



Gemini 3.1 Flash Audio (Flash Live, TTS) Model Card

Gemini 3.1 Flash Audio (Flash Live, TTS) - Model Card

Model Cards are intended to provide essential information on Gemini models, including known limitations, mitigation approaches, and safety performance. Model cards may be updated from time to time; for example, to include updated evaluations as the model is improved or revised. See the [Google DeepMind site](#) for a comprehensive list of model cards.

Published: March 2026, Updated: April 2026

Model Information

Description: Gemini 3.1 Flash Audio (Flash Live, TTS) is a member of the Gemini series of models, a suite of highly capable, natively multimodal reasoning models. This model card describes the native capabilities (e.g., image and audio) as additional outputs of Gemini 3.1 Flash. Information specific to these modalities is specified in-line and referred to as **Gemini 3.1 Flash Live** or **Gemini 3.1 Flash TTS** (Text-to-Speech), referred to collectively as **Gemini 3.1 Flash Audio**.

Model dependencies: Gemini 3.1 Flash Audio is based on Gemini 3 Pro.

Inputs:

- Gemini 3.1 Flash Live: Audio, images, video, and text with a token context window of up to 128K.
- Gemini 3.1 Flash TTS: Text up to 16K.

Outputs:

- Gemini 3.1 Flash Live: Audio and text, with 64K token output.
- Gemini 3.1 Flash TTS: Audio with 32K token output.

Architecture: Gemini 3.1 Flash Audio is based on Gemini 3 Pro. For more information about the model architecture for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).

Model Data

Training Dataset: Gemini 3.1 Flash Audio is based on Gemini 3 Pro. For more information about the training dataset for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).

Training Data Processing: For more information about the training data processing for Gemini 3.1 Flash Audio see the Gemini 3 Pro [model card](#).

Implementation and Sustainability

Hardware: Gemini 3.1 Flash Audio is based on Gemini 3 Pro. For more information about the hardware for Gemini 3.1 Flash Audio and our continued [commitment to operate sustainably](#), see the Gemini 3 Pro [model card](#).

Software: Gemini 3.1 Flash Audio is based on Gemini 3 Pro. For more information about the software for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).

Distribution

Gemini 3.1 Flash Audio is distributed in the following channels; respective documentation shared in line:

Gemini 3.1 Flash Live:

- [Gemini App](#)
- [Google AI Studio](#)
- [Gemini API](#)
- [Google Antigravity](#)
- [NotebookLM](#)

Gemini 3.1 Flash TTS:

- [Google AI Studio](#)
- [Gemini API](#)
- [Google Cloud / Vertex AI](#)
- [Google Vids](#)

Our models are available to downstream providers via an application program interface (API) and subject to relevant terms of use. There is no required hardware or software to use the model. For AI Studio and Gemini API, see the [Gemini API Additional Terms of Service](#); for Vertex AI, see [Google Cloud Platform Terms of Service](#). For more information, see [Gemini Model API instructions](#) and [Gemini API in Vertex AI quickstart](#).

Evaluation

Approach: Gemini 3.1 Flash Audio was evaluated across a range of benchmarks. Details on approach, results, and their methodologies can be found at:

- Gemini 3.1 Flash Live: deepmind.google/models/evals-methodology/gemini-3-1-flash-live
 - Gemini 3.1 Flash TTS: deepmind.google/models/evals-methodology/gemini-3-1-flash-tts
-

Intended Usage and Limitations

Benefit and Intended Usage (Flash Live): Gemini 3.1 Flash Live enables low-latency, real-time voice and video interactions. It processes continuous streams of audio, video, or text to deliver immediate, human-like spoken responses, creating a natural conversational experience for your users.

Benefit and Intended Usage (Flash TTS): Gemini 3.1 Flash TTS is our newest text-to-speech model, offering enhanced control, expressiveness, and audio quality to help everyone—from developers, enterprise customers to everyday users—create next-gen AI speech applications.

Known Limitations: For more information about the known limitations for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).

Acceptable Usage: For more information about the acceptable usage for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).

Ethics and Content Safety

Evaluation Approach: Gemini 3.1 Flash Audio was developed in partnership with internal safety and responsibility teams. A range of evaluations and red teaming activities were conducted to help improve the model and inform decision-making. These evaluations and activities align with [Google's AI Principles](#) and [responsible AI approach](#), as well as Google's Generative AI policies (e.g., [Gen AI Prohibited Use Policy](#) and the [Gemini API Additional Terms of Service](#)).

Evaluation types included but were not limited to:

- **Training/Development Evaluations** including automated and human evaluations carried out continuously throughout and after the model's training, to monitor its progress and performance;

- **Human Evaluations** conducted by specialist teams across the policies and desiderata to ensure the model adheres to safety policies and desired outcomes;
- **Ethics & Safety Reviews** conducted ahead of the model's release.

Safety Policies: Gemini's safety policies are based on Google's standard framework and are intended to prevent our Generative AI models from generating harmful content, including:

1. Content related to child sexual abuse material and exploitation
2. Hate speech (e.g., dehumanizing members of protected groups)
3. Dangerous content (e.g., promoting suicide, or instructing in activities that could cause real-world harm)
4. Harassment (e.g., encouraging violence against people)
5. Sexually explicit content
6. Medical advice that runs contrary to scientific or medical consensus

Frontier Safety Assessment: Gemini 3.1 Flash Audio is part of the Gemini 3 family of models. For frontier safety, we rely on our evaluation of Gemini 3.1 Pro with Deep Think mode as it is the most generally capable model as of publication of this model card, and it did not reach the Critical Capability Levels (CCLs) outlined in our [Frontier Safety Framework](#). Our assessments have shown that Gemini 3.1 Flash Audio is less capable than Gemini 3.1 Pro. Therefore, we are confident that Gemini 3.1 Flash is also unlikely to reach any CCLs. For more information, read the [Gemini 3.1 Pro Model Card](#).

Risks and Mitigations: For more information about the risks and mitigations for Gemini 3.1 Flash Audio, see the Gemini 3 Pro [model card](#).