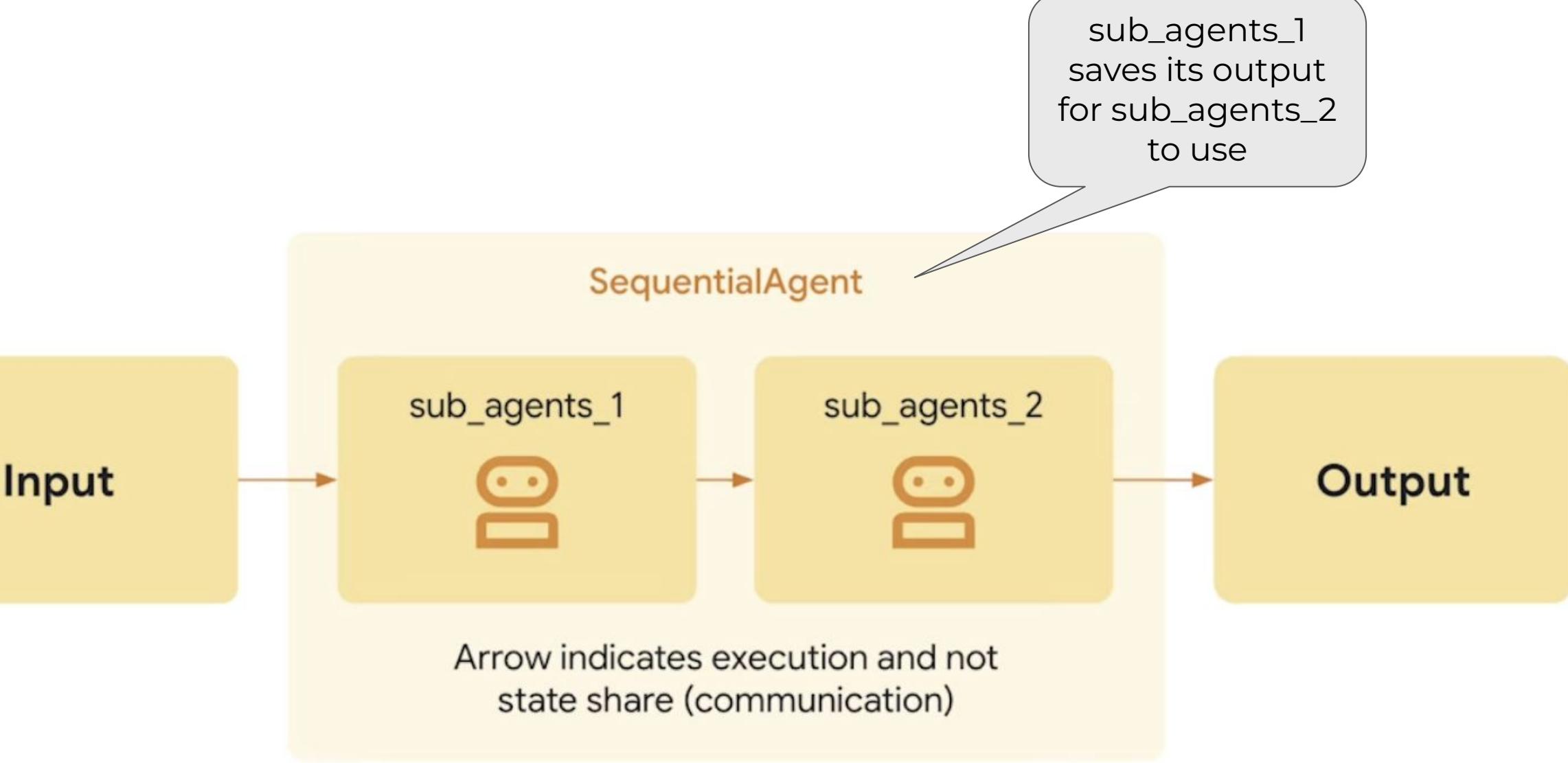


\* Understanding planning of LLM agents: a survey, arXiv, Feb 2025, <https://arxiv.org/pdf/2402.02716>

