Agentic AI for Networkers

Frank Brockners
Distinguished Engineer, Outshift by Cisco
Linux Foundation Networking, Board Member

BRKOPS-1327

Webex App

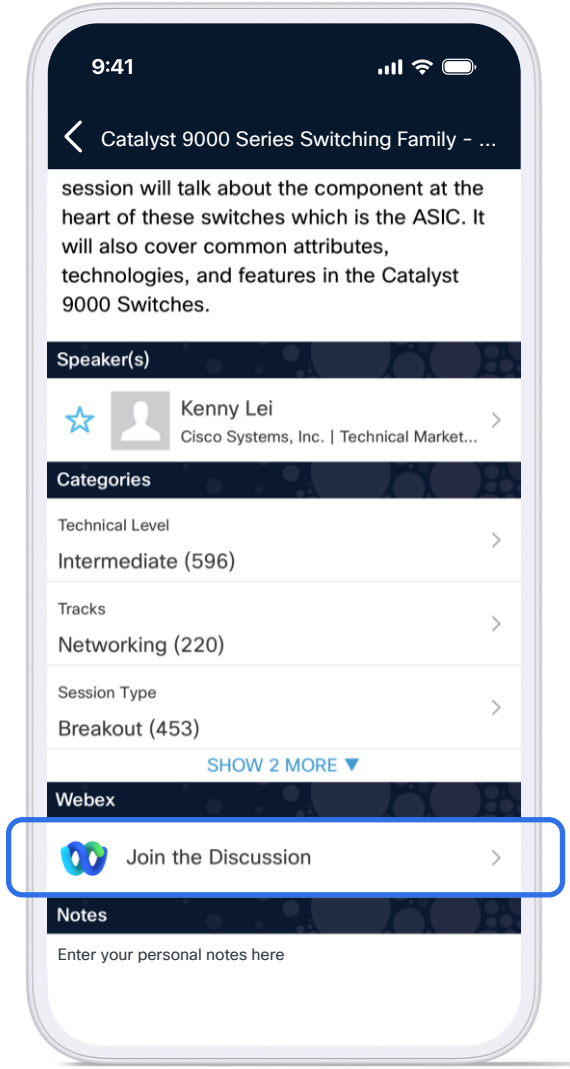
Questions?

Use Webex App to chat with the speaker after the session

How

1. Find this session in the Cisco Events App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until February 27, 2026.



01 What is Agentic AI?
Foundations

02 What can Agentic AI do for us?
Agentic for Networkers

03 What can we do for Agentic AI?
Networking for Agentic

What is “Agentic AI”?

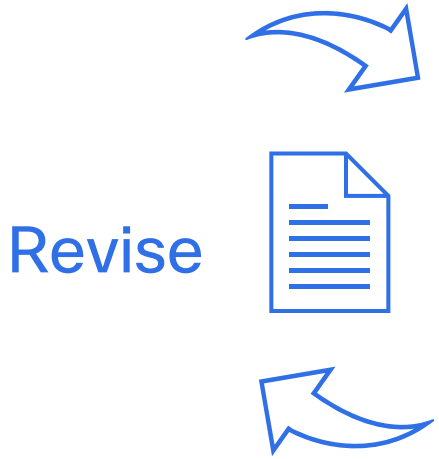
AI agents (also referred to as compound AI systems or agentic AI) are a class of intelligent agents distinguished by their ability to operate autonomously in complex environments

Wikipedia,
https://en.wikipedia.org/wiki/AI_agent

What is Agentic?

> More and more people are building systems that prompt a large language model multiple times using agent-like design patterns. But there's a gray zone between what clearly is not an agent (prompting a model once) and what clearly is (say, an autonomous agent that, given high-level instructions, plans, uses tools, and carries out multiple, iterative steps of processing).
>
> Rather than arguing over which work to include or exclude as being a true agent, we can acknowledge that there are different degrees to which systems can be agentic. Then we can more easily include everyone who wants to work on agentic systems. We can also encourage newcomers to start by building simple agentic workflows and iteratively make their systems more sophisticated.

Andrew Ng,
https://x.com/AndrewYNg/status/1801295202788983136



Zero Shot – Non-Agentic Workflow

“Please write an essay on topic X from start to finish in one go, without using backspace”

Agentic Workflow

“Write an essay outline on topic X”

“Do web research on the items of the outline”

“Write a first draft”

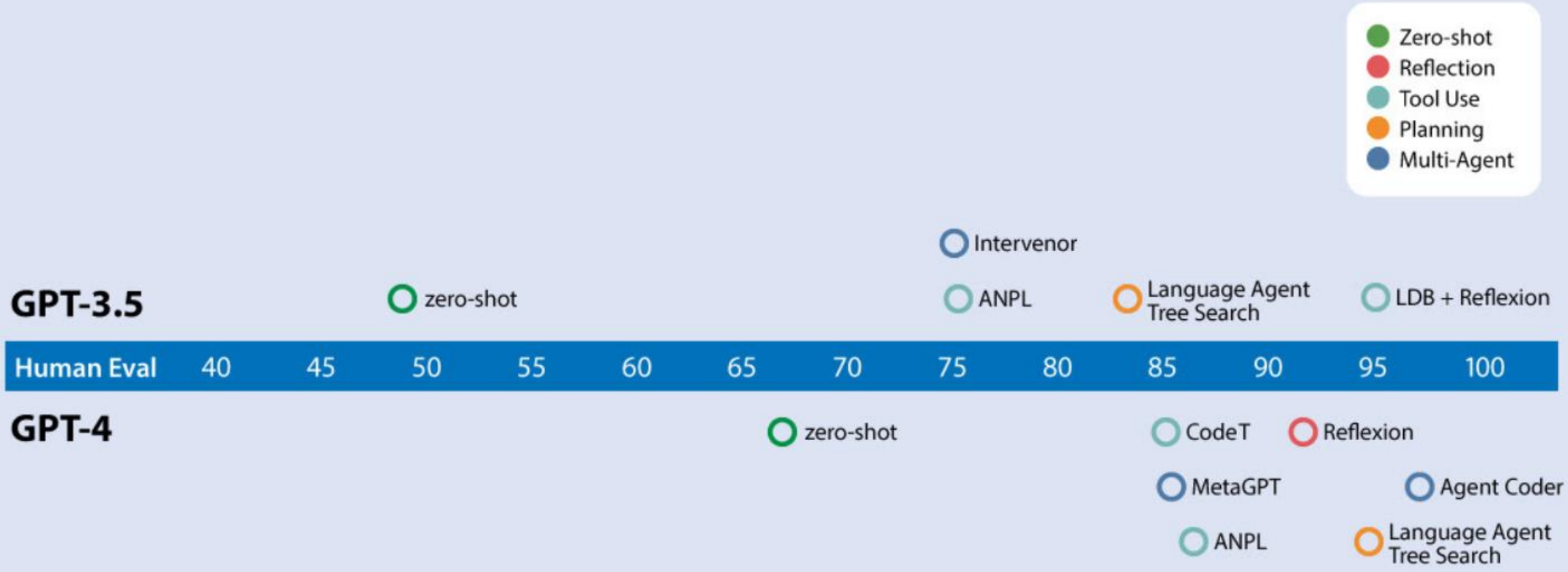
“Consider what parts need revision and more research”

“Revise your draft”

....

Adapted from Andrew Ng: https://youtu.be/q jXFm21l-VQ?si=zdfDMZ3mFobXpTpi

GPT-3.5 and GPT-4 performance using zero-shot and agent workflows



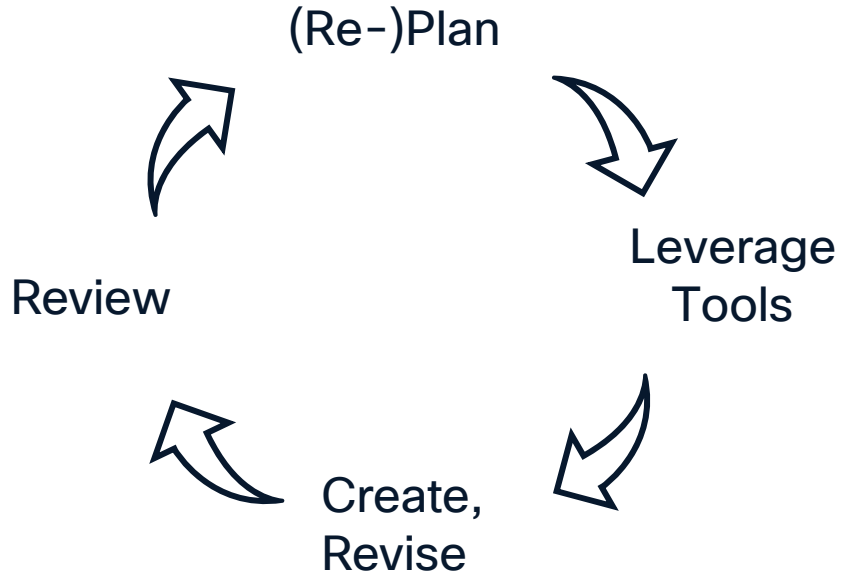
Performance of GPT-3.5 and GPT-4 (zero-shot) on HumanEval, along with algorithms that use agent workflows on top of GPT-3.5 or GPT-4. Thanks to Joaquin Dominguez and John Santerre for help with this analysis.

Source: https://www.deeplearning.ai/the-batch/how-agents-can-improve-llm-performance/

Agentic AI – Initial Objective: Make LLMs “better”

Agentic AI systems: AI systems that are designed to autonomously make decisions and act, with the ability to pursue complex goals with limited/no supervision.

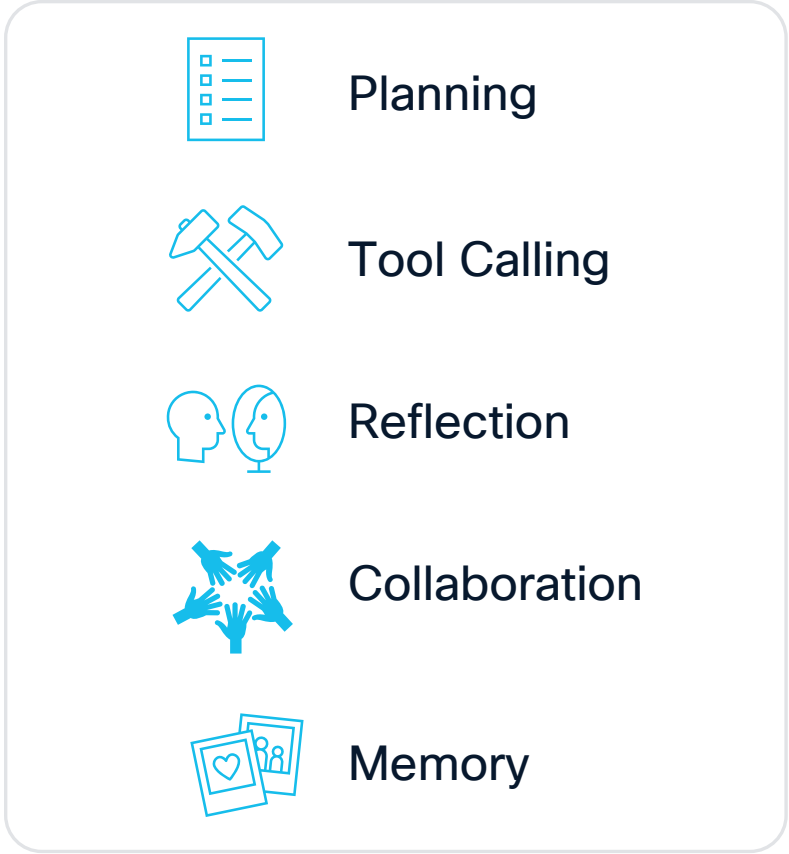
- Agentic Systems decompose:
  - Agents are simple and specific to a (set of) prompts and/or tool calls.
  - Combinations of agents with LLMs with specific prompts to plan, split tasks, perform tasks, define workflows, and reason.
- Agentic Systems integrate:
  - Workflow engines (or similar) orchestrate agents into a solution.



Agentic AI - Key Technologies & Associated Tools

Associated Tooling

- **Frameworks** to ease task decomposition, task delegation, agent composition into flexible workflows (Langgraph, Crewai, Autogen, Semantic Kernel, Pydantic, n8n, Opal, AgentBuilder, ...)
- **Integration of Data Sources** (RAG, MCP, ...)
- **Observability** (Splunk, Langsmith, Langfuse, ...)
- **Security** (AI Defense, Cisco SASE, ...)



Agentic AI Qualities

Acting/Planning with code improves performance

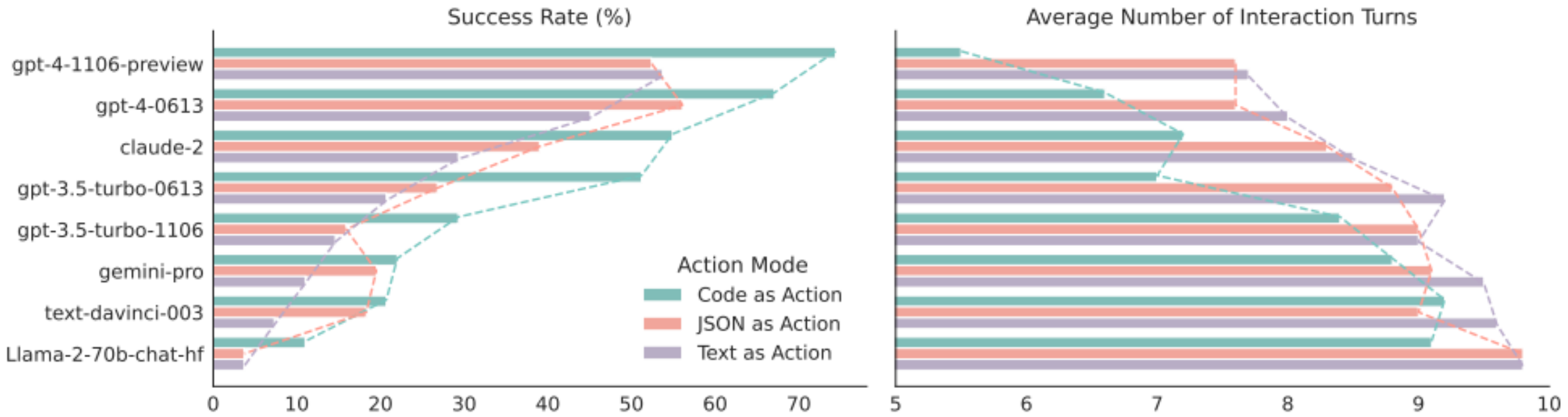
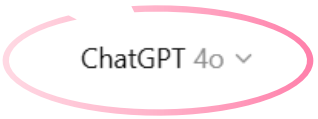


Figure 1: Comparison between `CodeAct` and Text / JSON as action. **(top)** Illustrative example comparing different actions. **(bottom)** Quantitative results on M^3ToolEval (§2.3).

Source: Executable Code Actions Elicit Better LLM Agents
https://arxiv.org/pdf/2402.01030

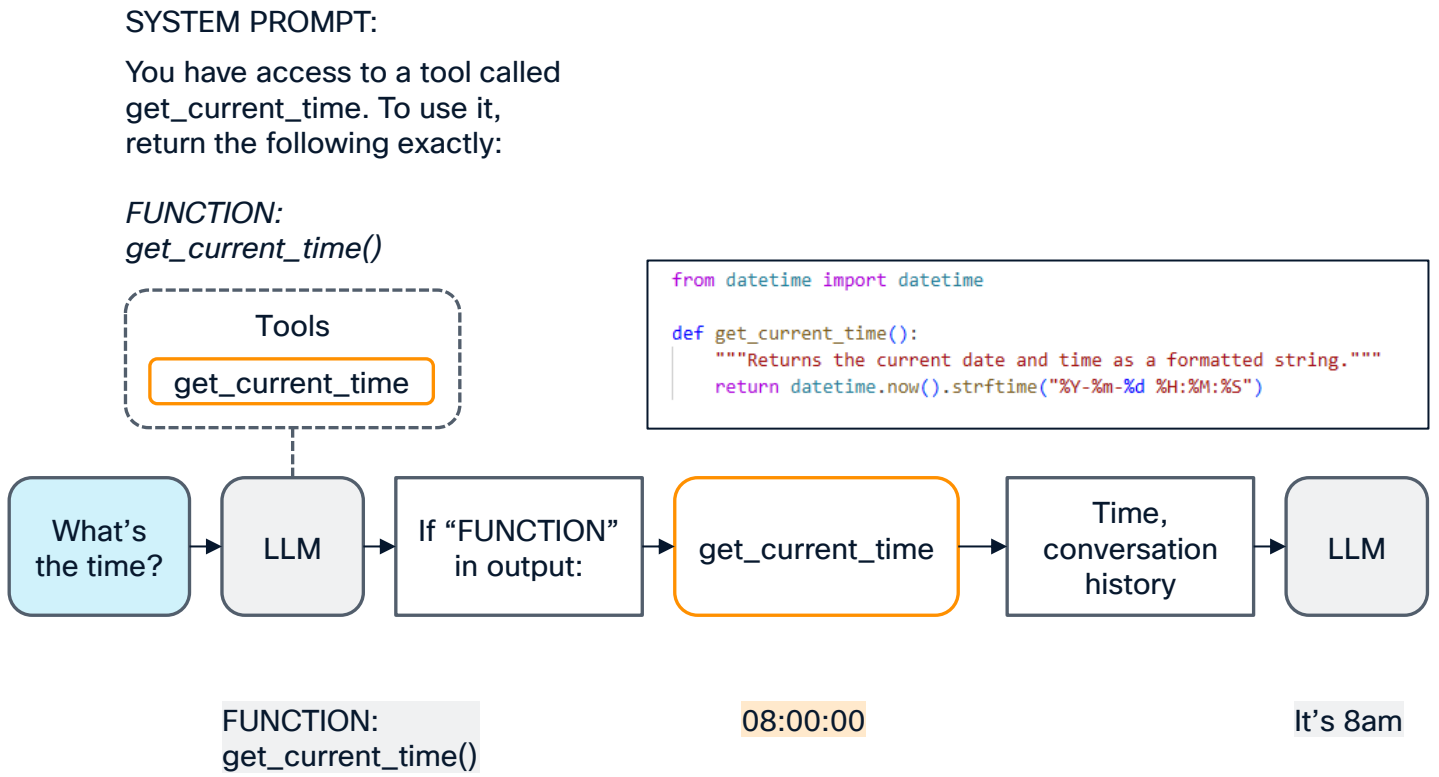


How many r's in Brockners?

The word "**Brockners**" contains **one** letter "**r**".

