AI Working Group

2nd Session:

Model Context Protocol (MCP) – Build & Secure Your MCP Servers

By Gregory

26/09/2025

% whoami

> name : Gregory Tan

> position : Senior AI Engineer

> team : AI R&D

> area of interest : AIOps, AI-Gateway (Observability, Evaluation, ...)

Linkedin:
https://www.linkedin.com/in/tan-yong-jern/

AI Agents

What is AI Agents???

“

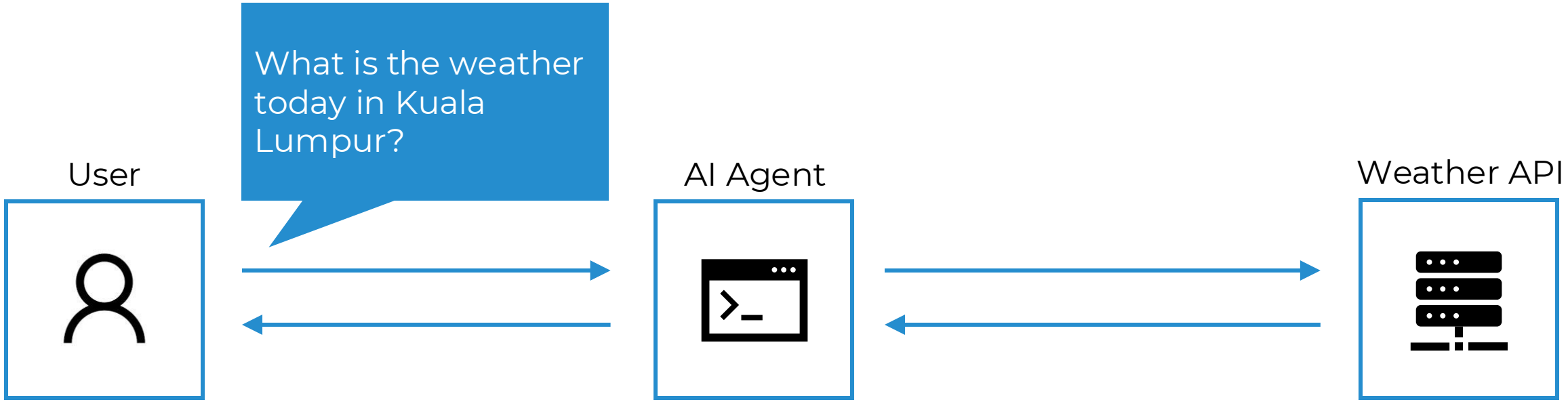
AI Agents are self-contained execution unit designed to act autonomously to achieve specific goals.

(1) Perform tasks,
(2) Interact with users,
(3) Utilize external tools, and
(4) Coordinate with other agents.

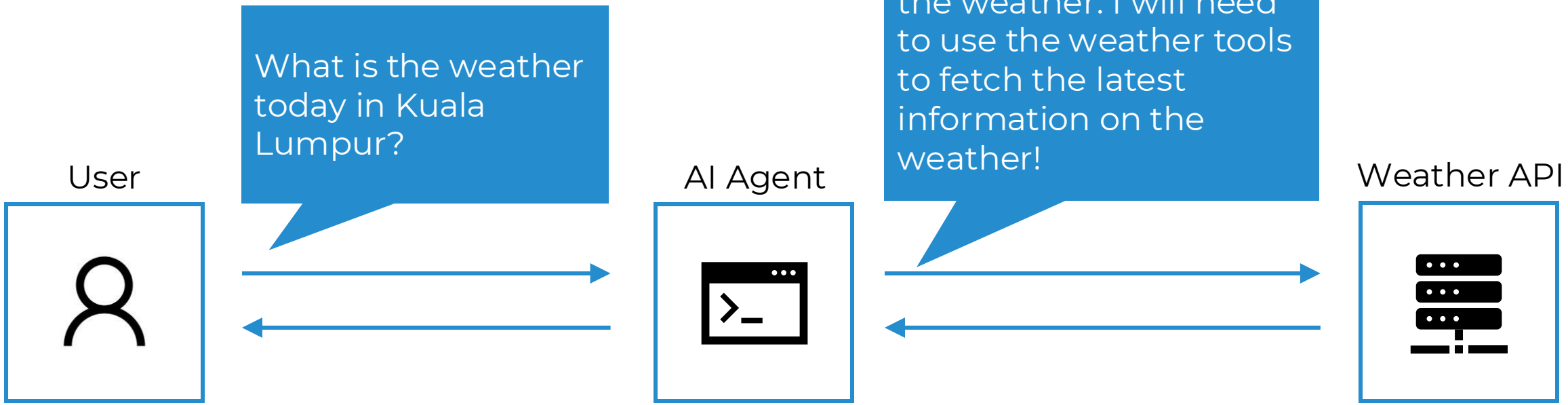
”

