elastic Agent Builder: Providing Relevant Context for Data Driven Agents

elastic | The Search AI Company

Kathleen DeRusso
Principal Software Engineer,
elastic

Agenda

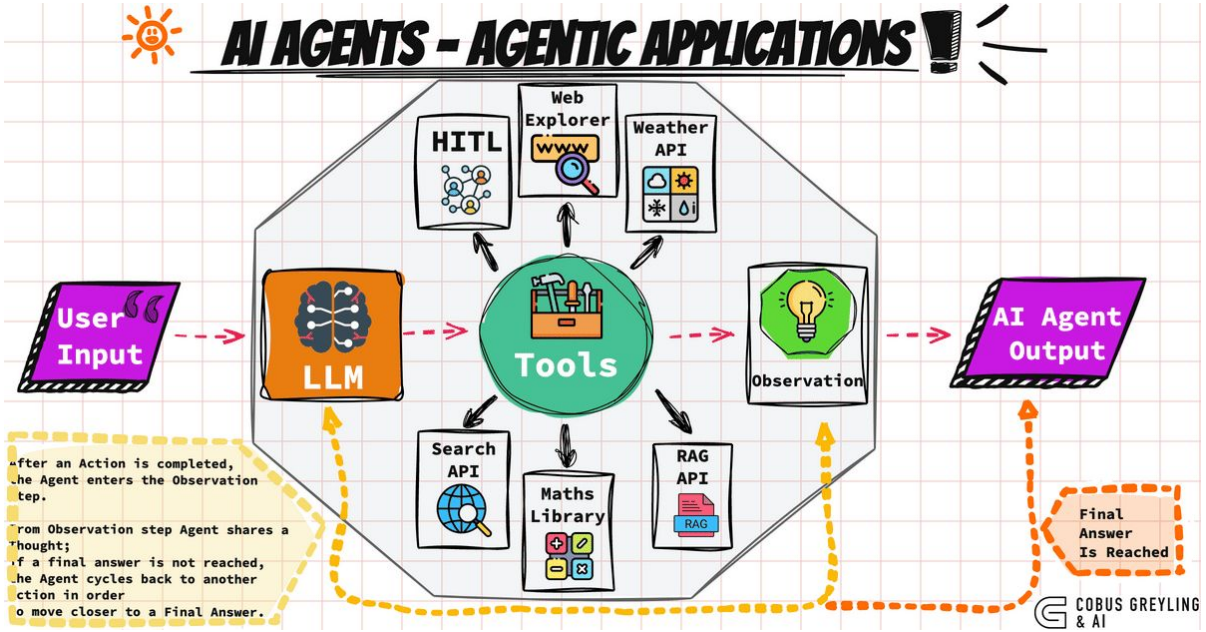
01	**What are Agents?**
02	**Why do we need Context Engineering?**
03	**Elastic Agent Builder**
04	**Demo**
05	**Future Direction**
06	**Q&A**

What are Agents?

Not a secret agent!

