Amazon Agentic AI for Email Triaging

Shanicus Yee

Solutions Architect

Agenda

- Overview of Agentic AI
- Deep Dive of Bedrock AgentCore
- Next Steps

The broadest choice for building and deploying agents

Agent Powered Applications

For Builders

For Business Users

Tools for Building AI Agents

Marketplace

Bedrock Agents

Strands

Nova Act SDK

Agentic AI Foundation

Amazon Bedrock AgentCore

Amazon Bedrock AgentCore

The prototype to production “chasm”

Excitement and potential

POC

Challenges on the path to production

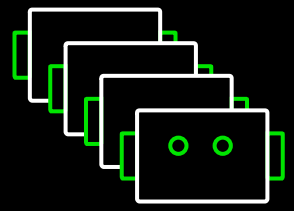
Performance

Scalability

Security

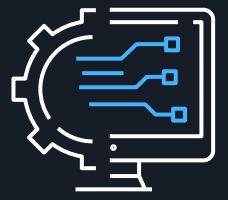
Governance

Meaningful business value

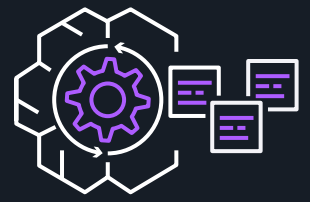


AI production agents

Foundational services for running highly capable agents, securely at scale



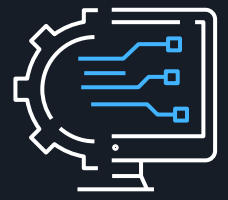
Deploy securely at scale



Enhance with tools and memory

Monitor

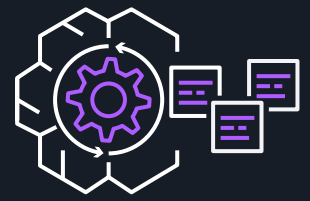
Secure, scalable runtime for agents



Deploy securely at scale

AgentCore Runtime

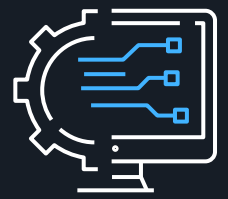
AgentCore Identity



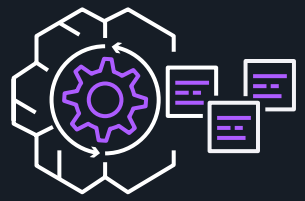
Enhance with tools and memory

Monitor

Essential tools and capabilities to build highly effective agents



Deploy securely at scale



Enhance with tools and memory

AgentCore Gateway

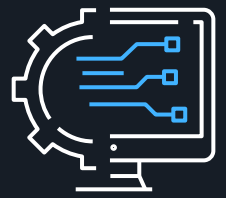
AgentCore Memory

AgentCore Browser

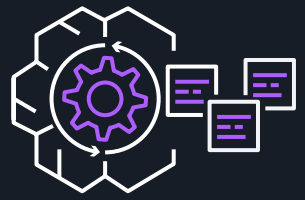
AgentCore Code Interpreter

Monitor

Visibility to operate agents you can trust



Deploy securely at scale



Enhance with tools and memory

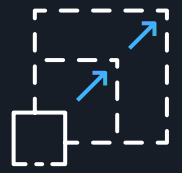
Monitor

AgentCore Observability

Amazon Bedrock AgentCore Services

AgentCore Runtime

Scale from real-time to multi-hour workloads



- Scale from real-time to multi-hour workloads with low latency and industry-leading extended runtime (up to 8 hours)
- Supports payloads across modalities

