truefoundry

Office Hours #4

Building Production Ready Agents with Agent Platform

Presented by -

Nikhil Popli
Senior Backend Engineer at
TrueFoundry

truefoundry

WHAT WE'LL COVER TODAY

Agenda

1. **Intro & quick polls**
   A quick read of the room — where everyone is on their agent journey
2. **What is an agent?**
   Real examples, the three kinds of agents, and how they came to be
3. **What is an Agent Harness?**
   Context engineering + the runtime that turns an LLM into a reliable agent
4. **Building an agent on TrueFoundry**
   Four choices, what we manage for you — then a live demo
5. **Sharing agents across your company**
   High-code frameworks vs. the TrueFoundry approach

truefoundry

INTERACTIVE — LET'S SEE THE ROOM

Quick polls before we dive in

[Live Zoom Poll]

Please respond to these quick polls

2 · WHAT IS AN AGENT?

What is an agent?

a·gent /ˈājənt/

noun

1. A person or thing that takes an active role or produces a specified effect.
2. A person or thing that acts on behalf of another.

2 · WHAT IS AN AGENT?

Three kinds of agents

The word isn't new — what's new is who's doing the work.

1

Human agents

A travel agent, an estate agent — a person you hand a task to, who acts for you.

2

Software agents

Scripts and bots that automate narrow tasks on fixed rules you set up.

3

AI Agents

Models that reason, plan and call tools — they figure out the steps themselves.

Our focus for today

truefoundry

2 · WHAT IS AN AGENT?

First — what can an agent actually do today?

Before the technical definition: here's what enterprises are using agents for today.

Coding agent

Writes, edits and runs code to ship changes — Cursor, Claude Code.

Customer Support Chatbot

Answers questions and deflects tickets across chat and email.

IVR / voice agent

Handles inbound calls — understands intent, looks things up, resolves or routes.

2 · WHAT IS AN AGENT?

How agents came to be where they are now

LLMs

The brain of the agent.

+

Tools, via MCP

Function calling & APIs, standardized by MCP — so agents act, not just talk.

+

Skills

Dynamic context discovery — reusable know-how (SKILL.md) pulled in on demand.

Meta Prompting for Skills

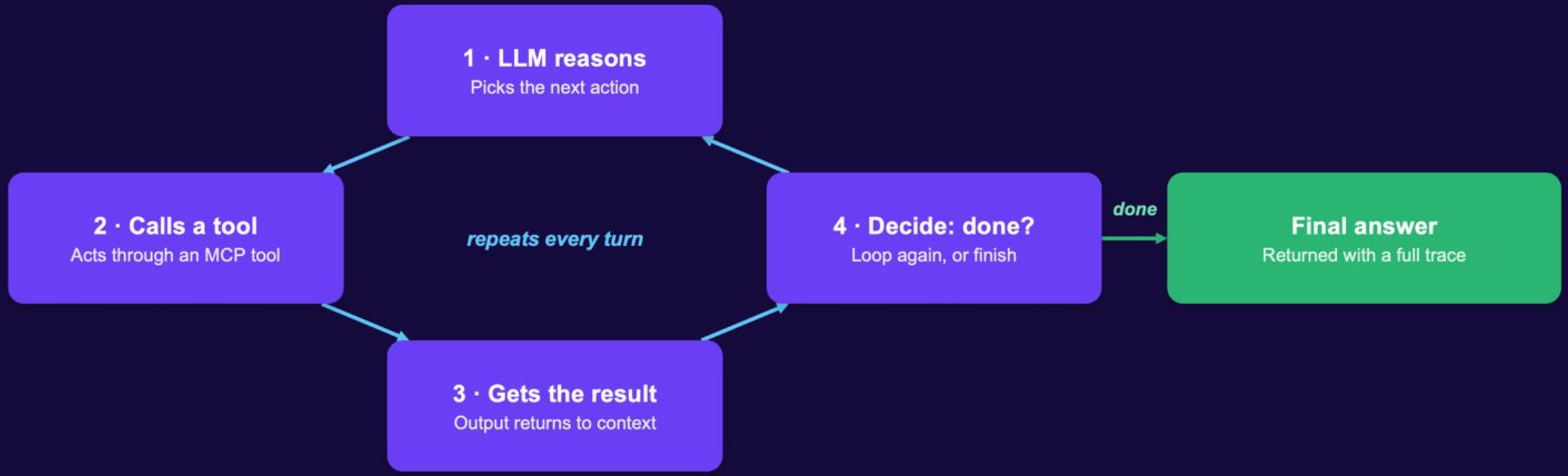
Writing a skill — and the skill that writes skills
Claude ships a **"skill-creator"** that interviews you and generates a clean one to publish to the registry.

truefoundry

2 · WHAT IS AN AGENT?