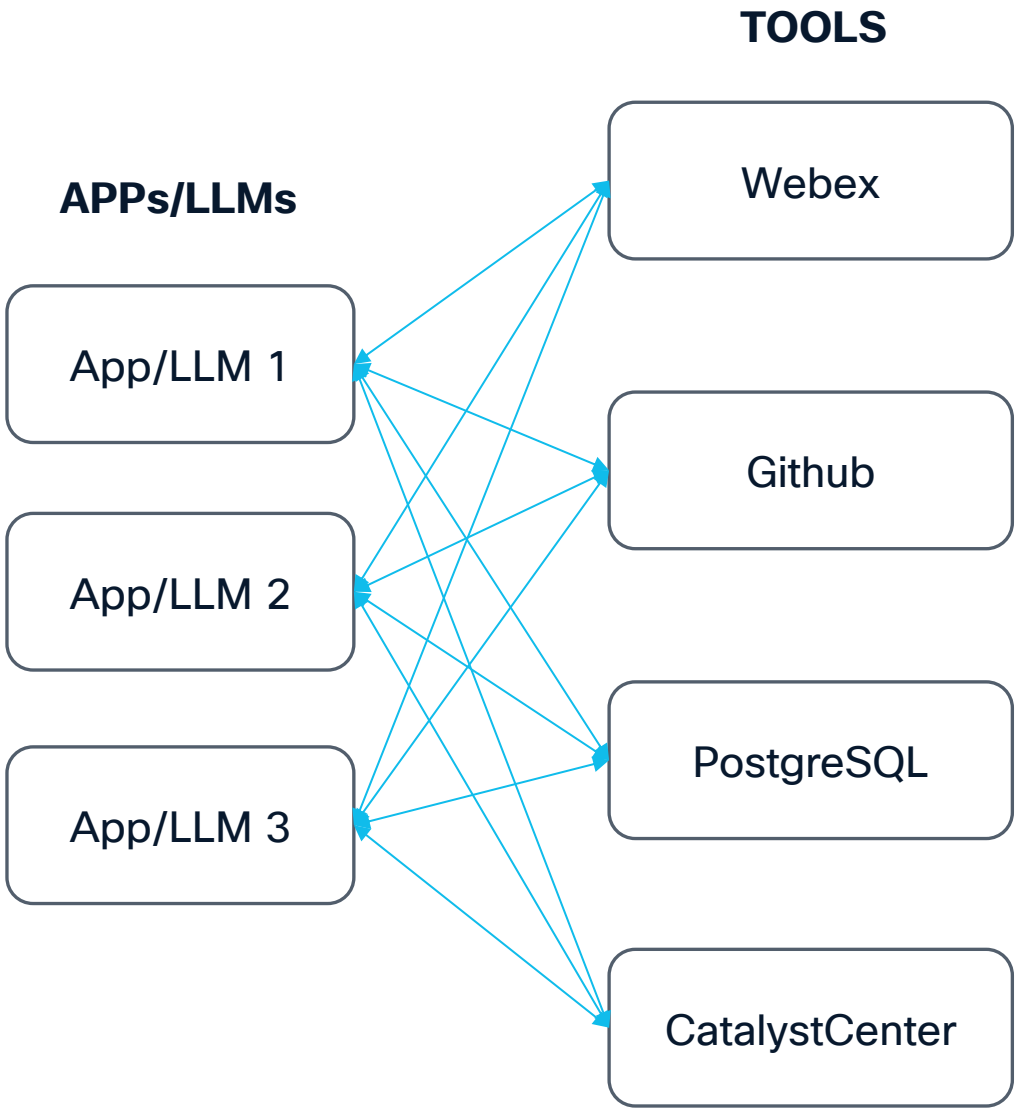
Tool Calling

- Tools are just code that the LLM can request to be executed
- LLMs can choose to use tools when appropriate
- Tools can write code – for enhanced flexibility
- Ensure secure execution of the code.



Tool Calling: Model Context Protocol

- Model Context Protocol (MCP)
  - Standardize how LLM-based agents access external tools, data, and services
  - Solve the 'N x M integration problem' between models and tools
  - Enable portable, reusable, and secure tool integrations
  - Decouple model providers from tool providers



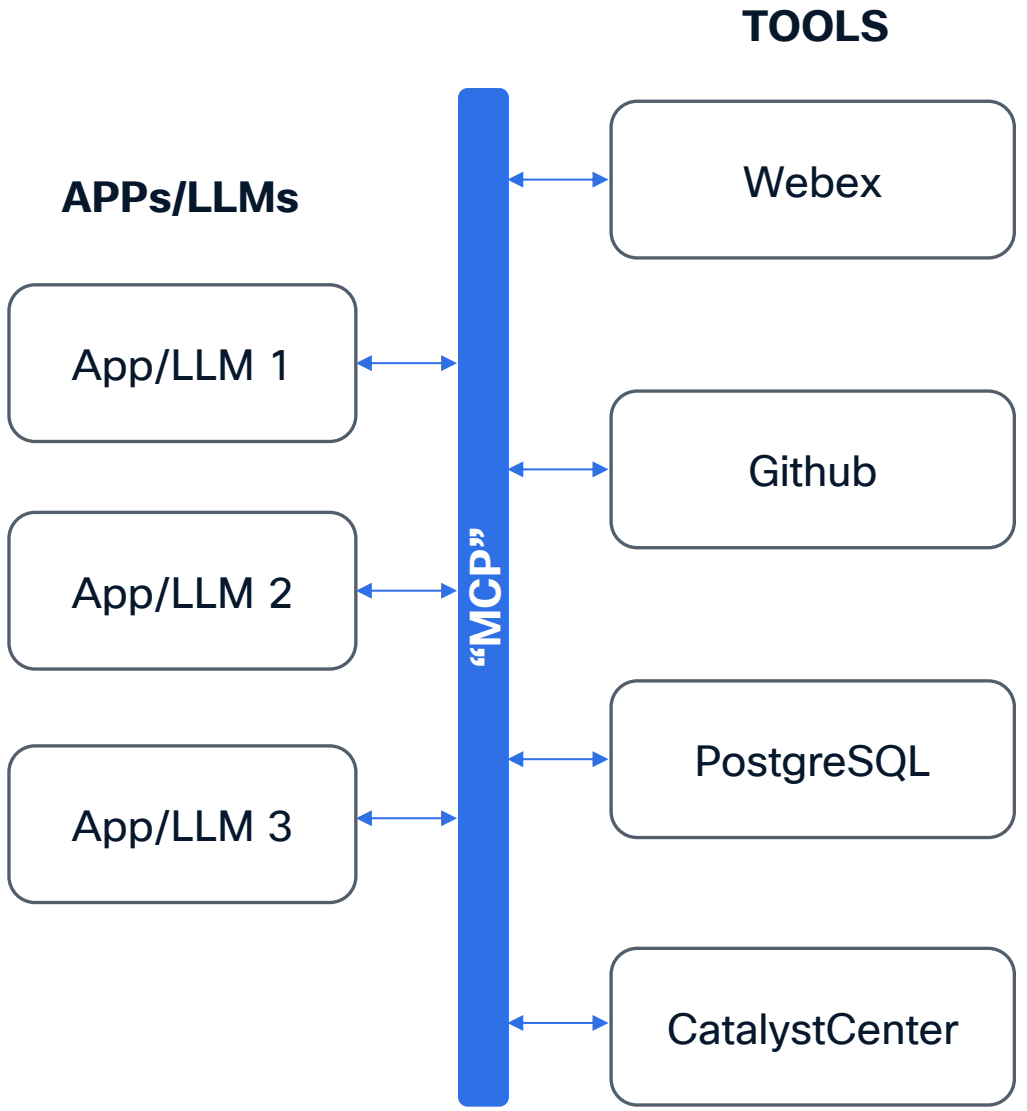
Each app creates its own tools

m x n

Tool Calling: Model Context Protocol

Tool Calling

- Model Context Protocol (MCP)
  - Standardize how LLM-based agents access external tools, data, and services
  - Solve the 'N x M integration problem' between models and tools
  - Enable portable, reusable, and secure tool integrations
  - Decouple model providers from tool providers



Each app creates its own tools

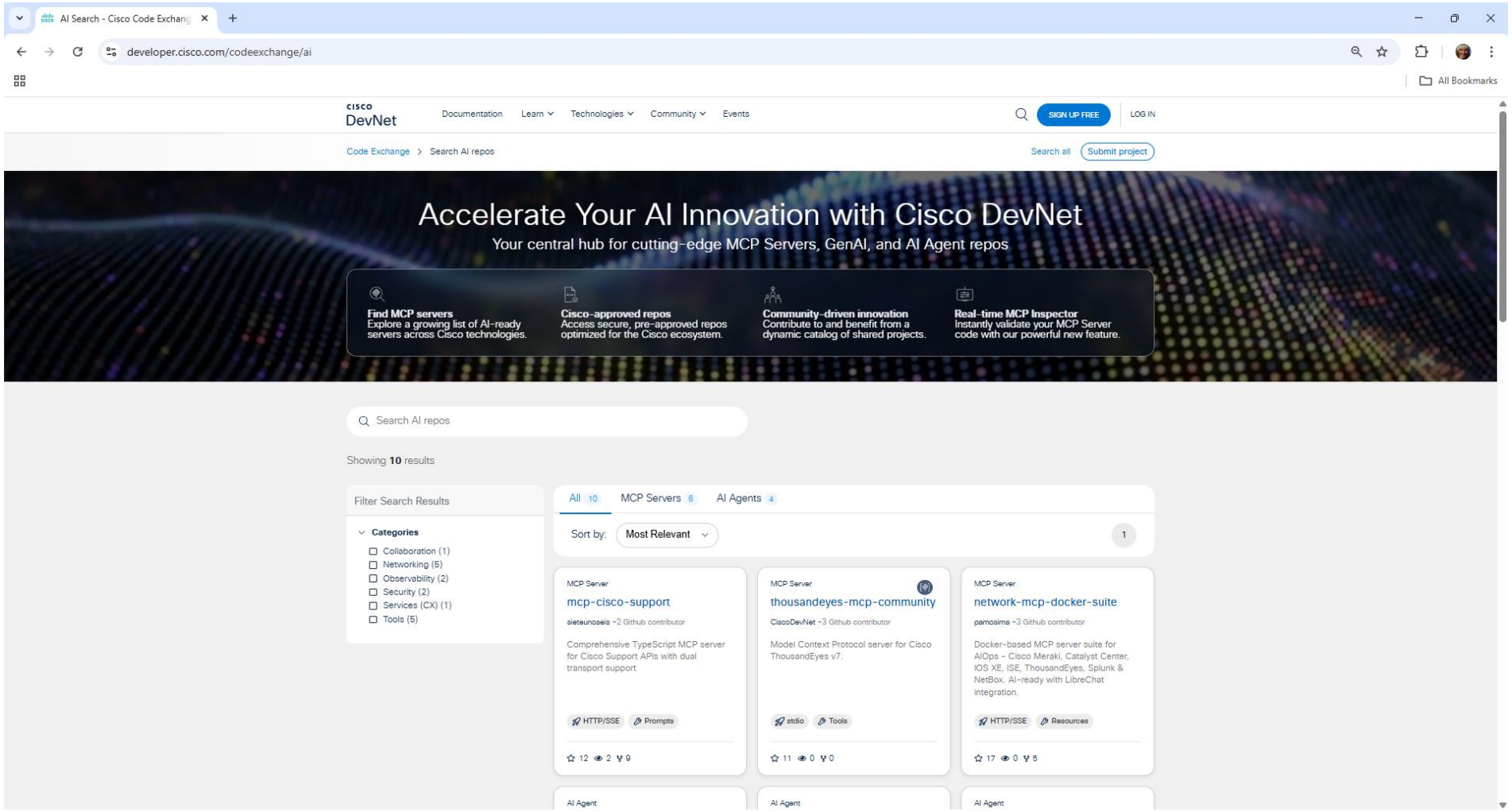
m x n

Each app uses shared MCP server

m+n

cisco

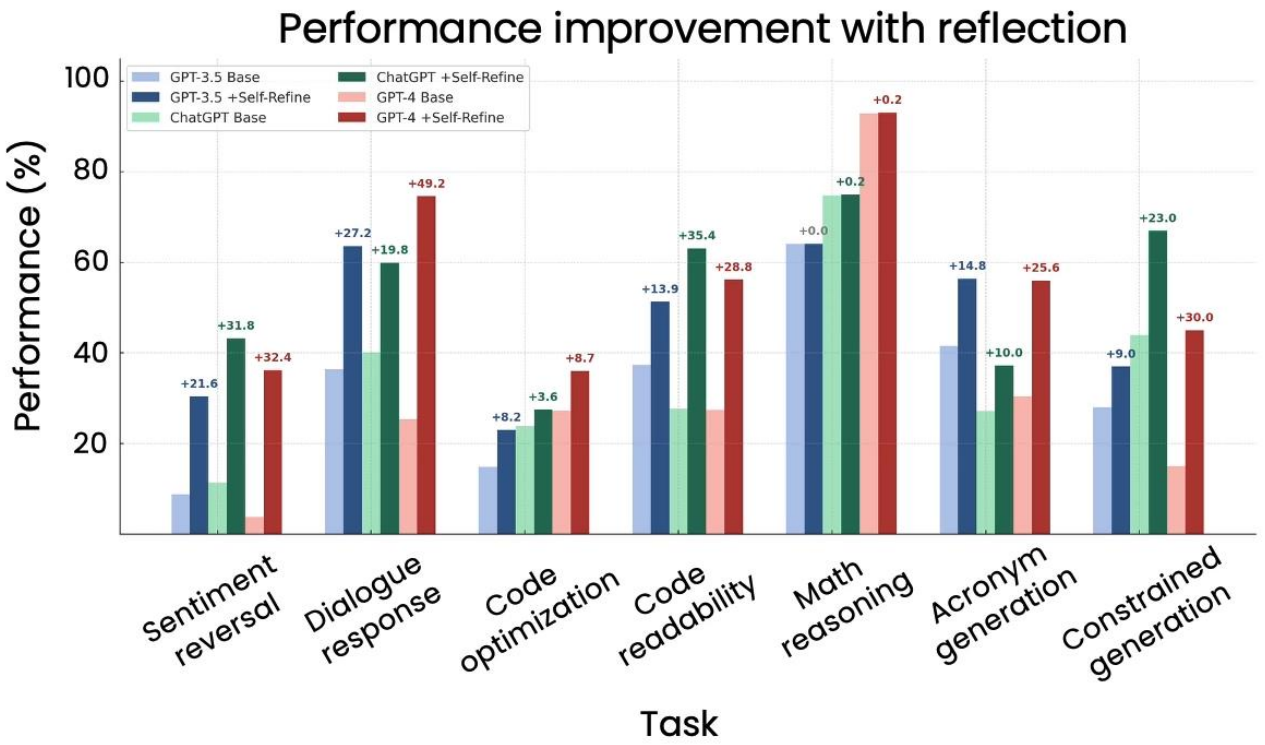
DevNet MCP-Servers and AI-Agents



https://developer.cisco.com/codeexchange/ai

Reflection has been tested

Reflection consistently outperforms direct generation on a number of tasks



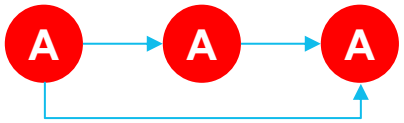
SELF-REFINE:
Iterative Refinement with Self-Feedback

Aman Madaan[1], Niket Tandon[2], Prakhar Gupta[1], Skyler Hallinan[3], Luyu Gao[1], Sarah Wiegreffe[2], Uri Alon[1], Nouha Dziri[2], Shrimai Prabhumoye[4], Yiming Yang[1], Shashank Gupta[2], Bodhisattwa Prasad Majumder[5], Katherine Hermann[6], Sean Welleck[2,3], Amir Yazdanbakhsh[6], Peter Clark[2]
[1]Language Technologies Institute, Carnegie Mellon University
[2]Allen Institute for Artificial Intelligence
[3]University of Washington [4]NVIDIA [5]UC San Diego [6]Google Research, Brain Team
amadaan@cs.cmu.edu, nikett@allenai.org

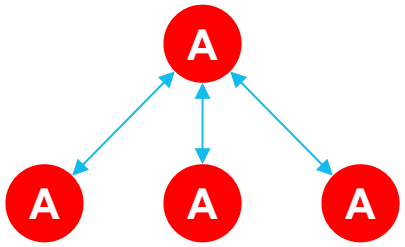
https://arxiv.org/pdf/2303.17651

Agentic Communication Patterns

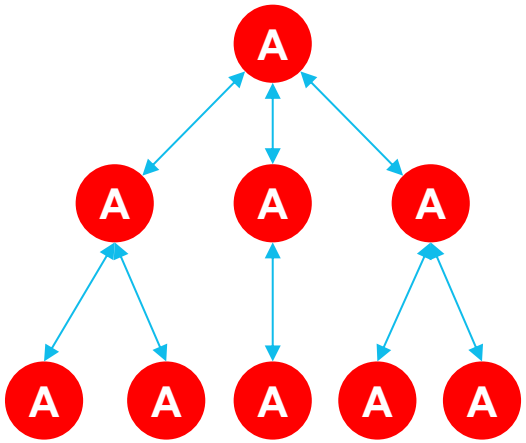
Linear Plan



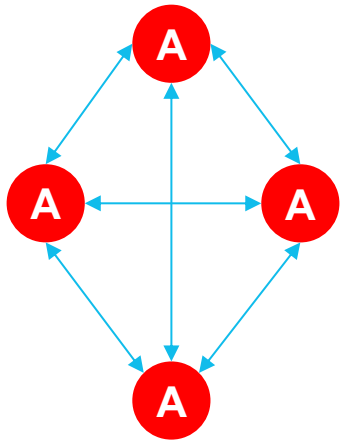
Hierarchical



Deeper Hierarchies



Many-to-Many



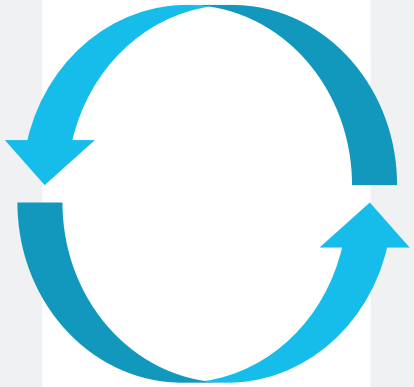
cisco

Agentic Development Process - Overview

BUILD

Build end-to-end system

Improve individual component



ANALYZE

Examine outputs, traces

Build evals, compute metrics

Error analysis

Component-level evals

01 What is Agentic AI? Foundations

02 What can Agentic AI do for us?

Agentic for Networkers

03 What can we do for Agentic AI?

Networking for Agentic

Assist

Augment

Offload

Assist

Assist the network engineer

- Human to Machine focus: Natural Language Interface
- Tool Calling, RAG

Assistant cross-product skills

Campus and Branch

Catalyst Center
Topology, client details, location, etc.

webex by cisco
Voice and video experience

Cisco Meraki
Topology, client details, location, etc.

ThousandEyes
WAN, Internet, App Insights

SD-WAN
WAN Details

Identity Intelligence
User trust level, identity checks & reasons

Data Center

Nexus Dashboard
Data center network management.

Hyperfabric
Data center network management.

Intersight
Unified management, automation, security.