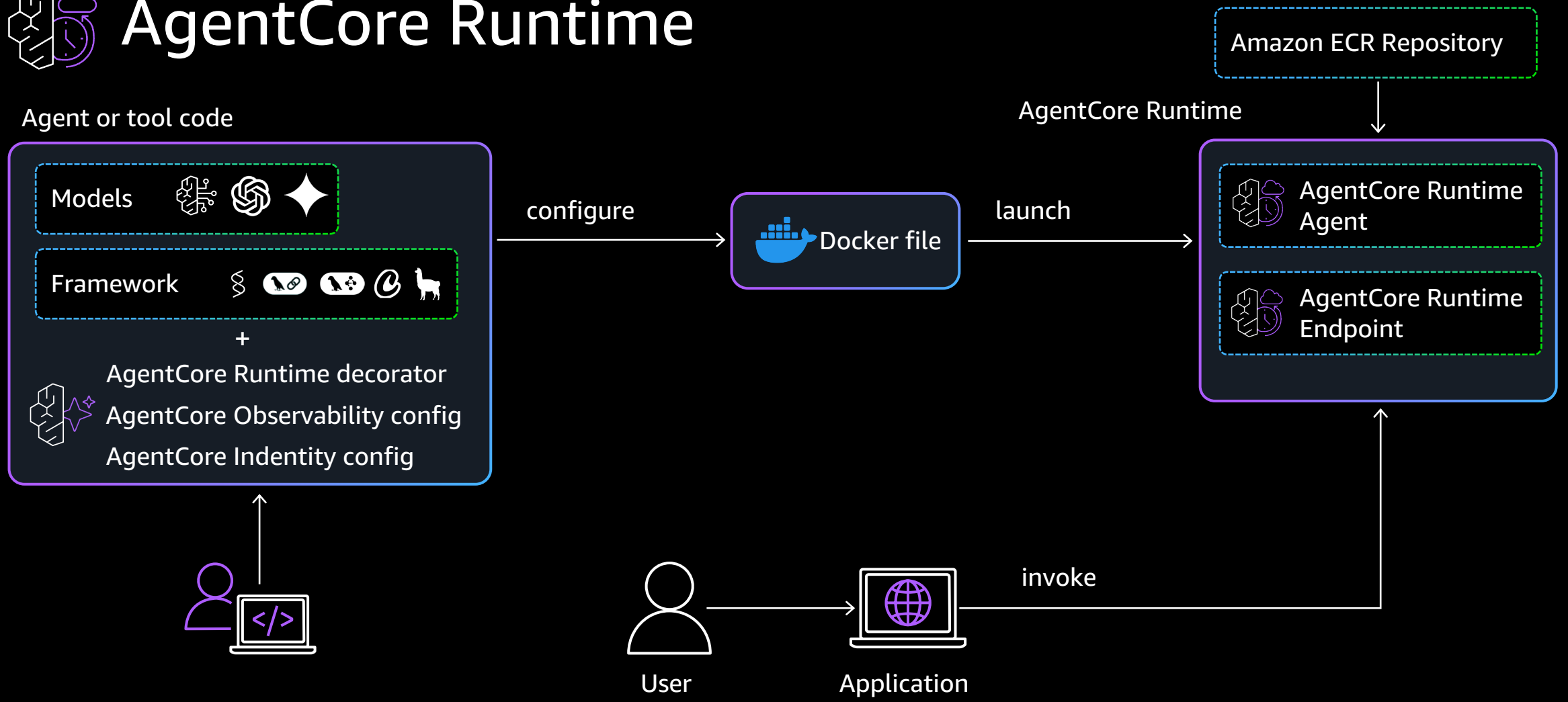
Accelerate time to market

- Deploy any AI agent using common open-source frameworks
- Deploy from POC to production in a few lines of code

Secure workload with enterprise-grade isolation

- True session isolation to protect your data
- Integrates with existing identity providers

AgentCore Runtime



AgentCore Identity

Secure, delegated access for AI agents

- Enables AI agents to securely access AWS resources and third-party tools such as GitHub, Google, Salesforce and Slack
- Robust access controls with just-enough access and secure permissions delegation