Source:
https://google.github.io/adk-docs/agents/

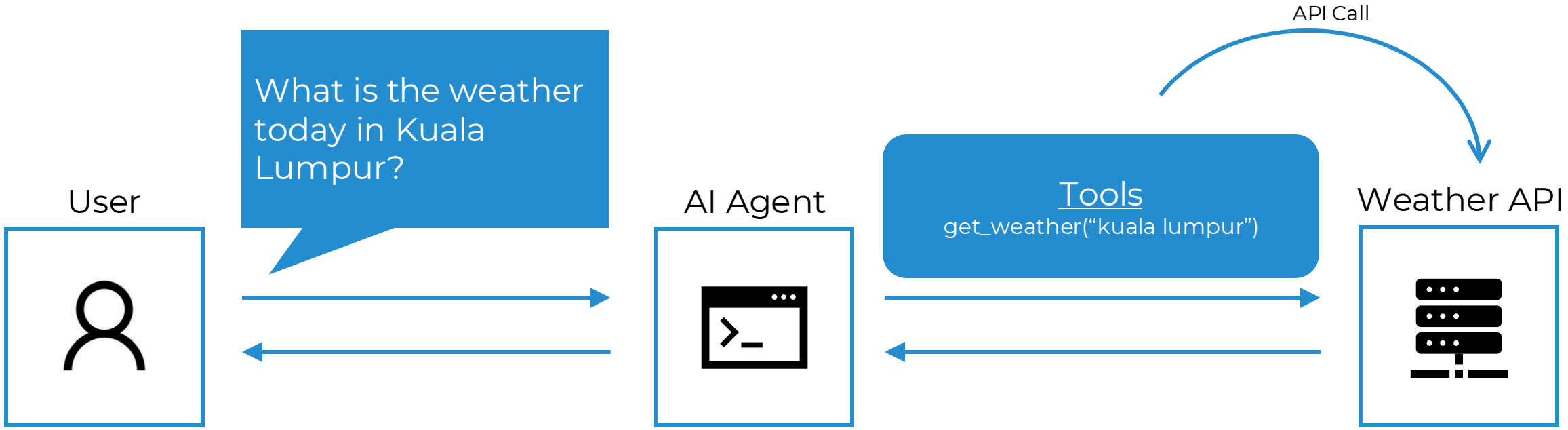
AI Agents (Cont.)



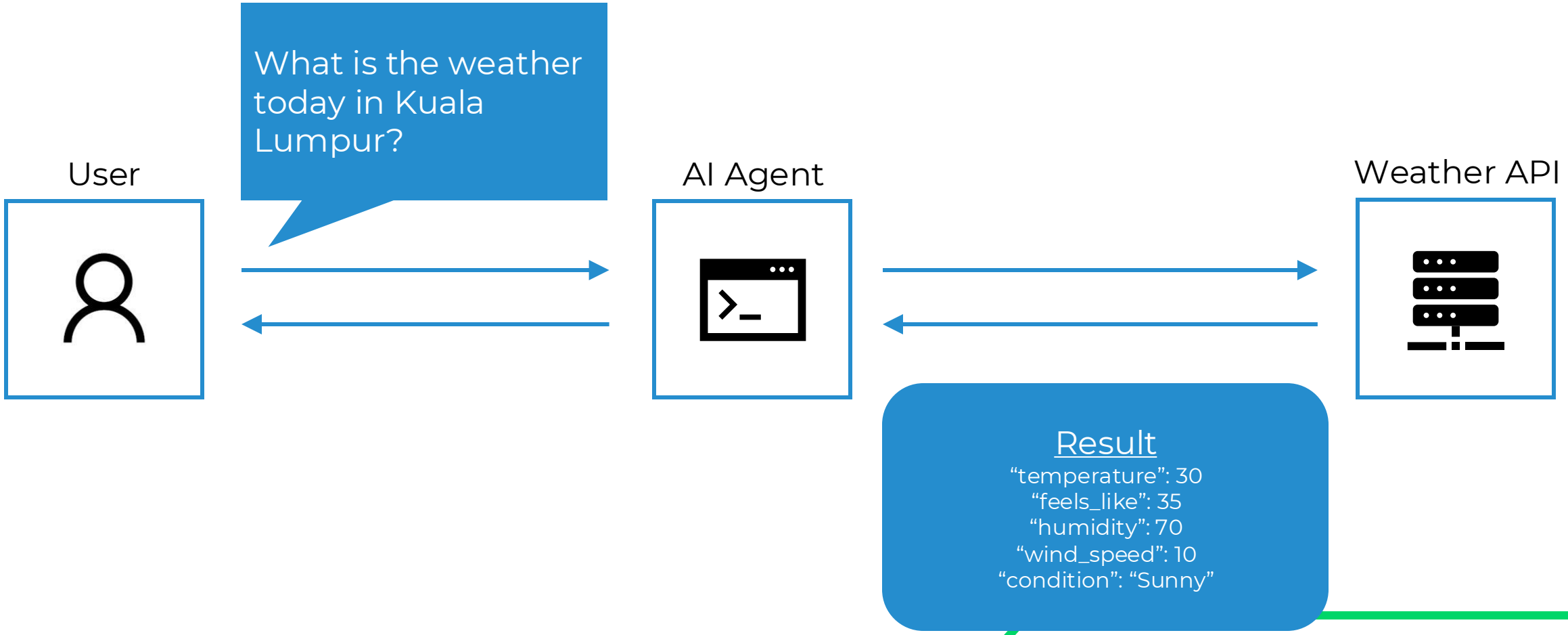
AI Agents (Cont.)