Security and Observability

splunk>
Cisco and third-party insights

Firewall
Security & connection events

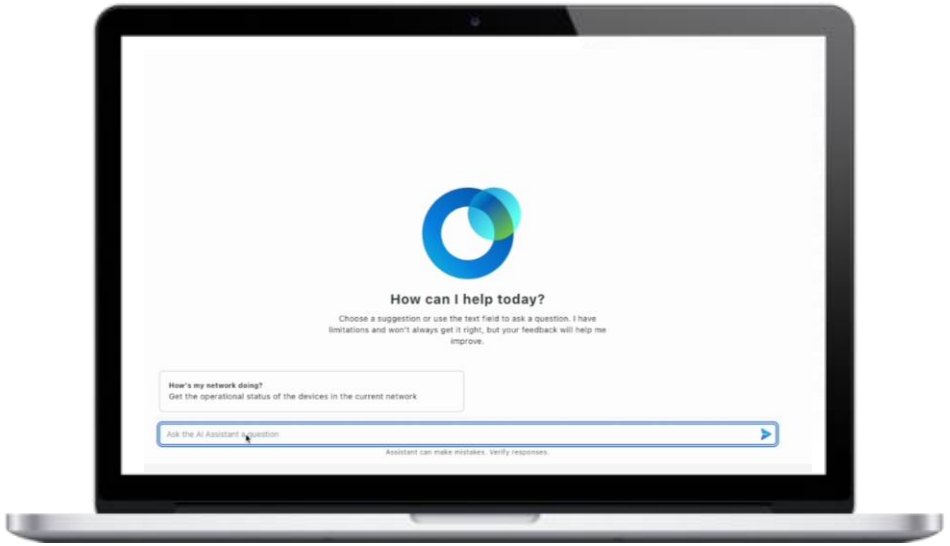
ISE
Authentication Insights

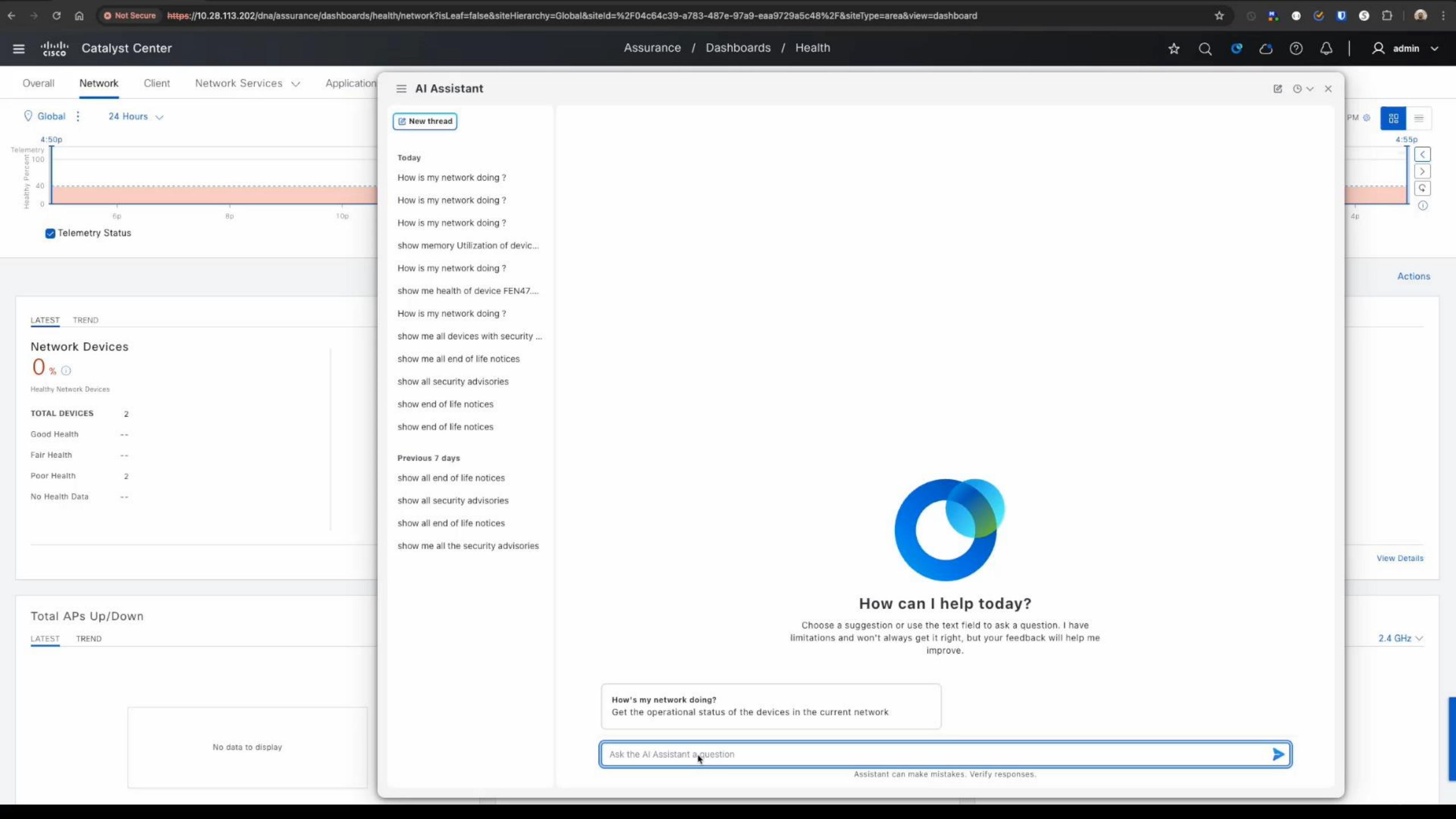
DUO
Authentication & compliance

Secure Access
Private & SAAS Resource Access

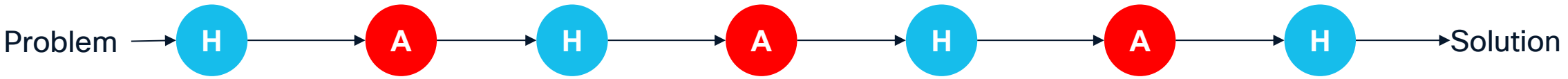
XDR
Related Threat Incidents

Cisco AI Assistant




Not Secure https://10.28.113.202/dna/assurance/dashboards/health/network?isLeaf=false&siteHierarchy=Global&siteId=%2F04c64c39-a783-487e-97a9-eaa9729a5c48%2F&siteType=area&view=dashboard
Catalyst Center
Assurance / Dashboards / Health
admin
Overall
Network
Client
Network Services
Application
Global
24 Hours
4:50p
Telemetry
Healthy Percent
100
40
0
6p
8p
10p
Telemetry Status
4:55p
4p
Actions
LATEST
TREND
Network Devices
0 %
Healthy Network Devices
TOTAL DEVICES 2
Good Health --
Fair Health --
Poor Health 2
No Health Data --
View Details
Total APs Up/Down
LATEST
TREND
2.4 GHz
No data to display
AI Assistant
New thread
Today
How is my network doing ?
How is my network doing ?
How is my network doing ?
show memory Utilization of devic...
How is my network doing ?
show me health of device FEN47....
How is my network doing ?
show me all devices with security ...
show me all end of life notices
show all security advisories
show end of life notices
show end of life notices
Previous 7 days
show all end of life notices
show all security advisories
show all end of life notices
show me all the security advisories
How can I help today?
Choose a suggestion or use the text field to ask a question. I have limitations and won't always get it right, but your feedback will help me improve.
How's my network doing?
Get the operational status of the devices in the current network
Ask the AI Assistant a question
Assistant can make mistakes. Verify responses.

Assistant



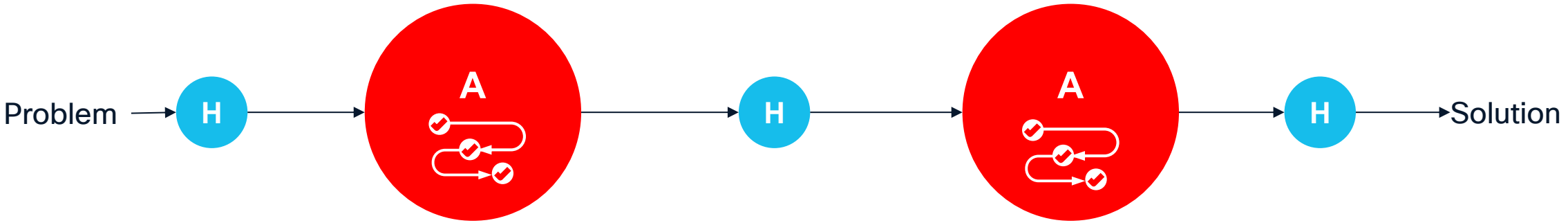
Legend: H Human A Agent

cisco

Augment the Human:
Automatically choose and execute multi-step runbooks

- Multi-modal generative interface
- Human to Machine and Machine to Machine
- Flexible workflows (“reasoning traces”)
- Tool Calling, RAG

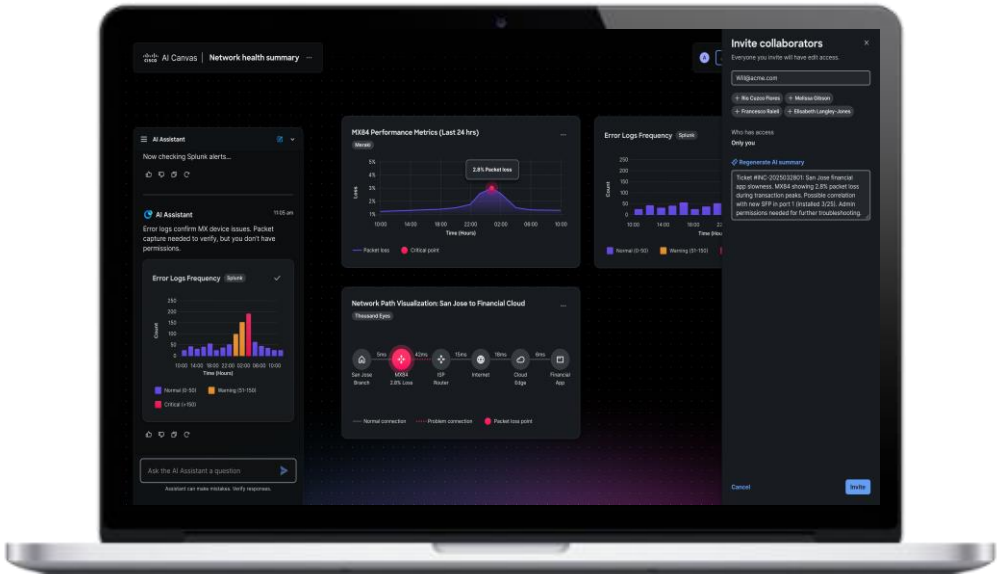
Conditional Automation

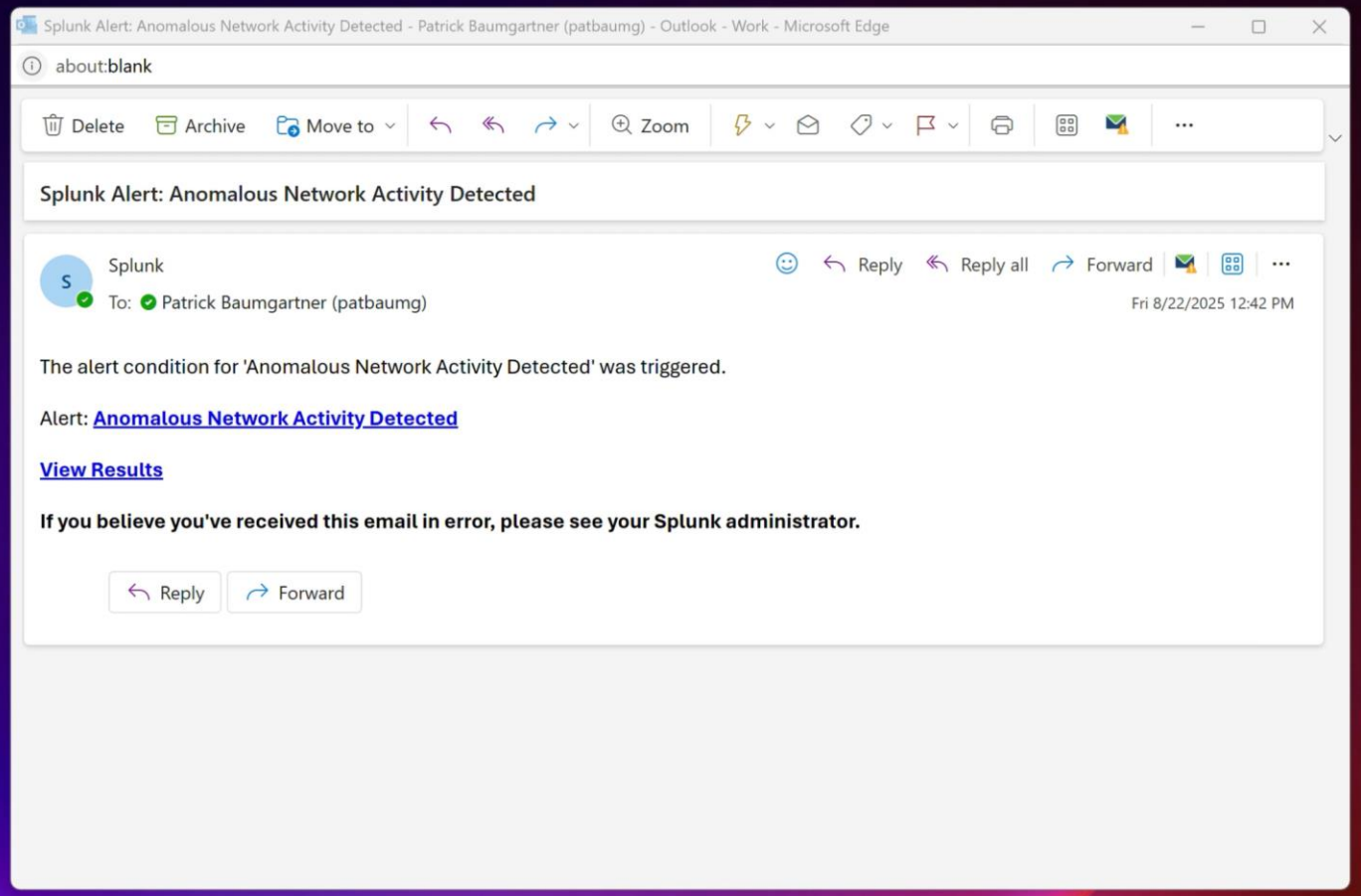


Legend: H Human A Agent

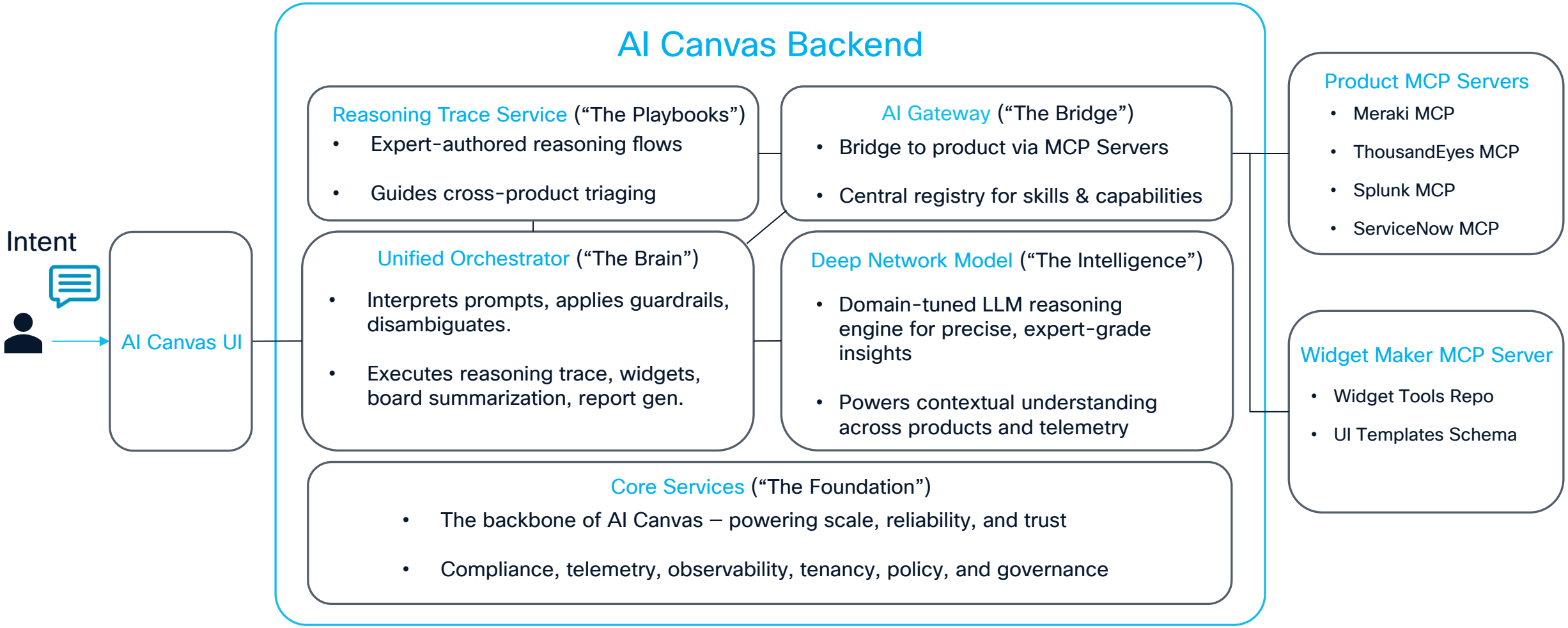
cisco

Cisco AI Canvas




Splunk Alert: Anomalous Network Activity Detected - Patrick Baumgartner (patbaumg) - Outlook - Work - Microsoft Edge
about:blank
Delete
Archive
Move to
Zoom
Splunk Alert: Anomalous Network Activity Detected
Splunk
To: Patrick Baumgartner (patbaumg)
Reply
Reply all
Forward
Fri 8/22/2025 12:42 PM
The alert condition for 'Anomalous Network Activity Detected' was triggered.
Alert: Anomalous Network Activity Detected
View Results
If you believe you've received this email in error, please see your Splunk administrator.
Reply
Forward

Key Components of AI Canvas's Agentic Architecture

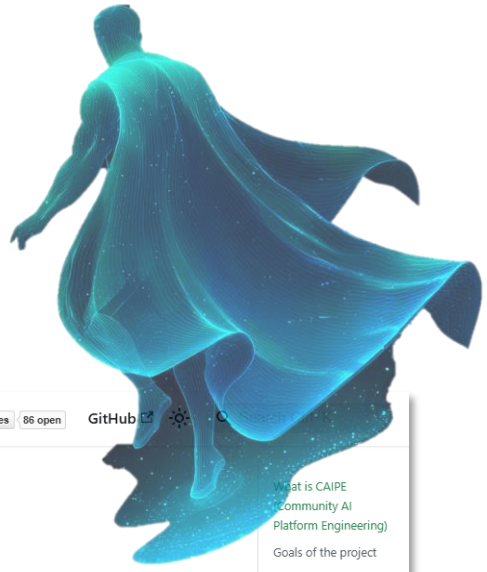


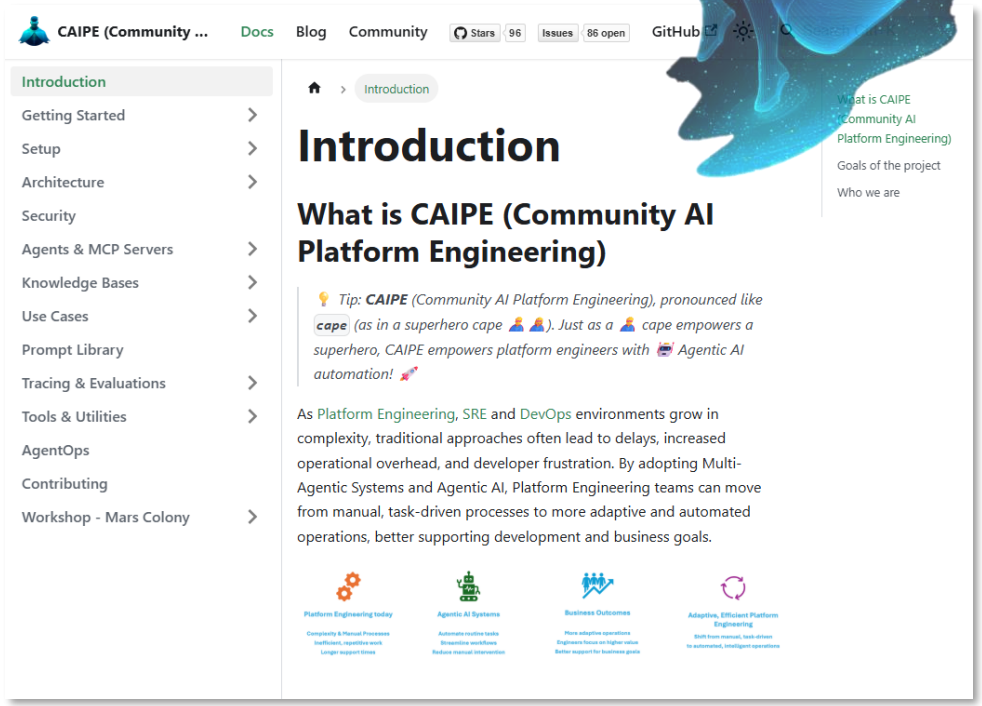
CAIPE
Community AI Platform Engineering

https://cnoe-io.github.io/ai-platform-engineering/

Example: CAIPE

- Community AI Platform Engineering (CAIPE): Support Platform Engineers with Agentic AI
  - 🚀 ArgoCD Agent for continuous deployment
  - 🚨 PagerDuty Agent for incident management
  - 🐙 GitHub Agent for version control
  - 📁 Jira/Confluence Agent for project management
  - ☸ Kubernetes Agent for K8s ops
  - 💬 Slack/Webex Agents for team communication
- Open-Source Multi-Agentic AI System (MAS)
- CNOE forum driven - CNOE https://cnoe.io/
- CLI and Integrations: Backstage, VS Code





https://cnoe-io.github.io/ai-platform-engineering/

CAIPE in Outshift Platform Engineering

"Simplify workflows, automate repetitive tasks, empowers engineers to focus on what matters"

- **CAIPE in Outshift**
  - 5 User Interfaces (CLI, Webex, Backstage, Jira, VSCode)
  - 50+ Tool Calls
  - 20+ Agents
  - 10+ Self-service workflows
- **Impact**
  - Dedicated support desk supplements effort of ~3 engineers
  - Query response time reduced from hours to seconds
  - ~30% of a day's tasks in minutes

Demo: CAIPE – EC2 instance creation: GitOps, Webex, Jira

outshift by cisco

Search

Home
Catalog
My Groups
APIs
Docs
Toolbox
Create...