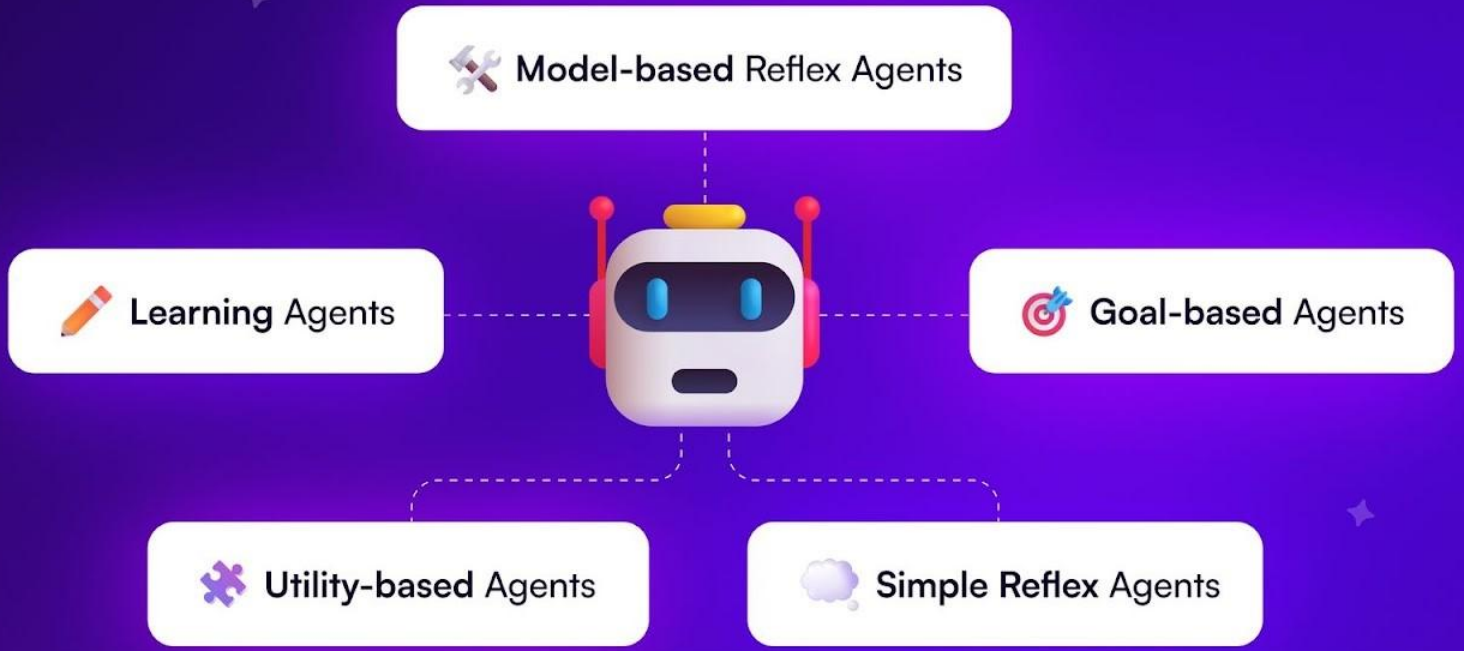
What is an Agent?

An AI agent is a software program that can autonomously perform complex tasks by making decisions, learning from its environment, and using tools to achieve goals



Credit: What's Your Definition Of An AI Agent? | Cobus Greyling on Medium

Types of AI Agents



Credit: What are AI Agents? A Complete Guide For Beginners | Vidhi Gupta

Context Engineering ensures agents have the correct information

Source: Context window overflow: Breaking the barrier | AWS Security Blog

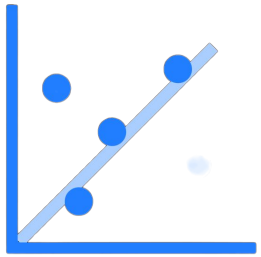
Context Window

The maximum number of tokens an LLM can process at once.

elastic

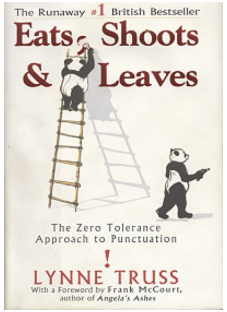
Why do LLMs Hallucinate?