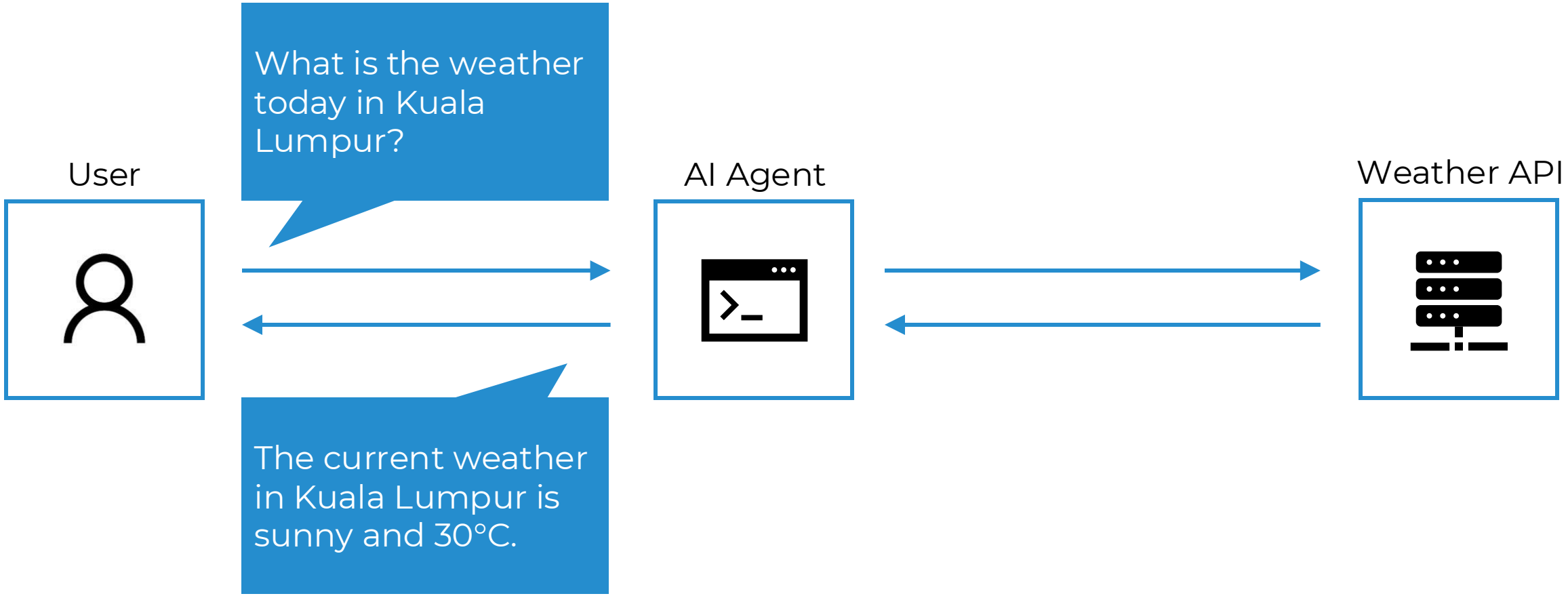
AI Agents (Cont.)