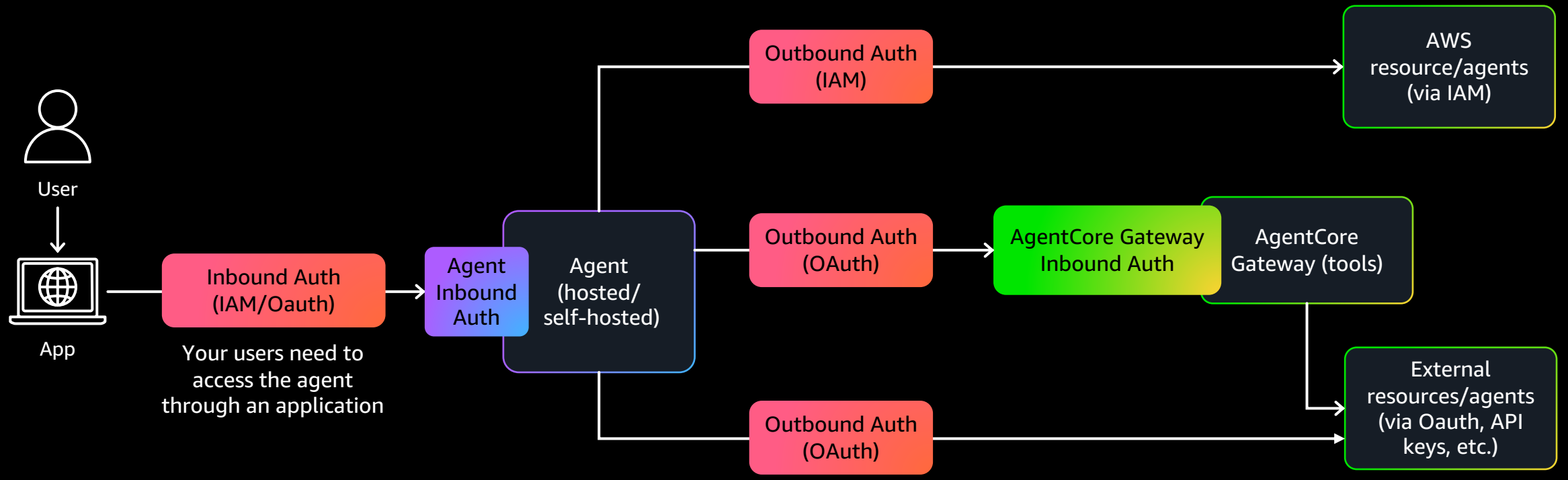
Build streamlined AI agent experiences

- Minimizes consent fatigue with a secure token vault
- Streamlines authentication flows

Accelerated AI agent development

- Preserves existing identity systems such as Okta, Azure AD, or Amazon Cognito
- Lowers custom development efforts without need for migrating users or rebuilding authentication flows

AgentCore Identity



AgentCore Gateway

Simplified tool development & integration

- Turn APIs, Lambda functions, and existing services into MCP-compatible tools
- Access thousands of tools through a standardized interface

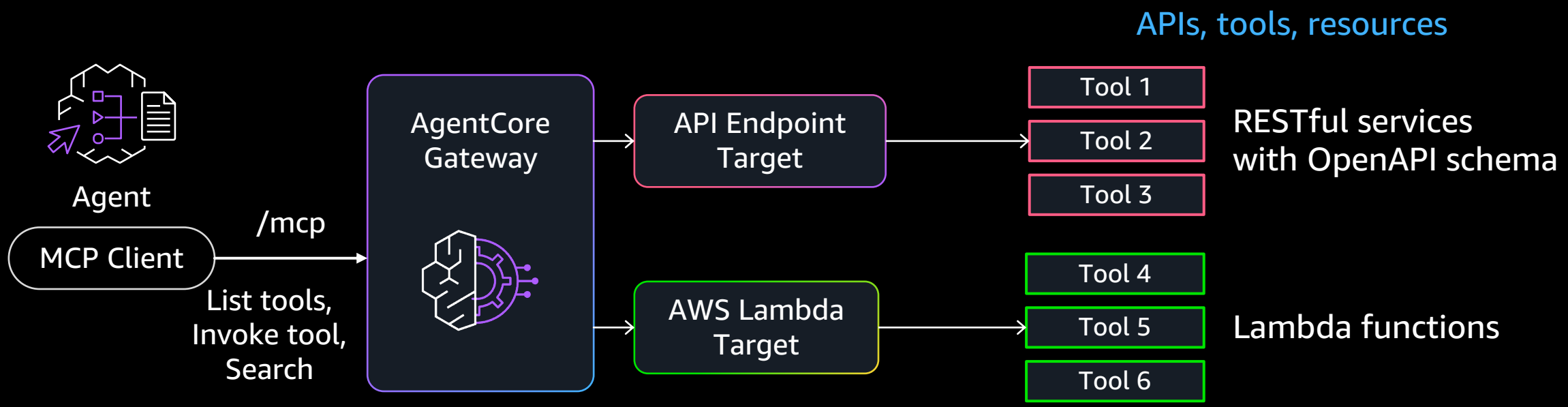
Secure and unified access

- Discover and use tools through a single, secure endpoint
- Combine multiple tools sources into one unified interface

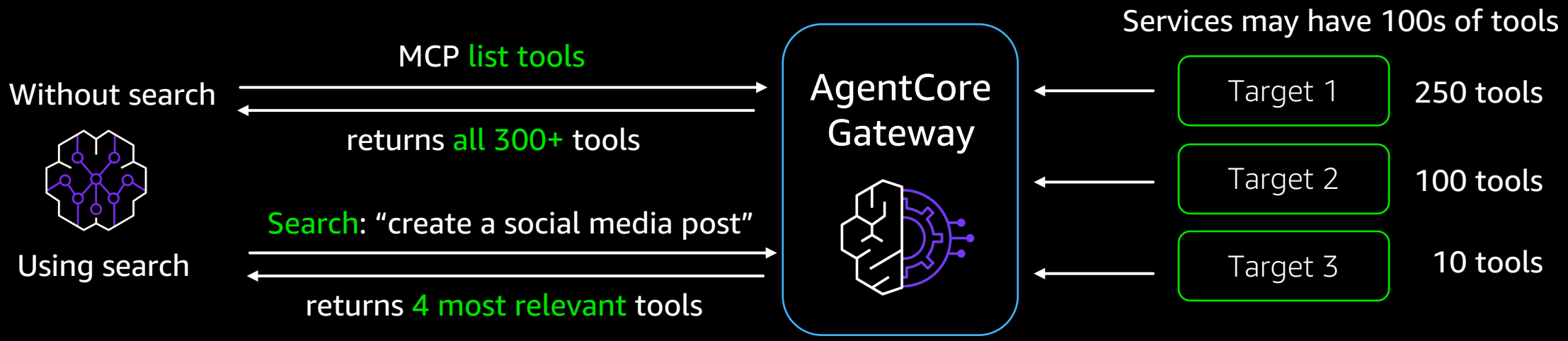
Intelligent tool discovery

- Enable agents to find the right tools with context aware discovery
- Curated tool collections with granular permissions