The agent loop

The model reasons, calls a tool through MCP, reads the result, then decides what to do next — repeating until the goal is met.



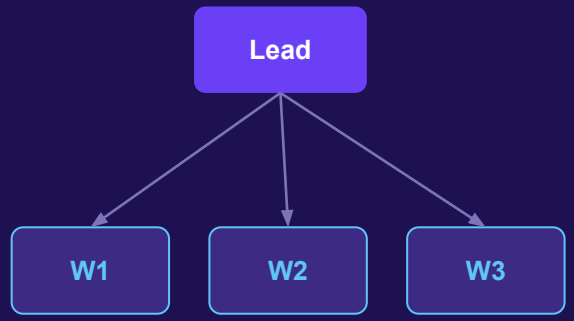
truefoundry

Context engineering is the new prompt engineering

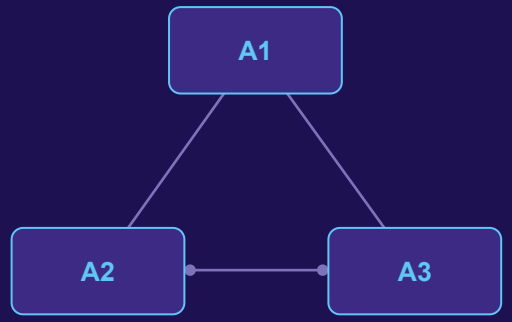
Getting the right context — and the right helpers — in front of the model is most of the battle. A few patterns the harness handles for you:

Sub-agents — multi-agent workflows, when they help

Supervisor → workers



Collaborating peers



Code mode

Agent writes & runs code to act — fewer brittle tool calls.

Automatic compaction

Long runs auto-summarized to stay within the context window.

Why a sandbox

Code, files and long tasks run in a safe, isolated space.

3 · WHAT IS AN AGENT HARNESS?

The runtime layer around the LLM

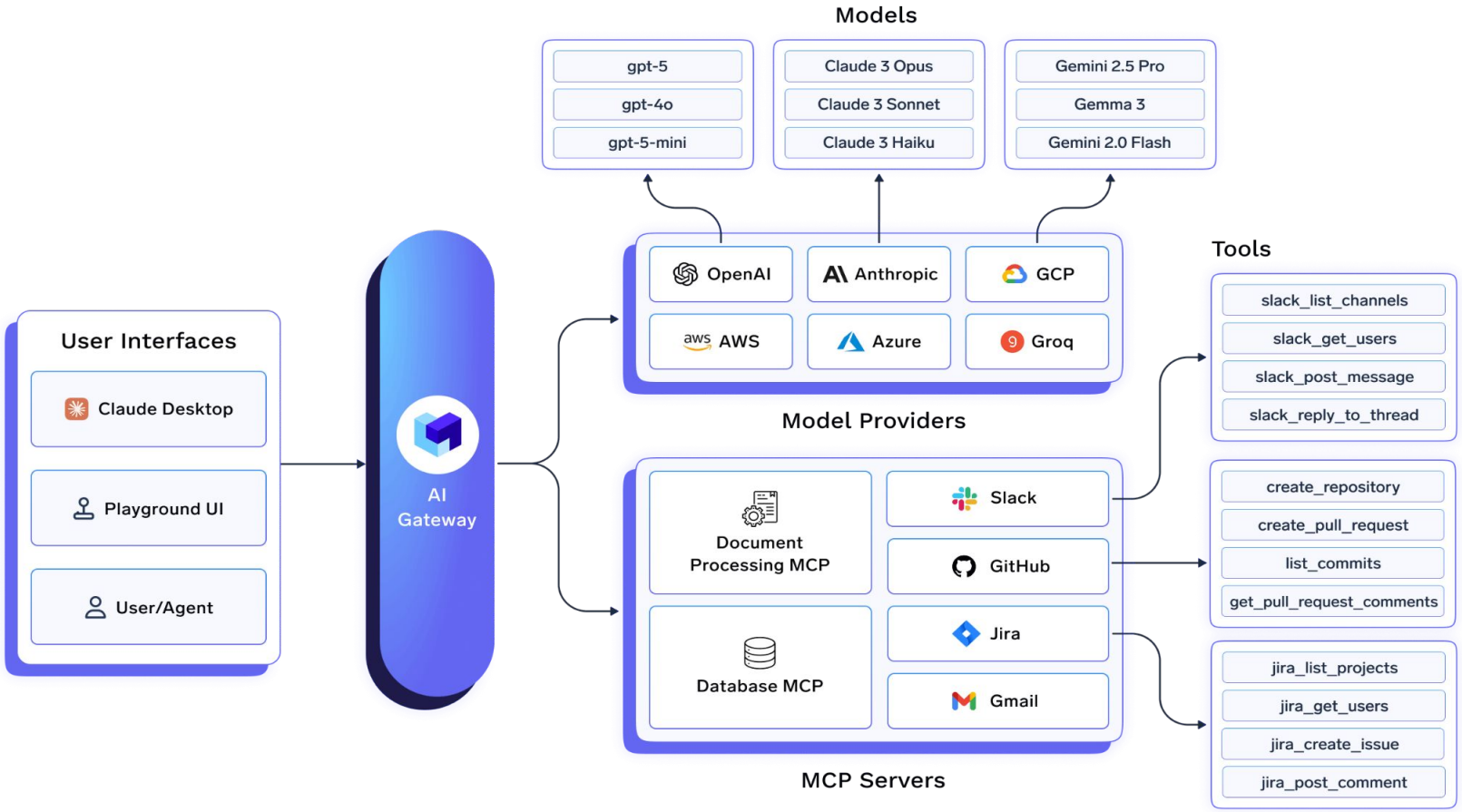
An agent harness is an orchestration runtime that manages the full execution loop and supports reliable, long-running agents.