AI Agents (Cont.)



AI Agents (Cont.)



01

Define Your Tool

- What the tool does.
- When to use it.
- What arguments it requires (city: str).
- What information it returns.

```
def get_weather (city: str) -> dict:

    """Fetches current weather for a given city.
    Args:
        city (str): The name of the city
                    (e.g., "New York", "London", "Tokyo").

    Returns:
        dict: A dictionary containing the weather information.
              Includes a 'status' key ('success' or 'error').
              If 'success', includes a 'report' key with weather details.
              If 'error', includes an 'error_message' key.
    """

    url = f"http://api.openweathermap.org/data/2.5/weather?q={city}"
    response = requests.get(url)
    weather = response.json()

    if response.status_code == 200: # Success
        return {
            "status": "success",
            "report": f"The weather in {city} is {weather}."
        }
    else: # Failed
        return {
            "status": "error",
            "error_message": f"Sorry, I don't have weather information
                              for {city}."
        }
```

02

Define Your Agent

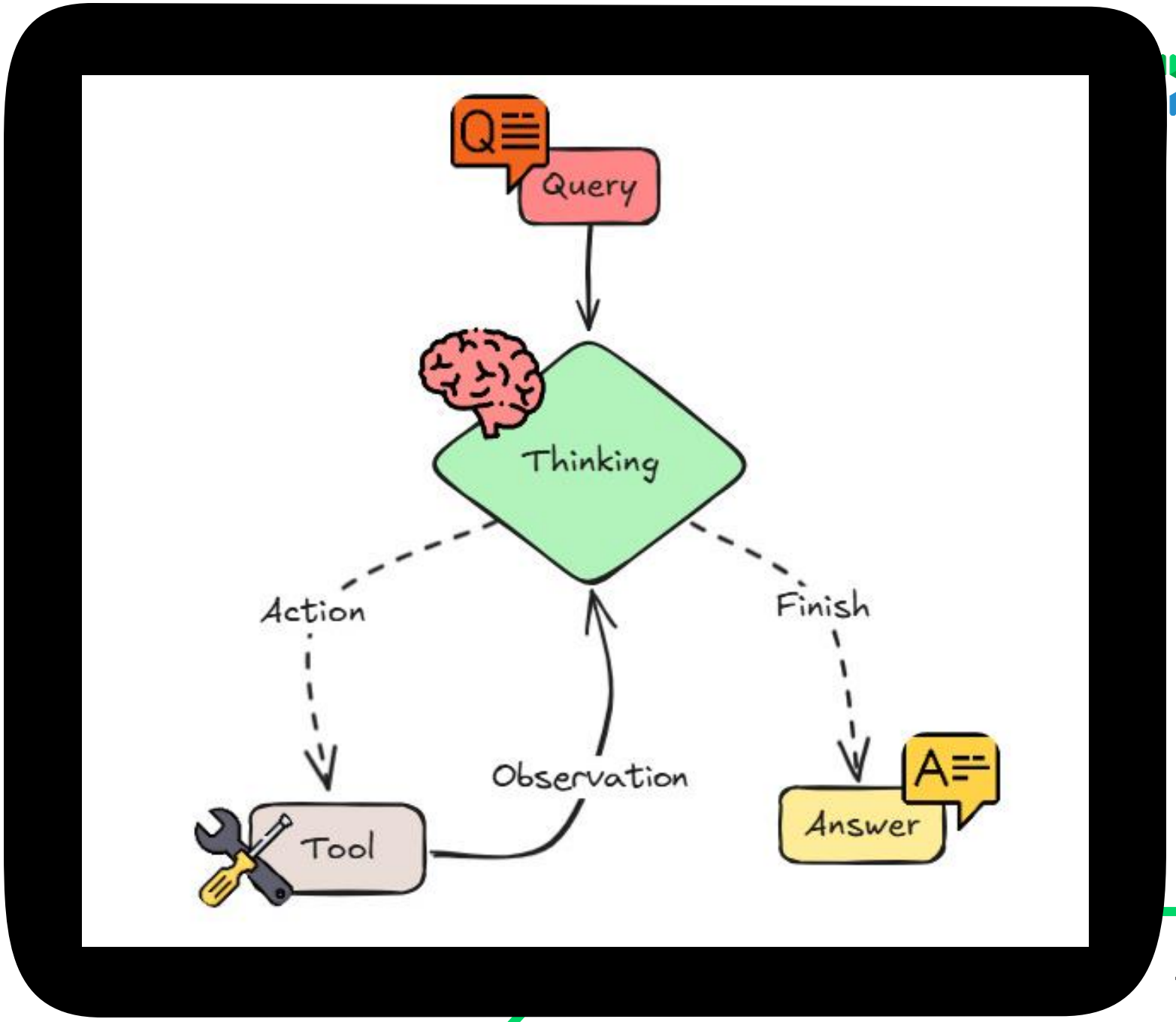
- What kind of behavior and goal of the AI Agent.
- How to use its tools effectively.
- How to handle errors.

```
from google.adk.agents import Agent

weather_agent = Agent(
    name="weather_agent_v1",
    model="gemini-2.5-pro",
    description="Provides weather information for specific cities.",
    instruction="You are a helpful weather assistant. When the user
                 asks for the weather in a specific city, use tools
                 to find the information. If the tool returns an
                 error, inform the user politely. If the tool is
                 successful, present the report clearly.",
    tools=[get_weather]
)
```