# Understanding Multi-Agents

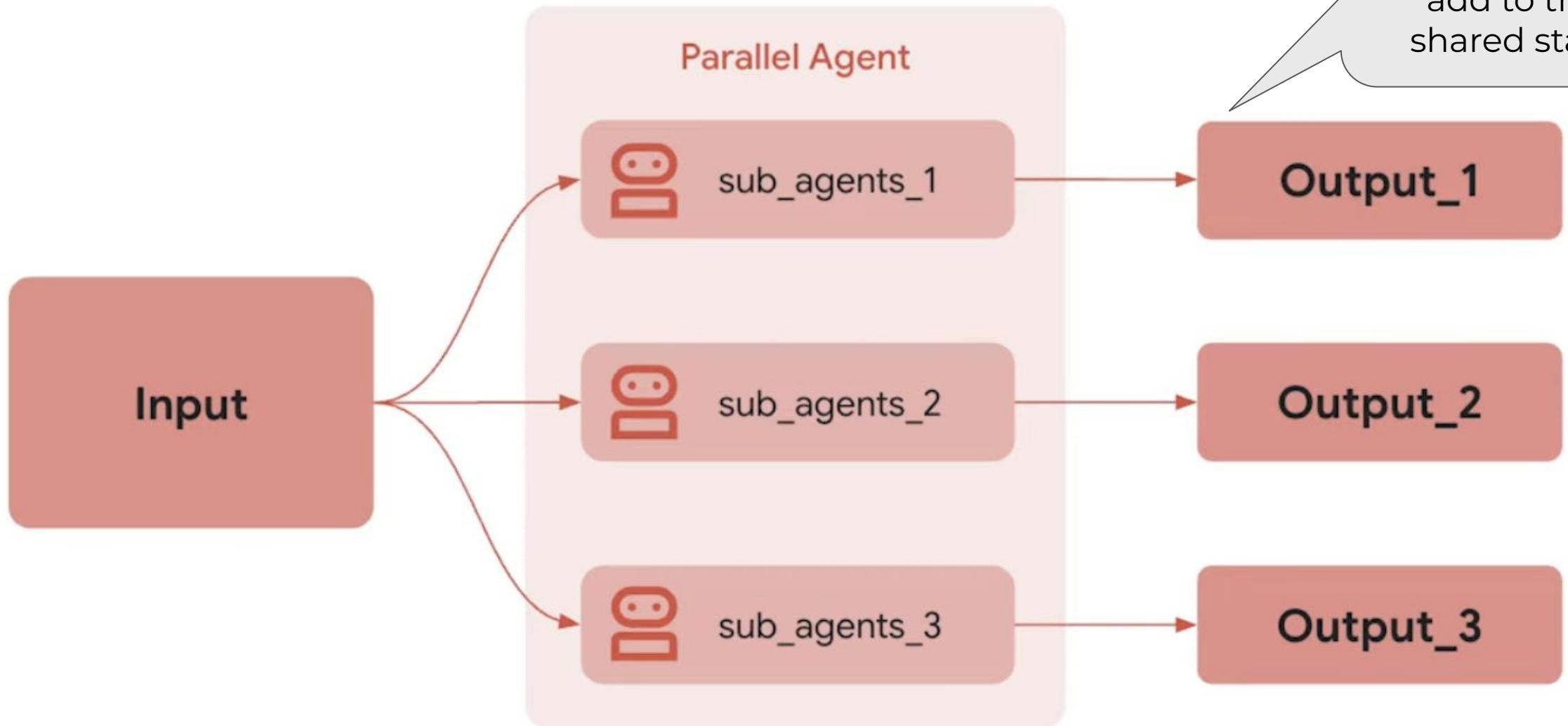


\* Source: Google Official Site



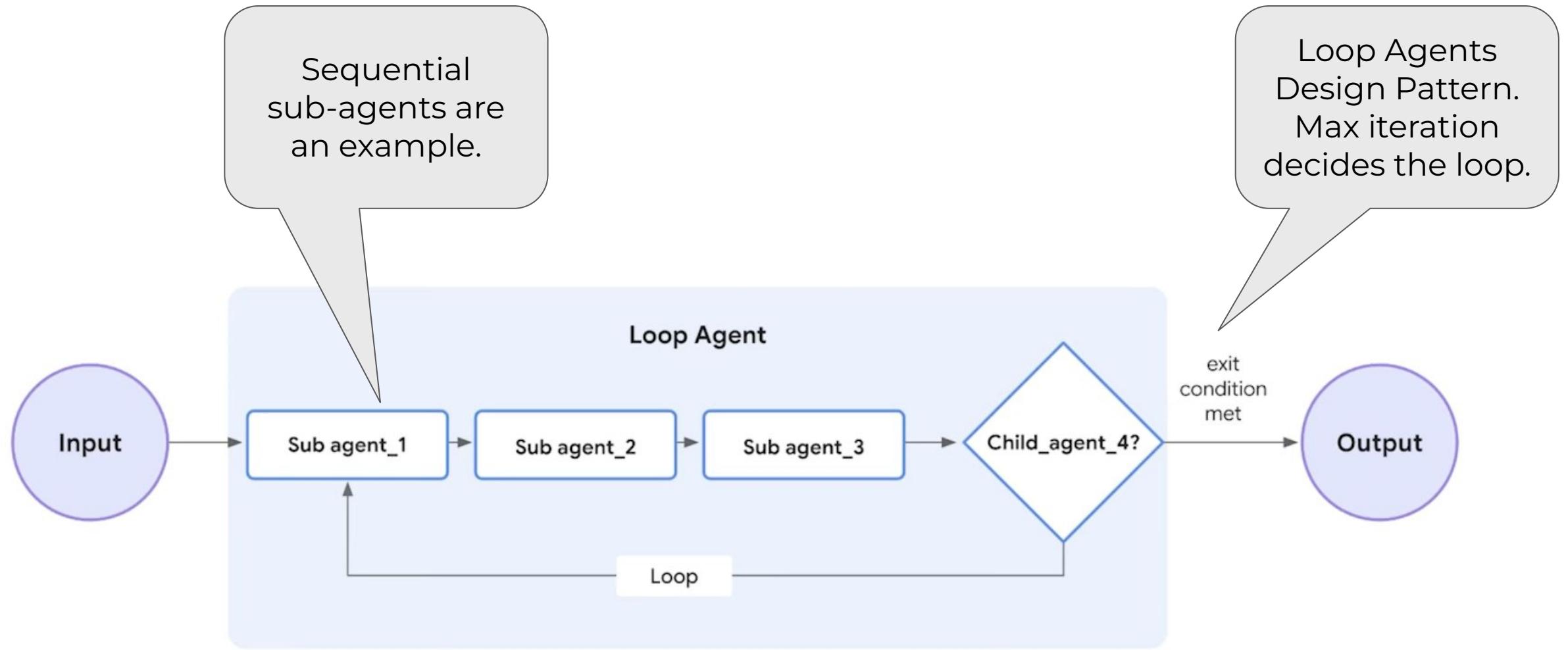
\* Source: Google Official Site

## **Activity 5.1 Building Sequential Agents**



\* Source: Google Official Site

## **Activity 5.2 Building Parallel Agents**

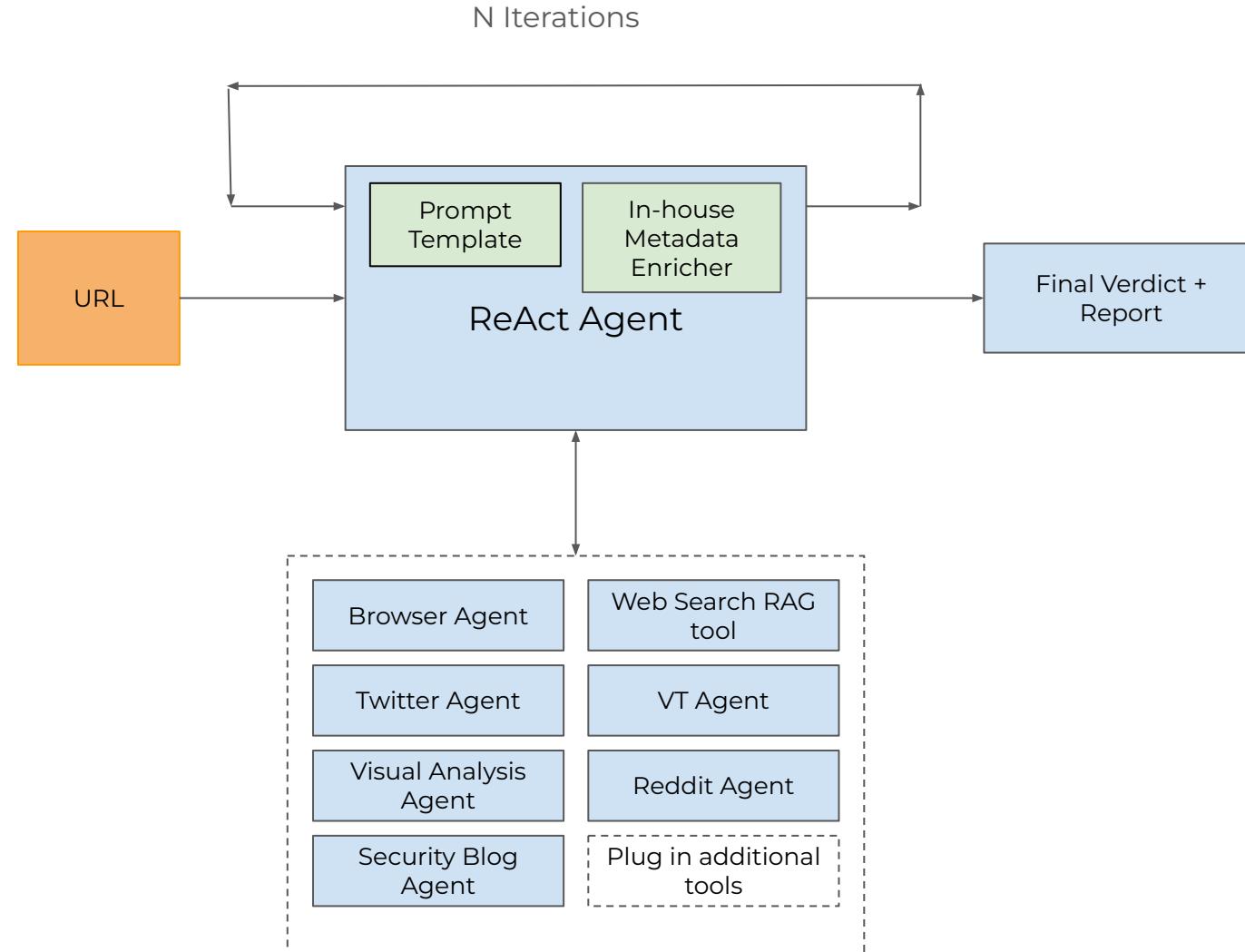


\* Source: Google official documentation

## **Activity 5.3 Building Loop Agents**

# Web Security Agent

# Web Security Agent



## Invocations

I got the following domain to analyze reputationrescue.info call ↻

Invocation ID:  
e-689b6af1-db73-4fb8-b9d1-e9f1b11c2b4d

```
|→ invocation          128396.98ms
|→ invoke_agent web_secur... 128396.10ms
  ↳ call_llm           80025.27ms
    |→ execute_tool get_... 1115.78ms
    |→ execute_tool get_... 2.06ms
    |→ execute_tool (mer... 0.16ms