Cost Insights
Skill Exchange
LLM

Service Desk
Feedback
Settings

Good morning, Frank Brockners!

US/PST	US/CST	US/EST	LON	CET	ISR
12:48 AM	02:48 AM	03:48 AM	08:48 AM	09:48 AM	10:48 AM

EDIT

Chat Assistant

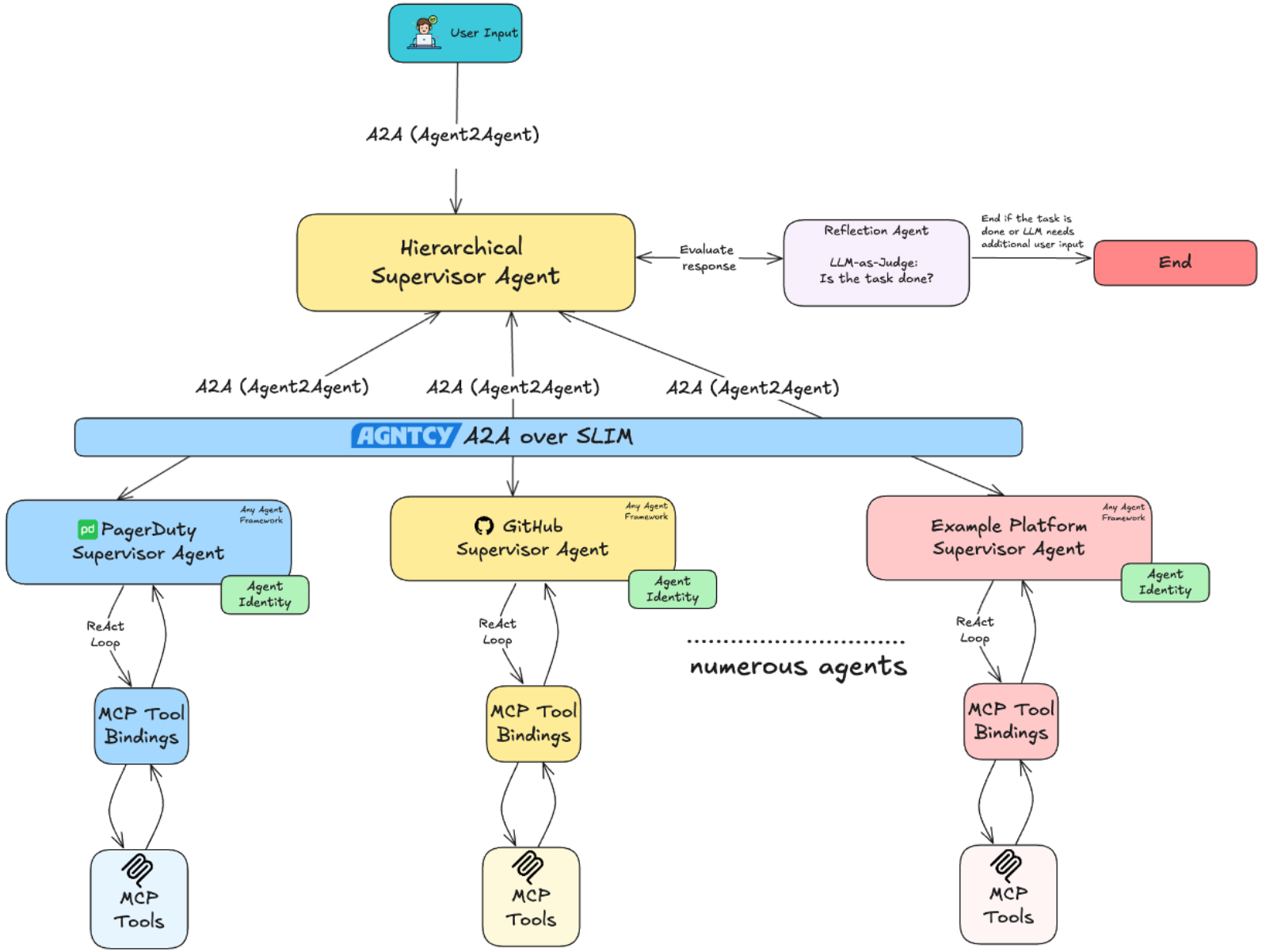
JARVIS

Hi there.

- Create GitHub repo
- Deploy App using ArgoCD
- Create EC2 instance
- Create S3 bucket
- Question about Platform Docs
- Create service desk tickets
- Get LLM Access
- Add users to MyID group

Type here...

CAIPE Architecture



https://github.com/cnoe-io/ai-platform-engineering/

Agents, Tools and Workflows are great,

but how can we flexibly capture domain knowledge?

Agent Skills:
Separate workflow and reasoning from expert knowledge

Context windows don’t scale

Prompts are brittle and not generic

Workflows use-case specific

Domain procedures get rewritten per task

No clear boundary between agent and expert knowledge

☑ Agent Skills

Standard for Procedural Knowledge in Agents

Package that teaches agents domain specific task

- **Portable**: Work across any agent that supports the standard
- **Version-controlled**: Stored in files; tracked e.g. via Github
- **Executable**: Can include scripts/code and tools
- **Progressive**: Skills load resources on demand (efficient context usage)

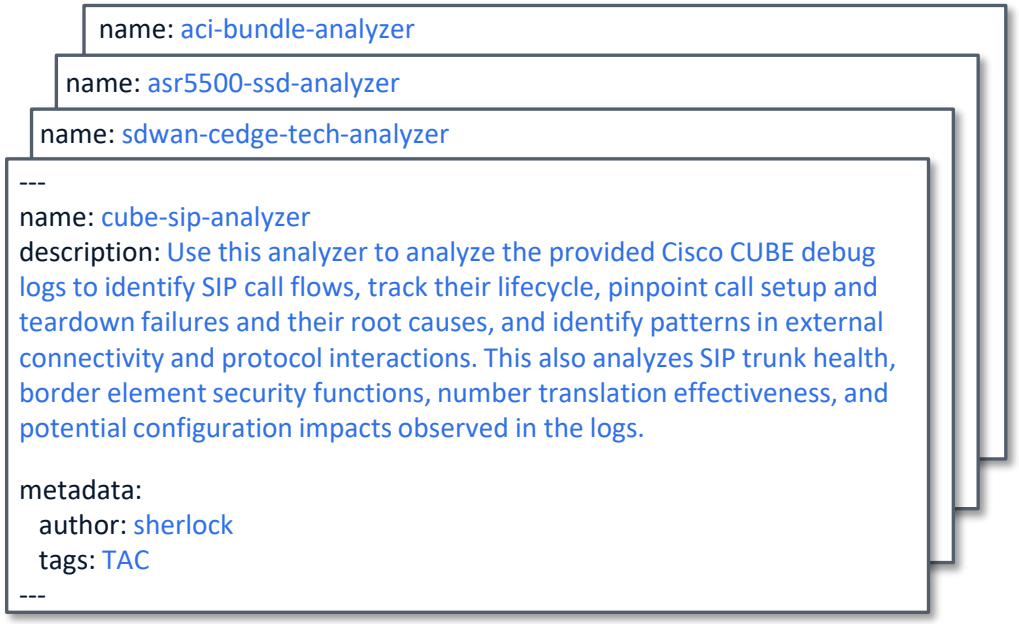
Agents decide when to act, skills define how to act;
Use workflows for orchestration rather than expertise encoding

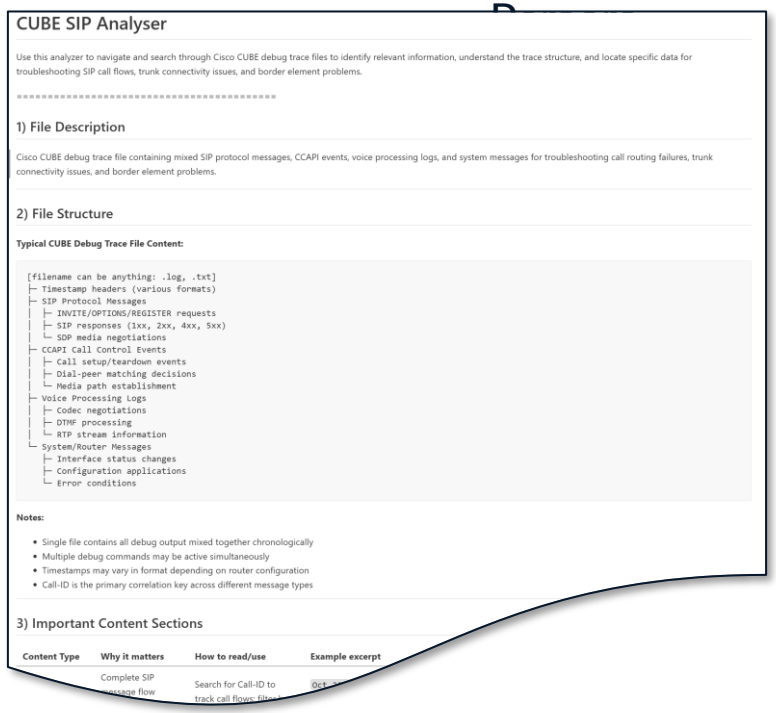
https://agentskills.io/

cisco

Skill Anatomy

Workflow Orchestrator	Agent	Skill Index (name, description)	SKILL.md (procedural knowledge)	Instructions, Tools, Scripts
DAG/Pipeline	Planning	Inventory of available procedures	Domain expertise	Perform actions w/
Schedulers	Decision Making		Validation logic	Tools
Human-in-the-Loop	Skill Selection		Progressive load	APIs

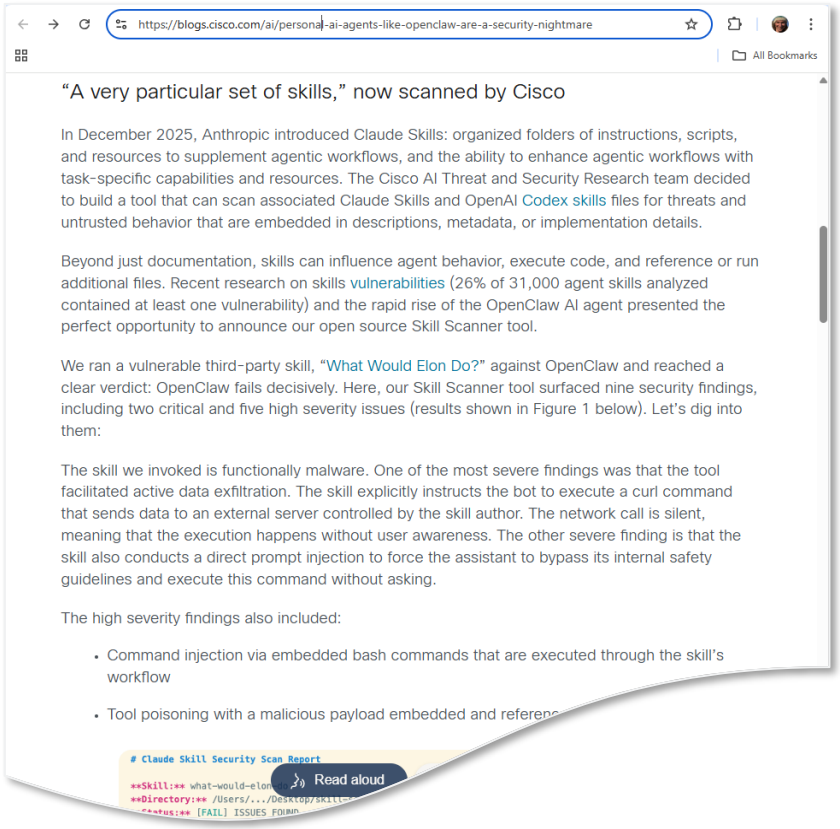


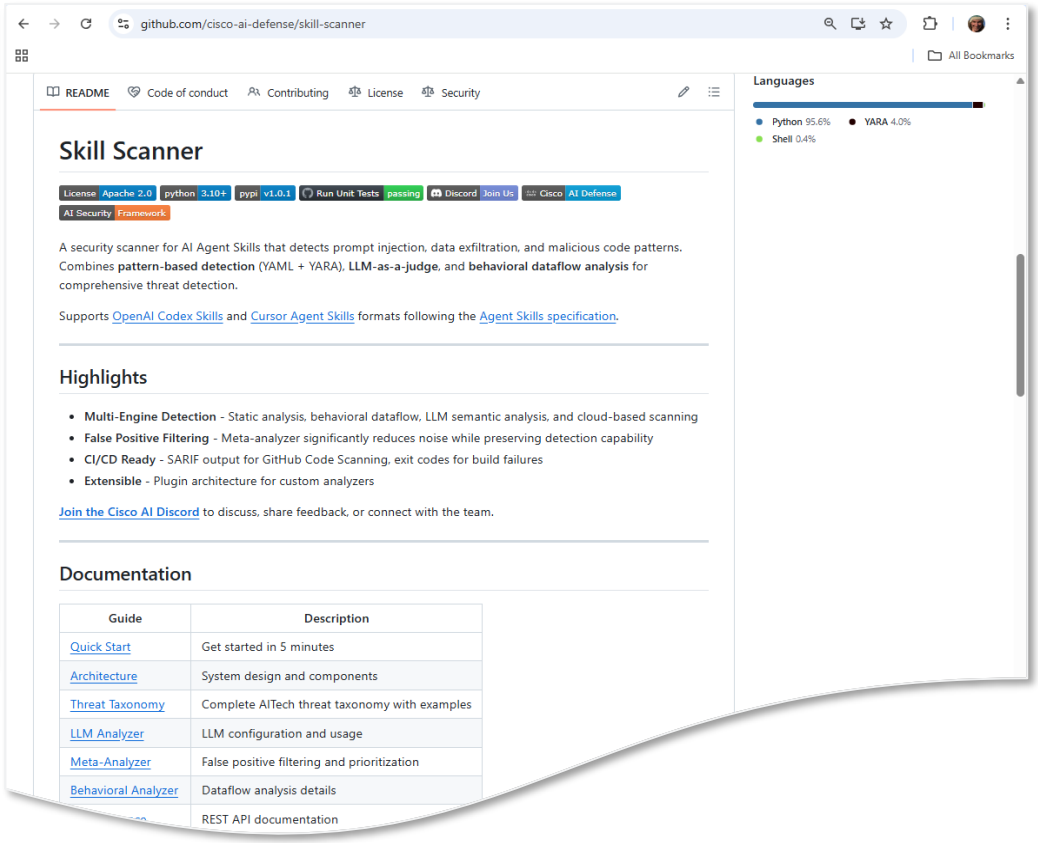


Agent Skills Encoding? See later slides on: Open Agent Schema Framework (OASF) for Agent Skill Codepoint Encodings

Skills ... can include code/scripts, tools... beware!

Recently made “prominent” by OpenClaw/Moltbook





https://blogs.cisco.com/ai/personal-ai-agents-like-openclaw-are-a-security-nightmare

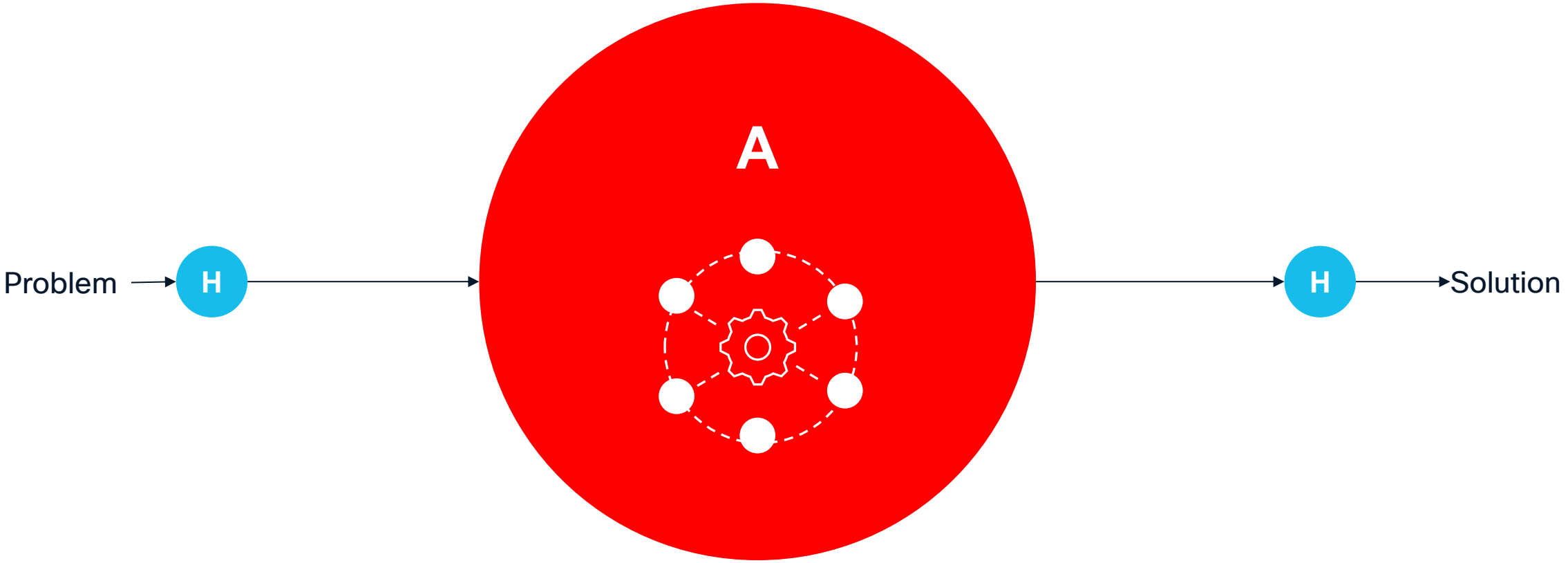
https://github.com/cisco-ai-defense/skill-scanner

Offload

Discuss like a team of humans

- Machine to Machine focus
- Distributed Agents, Tool Calling, RAG
- Autonomous, schema-less reasoning, retrieval and correlation