AgentCore Gateway



AgentCore Gateway semantic search



Benefits

- AgentCore Gateway automatically indexes tools, and gives serverless semantic search
- Reduces context passed to the agent’s LLM, improving accuracy, speed, and cost
- Lets agent focus on tools relevant for a given task

AgentCore Memory

Simplify memory management

- Abstracts memory infrastructure
- Scales automatically with serverless architecture
- Automatically stores and manages agent context across sessions

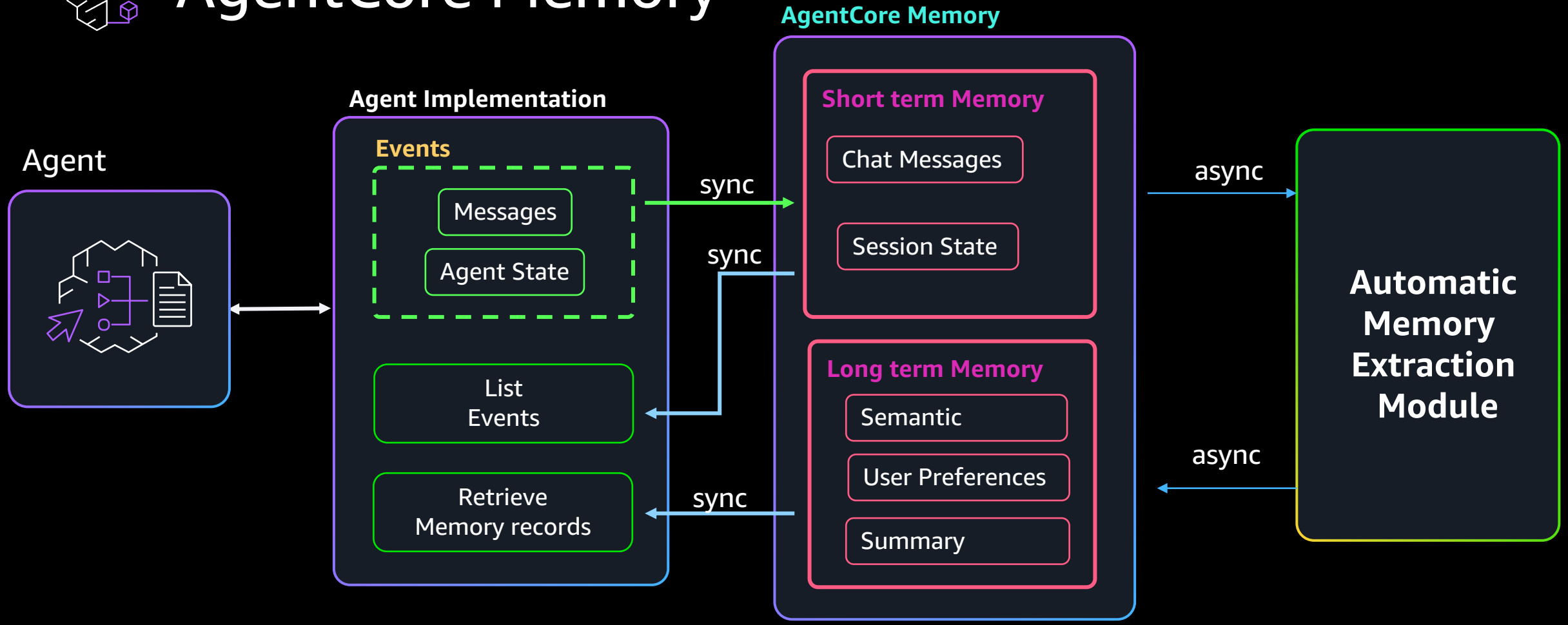
Enterprise-grade

- Complete data privacy with dedicated storage for each customer
- Enterprise security with encryption and regional data storage options

Deep customization

- Define memory patterns based on your use case
- Configure extraction rules
- Choose models and customize prompts for memory extraction

AgentCore Memory



AgentCore Browser

Serverless browser infrastructure

- Low latency browser sessions
- Auto-scales from 0 to hundreds of concurrent sessions