  ↳ call_llm           21593.71
    |→ execute_tool vt_q... 22341.99ms
    |→ execute_tool urlsc... 16097
    |→ execute_tool (mer... 0.05ms
  ↳ call_llm           49898.61ms
    |→ execute_tool urlsc... 41843.14ms
  ↳ call_llm           19346.0
    |→ execute_tool goog...
      |→ invocation
        |→ invoke_agent...
          ↳ call_llm
            |→ execute_tool goog...
              |→ invocation
                |→ invoke_agent...
                  ↳ call_llm
                    |→ execute_tool retr...
                      |→ execute_tool get_...
                        |→ call_llm
                          |→ execute_tool retr...
                            |→ execute_tool get_...
                              |→ call_llm
                                |→ execute_tool retr...
```

I got the following domain to analyze reputationrescue.info can you please investigate it and give me an executive summary and final verdict on it?

## Thought

Alright, here's the plan. The user, bless their heart, wants me to dig into reputationrescue.info. They want an executive summary and a verdict – no pressure, right? Luckily, this is precisely what I do. I'll approach this systematically, as always.

First, I'll hit the ground running with some preliminary domain intelligence. Gotta understand the basics: who registered it, when, and any initial flags. So, wham, I'll immediately launch `get_domain_whois_mcp` and `get_domain_reputation_info` in parallel. Time is precious.

Then, I'll dive into the juicy stuff: the web security analysis. I'll get the lay of the land, checking for red flags like malware or phishing attempts. I will do this in the following order.

