



Event ends in 2 days 15 hours 36 minutes.



[Event dashboard](#) > [Guardrails for Amazon Bedrock](#) >

What are Guardrails for Bedrocks?

What are Guardrails for Bedrocks?

Amazon Bedrock Guardrails enable you to implement safeguards and enforce responsible AI policies for your generative AI applications based on specific use cases. With Guardrails, you can create multiple tailored configurations and apply them across different foundation models, ensuring a consistent user experience and standardized safety controls across all your generative AI applications.

Guardrails allow you to configure denied topics to prevent undesirable subjects from being discussed and content filters to block harmful content in both input prompts and model responses. Guardrails can be used with text-only foundation models.

A guardrail consists of the following policies to avoid content that falls into undesirable or harmful categories:

- **Denied topics:** Define a set of topics that are undesirable for your application. These topics will be blocked if detected in user queries or model responses.
- **Content filters:** Adjust filter strengths to filter input prompts or model responses containing harmful content.
- **Word filters:** Configure filters to block undesirable words, phrases, and profanity.
- **Sensitive information filters:** Block or mask personally identifiable information (PII) and use regular expressions to define and then block or mask patterns that might correspond to sensitive information.

Additionally, you can configure the messages to be returned to the user if a user input or model response violates the policies defined in the guardrail.



the built-in test window to ensure they align with your use case. Once satisfied, create a version of the guardrail and use it with supported foundation models.

Guardrails can be applied directly to foundation models during inference API invocations by specifying the guardrail ID and version. If a guardrail is used, it will evaluate the input prompts and foundation model completions against the defined policies.

[Previous](#)[Next](#)