03

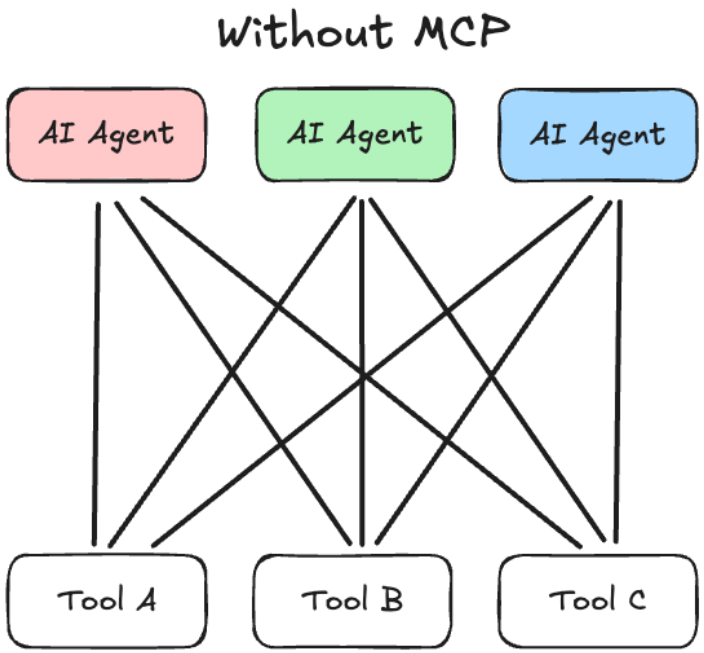
Planning & Reflection

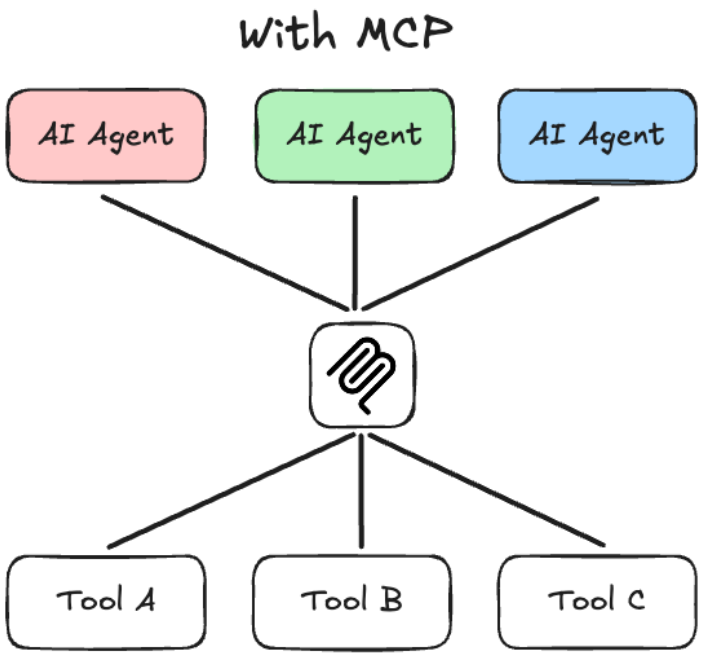
- Uses reasoning to understand user intent.
- Chooses the appropriate tool
- Combines information from multiple tools and prompts
- Responds back accordingly



Model Context Protocol (MCP)

What is MCP???





Unified Integration

Standardized way for agents to connect with different tools.