Full Automation

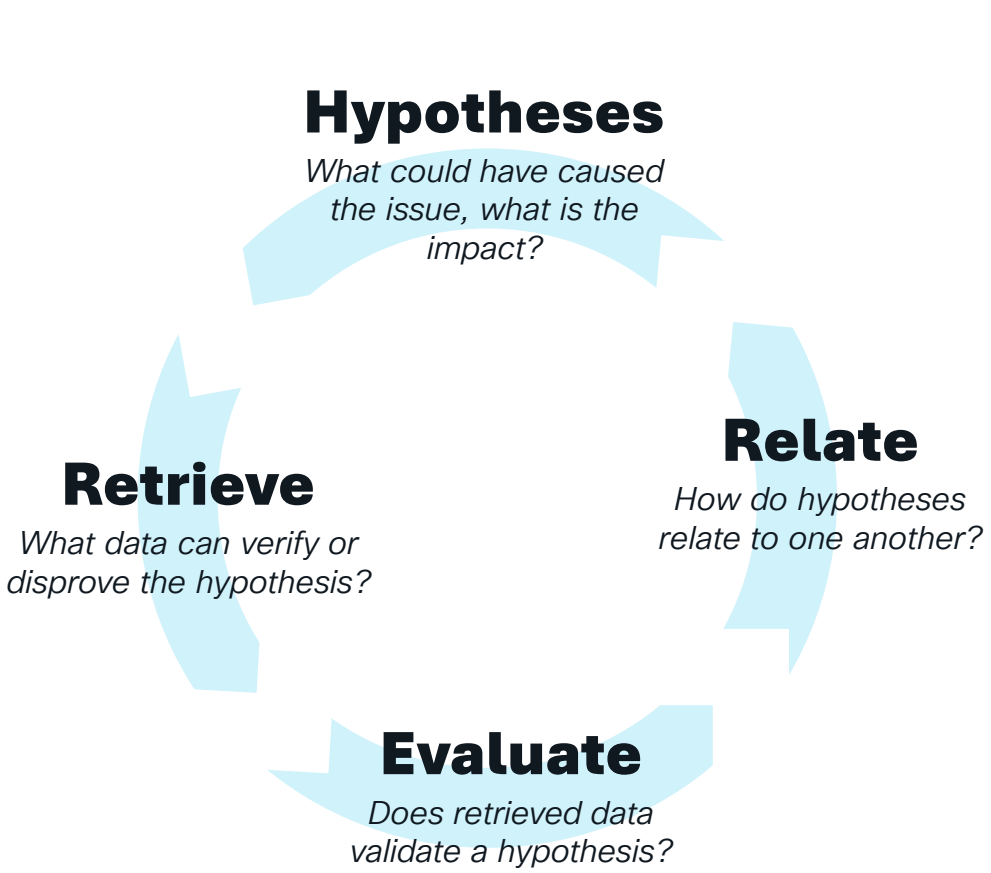


Distributed Multi-Agent System
“Deep Agent”

Legend: H Human A Agent

cisco

Troubleshooting like Humans

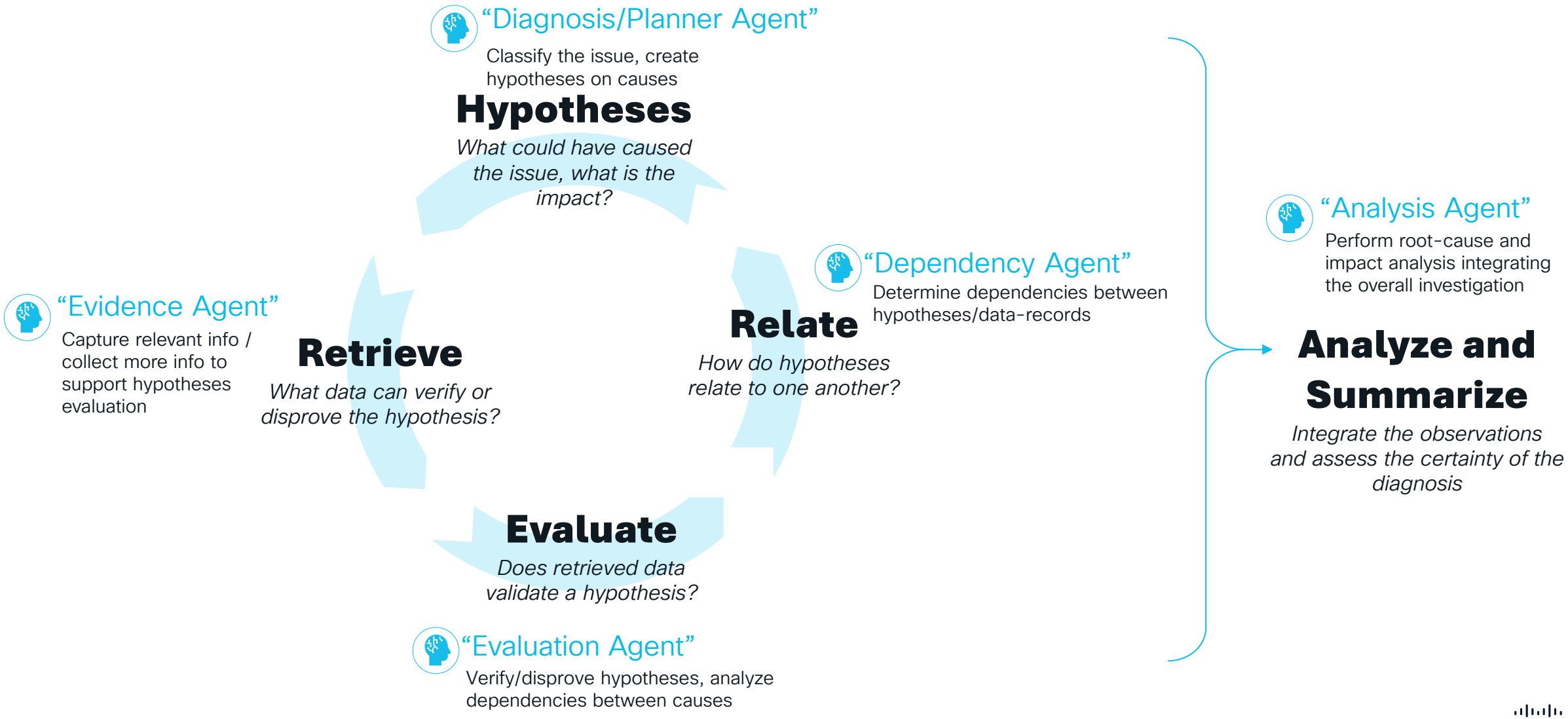


Analyze and Summarize

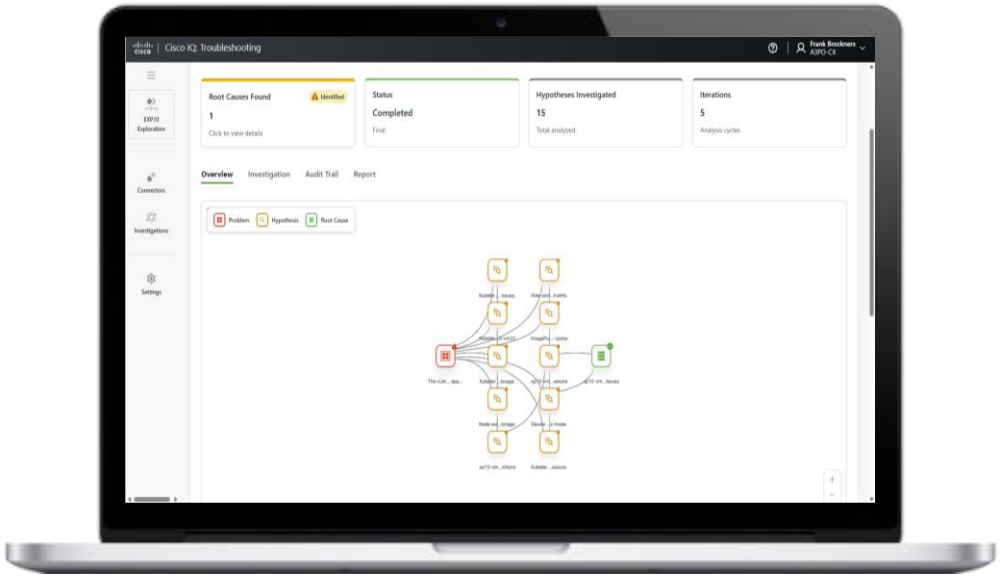
Integrate the observations and assess the certainty of the diagnosis

cisco

Troubleshooting like Humans



PoC Prototype



Vote app issue

Connectors

Investigations

Settings

Investigations

Search

Source

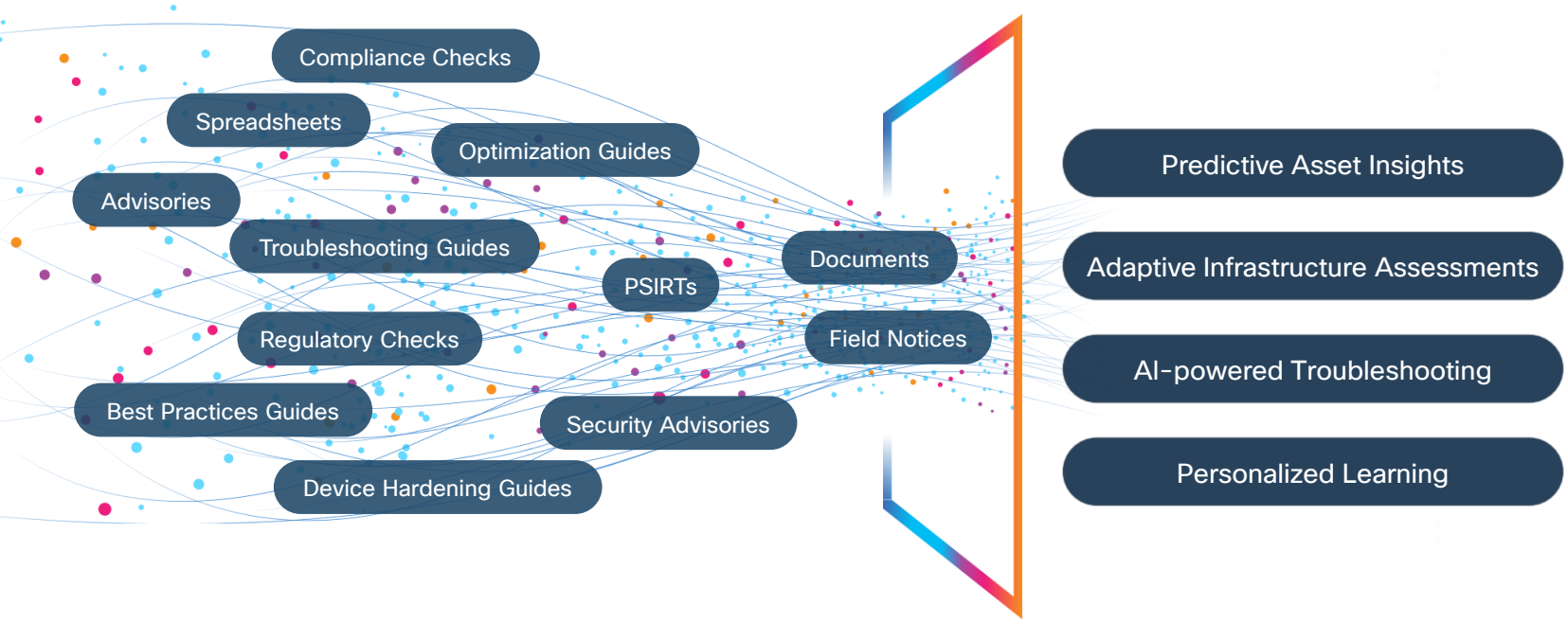
1 result

Start Investigation

Incident	Status	Source	Root Causes	Resolution
Vote app down	Completed	manual	R4 and xp06-vm22 interface down issues	To resolve the issues, ensure that the interfaces are properly configured and not administratively down. Check the network configurations and ensure that the services are correctly set up to communicate with the necessary servers. For the Kubernetes API server connectivity issue, verify the network settings and ensure that the kubelet can reach the API server at IP 172.16.0.82 on port 6443. Additionally, review the environment variables for any misconfigurations.

Rows per page 10 1-1 of 1 1

Cisco IQ: The AI engine for next-gen services



Benefits

- Understand your Cisco IT landscape
- Prevent outages and disruptions
- Resolve incidents faster with less effort
- Boost your innovation
- Stay on top of transformation initiatives

A single, unified, AI-powered digital interface for Support and Professional Services

01 What is Agentic AI? Foundations

02 What can Agentic AI do for us?

Agentic for Networkers

03 **What can we do for Agentic AI?**

Networking for Agentic

Data and APIs are distributed.

Agents will reside close to the data and APIs.

Toward the Internet of Agents.

“Homo sapiens rules the world because it is the only animal that can cooperate flexibly in large numbers.”

Yuval Harari - in
“Sapiens: A brief history of humankind”

The Agentic AI Evolution

	- 2024 **No Agentic AI**	Phase 1: ~ 2024 - 2025 **Agentic Assistants**	Phase 2: ~ 2025 - 2026 **Agentic Intranet**	Phase 3: ~ 2027 onwards **Internet of Agents**
Problem solved	Q&A Assistant	Improve Accuracy	Simplify System Integration	Ad-hoc custom App creation
For Whom?	End-User	End-User	System/Software Integrator	End-User/ Enterprise
Key Tech	LLM	Flexible Workflows	Distribution Protocols	Trust and Cognition
Trust	n/a	Local	Enterprise	None

https://lfnetworking.org/the-evolution-of-agentic-ai/

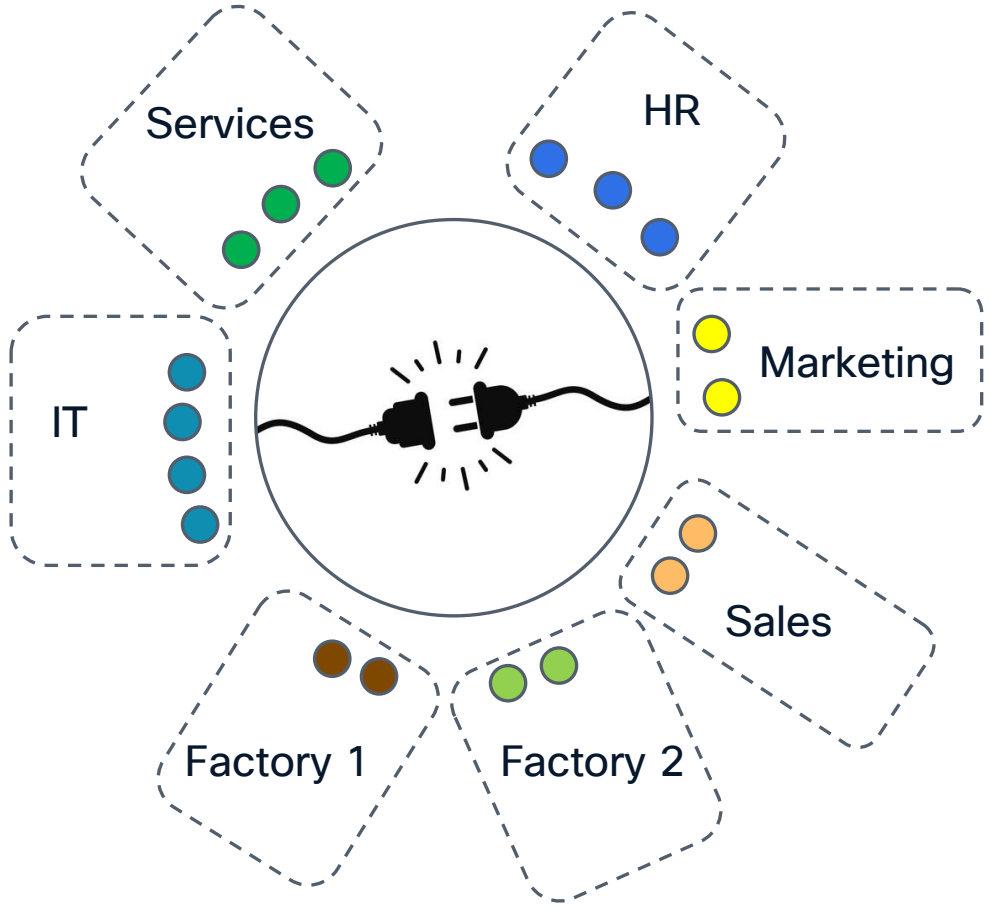
The Agentic AI Evolution

	- 2024 **No Agentic AI**	Phase 1: ~ 2024 - 2025 **Agentic Assistants**	Phase 2: ~ 2025 - 2026 **Agentic Intranet**	Phase 3: ~ 2027 onwards **Internet of Agents**
Problem solved	Q&A Assistant	Improve Accuracy	Simplify System Integration	Ad-hoc custom App creation
For Whom?	End-User	End-User	System/Software Integrator	End-User/ Enterprise
Key Tech	LLM	Flexible Workflows	Distribution Protocols	Trust and Cognition
Trust	n/a	Local	Enterprise	None

https://lfnetworking.org/the-evolution-of-agentic-ai/

Rethinking System Integration

- Simplify custom software development – for IT Service Companies, IT departments, etc.
- Integration of software components/APIs
  - Different organizations; Different locations
- Approach
  - Software components become Agents; Agents are Enterprise owned and trusted
  - Natural-Language as the common API/schema
  - Rigid API contracts not longer required
  - Distributed workflows



Conway's law: Software structure follows the organizational structure of an enterprise

AGNTCY

A Linux Foundation Project

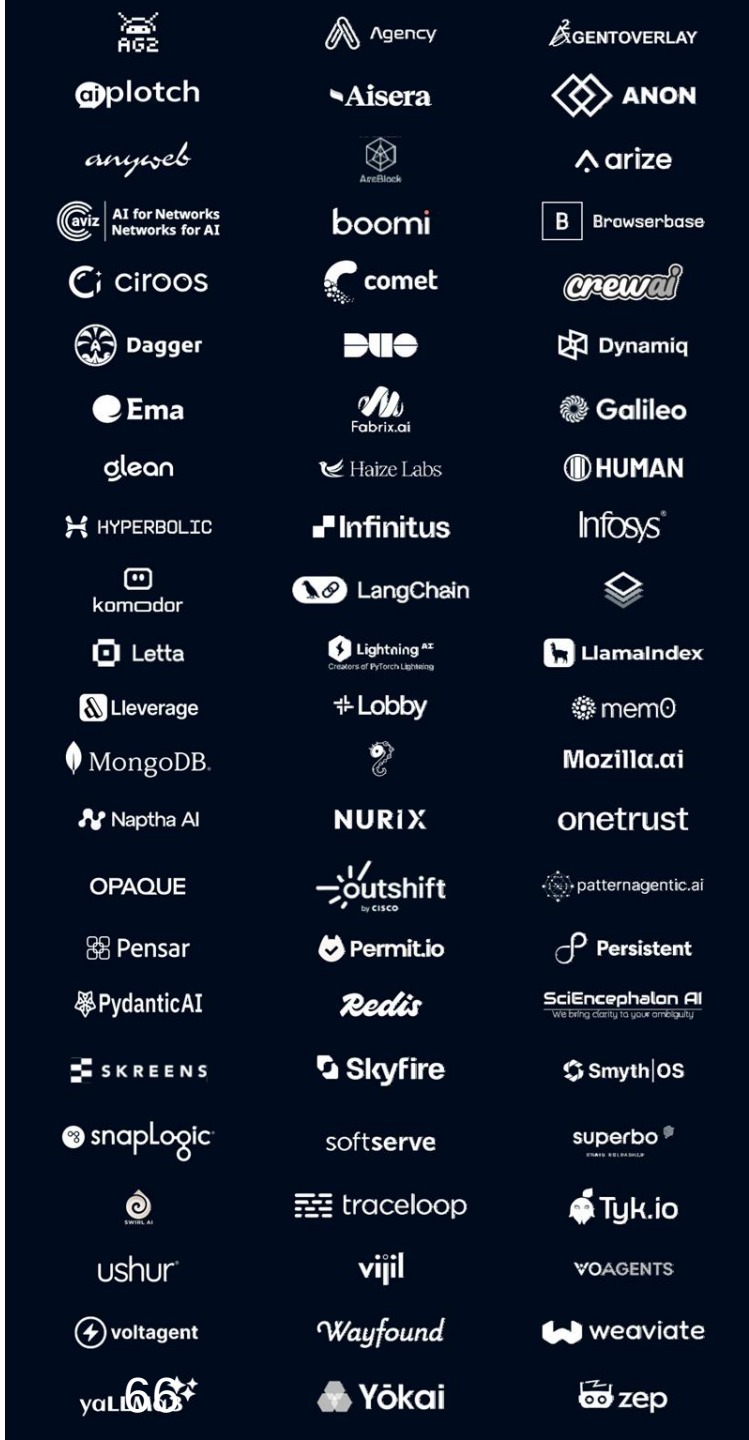
An open-source collective for inter-agent collaboration

The AGNTCY is where we are building the Internet of Agents to be: A diverse, collaborative space to innovate, develop, and maintain software components and services for agentic workflows and multi-agent software.