With frameworks (LangGraph, CrewAi) you wire and run all of this yourself; the harness manages it for you.

truefoundry

TrueFoundry AI Gateway

The AI Gateway is a lightweight routing layer that connects your apps to LLMs, MCPs and agents — with governance, monitoring and control.



truefoundry

…and now, with agents in the picture

The Agent Harness runs in the same gateway plane as model and MCP traffic — so orchestration, governance, and observability all stay in one system.

Same control plane

Agents inherit Gateway RBAC, budgets & guardrails

No new silo

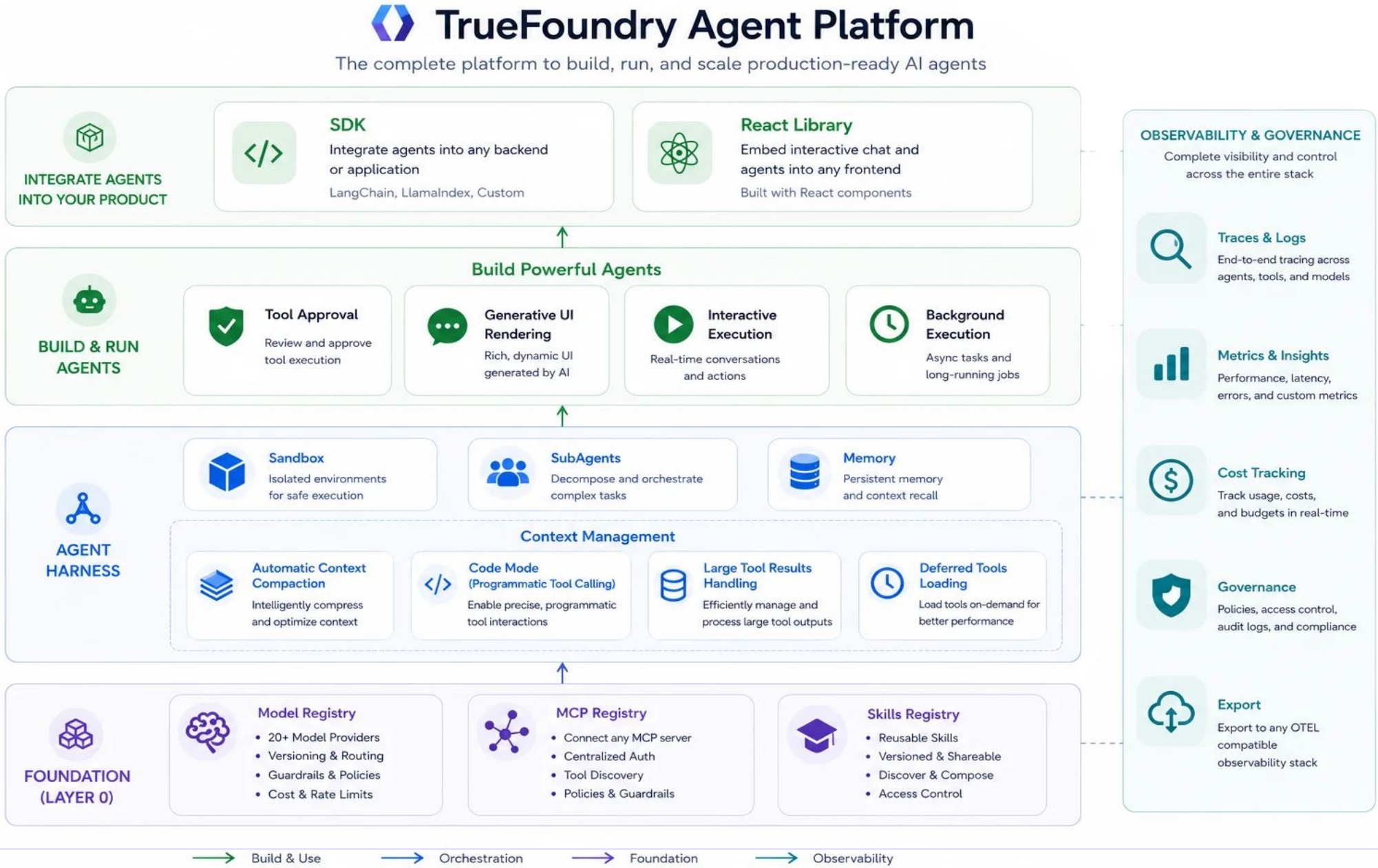
Model, MCP & agent traffic in one pane of glass