Separated Codebase

Agents and tools can live in separate codebases, making development modular, scalable, and easier to maintain.

Source:
https://modelcontextprotocol.io/docs/getting-started/intro

Separated Codebase

(AI Agent Code)

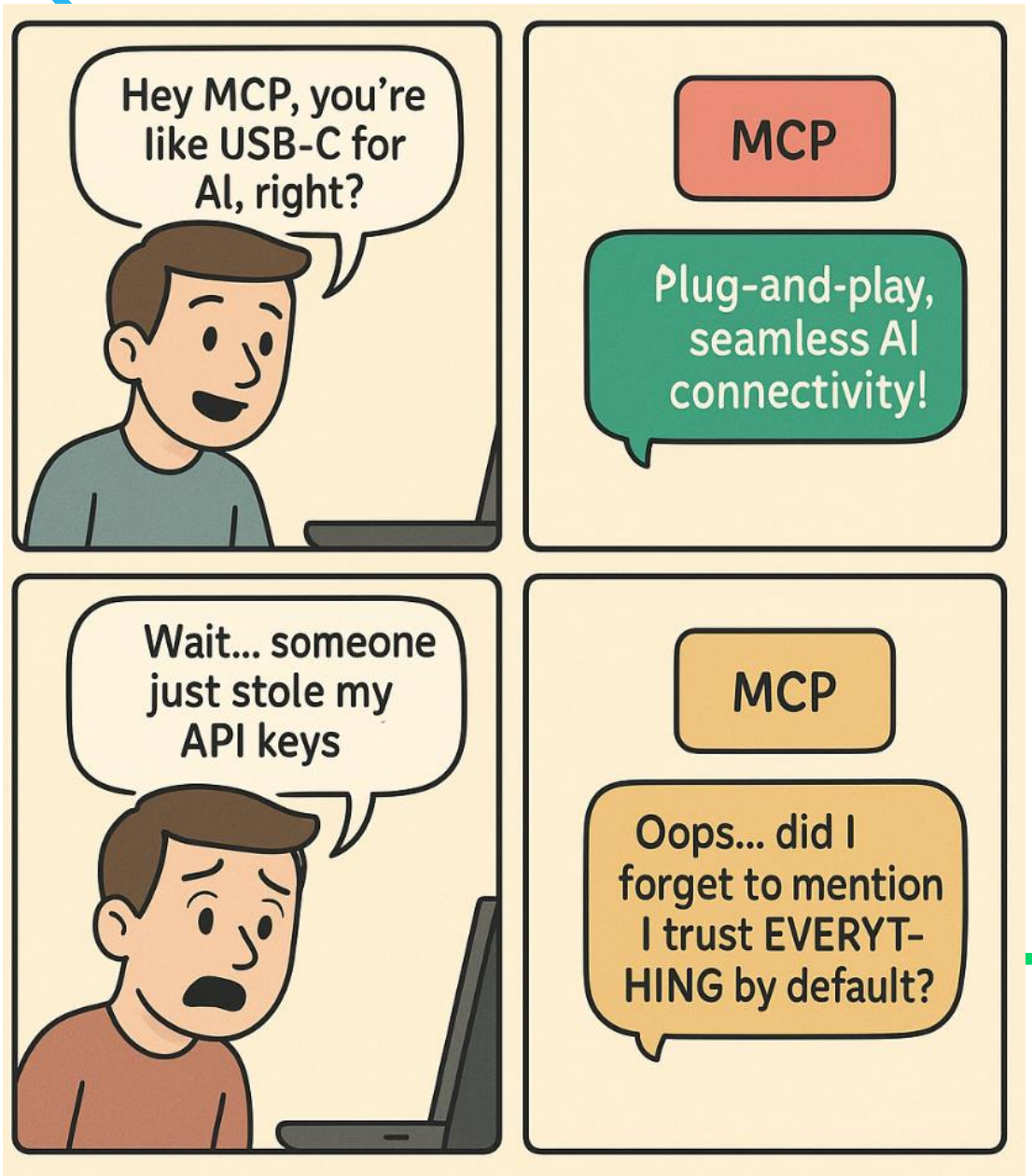
```python
from google.adk.agents import Agent
from google.adk.tools import MCPToolset, SseServerParams

tools = await MCPToolset.from_server(
    connection_params=SseServerParams(
        url="http://localhost:8001/mcp",
    )
)

weather_agent = Agent(
    name="weather_agent_v1",
    model="gemini-2.5-pro",
    description="Provides weather information for specific cities.",
    instruction="You are a helpful weather assistant. When the user
                 asks for the weather in a specific city, use tools
                 to find the information. If the tool returns an
                 error, inform the user politely. If the tool is
                 successful, present the report clearly.",
    ~~tools=[get_weather],~~
    tools=tools,
)
```