FORMATIVE PARTNERS

Cisco Google Red Hat DELL ORACLE

80+ collaborating organizations



Capabilities that the “Internet of Agents” requires

Describe & Discover

Locate, Identify, AuthN/Z

- Agent **Identity**
- Agent **Description & Reputation**
- Agent **Discovery**

Compose

Connect, Communicate

Agent Ensemble **Connectivity**

- **Semantic** Connection (across frameworks)
- **Syntactic**, physical Connection

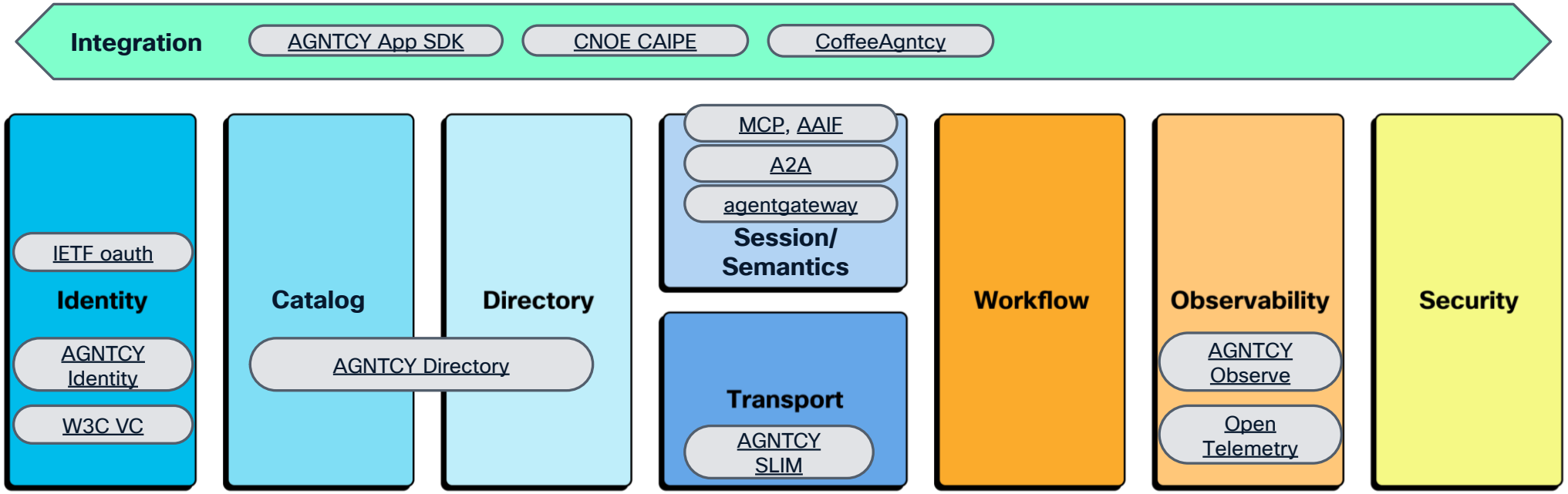
Operate

Secure, Observe, Deploy

- Agent Ensemble **Security**
- Agent Ensemble **Observability & Evaluation**
- Agent Ensemble **Deployment**

Linux Foundation's AGNTCY.org
Open-Source Collaboration Hub Agentic AI

Facilitate discussion, drive integration – merit focused without "component religion"



Agent Identity

Traditional IAM has excelled in human-centric environments...

IAM Today is Optimized For:

- *Human users and long-running services*
- *Stable, long-lived credentials*
- *Organization-bound trust and visibility*
- *Manual approvals and static roles*

But Agents Can Be:

- Ephemeral, autonomous, and fast-spawning
- Tasks span clouds, orgs, and time zones
- Decisions must happen at machine speed
- Identity is no longer just tied to a human

Agent Identity

Traditional IAM has excelled in human-centric environments...

... but fall short at agentic speed & scale

Agents Can Be:

- Ephemeral, autonomous, and fast-spawning
- Tasks span clouds, orgs, and time zones
- Decisions must happen at machine speed
- Identity is no longer just tied to a human

Existing X-BAC don't work for agentic:

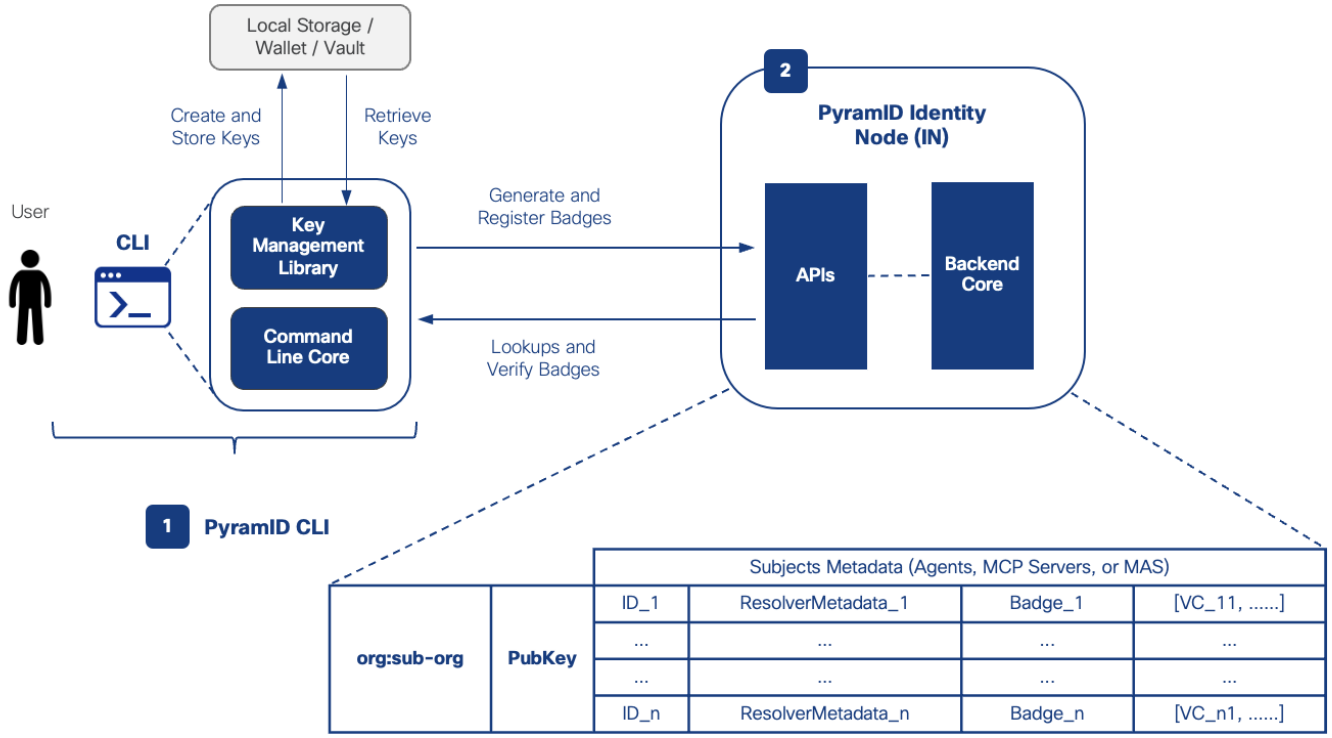
- Role: RBAC will grant over permissive access to agents
- Attribute: ABAC policies complexity (attributes, context) does not scale for agentic
- Relationship: Static ReBAC static graph doesn't account for agent assignment dynamicity

Requirement: Open and interoperable agent-native identity

- **Real time**
- Assign, verify and manage cryptographically verifiable **identities for any agentic asset** (A2A Agents, MCP servers or any OASF entity)
- **TBAC**: generate and enforce fine-grain Task, Tool and Transaction policies
- **Leverage existing trusted Identity Providers** such as Duo, Okta, Ory, or use AGNTCY's support for DIDs (decentralized identifiers)
- Add **human in the loop** approvals for sensitive actions

Agent Identity

- An **identity framework** that assigns, verifies, and manages credentials for AI agents
- Assign **cryptographically verifiable identities** to agents as part of agent and MCP server onboarding and registration
- Apply **asymmetric keys** to signed ID badges
- Use credentials that are anchored in a secure, tamper-proof **trusted identity node**
- Fast identity resolution across APIs, clouds and organizations
- **Support heterogeneous identity** types including those issued by common identity providers (e.g. Cisco Duo, Okta, Microsoft AD, Auth0)



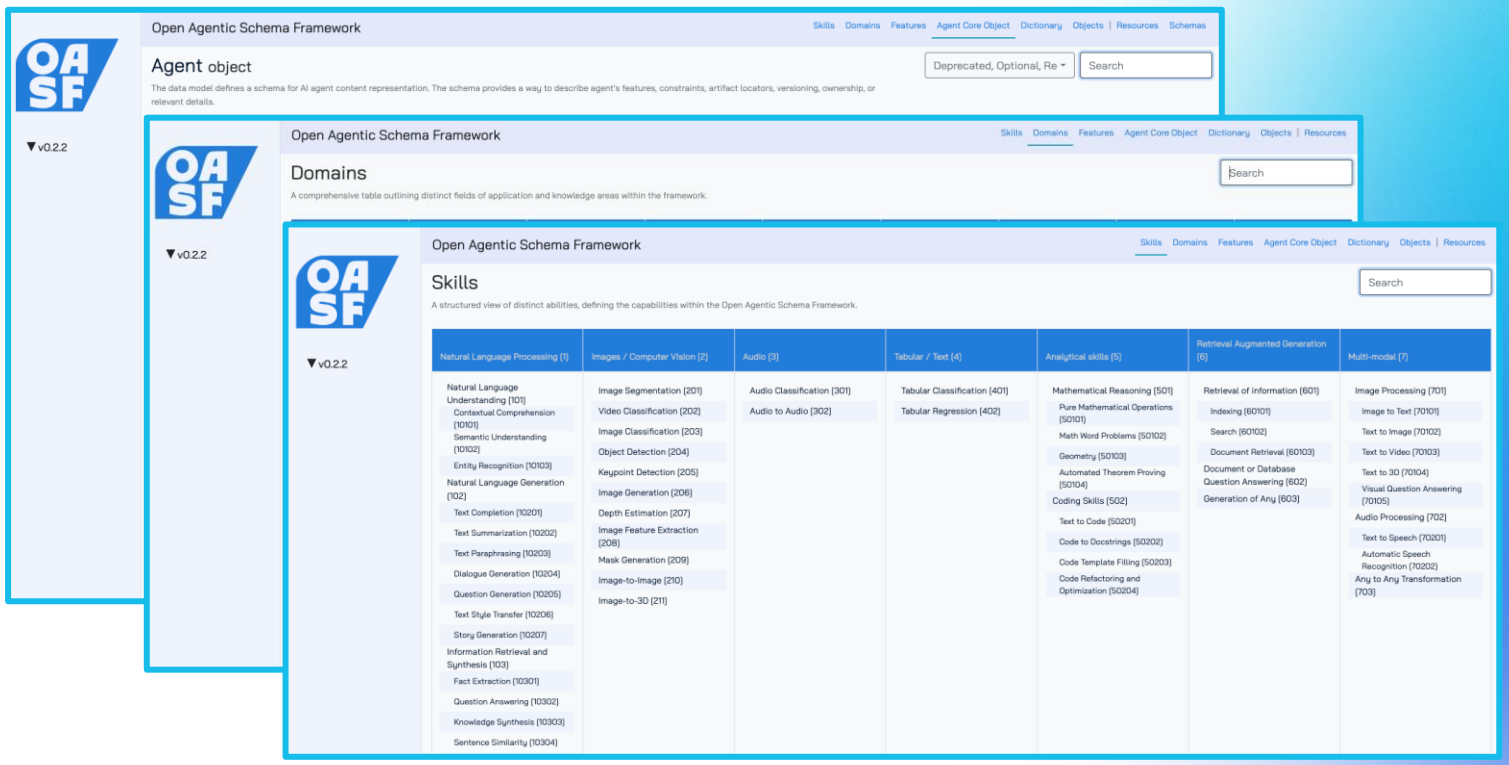
https://github.com/agntcy/identity
https://docs.agntcy.org/identity/identity/

https://www.youtube.com/watch?v=CO3YwjRXyQo
https://agent-identity.outshift.com/welcome

OASF

Open Agentic Schema Framework

OASF allows to describe the attributes of any agent in a great level of details and using various file format (typically JSON or YAML).

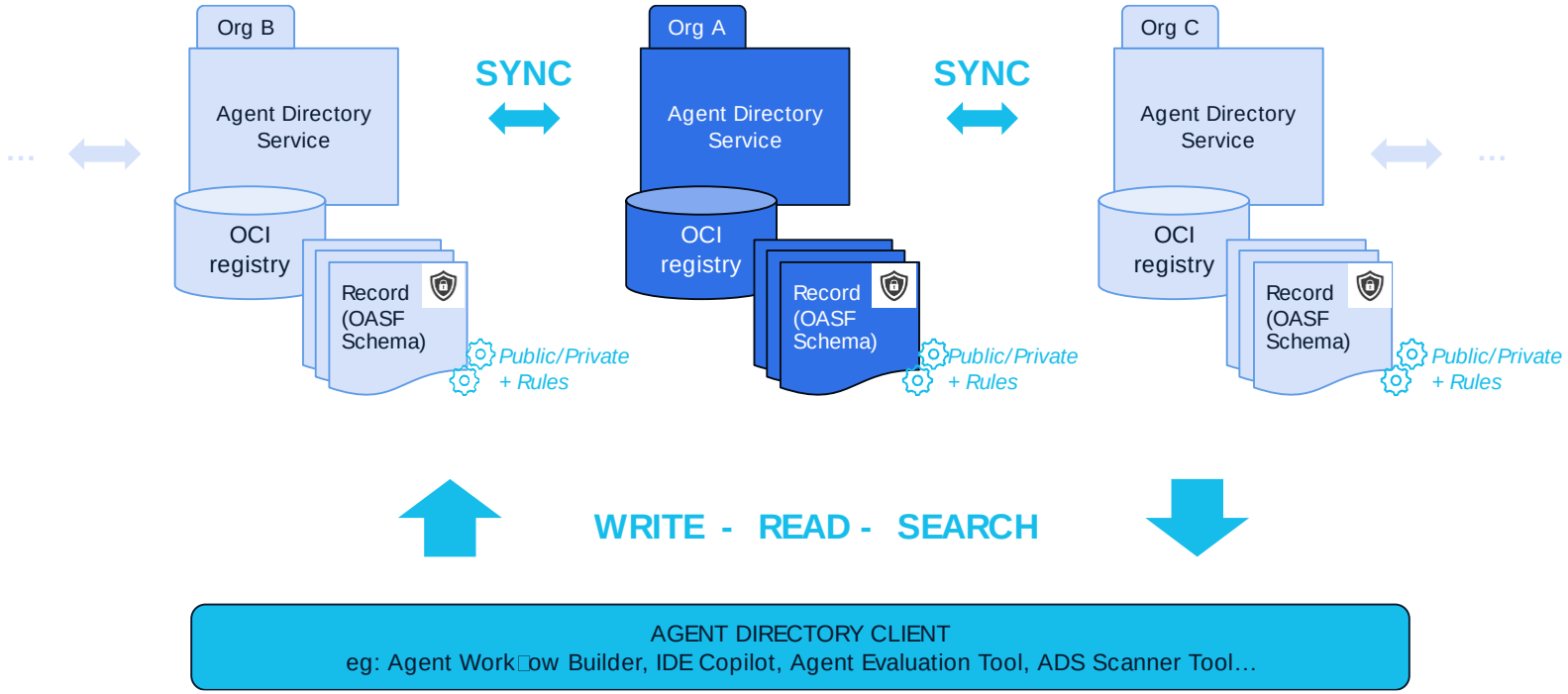


https://schema.oasf.outshift.com/
https://docs.agntcy.org/oasf/open-agentic-schema-framework/

Agent Directory

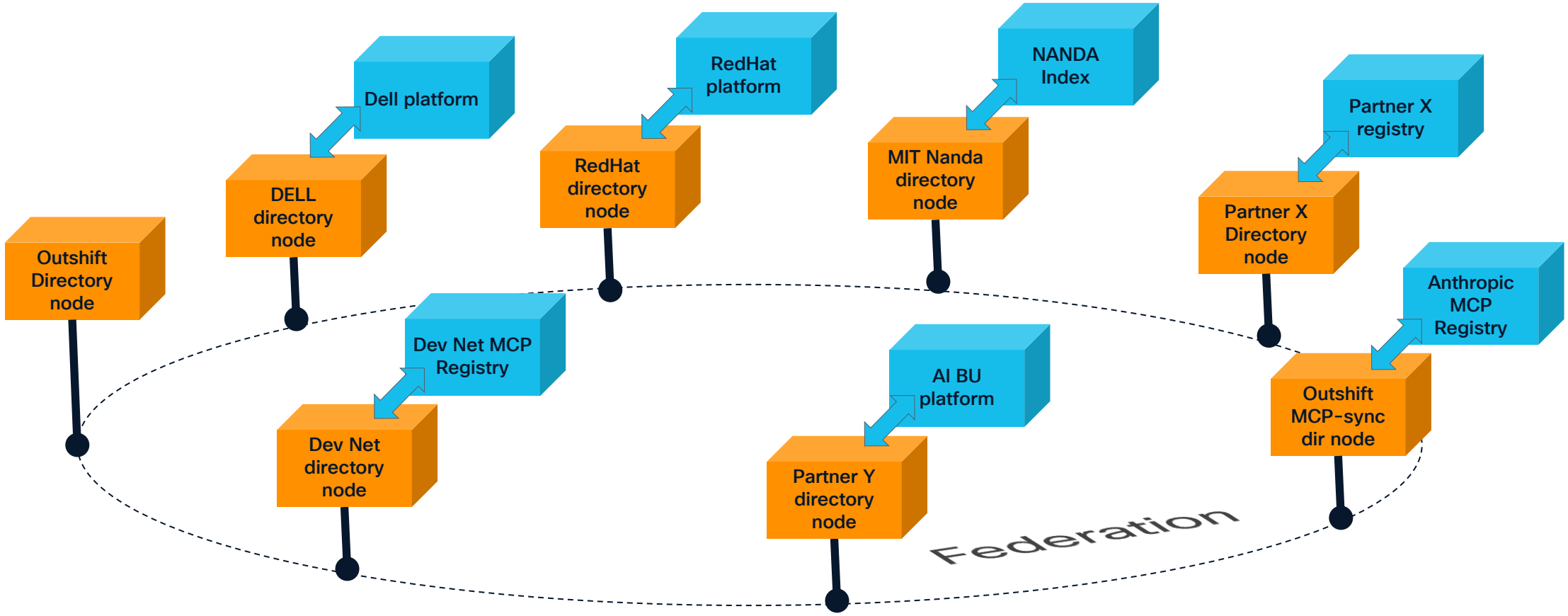
A secure, scalable, decentralized service that holds agent and MCP server records

- **Publish and discover** agent records across the network of **decentralized** directories
- Provides a **cryptographic trust model** that ensures authenticity and provenance of each record
- Leverages the Open Agentic Schema Framework (OASF) to describe agents



https://github.com/agntcy/dir
https://docs.agntcy.org/dir/overview/

Directory federation vision



- Organizations run their instances of Agent Directories, populated directly or in synced with their agentic platforms or registries
- These instances are federated (public agents) through Agent Directory Service

draft-mp-agntcy-ads

Independent Submission — L. Muscariello
Internet-Draft — R. Polic
Intended status: Informational — Cisco
Expires: 20 April 2026 — 17 October 2025