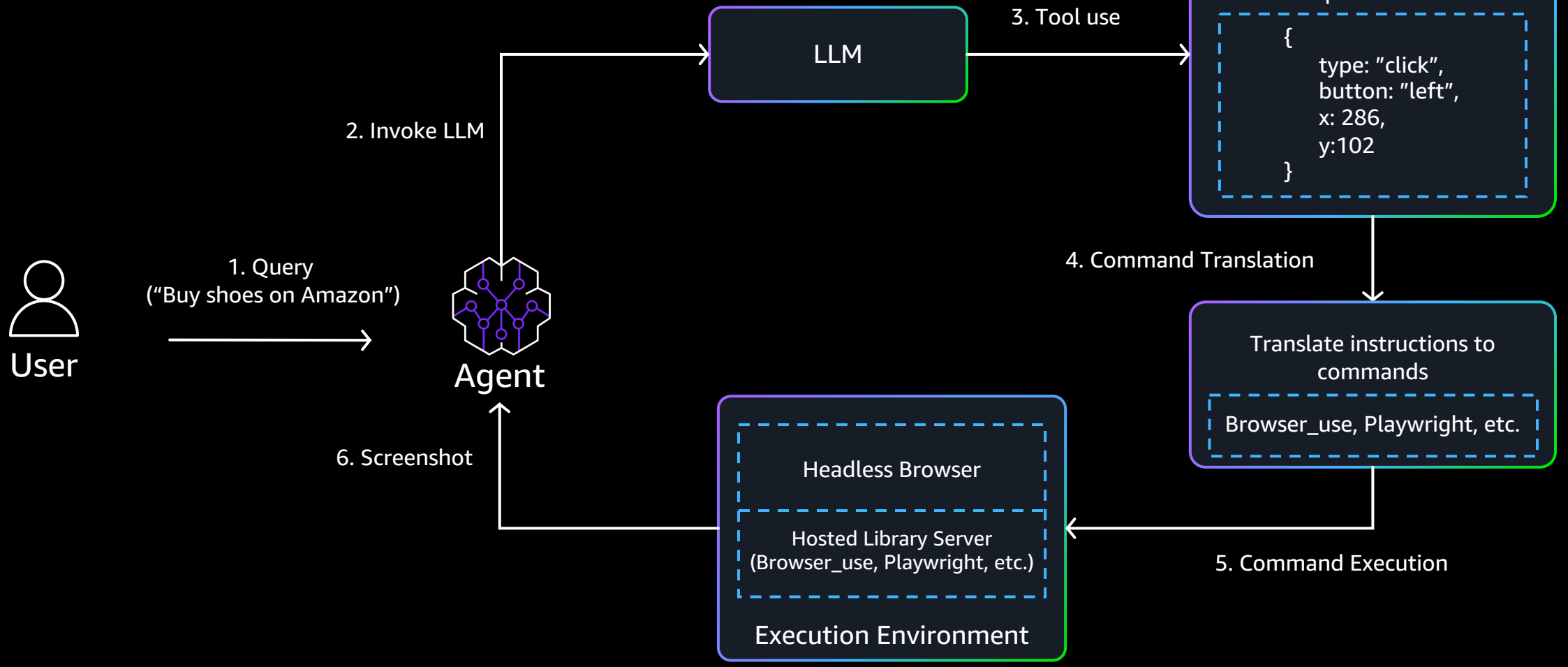
Enterprise-grade security

- Session isolated compute with VM-level isolation per user
- Secure credential handling

Enterprise observability

- Live streaming URLs for real-time monitoring
- Session replays for debugging
- Extensive logging of all browser commands to CloudTrail

AgentCore Browser



AgentCore Code Interpreter

Execute code securely

- Execute complex workflows and data analysis in isolated sandbox environments
- Access internal data sources securely without exposing sensitive data