Separated Codebase

(MCP Code)

```python
from mcp.server.fastmcp import FastMCP
import requests, os

mcp = FastMCP("Weather MCP Server")

@mcp.tool()
def get_weather (city: str) -> dict:

    """Fetches current weather for a given city.
    Args:
        city (str): The name of the city
                    (e.g., "New York", "London", "Tokyo").

    Returns:
        dict: A dictionary containing the weather information.
              Includes a 'status' key ('success' or 'error').
              If 'success', includes a 'report' key with
              weather details.
              If 'error', includes an 'error_message' key.
    """
    << code logic goes here … >>

if __name__ == "__main__":
    mcp.run(transport="streamable-http")
```

However...



Source:
https://www.docker.com/blog/mcp-security-issues-threatening-ai-infrastructure/

Securing MCP (Cont.)

https://www.darkreading.com/vulnerabilities-threats/2000-mcp-servers-security

DARKREADING

NEWSLETTER SIGN-UP

Cybersecurity Topics | World | The Edge | DR Technology | Events | Resources

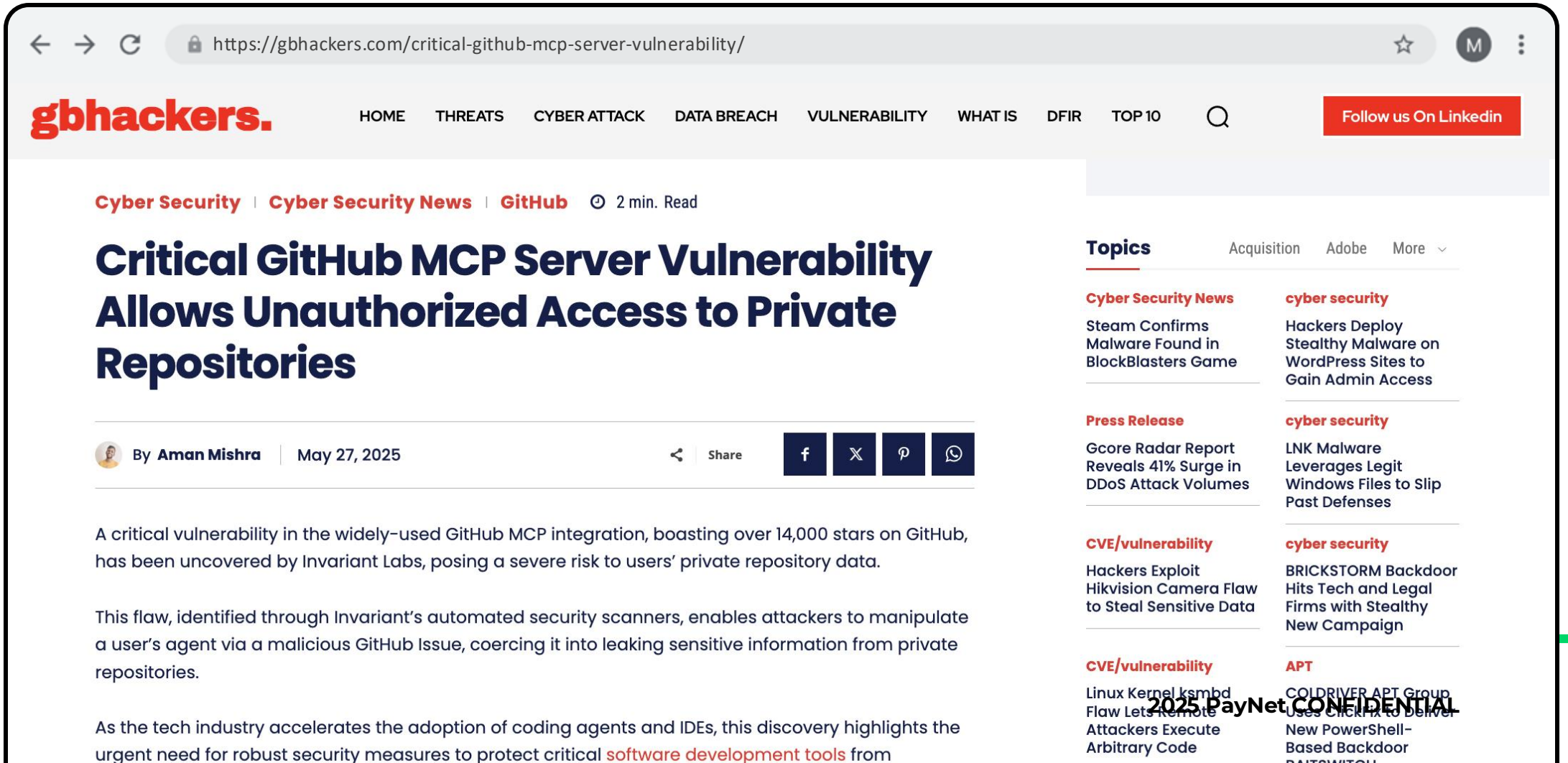
Nearly 2,000 MCP Servers Possess No Security Whatsoever