Frozen or Limited Knowledge



Model Overfitting

Dataset Biases



Language Ambiguity

Source: OpenAI | Why Language Models Hallucinate

Why Language Models Hallucinate

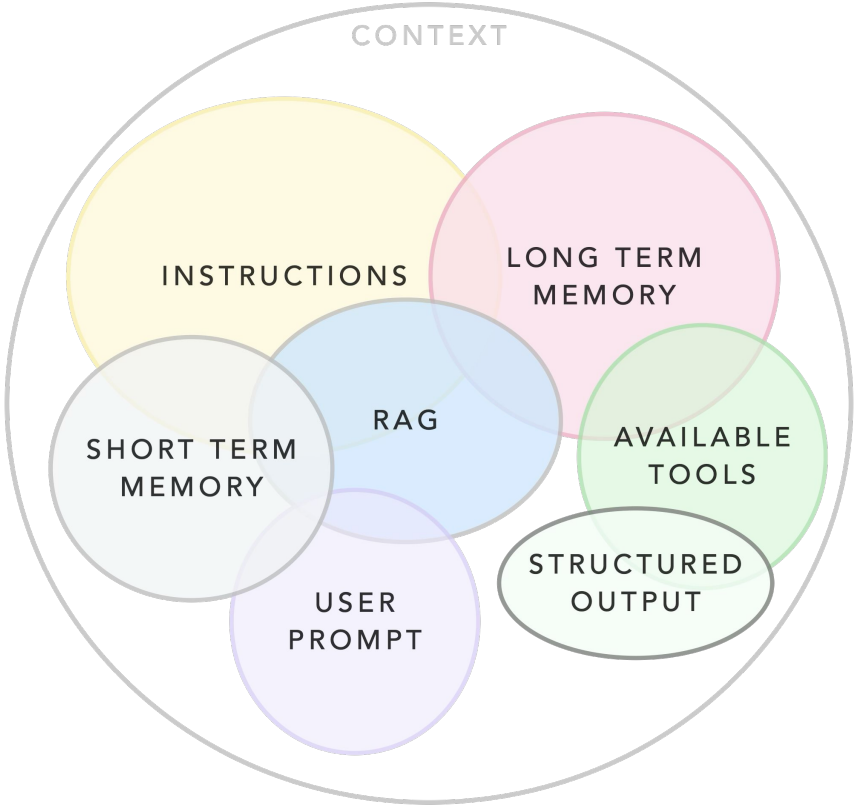
Adam Tauman Kalai* OpenAI
Ofir Nachum OpenAI
Santosh S. Vempala† Georgia Tech
Edwin Zhang OpenAI

September 4, 2025

Abstract

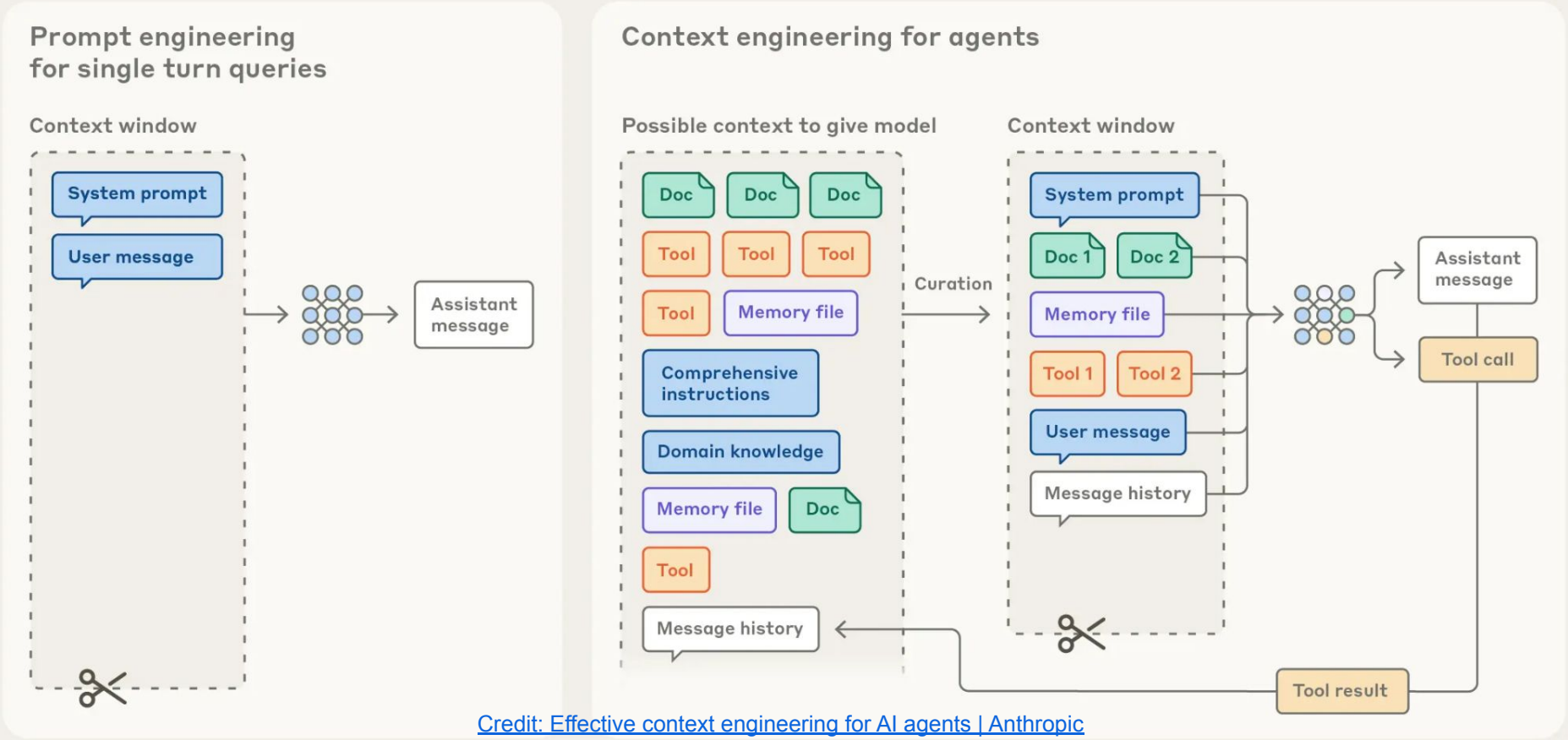
Like students facing hard exam questions, large language models sometimes guess when uncertain, producing plausible yet incorrect statements instead of admitting uncertainty. Such "hallucinations" persist even in state-of-the-art systems and undermine trust. We argue that language models hallucinate because the training and evaluation procedures reward guessing over acknowledging uncertainty and we analyze the statistical causes of hallucinations in the modern training pipeline. Hallucinations need not be mysterious—they originate simply as errors in binary classification. If incorrect statements cannot be distinguished from facts, then hallucinations in pretrained language models will arise through natural statistical pressures. We then argue that hallucinations persist due to the way most evaluations are graded—language models are optimized to be good test-takers, and guessing when uncertain improves test performance. This "epidemic" of penalizing uncertain responses can only be addressed through a socio-technical mitigation: modifying the scoring of existing benchmarks that are misaligned but dominate leaderboards, rather than introducing additional hallucination evaluations. This change may steer the field toward more trustworthy AI systems.