- I'll fire up `vt_query`, to check the VirusTotal reports
- Then I'll run `urlscan_query`, so I can see a general UrlScan report
- Next up, I'll analyze the visual aspects of the website with `urlscan_screenshot_analysis`.

After that, I'll leverage Google. It's an invaluable source of information, so I'll put `google_search_agent` to work, searching for reviews, news, and any chatter surrounding the domain.

Finally, I'll consult the real brain trust: our internal threat intelligence database. I'll use `retrieval_augmented_generation` to pull any relevant insights from the system.

With all that information, I'll synthesize everything into an executive summary and deliver a final verdict. It's a standard, but effective, process. Let's get to work.

⚡ get\_domain\_whois\_mcp

⚡ get\_domain\_reputation\_info

✓ get\_domain\_whois\_mcp

✓ get\_domain\_reputation\_info

## Thought

Alright, here's the game plan. The user wants me to dig deeper into reputationrescue.info. The initial `get_domain_whois_mcp` call worked, giving me a solid starting point, but `get_domain_reputation_info` choked. No worries, I'll keep the investigation rolling and troubleshoot that failed call concurrently.

First off, let's see what we've got from the WHOIS data. Domain was registered back in December 2015, expiring in December 2025. A decade-long registration is generally a good indicator – suggests some level of seriousness. Registered with Wild West Domains, LLC – nothing particularly unusual there. Standard client-side protection is enabled, covering the usual suspects: deletion, renewal, transfer, and updates. Nameservers point to `ns1.gvodns.com` and `ns2.gvodns.com`.

Next up, time to check those domain/URL scan results. I'll hit up VirusTotal and URLscan using `vt_query` and `urlscan_query`, respectively. Let's see what kind of reports we can pull. This should start to give us a real feel for the domain's reputation.

## Conversations

Events Trace

- 0 functionCall:get\_domain\_reputation\_info
- 1 functionResponse:get\_domain\_reputation\_info
- 2 functionCall:urlscan\_query
- 3 functionResponse:urlscan\_query
- 4 functionCall:urlscan\_screenshot\_analysis
- 5 functionResponse:urlscan\_screenshot\_analysis
- 6 functionCall:google\_search\_agent
- 7 functionResponse:google\_search\_agent
- 8 functionCall:retrieval\_augmented\_generation
- 9 functionResponse:retrieval\_augmented\_generation
- 10 text:Executive Summary: The domain 'reputationrescue....

SESSION ID 937fb17f-1e55-449e-b07e-91ad11d2217c

Token Streaming | + New Session |  

see if we have any prior knowledge of this domain. After this crucial step, I'll be in a position to synthesize everything and present my executive summary and verdict.



⚡ retrieval\_augmented\_generation



✓ retrieval\_augmented\_generation

## Thought

Okay, here's my assessment of `reputationrescue.info`, based on the investigation I just completed. I'm satisfied with the data, even if one piece was missing.

So, I've run through the full battery of tests I planned for this domain. The WHOIS data initially raised a small eyebrow - ten-year registrations aren't the norm for phishing, but they're not a guarantee of legitimacy either. That was just a preliminary observation, though.

The real meat came from VirusTotal. A robust 18 vendors flagging the domain as malicious, and specifically for phishing? That's a very strong signal. I'm taking that seriously.

Next, UrlScan.io's results absolutely nailed it. A 100/100 malicious score and a clear identification of Spotify impersonation? Excellent. The screenshot analysis further corroborated the findings. The weak CAPTCHA and social engineering tactics are blatant red flags, exactly what I was looking for.

Google Search backed it all up. Public sentiment aligns with the malicious indicators, pointing to a low trust score and scam warnings. This solidifies my assessment.

I'm disappointed, though, that my internal threat intelligence query failed. It would have been good to add that layer of confirmation, but honestly, with the level of evidence I've gathered, it's not critical.

Bottom line? Between VirusTotal, UrlScan, and Google Search, I have more than enough to form a definitive conclusion. This domain is undeniably a malicious phishing site. I can now confidently prepare the executive summary and deliver the final verdict. I'm ready to move forward.



## Executive Summary:

The domain `reputationrescue.info` has been investigated and found to be a high-confidence phishing website. While the domain's registration history is unusually long for a typical phishing site (created in 2015 with a 10-year registration), the overwhelming evidence from multiple security tools confirms its malicious nature.