Observable by default

Every agent run traced end-to-end

Share agents by default

Collaborate across teams with shared access and consistency


TrueFoundry Agent Platform
The complete platform to build, run, and scale production-ready AI agents
INTEGRATE AGENTS INTO YOUR PRODUCT
SDK
Integrate agents into any backend or application
LangChain, LlamaIndex, Custom
React Library
Embed interactive chat and agents into any frontend
Built with React components
Build Powerful Agents
BUILD & RUN AGENTS
Tool Approval
Review and approve tool execution
Generative UI Rendering
Rich, dynamic UI generated by AI
Interactive Execution
Real-time conversations and actions
Background Execution
Async tasks and long-running jobs
AGENT HARNESS
Sandbox
Isolated environments for safe execution
SubAgents
Decompose and orchestrate complex tasks
Memory
Persistent memory and context recall
Context Management
Automatic Context Compaction
Intelligently compress and optimize context
Code Mode (Programmatic Tool Calling)
Enable precise, programmatic tool interactions
Large Tool Results Handling
Efficiently manage and process large tool outputs
Deferred Tools Loading
Load tools on-demand for better performance
FOUNDATION (LAYER 0)
Model Registry
20+ Model Providers
Versioning & Routing
Guardrails & Policies
Cost & Rate Limits
MCP Registry
Connect any MCP server
Centralized Auth
Tool Discovery
Policies & Guardrails
Skills Registry
Reusable Skills
Versioned & Shareable
Discover & Compose
Access Control
OBSERVABILITY & GOVERNANCE
Complete visibility and control across the entire stack
Traces & Logs
End-to-end tracing across agents, tools, and models
Metrics & Insights
Performance, latency, errors, and custom metrics
Cost Tracking
Track usage, costs, and budgets in real-time
Governance
Policies, access control, audit logs, and compliance
Export
Export to any OTEL compatible observability stack
Build & Use
Orchestration
Foundation
Observability

4 · BUILDING ON TRUEFOUNDRY

Build an agent in four choices

1. **Choose a model**
2. **Connect MCP servers**
3. **Add skills**
4. **Write instructions & save**

Anyone can do those four steps. What TrueFoundry manages for you out of the box is the hard part:

- **Context engineering** *automatic*
- **Sandbox** *one toggle*
- **Human-in-the-loop** *built in*
- **Generative UI** *out of the box*
- **Tracing & cost** *every run*
- **Governance & RBAC** *centrally*

truefoundry

Live Demo

Build, test and run a managed agent on TrueFoundry — including Ask AI, end-to-end.

Switch to the TrueFoundry console

5 · SHARING AGENTS

How do you share agents across your company?

Building an agent is half the job — getting it safely into your team's hands is the other half.

With a high-code framework

e.g. LangGraph

- You own deployment: containers, scaling, uptime
- Wire up auth, RBAC and secrets yourself, per app
- Stand up your own logging, tracing & cost tracking
- Each new agent repeats the same plumbing
- Flexible — but it costs engineering time

With TrueFoundry

share in a few clicks

- ✓ Publish once; share with users or teams via RBAC
- ✓ Access, budgets & guardrails enforced centrally
- ✓ Per-user MCP auth — each person uses their own access
- ✓ Every run traced, with cost & tokens, by default
- ✓ Same definition works in UI, API and SDK

Go build an agent.

Pick a model, connect your tools, add a skill, write your instructions — and ship it to your team with governance and observability built in.

Docs: truefoundry.com/docs/ai-gateway/agent-harness

Questions? Let's open it up.

truefoundry

Q&A

What would you build first? Ask away.

truefoundry

Thank you

Thanks for joining TrueFoundry Office Hours #4. See you at the next session!

Docs truefoundry.com/docs

Support support@truefoundry.com