

10 essential LLMOps tools every AI team should have.

Discover 10 essential LLMOps tools to help build a complete AI production system in 2026.

The operation of large language model implementations (LLMOps) in 2026 has changed significantly compared to a few years ago. While previously AI deployment revolved around selecting a model and adding a few tracking steps, development teams now need a much more complete ecosystem.

Modern AI systems require orchestration, routing, system observation, evaluation, guardrails, long-term memory, user feedback, and practical tool packaging and integration. In other words, LLMOps have evolved into a complete production stack, where each component plays a crucial role in running AI at enterprise scale.

Below are 10 representative tools, each representing a part of a modern LLMOps system.

PydanticAI: A platform that enables AI systems to function like software.

PydanticAI is becoming a popular choice for teams looking to build well-structured LLM systems. This tool focuses on output with defined data types, supports multiple models, and handles lengthy workflows with the ability to recover from errors.

PydanticAI's strength lies in its ability to mitigate runtime risks as systems become increasingly complex. As the number of tools, schemas, and agents increases, ensuring stable output becomes more crucial than ever.

Bifrost: Gateway for multi-model systems

Bifrost acts as a gateway layer for systems using multiple AI models or vendors. It provides a single API to route through over 20 different providers, resulting in cleaner application code.

Additionally, Bifrost supports failover, load balancing, caching, and access control. The tool also integrates OpenTelemetry for system monitoring during production runs, making operation easier.

Traceloop / OpenLLMetry: Observing LLM systems

OpenLLMetry is suitable for teams already using OpenTelemetry. This tool records prompt, completion, token usage, and system trace in a format consistent with existing logs.

This allows development teams to debug AI behavior more easily and monitor the system just like other backend components.

Promptfoo: Testing and Evaluating AI Systems

Promptfoo is an open-source tool that helps integrate testing into the AI development process. It allows for the creation of iterative test cases, running evaluations, and red-teaming applications.

The key point is that Promptfoo can be integrated into CI/CD, enabling automated testing before deployment. This helps transform prompt changes into a measurable and controllable process.

Invariant Guardrails: Agent Behavior Control

When the AI agent starts calling APIs or manipulating the actual system, guardrails become extremely important. Invariant Guardrails allow the establishment of runtime rules between the application and the model.

This helps control the agent's behavior without constantly changing the application code, which is especially useful as the system scales.

Letta: Long-term memory for AI agents

Letta is designed for agents that need long-term memory retention. This tool stores state in a Git-like structure, making it easy to track changes, debug, and rollback when needed.

This is a crucial component for agents that run long-term or perform complex workflows.

OpenPipe: Learning from real-world data

OpenPipe helps systems learn from production data. This tool supports logging requests, creating datasets, running evaluations, and fine-tuning models.

This allows development teams to build continuous improvement loops from real-world data.

Argilla: Gathering feedback from people

Argilla focuses on user feedback and data processing. This tool helps collect feedback, label data, and analyze errors.

This is a crucial component if you want to improve model quality over time, especially when using RLHF.

KitOps: Packaging a complete AI system

KitOps addresses a common problem in AI where models, datasets, prompts, and configurations are scattered. This tool packages everything into a single, clearly versioned artifact.

This makes deployment easier, while also supporting rollback and sharing between development teams.

Composio: Connecting AI with Real-World Applications

Composio helps agents connect with external applications such as Slack, Gmail, GitHub, or CRM. This tool handles authentication, permissions, and execution.

This is a crucial step as AI transitions from demos to real-world workflows within businesses.

LLMOps are entering adulthood.

LLMOps are no longer just about choosing a model. Instead, businesses need to build a complete system encompassing testing, observability, memory, guardrails, and integration.

In 2026, the crucial question will no longer be which model to use, but how to build a system around that model. This is the major shift for LLMOps in the age of AI agents.

You finished reading the article "**10 essential LLMOps tools every AI team should have.**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.