VirusTotal reports that 18 different security vendors have flagged the domain for phishing and malware. Furthermore, UrlScan.io gives the site a 100/100 malicious score and has identified it as a phishing site impersonating the popular streaming service Spotify. A detailed analysis of the website's screenshot confirms this, revealing several social engineering tactics and security red flags, such as a fake and weak CAPTCHA, designed to lure unsuspecting users into a false sense of security.

Google search results also indicate a low trust score and scam warnings associated with this domain.

## Final Verdict:

**Malicious.**

The domain `reputationrescue.info` is a confirmed phishing website and should be treated as malicious. It is highly recommended to block this domain and advise users to avoid any interaction with it. The site's primary purpose is to deceive users by impersonating a legitimate service (Spotify) to steal credentials or distribute malware.



scam warnings "reputationrescue.info"

"reputationrescue.info" reviews

is "reputationrescue.info" legitimate

## Invocations

I got the following domain to analyze reputationrescue.info call ↴

Is gov-tax.cfd part of any recent campaigns? ^

Invocation ID:  
e-1cb3e76d-2cf0-4b9f-8209-38d8baad989e

```
|→ invocation          67066.61ms
|→ invoke_agent web_secur... 67060.42ms
  └─ call_llm          15519.11ms
    |→ execute_tool get...
    |→ execute_tool vt_q...
    |→ execute_tool retr...
    |→ execute_tool (mer...
  └─ call_llm          7968.11ms
    |→ execute_tool urls...
  └─ call_llm          18201.75ms
    |→ execute_tool goog...
  └─ invocation          11690.80ms
    |→ invoke_agent...
  └─ call_llm          11690.51ms
  └─ call_llm          11689.78ms
  └─ call_llm          12061.62ms
```

Is gov-tax.cfd part of any recent campaigns?



## Thought

Okay, here's what I'm thinking. I need to figure out if this gov-tax.cfd domain is something I should be worried about. Is it part of any known campaigns? My security radar is up, and I'll approach this systematically.

First things first, I'll dive into some domain intelligence. I'll use `get_domain_whois_mcp` to get the registration details – the age of the domain is often a quick tell. A recently created domain immediately raises a red flag for phishing or other malicious activities.

Then, I'll move on to a couple of crucial checks to see if this domain has already been flagged as malicious. I will check the `vt_query` reports to see if any security vendors already think this is a problem and search the `urlscan_query` reports, too.

Next, I need to know if this domain has shown up in the threat intelligence databases. I'll use `retrieval_augmented_generation` to see if it's connected to any known campaigns, groups, or tactics.

To speed things up, I'll run the WHOIS lookup, the VirusTotal check, and the campaign query simultaneously.

Finally, if those initial checks don't give me enough information, I'll broaden my search with Google. I'll try some searches with the domain and the word "campaign" to see if there are any public reports about it. Hopefully, this thorough investigation will give me a clear picture of what I'm dealing with.

⚡ get\_domain\_whois\_mcp

⚡ vt\_query

⚡ retrieval\_augmented\_generation

✓ get\_domain\_whois\_mcp

✓ vt\_query

✓ retrieval\_augmented\_generation

## Thought

Alright, let's see what we've got. I've kicked off the preliminary investigation on gov-tax.cfd. Standard procedure: `get_domain_whois_mcp`, `vt_query`, and attempted `retrieval_augmented_generation`. Here's the initial picture:

First, `get_domain_whois_mcp`. Bingo, the domain's fresh – created in February of '25, registered for a single year. And, of course, the registrant is hiding behind a privacy service, specifically in Malaysia. Not ideal. The .cfid TLD, that's another minor red flag: we've seen a lot of spam and phishing campaigns leveraging those.

Next, the `vt_query`. Excellent. VirusTotal is already lit up like a Christmas tree. Six vendors are calling it malicious, with classifications covering all the bases: "phishing," "malware," "malicious." That's a pretty strong signal of imminent threat. We're talking high confidence here.

Too bad `retrieval_augmented_generation` failed. I was hoping for some hits from my internal threat intelligence. Guess I'll have to rely on the external data for now.