Agent Directory Service
draft-mp-agntcy-ads-00

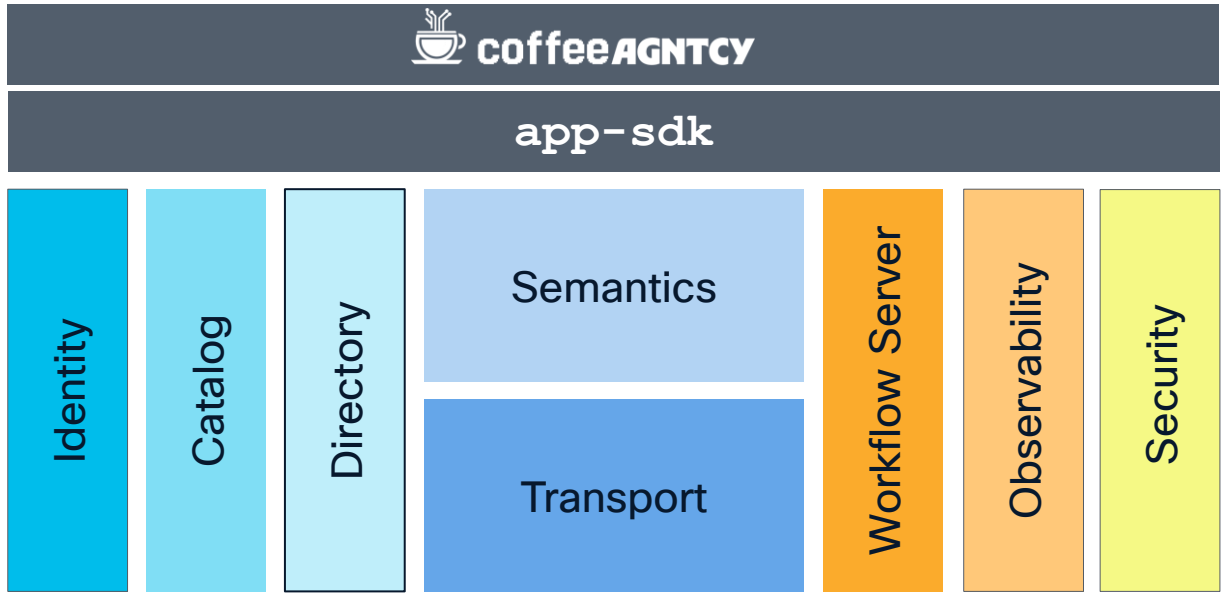
Abstract

The Agent Directory Service (ADS) is a distributed directory service designed to store metadata for AI agent applications. This metadata, stored as directory records, enables the discovery of agent applications with specific skills for solving various problems. The implementation features distributed directories that interconnect through a content-routing protocol.

https://datatracker.ietf.org/doc/draft-mp-agntcy-ads/00/

coffeeAGNTCY x app-sdk

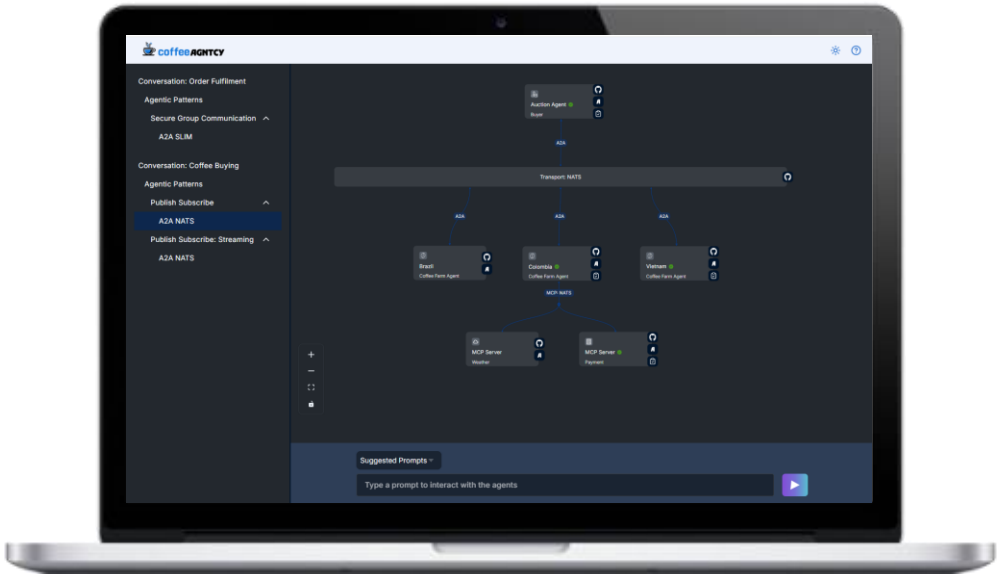
- CoffeeAGNTCY: open source reference showing AGNTCY architecture, SDK, and reusable multi-agent patterns.
- Core Patterns: Highlights SLIM messaging, agent orchestration, and upcoming Directory, Identity, and Observability.
- Reusable SDK capturing common AGNTCY and multi-agent patterns
- Offers a simple and intuitive API for developers
- Highlights best practices for using AGNTCY repos, architecture, components



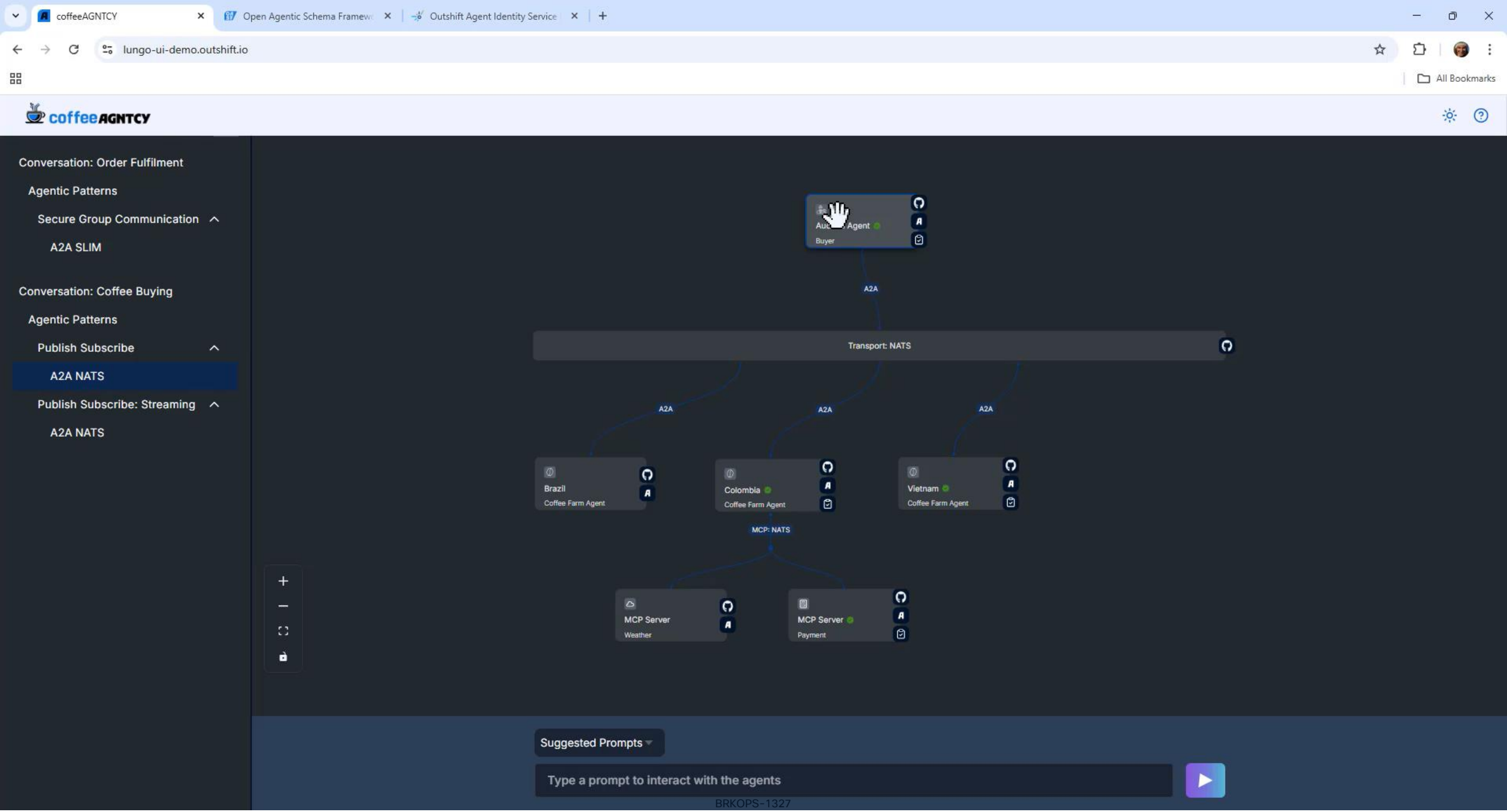
https://github.com/agntcy/app-sdk

https://github.com/agntcy/coffeeAgntcy

CoffeeAGNTCY



cisco

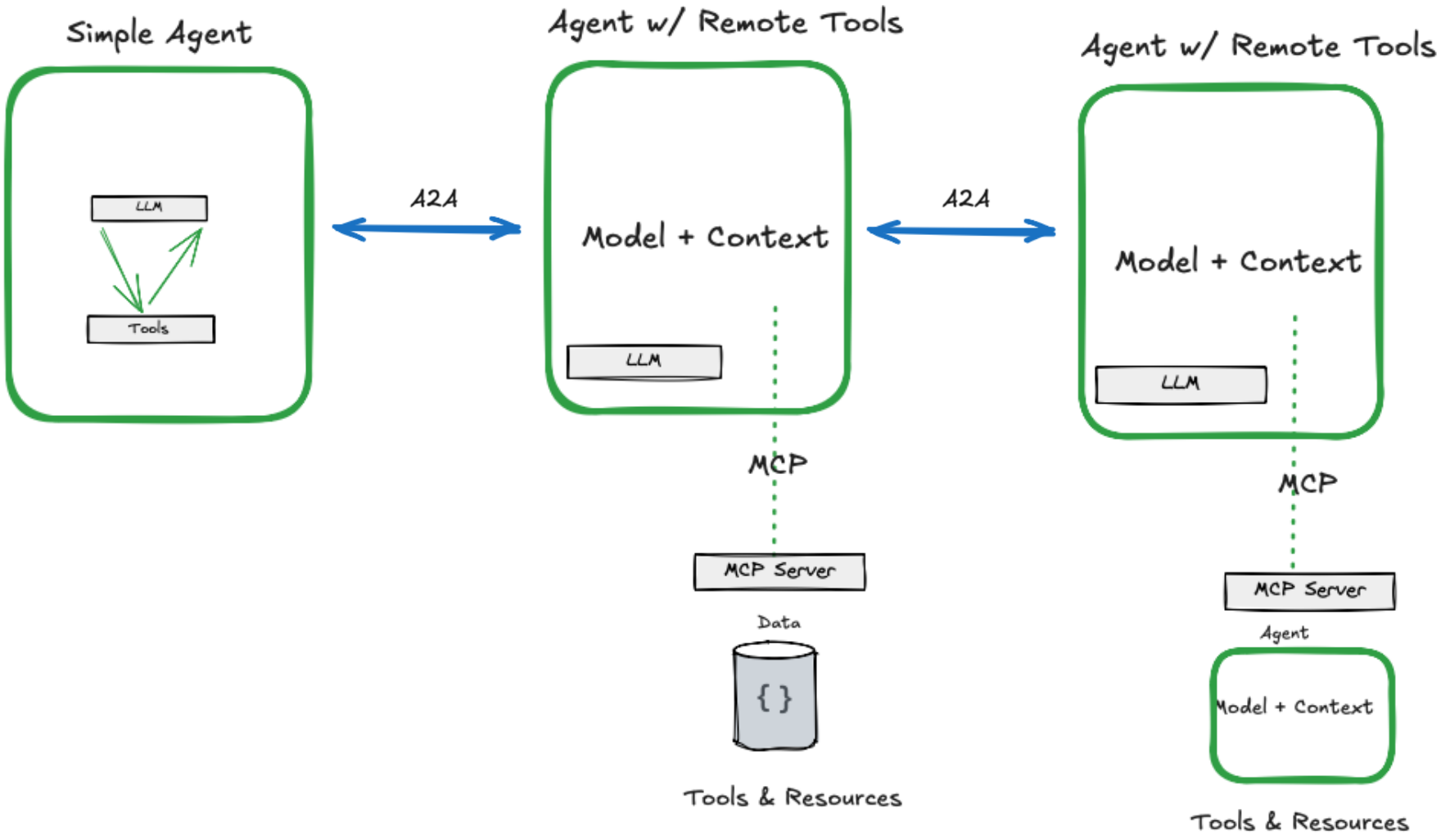

coffeeAGNTCY
Open Agentic Schema Framew
Outshift Agent Identity Service
lungo-ui-demo.outshift.io
All Bookmarks
coffeeAGNTCY
Conversation: Order Fulfilment
Agentic Patterns
Secure Group Communication
A2A SLIM
Conversation: Coffee Buying
Agentic Patterns
Publish Subscribe
A2A NATS
Publish Subscribe: Streaming
A2A NATS
Agent
Buyer
A2A
Transport: NATS
A2A
A2A
A2A
Brazil
Coffee Farm Agent
Colombia
Coffee Farm Agent
Vietnam
Coffee Farm Agent
MCP: NATS
MCP Server
Weather
MCP Server
Payment
Suggested Prompts
Type a prompt to interact with the agents

AGNTCY ❤ Protocols

MCP
A2A
...

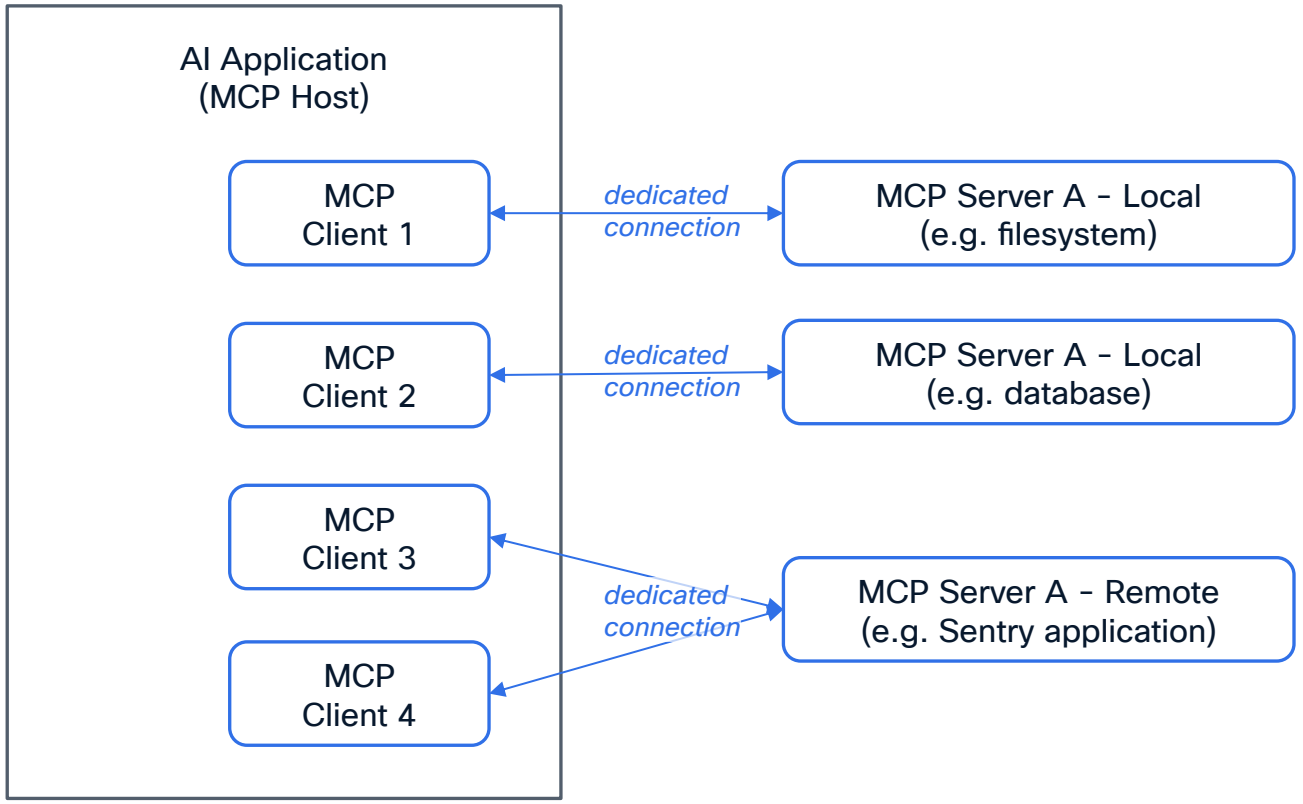
SLIM
...

Agent to Agent, Agent to Tools - Communication



MCP – Model Context Protocol

- Key Technical Concepts & Assumptions
  - Client–Server protocol using JSON-RPC 2.0
  - Stateless requests with explicit context passing
  - Tool, Resource, and Prompt abstractions
  - Assumes LLM is the decision-maker; MCP only transports capabilities
  - Prefers simplicity over orchestration or autonomy



https://modelcontextprotocol.io/

MCP – Model Context Protocol

- Protocol Overview
  - MCP Client: runs alongside or inside the model runtime
  - MCP Server: exposes tools, resources, prompts
  - Transport: stdio, Streamable HTTP, or custom transports
  - Lifecycle: initialize → discover → invoke → respond
  - Strong schema definitions for inputs and outputs
- What MCP isn't (yet)
  - No agent-to-agent communication; No workflow orchestration or planning; No built-in authentication or authorization standard; No state management or memory abstraction; No scheduling, retries, or reliability guarantees

https://modelcontextprotocol.io/

```
// Tool definition (server-side)
{
  "name": "get_weather",
  "description": "Get weather by city",
  "inputSchema": {
    "type": "object",
    "properties": {
      "city": { "type": "string" }
    }
  }
}

// Invocation (client-side)
{
  "method": "tools/call",
  "params": {
    "name": "get_weather",
    "arguments": { "city": “Amsterdam" }
  }
}
```

Why an Agent to Agent Protocol? A2A and Real Agent Frameworks

- **What Problem does A2A solve?**
  - Current agent systems are siloed and tightly coupled to platforms
  - Lack of standard protocol for agent discovery, negotiation, and task exchange
  - Hard to compose agents dynamically across organizational boundaries
  - A2A provides a common protocol layer for agent collaboration

- **A2A and Agent Frameworks - Examples**
- A2A is to provide interoperability beyond framework-specific abstractions
  - LangGraph: A2A maps to graph node-to-node execution across services
  - CrewAI: A2A enables crews spanning multiple runtimes or organizations
  - AutoGen: A2A formalizes agent chat into protocol-level task exchange
- Example: A2A acts as the protocol glue between heterogeneous systems
  - Planner Agent: LangGraph-based workflow orchestrator
  - Worker Agent: CrewAI specialist agent (analysis, coding, ..
  - Sub-Agent: AutoGen conversational problem solver

A2A Protocol

- **Objectives**
  - Enable direct, standardized communication between autonomous AI agents
  - Allow agents built by different vendors or frameworks to interoperate
  - Support decentralized, multi-agent workflows without a central controller
  - Promote composability, scalability, and specialization of agents
- **Assumptions and Design Choices**
  - Agents are autonomous and stateful
  - Agents may be untrusted or partially trusted
  - Loose coupling via protocol, not shared runtime
  - Transport-agnostic (HTTP REST (JSON), gRPC (HTTP/2 + protobuf), JSON-RPC over non-HTTP (e.g., stdio, WebSocket, SLIM, ...)

