

# Troubleshooting guide

Use this guide to help you diagnose and resolve common issues that arise when you call the Gemini API. You may encounter issues from either the Gemini API backend service or the client SDKs. Our client SDKs are open sourced in the following repositories:

- [python-genai](https://github.com/googleapis/python-genai) (https://github.com/googleapis/python-genai)
- [js-genai](https://github.com/googleapis/js-genai) (https://github.com/googleapis/js-genai)
- [go-genai](https://github.com/googleapis/go-genai) (https://github.com/googleapis/go-genai)

If you encounter API key issues, ensure you have set up your API key correctly per the [API key setup guide](/gemini-api/docs/api-key) (/gemini-api/docs/api-key).

## Gemini API backend service error codes

The following table lists common backend error codes you may encounter, along with explanations for their causes and troubleshooting steps:

HTTPStatus Code	Description	Example	Solution
400 INVALID_ARGUMENT	The request body is malformed.	There is a typo, or a missing required field in your request.	Check the <a href="/api">API reference</a> (/api) for request format, examples, and supported versions. Using features from a newer API version with an older endpoint can cause errors.

400	FAILED_PRECONDITION	Gemini API free tier is not available in your country. Please enable billing on your project in Google AI Studio.	You are making a request in a region where the free tier is not supported, and you have not enabled billing on your project in Google AI Studio.	To use the Gemini API, you will need to setup a paid plan using <a href="#">Google AI Studio</a> ( <a href="https://aistudio.google.com/app/apikey">https://aistudio.google.com/app/apikey</a> ) .
403	PERMISSION_DENIED	Your API key doesn't have the required permissions.	You are using the wrong API key; you are trying to use a tuned model without going through <a href="#">proper authentication</a> ( <a href="/docs/model-tuning/tutorial?lang=python#set_up_authentication">/docs/model-tuning/tutorial?lang=python#set_up_authentication</a> ) .	Check that your API key is set and has the right access. And make sure to go through proper authentication to use tuned models.
404	NOT_FOUND	The requested resource wasn't found.	An image, audio, or video file referenced in your request was not found.	Check if all <a href="#">parameters in your request are valid</a> ( <a href="/docs/troubleshooting#check-api">/docs/troubleshooting#check-api</a> ) for your API version.
429	RESOURCE_EXHAUSTED	You've exceeded the rate limit.	You are sending too many requests per minute with the free tier Gemini API.	Ensure you're within the model's <a href="#">rate limit</a> ( <a href="/models/gemini#model-variations">/models/gemini#model-variations</a> ) . <a href="#">Request a quota increase</a> ( <a href="/docs/increase_quota">/docs/increase_quota</a> ) if needed.

500	INTERNAL	An unexpected error occurred on Google's side.	Your input context is too long.	Reduce your input context or temporarily switch to another model (e.g. from Gemini 1.5 Pro to Gemini 1.5 Flash) and see if it works. Or wait a bit and retry your request. If the issue persists after retrying, please report it using the <b>Send feedback</b> button in Google AI Studio.
503	UNAVAILABLE	The service may be temporarily overloaded or down.	The service is temporarily running out of capacity.	Temporarily switch to another model (e.g. from Gemini 1.5 Pro to Gemini 1.5 Flash) and see if it works. Or wait a bit and retry your request. If the issue persists after retrying, please report it using the <b>Send feedback</b> button in Google AI Studio.
504	DEADLINE_EXCEEDED	The service is unable to finish processing within the deadline.	Your prompt (or context) is too large to be processed in time.	Set a larger 'timeout' in your client request to avoid this error.

## Check your API calls for model parameter errors

Ensure your model parameters are within the following values:

Model parameter	Values (range)

Candidate count	1-8 (integer)
Temperature	0.0-1.0
Max output tokens	Use <code>get_model</code> (Python ( <a href="/api/python/google/generativeai/get_model">/api/python/google/generativeai/get_model</a> )) to determine the maximum number of tokens for the model you are using.
TopP	0.0-1.0

In addition to checking parameter values, make sure you're using the correct [API version](/gemini-api/docs/api-versions) (`/gemini-api/docs/api-versions`) (e.g., `/v1` or `/v1beta`) and model that supports the features you need. For example, if a feature is in Beta release, it will only be available in the `/v1beta` API version.

## Check if you have the right model

Ensure you are using a supported model listed on our [models page](/gemini-api/docs/models/gemini) (`/gemini-api/docs/models/gemini`).

## Safety issues

If you see a prompt was blocked because of a safety setting in your API call, review the prompt with respect to the filters you set in the API call.

If you see `BlockedReason.OTHER`, the query or response may violate the [terms of service](/terms) (`/terms`) or be otherwise unsupported.

## Recitation issue

If you see the model stops generating output due to the `RECITATION` reason, this means the model output may resemble certain data. To fix this, try to make prompt / context as unique as possible and use a higher temperature.

## Improve model output

For higher quality model outputs, explore writing more structured prompts. The [prompt engineering guide](/gemini-api/docs/prompting-strategies) (/gemini-api/docs/prompting-strategies) page introduces some basic concepts, strategies, and best practices to get you started.

If you have hundreds of examples of good input/output pairs, you can also consider [model tuning](/gemini-api/docs/model-tuning) (/gemini-api/docs/model-tuning).

## Understand token limits

Read through our [Token guide](/gemini-api/docs/tokens) (/gemini-api/docs/tokens) to better understand how to count tokens and their limits.

## Known issues

- The API supports only a number of select languages. Submitting prompts in unsupported languages can produce unexpected or even blocked responses. See [available languages](/gemini-api/docs/models#supported-languages) (/gemini-api/docs/models#supported-languages) for updates.

## File a bug

Join the discussion on the [Google AI developer forum](https://discuss.ai.google.dev) (https://discuss.ai.google.dev) if you have questions.

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (https://www.apache.org/licenses/LICENSE-2.0). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (https://developers.google.com/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2025-05-09 UTC.