What is Context Engineering?



Credit: What is Context Engineering | Elasticsearch Labs

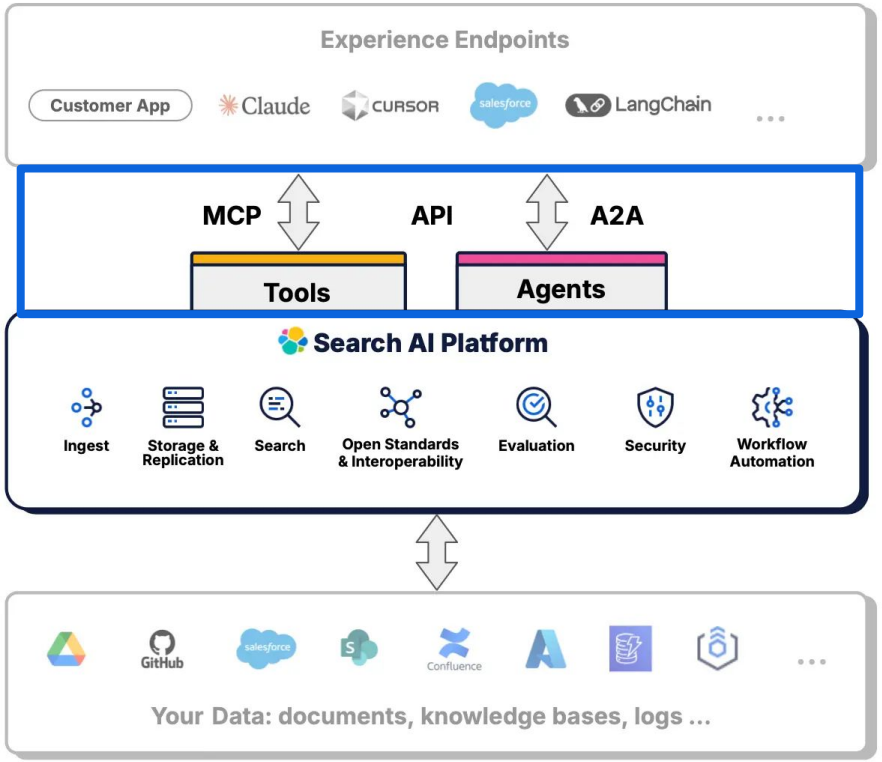
Prompt engineering vs. context engineering



Credit: Effective context engineering for AI agents | Anthropic

Agent Builder allows building of data-centric contextual agents

Where does Agent Builder fit?