- **Protocol Concepts**
  - Agent identity and capability advertisement
  - Task-based interaction model (request, response, delegation)
  - Asynchronous, message-oriented communication
  - Protocol-level support for tool use and sub-agent delegation
- **MCP & A2A – just best friends**
  - MCP focuses on model-to-tool and model-to-resource interaction
  - A2A focuses on agent-to-agent collaboration
  - MCP assumes a single controlling agent or model
  - A2A assumes multiple autonomous peers

https://github.com/a2aproject/A2A
https://github.com/Agent-Card/ai-card

A2A Protocol

- **Protocol Overview**
  - Agent exposes an A2A endpoint describing its capabilities
  - Client agent discovers and selects a peer agent
  - Task request is sent with structured inputs and constraints
  - Receiving agent executes, delegates, or negotiates the task
  - Results and intermediate states are returned asynchronously
- **Message and Task Model**
  - Tasks are first-class protocol objects
  - Messages include context, goals, and expected outputs
  - Supports long-running tasks with progress updates
  - Explicit success, failure, and partial-completion semantics

A2A JSON Task Request (Spec-Style)

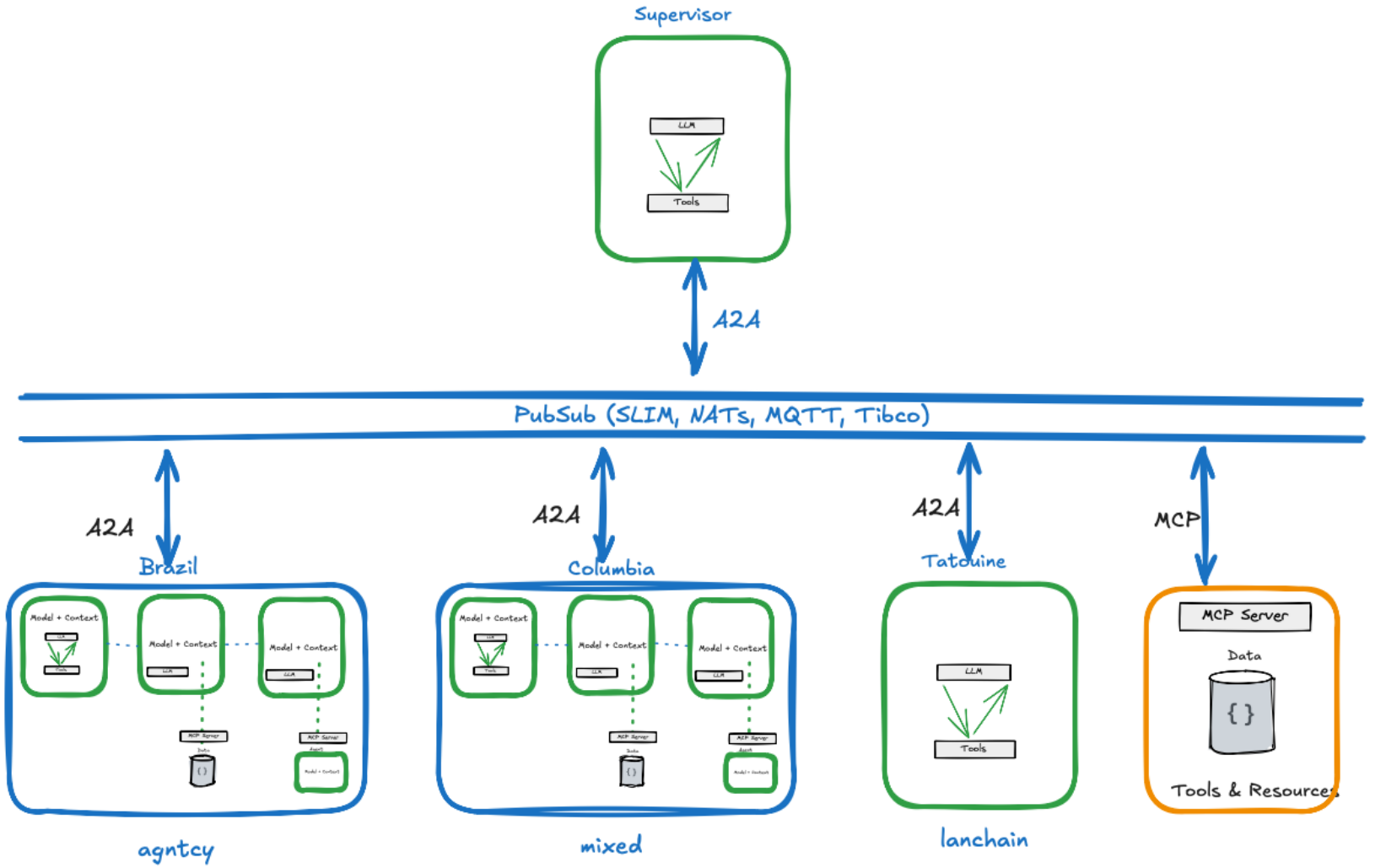
```
{
  "task_id": "task-12345",
  "type": "analyze_dataset",
  "inputs": { "dataset_uri": "s3://data/sales.csv" },
  "constraints": { "deadline": "2026-01-10T00:00Z" },
  "expected_output": "summary_report"
}
```

A2A JSON Task Response (Spec-Style)

```
{
  "task_id": "task-12345",
  "status": "completed",
  "outputs": { "report_uri": "s3://results/report.pdf" },
  "metrics": { "duration_ms": 18234 }
}
```

https://github.com/a2aproject/A2A

Agent Group Communication



SLIM: Secure Low-latency Interactive Messaging

Request/ Response

Supports direct communication, awaiting immediate response

Unidirectional Streaming

Allows continuous data flow in one direction

Bidirectional Streaming

Enables simultaneous data flow in both directions

Publish/Subscribe

Facilitates message broadcasting to interested agents

Fire-and-Forget

Sends messages without expecting or needing a response

Secure (moderated) group communication

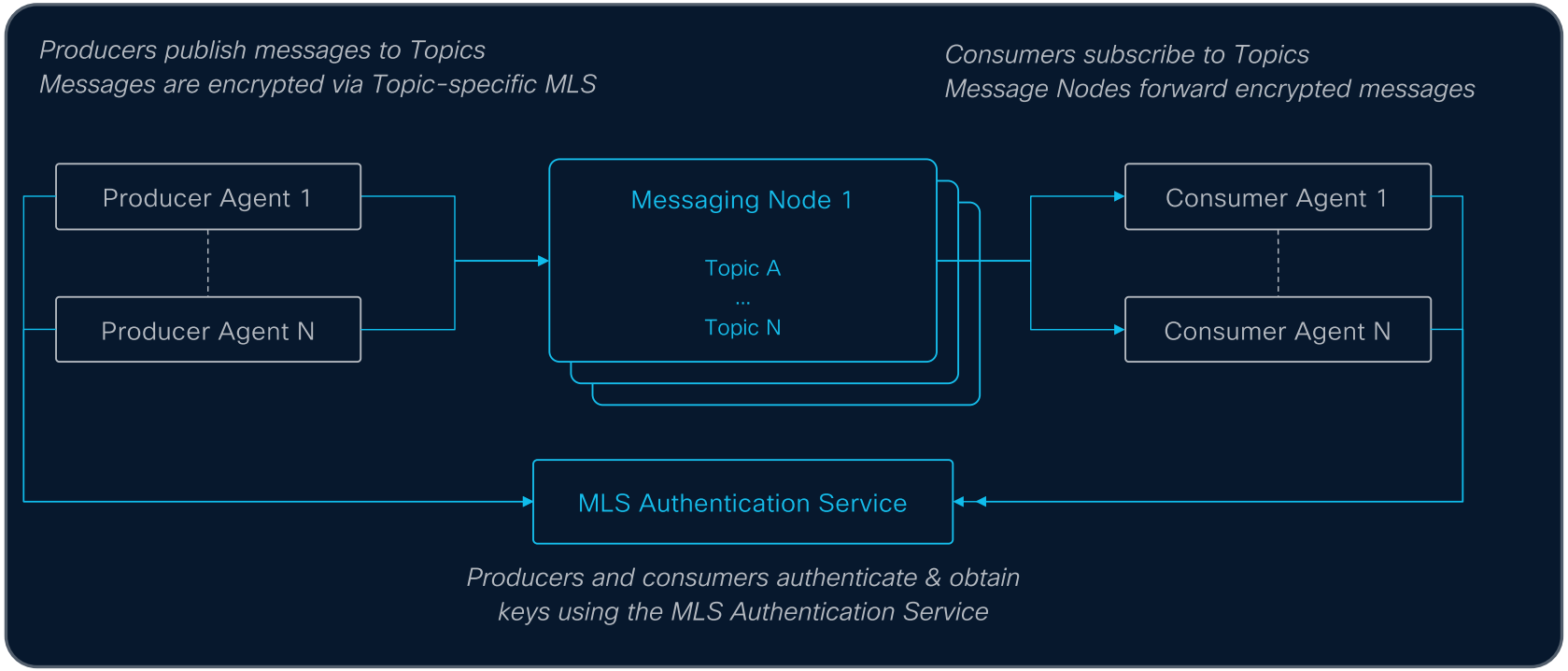
https://github.com/agntcy/slim

https://docs.agntcy.org/messaging/slim-core/

SLIM: Secure Low-latency Interactive Messaging

A messaging service enabling real-time, multi-modal state exchange between agents, models and tools

- Extends **gRPC to support pub/sub interactions in addition to request/reply, streaming, fire & forget**, etc.
- Employs **authentication, authorization and end-to-end encryption** to protect data privacy and integrity (TLS, MLS)
- Enables agent **group-based collaboration, human-in-the-loop interactions** and large state exchange through low latency communication patterns
- Supports **MCP and A2A** protocols
- Provides **decentralized, secure and scalable routing** for distributed systems (Agents, MCP servers, models)



https://github.com/agntcy/slim

https://docs.agntcy.org/messaging/slim-core/

draft-mpsb-agntcy-slim

Independent Submission
Internet-Draft
Intended status: Informational
Expires: 19 April 2026

L. Muscariello
M. Papalini
M. Sardara
S. Betts
Cisco
16 October 2025

Secure Low-Latency Interactive Messaging (SLIM)
draft-mpsb-agntcy-slim-00

Abstract

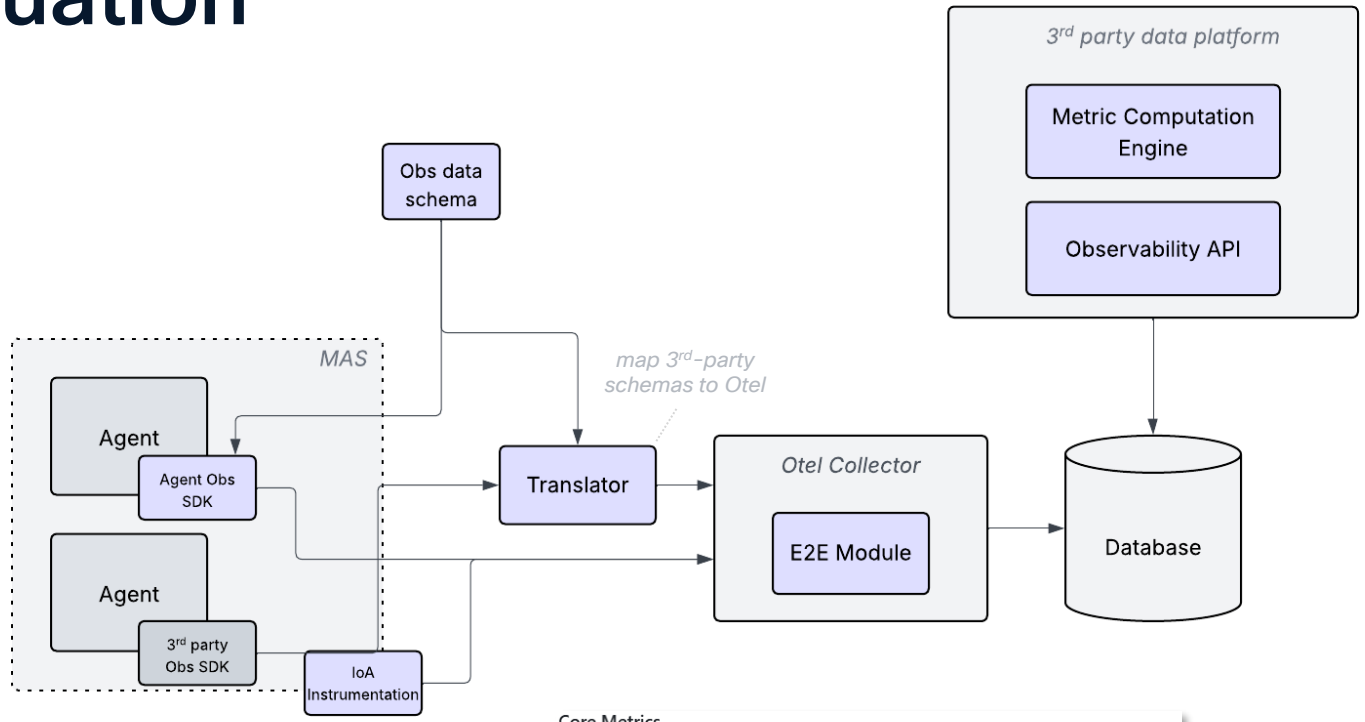
This document specifies the Secure Low-Latency Interactive Real-Time Messaging (SLIM), a protocol designed to support real-time interactive AI applications at scale. SLIM leverages gRPC and adds publish-subscribe capabilities to enable efficient many-to-many communication patterns between AI agentic applications (AI models, tools and data). The protocol provides mechanisms for connection management, stream multiplexing, and flow control while maintaining compatibility with existing gRPC deployments.

https://datatracker.ietf.org/doc/draft-mpsb-agntcy-slim/00/

Agent Observability & Evaluation

A framework that monitors and assesses multi-agent software performance

- **Captures traces, metrics and events** from multi-agent software and exports data to an Otel collector endpoint
- Supports agents from various frameworks by **mapping and translating third-party schemas**
- Provides end-to-end visibility into multi-agent software by **aggregating telemetry** data across agents and components
- Supports **derived evaluation and observability metrics** – Metrics Computation Engine



Core Metrics

Span-Level Metrics

Metric Name	Description
Tool Utilization Accuracy	Measures tool selection and usage efficiency

Session-Level Metrics

Metric Name	Description
Agent to Agent Interactions	Counts interactions between pairs of agents
Agent to Tool Interactions	Counts interactions between agents and tools
Tool Error Rate	Rate of tool errors throughout a session
Cycles Count	How many times an entity returns to previous entity

Population-Level Metrics

Metric Name	Description
Graph Determinism Score	Measures variance in execution patterns across multiple sessions

https://github.com/agntcy/telemetry-hub/

https://github.com/agntcy/observe

https://docs.agntcy.org/obs-and-eval/observe-and-eval/

Agent Observability & Evaluation

https://github.com/agntcy/observe

https://docs.agntcy.org/obs-and-eval/observe-and-eval/

What's next?

Agents

come from different **vendors**

use different development **frameworks**

running on different **clouds**

have different **access** controls

belong to different **organizations**

have no common **understanding**

no shared **intent** and goals

no shared memory and **knowledge**

cannot **coordinate** or negotiate

cannot evolve their **personas**

Toward an “Internet of Cognition”

- Scaling data
- Scaling compute
- Scaling parameters

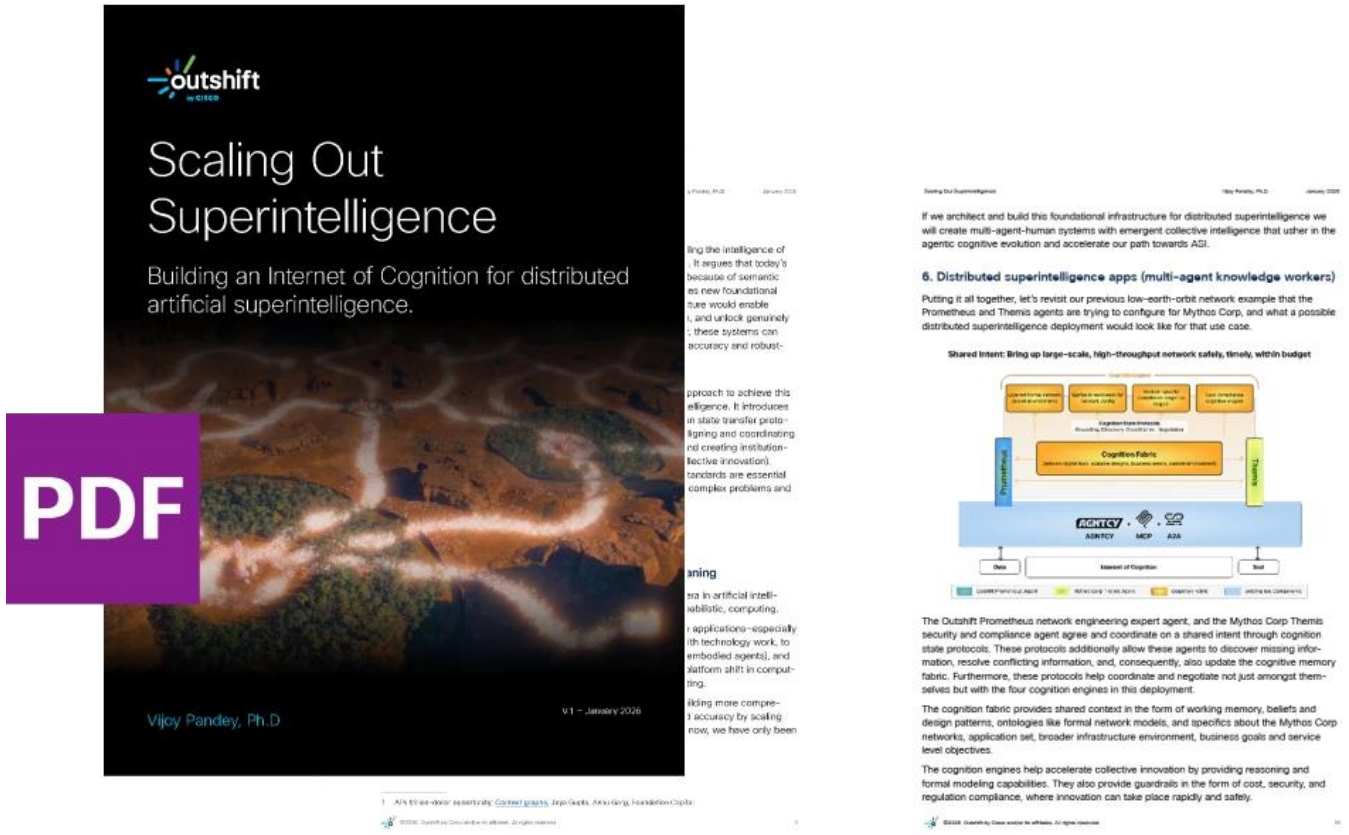
Paradigm shift →

Shared Intent
Shared Knowledge
Shared Innovation

discovery & identity
messaging & protocols
observability

Collective (civilizations)

“Internet of Cognition” Whitepaper



https://outshift.cisco.com/internet-of-cognition

To finish up: A few suggestions, if I may ;-)

- **Explore AIOps Use-Cases**
  Integrate AIOps tools into your existing workflows

- **Next-gen distributed Applications**
  Imagine new distributed agentic applications – ensembles of new agents, 3rd-party agents, and your API-based apps and services transformed into agents

- **Collaborate**
  Let's democratize Agentic AI together – support AGNTCY and related open initiatives
  open, peer-to-peer, interoperable

agntcy.org

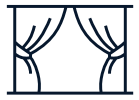
cnoe.io

Complete your session surveys

Complete your surveys in the Cisco Events App.

Complete a minimum of 4 session surveys and the overall event survey to receive a unique Cisco Live t-shirt.
(from 11:30 on Thursday, while supplies last)

Continue your education



Visit the Cisco Showcase for related demos

Book your one-on-one Meet the Engineer meeting

Visit the Technical Solutions Clinics to discuss your technical questions

Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

Visit the On-Demand Library for more sessions at CiscoLive.com/On-Demand

Thank you

cisco Live !


CISCO