AGENTOPS WORKSHOP

INSTRUCTOR DELIVERY GUIDE

AgentOps Value Delivery Workshop

How to facilitate the full-day hands-on VBD Workshop

Format: ~8 hours | Hands-on labs

This guide helps you facilitate the full-day AgentOps Value Delivery Workshop. The day is hands-on: attendees build, break, and fix a real agent. Your job is to keep the continuity spine intact, protect the timeboxes, and make the two deliberate failure moments land as teaching peaks rather than support tickets.

Room and logistics

- Plan for two breaks and a meal break; the table below shows checkpoints.
- Pair attendees who are new to the Azure CLI with someone comfortable in a terminal.
- Keep a shared channel open for pasting commands and error output.
- Have a backup Foundry project ready in case an attendee's access fails.

Verify prerequisites before Lab 1

- Azure subscription with a Microsoft Foundry project and a gpt-4o-mini deployment.
- Python 3.10+, Azure CLI, and git installed and on PATH.
- A GitHub repository each attendee can push to.
- An Application Insights resource and its connection string.
- Run the pre-flight check so no one starts Lab 1 blocked.

Pacing and checkpoints

Block	Lab	Checkpoint
Morning	Lab 1 - Foundations	Everyone has travel-agent:1 live in Foundry.
Morning	Lab 2 - Evaluation	Each attendee has a passing eval suite.
Late morning	Lab 3 - Release gates	A regression is blocked by the gate (exit code 2).
Afternoon	Lab 4 - Observability	Traces correlate to a specific release.
Afternoon	Lab 5 - Safety	Safety checks and red-team follow-through run.
Late afternoon	Lab 6 - Improvement	A trace feeds a new eval case.
Close	Lab 7 - Capstone	A red PR is driven to green through the full loop.

Lab-by-lab facilitation notes

Lab 1 - Foundations and control plane

- Objective: deploy travel-agent:1 and set up the agentops-vbd workspace and CLI.
- Common stumble: wrong Foundry project or model deployment name. Verify early.
- Produces: a live agent and a configured workspace the rest of the day depends on.

Lab 2 - Evaluation suites

- Objective: author an eval suite that scores the agent on real travel criteria.
- Common stumble: vague pass criteria. Push attendees to concrete, checkable rules.
- Produces: an eval suite that Lab 3 will wire into a gate.

Lab 3 - Release gates and CI

- Objective: turn the eval suite into a CI gate, then watch it block a regression.
- Teaching peak: the gate exits non-zero (exit code 2) on a deliberately weakened travel-agent:2. Let attendees feel the block before fixing it.
- Produces: a passing, gated pipeline and travel-agent:3.

Lab 4 - Observability deep dive

- Objective: emit traces and correlate runtime telemetry back to a release.
- Reserve the full 90 minutes; this is the deepest lab and the core of AgentOps.
- Produces: dashboards and traces that the capstone reuses as release evidence.

Lab 5 - Safety and red-team

- Objective: add safety checks and a red-team pass with follow-through actions.
- Common stumble: treating safety as a one-time scan. Frame it as a standing gate.

Lab 6 - Continuous improvement

- Objective: turn a real trace into a new eval case, closing the loop.
- This is where Evaluate, Ship, Observe, and Operate visibly connect.

Lab 7 - Capstone

- Objective: drive a red pull request to green through the entire operating loop.
- Teaching peak: the PR starts failing the gate. Attendees run the full loop to ship it.
- Produces: an evidence pack proving the agent is production ready.

If attendees lack live access

If an attendee cannot reach Azure, pair them with a working attendee as navigator and have them complete the artifacts locally where possible. Every lab states what it consumes and produces, so a pair can stay on the spine even with one live environment.

> **After the session**
>
> Have each attendee keep their agentops-vbd workspace and evidence pack. The fastest path to value is to point the same loop at one of their own agents within two weeks, while the muscle memory is fresh.