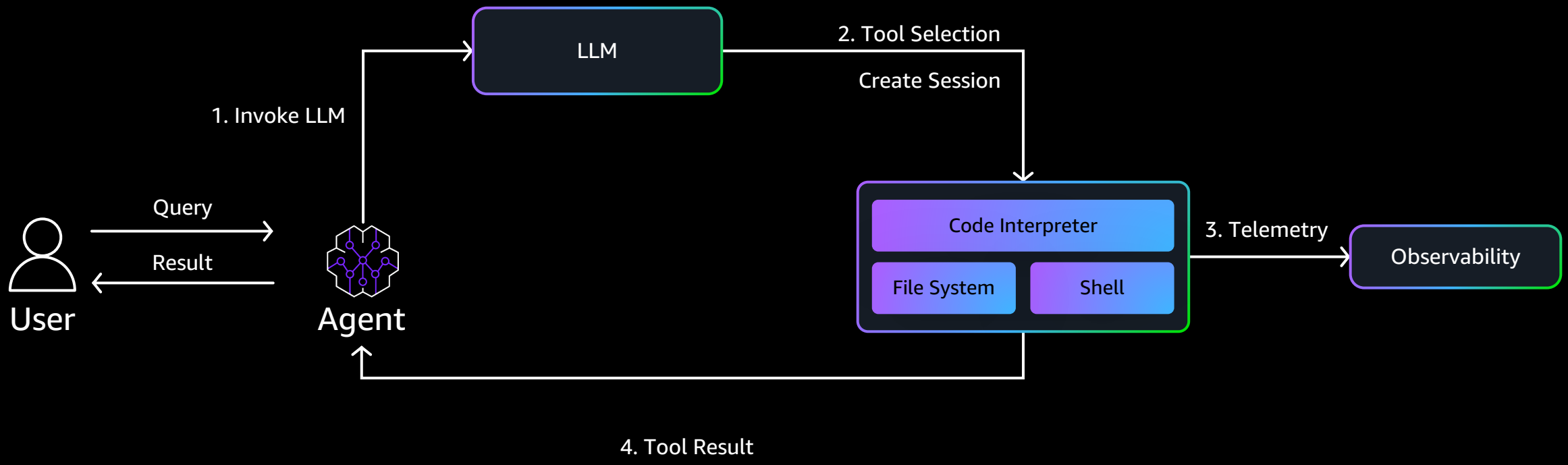
Large-scale data processing

- Process gigabyte-scale datasets efficiently using Amazon S3 integration, without API limitations

Ease of use

- Quick start with pre-built execution runtimes for JavaScript, TypeScript, and Python with common libraries pre-installed

AgentCore Code Interpreter



AgentCore Observability

Maintain quality and trust

- Comprehensive end-to-end visibility into agent behavior
- Accelerated debugging and quality audits
- Quickly detect issues and assess performance trends

Integrate with 3P observability tools

- Integration with a wide range of monitoring and observability tools, including CloudWatch
- Flexibility to leverage your existing observability stack

AgentCore Observability

Agent deployed on AgentCore Runtime

Agent memory using AgentCore Memory

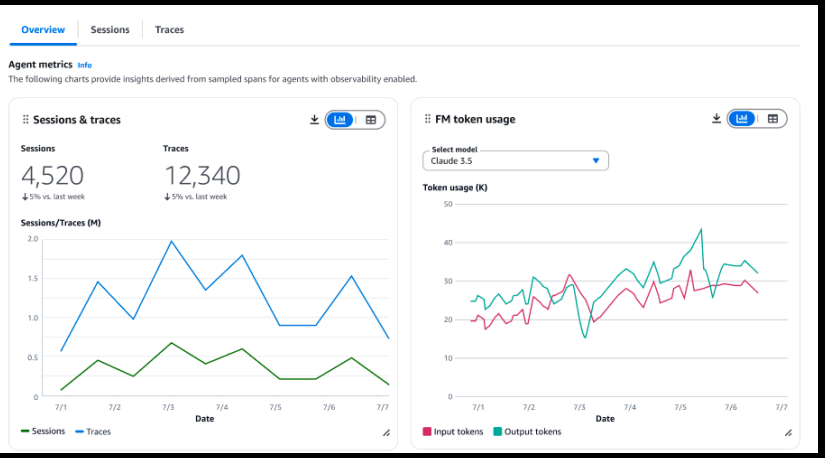
Tools from AgentCore Gateway

...