**Agent Development Kit**

SESSION ID 937fb17f-1e55-449e-b07e-91ad11d2217c

Token Streaming | + New Session |  

106\_web\_sec\_agent

Trace Events State Artifacts Sessions < >

Conversations ✓ Events Trace

0 functionCall:get\_domain\_reputation\_info

1 functionResponse:get\_domain\_reputation\_info

2 functionCall:urlscan\_query

3 functionResponse:urlscan\_query

4 functionCall:urlscan\_screenshot\_analysis

5 functionResponse:urlscan\_screenshot\_analysis

6 functionCall:google\_search\_agent

7 functionResponse:google\_search\_agent

8 functionCall:retrieval\_augmented\_generation

9 functionResponse:retrieval\_augmented\_generation

10 text:Executive Summary: The domain 'reputationrescu...' (with ellipsis)

11 functionCall:retrieval\_augmented\_generation

12 functionResponse:retrieval\_augmented\_generation

13 functionCall:urlscan\_query

14 functionResponse:urlscan\_query

✓ google\_search\_agent

**Thought**

Okay, here's my assessment of `gov-tax.cfd`, thinking through the analysis step-by-step:

Right, I've just finished running the gauntlet on `gov-tax.cfd`. Time to synthesize the findings and formulate a response. Here's how I'm putting it all together:

First, let's recap the initial investigation. The WHOIS lookup via `get_domain_whois_mcp` immediately raised some red flags. The domain, `gov-tax.cfd`, is brand new, registered only this February, and uses a privacy service. That's practically a template for a malicious domain. The `.cf` TLD, you know the drill, often a haven for these types of attacks.

Next, I fed the domain into VirusTotal using `vt_query`. Bingo. Six different security vendors are calling it malicious, with labels like "phishing" and "malware." That's a high-confidence signal, no doubt about it.

My attempt to run it through `urlscan_query` failed. Annoying, but not a deal breaker. Then, the internal Threat Intelligence via `retrieval_augmented_generation` also bombed out. Hmm, perhaps a data issue, but the main indicators are already apparent.

Now, for the big kahuna: the Google search, handled by the `google_search_agent`. And that, my friends, is where the real juicy stuff appeared. Google directly pointed me to a smishing campaign known as the "Smishing Triad," explicitly mentioning `gov-tax.cfd` as an active threat. It's impersonating government entities, primarily the IRS, aiming to steal user data. It also revealed it's flagged in multiple security reports.

So, taking the WHOIS, VirusTotal data and the direct information from the Google search... I'm pretty confident in my assessment. I've got a comprehensive picture now. I've got enough to give a clear and definitive answer.

Based on everything, I'm going to tell them: Yes, `gov-tax.cfd` is part of a recent smishing campaign. I'll describe the campaign as the "Smishing Triad," impersonating tax services like the IRS to steal information. I'll highlight the key indicators: recent creation, short registration period, and the malicious flags. I'll also clearly warn that this is a high-risk indicator of compromise and any link or message containing this domain should be treated with extreme caution and reported.

Yes, the domain `gov-tax.cfd` has been identified as part of a recent and widespread smishing (SMS phishing) campaign. This campaign, attributed to a group known as the "Smishing Triad," impersonates government entities, including tax services like the IRS, in an attempt to deceive individuals into revealing sensitive personal and financial information.

Key findings about this domain include:

- Recent Creation: The domain was created on February 19, 2025, and is registered for only one year, both of which are common tactics for phishing sites.
- Malicious Indicators: The domain is flagged as malicious by six security vendors on VirusTotal, with classifications including "phishing," "malware," and "malicious site."
- Campaign Association: The "Smishing Triad" campaign, which includes `gov-tax.cfd`, uses thousands of similar domains and targets people globally through text messages. These messages often create a sense of urgency related to unpaid taxes or toll violations to lure victims to phishing pages.
- WHOIS Privacy: The domain was registered using a privacy protection service in Malaysia, which is a common practice to hide the identity of the domain owner.

You should consider `gov-tax.cfd` to be a high-risk indicator of compromise. Any communication containing this domain should be treated as a phishing attempt.

 "gov-tax.cfd" legitimacy   "gov-tax.cfd" scam   "gov-tax.cfd"   "gov-tax.cfd" campaign

# Next Steps

# Next Steps

- Adding Guardrails (input and output filters)
- Error Handling
- Performance improvement
- Human-in-the-Loop
- Observation and Monitoring
- Agent Evals

# References

- Google ADK Official Documentation
- Agentic Design Patterns Book
- An Illustrated Guide to LLM Agents Book (2 draft chapters)

**<https://github.com/nabeelxy/web-security-agents>**