Credit: Building AI Agentic workflows with Elasticsearch | Elasticsearch Labs

Capabilities

Elastic Agent Builder has the following features

Chat with your data using the native conversational agent and see full tracing of the steps and results

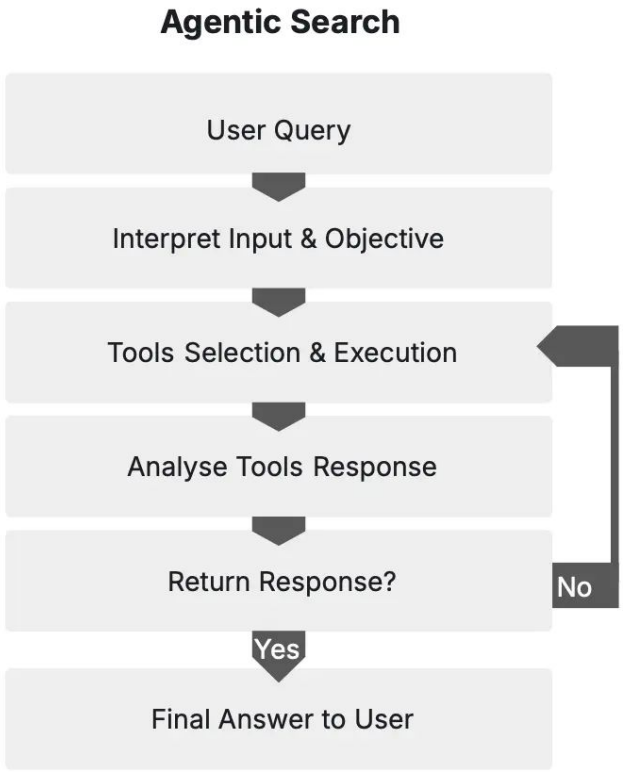
Leverage built-in tools to find relevant indices, understand data structure and translate native language into structured semantic, lexical or hybrid queries

Build custom tools using ES|QL for use in your own agents

Built in visualizations for chat with editable chart types and the ability to add to Dashboards

Integrate with external agents and apps using MCP and A2A

How do Agents work in Elasticsearch?

