Feature flags and AI Configs

AI Configs

Overview

This category explains how to use LaunchDarkly to manage your AI Configs. An **AI Config** is a resource that you create in LaunchDarkly. You can use AI Configs to customize, test, and roll out new large language models (LLMs) within your generative AI applications.

With AI Configs, you can:

- Manage your model configuration outside of your application code. This means you can update model details and messages at runtime, without deploying changes. Teammates who have LaunchDarkly access but no codebase familiarity can collaborate and iterate on messages.

- Upgrade your app to the latest model version as soon as it's available, then roll out the change gradually and safely.

- Add a configuration for a new model provider and progressively shift production traffic to that provider.

- Compare variations to determine which performs better based on satisfaction, cost, or other metrics. You can compare messages across models or compare the same model across providers.

- Run experiments to measure the impact of generative AI features in your application.

AI Configs also support advanced use cases such as retrieval-augmented generation (RAG) and evaluation in production. You can:

- Track which knowledge base or vector index is active for a given model or audience.

- Experiment with different chunking strategies, retrieval sources, or prompt structures.

- Build guardrail logic into your runtime configs, such as models, prompts, and filters, and use targeting rules to block risky generations or switch to fallback behavior.

- Apply different safety filters by user type, geography, or application context.

- Use live metrics for satisfaction, factuality, and hallucination detection to guide rollouts.

These capabilities let you build safety controls, run targeted experiments, and evaluate model behavior in production without being locked into a single model provider or manual workflow.

How AI Configs work

Every AI Config contains one or more variations. Each variation includes a model configuration and, optionally, one or more messages. You can also define targeting rules, just like you do with feature flags, to make sure that particular messages and model configurations are served to particular end users of your application.

Then, within your application, you use one of LaunchDarkly's AI SDKs. The SDK determines which messages and model your application should serve to which contexts. The SDK can also customize the messages based on context attributes and other variables that you provide. This means both the messages and the model evaluation are modified to be specific to each end user at runtime. You can update your messages, specific to each end user, without redeploying your application.

After you use this customized config in your AI model generation, you can use the SDK to record various metrics, including generation count, tokens, and satisfaction rate. These appear in the LaunchDarkly user interface for each AI Config variation.

The topics in this category explain how to create AI Configs and variations, update targeting rules, monitor related metrics, and incorporate AI Configs in your application.

Additional resources

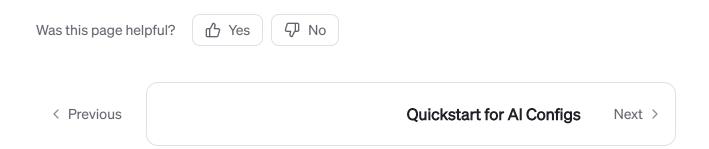- Quickstart for AI Configs

- Create AI Configs and variations

- Target with AI Configs

- Monitor AI Configs

- Manage AI Configs

- Run experiments with AI Configs

- AI Configs and information privacy

- Agents in AI Configs

In our guides:

- Managing AI model configuration outside of code

- Using targeting to manage AI model usage by tier

In our SDK documentation:

- .NET AI SDK reference

- Go AI SDK reference

- Node.js (server-side) AI SDK reference

- Python AI SDK reference

- Ruby AI SDK reference

Was this page helpful?   👍 Yes   👎 No

Previous   Quickstart for AI Configs   Next

Built with ❄ fern