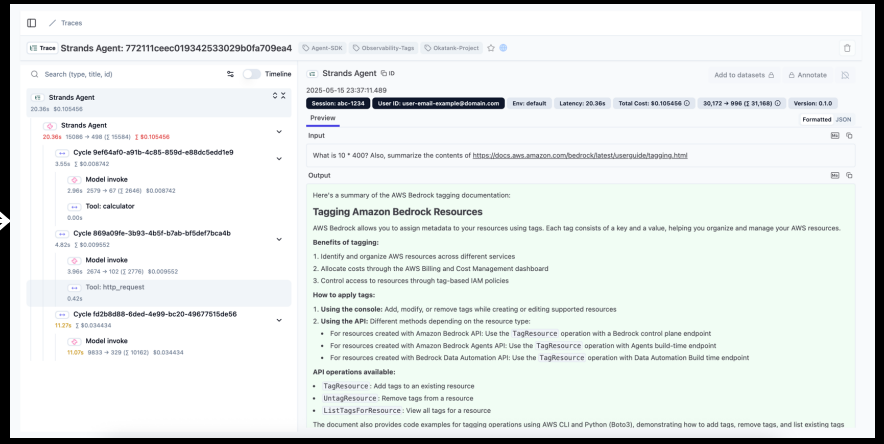
Amazon Bedrock AgentCore Observability

OTEL logs

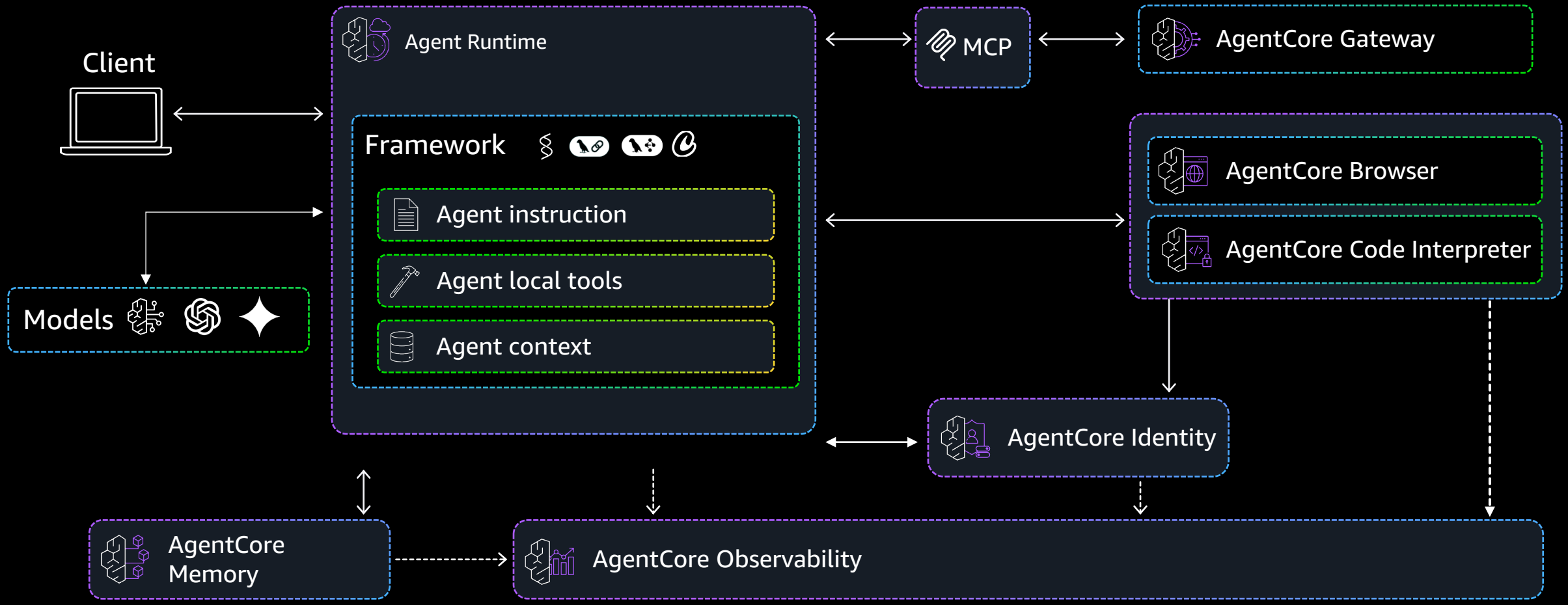
AgentCore Observability dashboards



Third-party observability dashboards



AgentCore capabilities enabling agents at scale



AI agents and tools in AWS Marketplace

PRE-BUILT AGENTS

ANTHROP\C

AUTOMATION ANYWHERE

karini

AGENT TOOLS

brave

CONFLUENT

Saviynt

tavily

AGENT DEV SOLUTIONS

accenture

Agentforce

lyzr

PROFESSIONAL SERVICES

Deloitte.