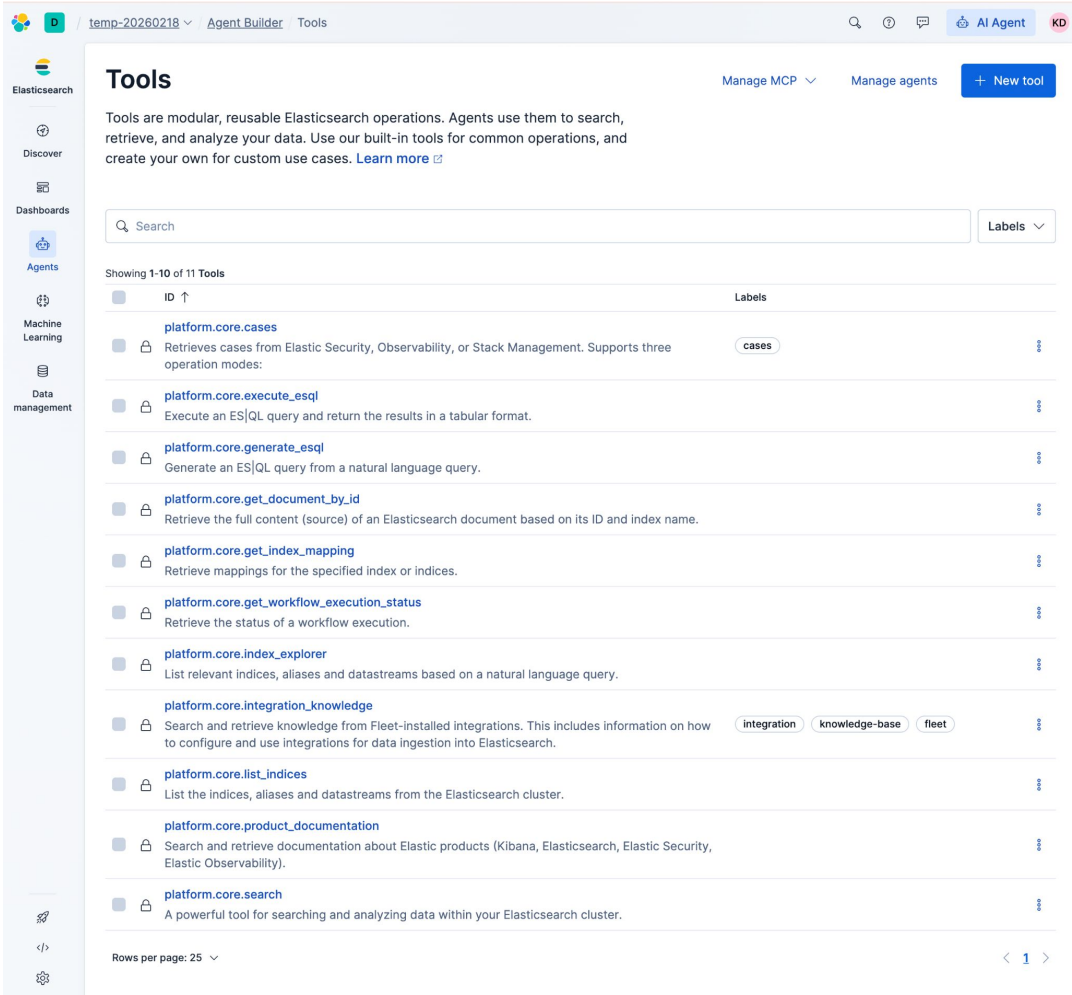


Credit: Building AI Agentic workflows with Elasticsearch | Elasticsearch Labs

In-built tools

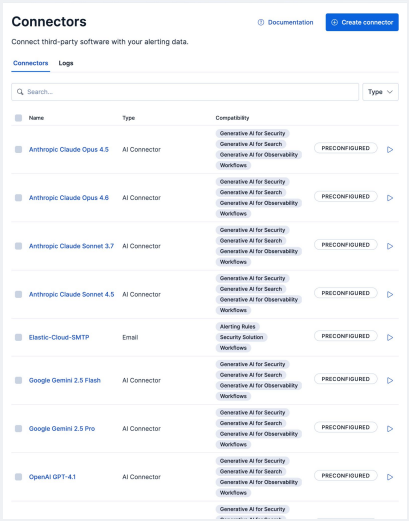
There are currently 11 tools available, prefixed by `platform.core`:

- Execute ES|QL queries
- Generate ES|QL query from natural language
- Get full document content based in ID
- Retrieve mappings for one or more specified indices
- List relevant indices, aliases and streams based on natural language query
- List indices, alises and datastreams in the cluster
- Core search tool
- ... and more!

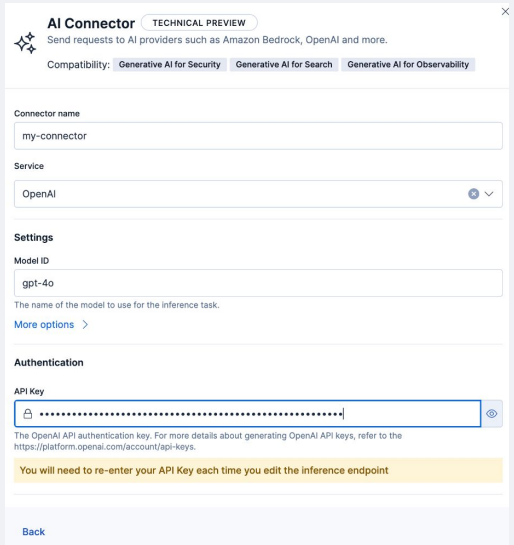


Which LLM?