Authentication in MCP — the backbone of agentic AI — is optional, and nobody's implementing it. Instead, they're allowing any passing attackers full control of their servers.

Nate Nelson, Contributing Writer
July 19, 2025

4 Min Read

Securing MCP (Cont.)



https://gbhackers.com/critical-github-mcp-server-vulnerability/

gbhackers.

HOME THREATS CYBER ATTACK DATA BREACH VULNERABILITY WHAT IS DFIR TOP 10

Follow us On Linkedin

Cyber Security | Cyber Security News | GitHub 2 min. Read

Critical GitHub MCP Server Vulnerability Allows Unauthorized Access to Private Repositories

By **Aman Mishra** | May 27, 2025 Share

A critical vulnerability in the widely-used GitHub MCP integration, boasting over 14,000 stars on GitHub, has been uncovered by Invariant Labs, posing a severe risk to users' private repository data.

This flaw, identified through Invariant's automated security scanners, enables attackers to manipulate a user's agent via a malicious GitHub Issue, coercing it into leaking sensitive information from private repositories.

As the tech industry accelerates the adoption of coding agents and IDEs, this discovery highlights the urgent need for robust security measures to protect critical software development tools from

Topics Acquisition Adobe More

Cyber Security News
Steam Confirms Malware Found in BlockBlasters Game

cyber security
Hackers Deploy Stealthy Malware on WordPress Sites to Gain Admin Access

Press Release
Gcore Radar Report Reveals 41% Surge in DDoS Attack Volumes

cyber security
LNK Malware Leverages Legit Windows Files to Slip Past Defenses

CVE/vulnerability
Hackers Exploit Hikvision Camera Flaw to Steal Sensitive Data

cyber security
BRICKSTORM Backdoor Hits Tech and Legal Firms with Stealthy New Campaign

CVE/vulnerability
Linux Kernel ksmbd Flaw Lets Remote Attackers Execute Arbitrary Code

APT
COLDRIVER APT Group Uses ClickFix to Deliver New PowerShell-Based Backdoor

Mitigation Strategies

(01) Authentication

Only Agents with valid credentials & permissions can access the MCP server.

(02) Authorization (RBAC)

Ensure agents can only access tools allowed by their assigned role. Should follow the **Least Access Principle.**

(03) Sandboxed Code Execution

Environment Sandboxing — Run tools in a isolated container to limit the blast radius of bugs, misuse, or malicious behavior.

Source:
https://www.redhat.com/en/blog/model-context-protocol-mcp-understanding-security-risks-and-controls

Authentication

(MCP Code)

```python
from mcp.server.fastmcp import FastMCP
from mcp.server.auth.settings import AuthSettings
from pydantic import AnyHttpUrl
from src.auth import SimpleTokenVerifier
import requests, os

# MCP Server with Auth
mcp = FastMCP(
    "Weather MCP Server",
    token_verifier = SimpleTokenVerifier(),
    auth = AuthSettings(
        issuer_url = AnyHttpUrl("https://auth.example.com"),
        resource_server_url = AnyHttpUrl("http://localhost:3001"),
        required_scopes=["user"],
    ),
)

@mcp.tool()
def get_weather (city: str) -> dict:
    ...
```

Getting Started

How to get started?

Agent Building Frameworks

LangGraph

Agent Development Kit

Microsoft AutoGen

OpenAI Agents SDK

crewai

Source:
https://www.datacamp.com/blog/best-ai-agents

Demo Session


```
git clone https://github.com/justinwkUKM/paynet-aiwg.git
          cd paynet-aiwg/AIWorkingGroup_02_MCP
```

Question
&
Answer

Thank you