genpact

IBM

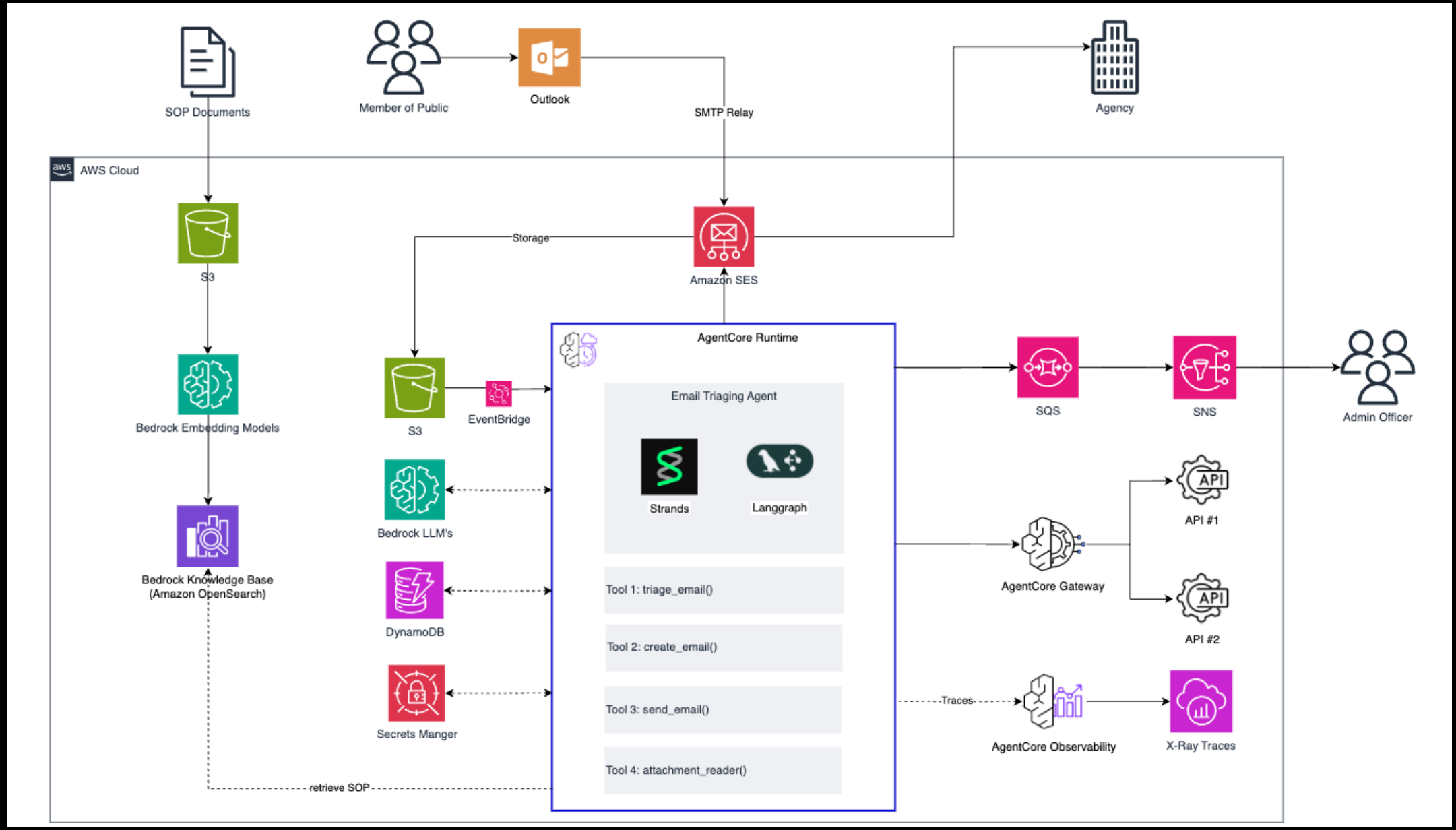
SOFTWARE SOLUTIONS

mlx

servicenow

snowflake

This is not a complete list. To view all products listed in AWS Marketplace, visit AWS Marketplace.
This list is current as of July 15, 2025.



Architecture flow

1. Member of public write email into general mailbox in Outlook.
2. Outlook forwards the email to Amazon SES through SMTP relay which stores the email in S3.
3. For batch processing, can trigger EventBridge on schedule to AgentCore Runtime.
4. For real-time process, S3 can send events to EventBridge.
5. You can use any open source AI famework to create agent and tools such as Strands, Langgraph.
6. Tool 1 triage_email(): this tool reads the email from S3, utitlizes Bedrock LLM's and knowledge base to triage, categorizes email. Metadata and raw email texts can be saved in DynamoDB. DynamoDB can serve as the storage for session data and auditing.
7. Tool 2 create_email(): if there is a need to summarize, add annotations, the agent can utilize this tool. Output from this tool can also be saved in DynamoDB.
8. Tool 3 send_email(): based on the output of Tool1 and/or Tool2, the agent can utilize this tool to send email to relevant agency. Secrets for Amazon SES is stored in AWS Secrets Manager.
9. Tool 4 attachment_reader(): if there is an attachment, this tools open the attachment.
10. If an email requires escalation/human in the loop, the agent writes the email draft into SQS and SNS notifies the Administrative Officer to review.
11. AgentCore Observability helps you trace, debug, and monitor agent performance in production environments. It offers detailed visualizations of each step in the agent workflow, enabling you to inspect an agent's execution path, audit intermediate outputs, and debug performance bottlenecks and failures.
12. AgentCore Gateway allows you to register WoG products using OpenAI specifications and authenticate using OAuth 2LO or API keys.
13. SOP Documents are stored in S3 and embedded into Bedrock Knowledge Bases (OpenSearch) to support the agent for email triaging procedures.

Thank you!