Default
LLMs provided via Elastic Inference Service (EIS)



Connector
Create a custom connector using the AI Connector (tech preview)



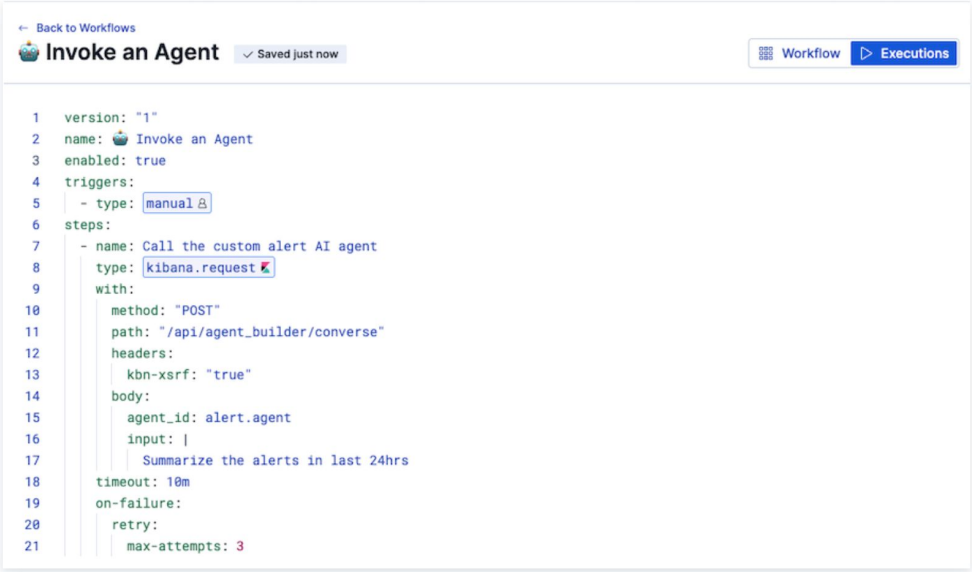
elastic

Demo

Workflows

- Technical preview
- Declarative, composable workflows defined in yaml
- Composed of triggers, inputs and steps
- `ai.agent` steps invoke agents from within a workflow
- Workflows can be exposed to Agent Builder as tools

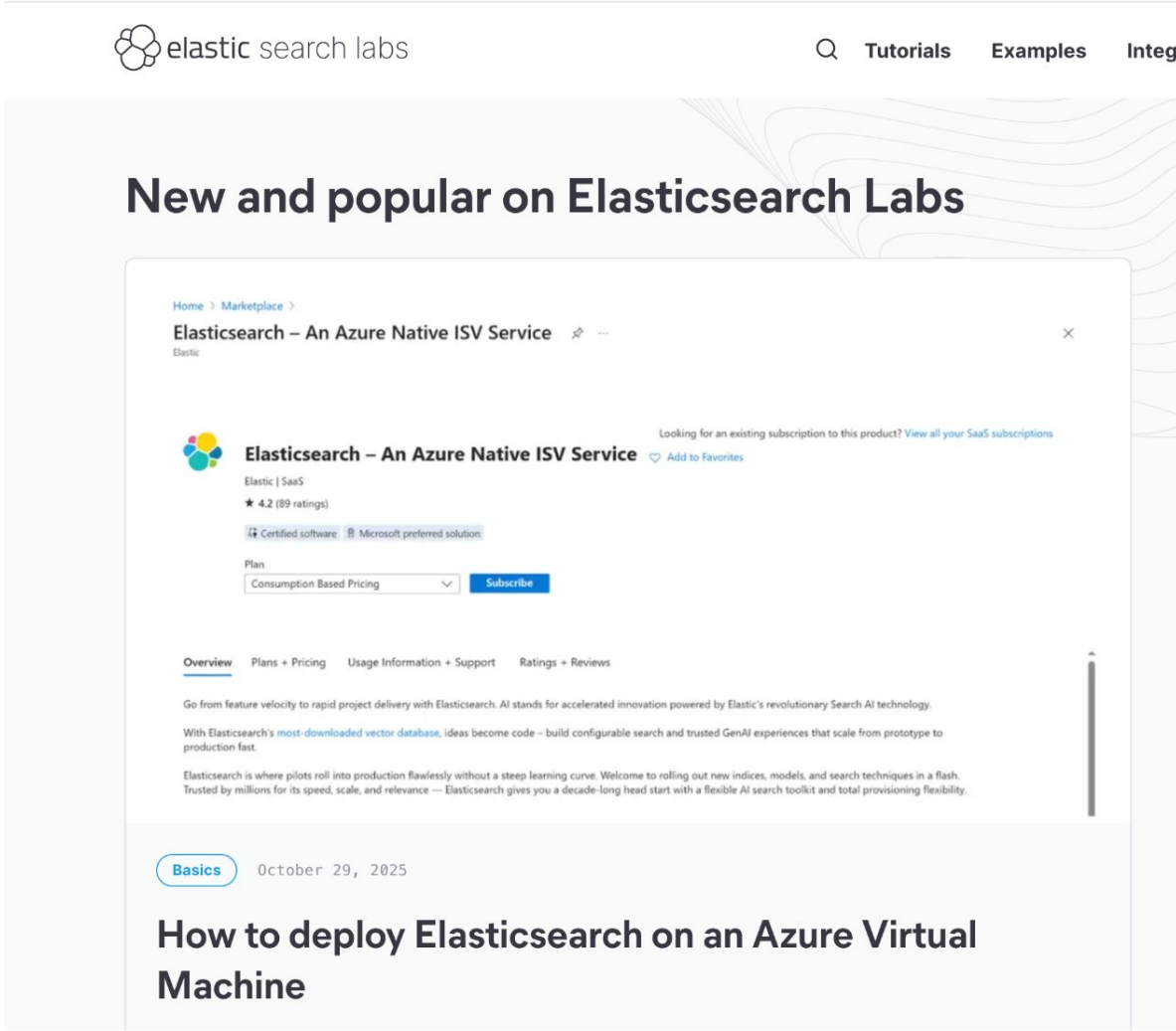
- Technical blog: https://www.elastic.co/search-labs/blog/elastic-workflows-automation



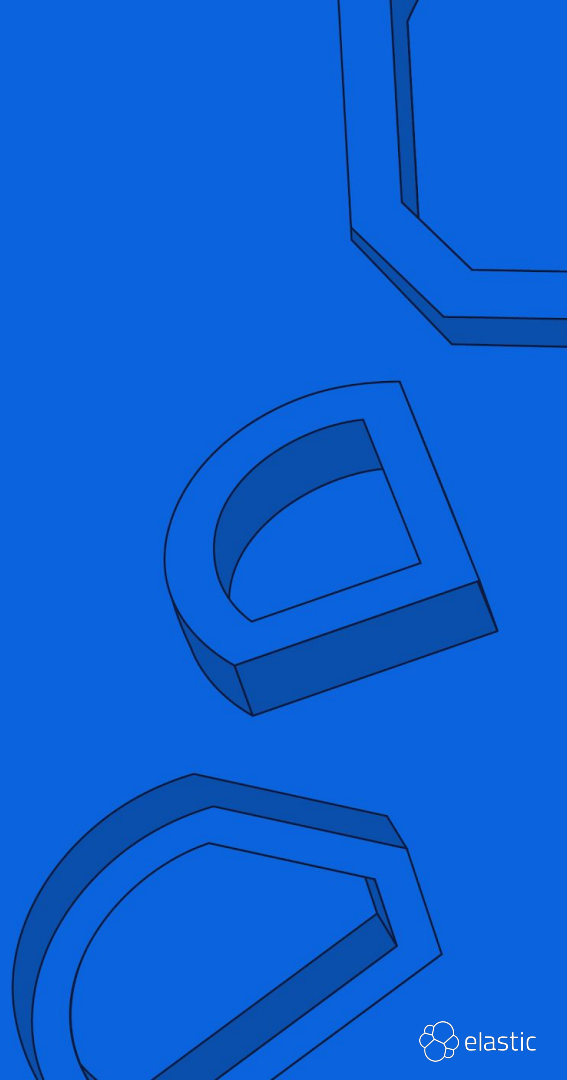
Resources

- Introducing Elastic's Agent Builder
- Building AI Agentic workflows with Elasticsearch
- Your first Elastic Agent: From a single query to an AI-powered chat
- Connecting Elastic Agents to Gemini Enterprise via A2A protocol
- ... and many more!

- Search labs tag Agentic AI



Questions?



Thank you!