



Gemini 3.5 Audio (Live Translate) Model Card

Gemini 3.5 Audio (Live Translate) - 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: June 2026

Model Information

Description: Gemini 3.5 Live Translate is a member of the Gemini series of models, a suite of highly-capable, natively multimodal reasoning models.

Model dependencies: Gemini 3.5 Live Translate is based on Gemini 3 Pro.

Inputs: Audio with a token context window of up to 128K.

Outputs: Audio and text, with up to 64K token output.

Architecture: Gemini 3.5 Live Translate is based on Gemini 3 Pro. For more information about the model architecture for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Model Data

Training Dataset: Gemini 3.5 Live Translate is based on Gemini 3 Pro. For more information about the training dataset for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Training Data Processing: For more information about the training data processing for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Implementation and Sustainability

Hardware: Gemini 3.5 Live Translate is based on Gemini 3 Pro. For more information about the hardware for Gemini 3.5 Live Translate and our continued [commitment to operate sustainably](#), see the Gemini 3 Pro [model card](#).

Software: Gemini 3.5 Live Translate is based on Gemini 3 Pro. For more information about the software for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Distribution

Gemini 3.5 Live Translate is distributed in the following channels; respective documentation shared in line:

- [Google Translate](#)
- [Google AI Studio](#)
- [Gemini API](#)

Our models are available to downstream providers via an Application Programming 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 more information, see [Gemini Model API instructions](#).

Evaluation

Approach: Gemini 3.5 Live Translate was evaluated across three main quality dimensions: translation quality, latency, and speech naturalness.

Capabilities / Benchmarks:

- **Translation Quality:** We evaluate translation quality using [AutoMQM](#) - an error-based automatic metric which identifies and categorizes translation errors to produce a fine-grained quality score. These metrics are applied across a variety of language pairs and content types to assess the accuracy and adequacy of the translated output.
- **Latency:** For a real-time translation system, latency (how far behind the translation lags during a live session) is a critical quality dimension. We measure latency at multiple granularities:
 - *Initial latency:* the time between the start of speech in the input stream and the start of speech in the translated output stream. This is measured using standard voice activity detection.
 - *Word-level latency:* to capture the most complete picture of translation delay, we align words in the input to their corresponding words in the output and measure the average time between the end of a source word and the start of its corresponding translated word.
- **Speech Naturalness:** We evaluate the naturalness and quality of the synthesized output audio using established speech synthesis quality metrics. These capture issues such as choppy or discontinuous audio, voice drift (where the output voice characteristics shift over the course of a session), and unintended artifacts. Maintaining natural-sounding, consistent output audio is essential for a usable real-time translation experience.

Methodology: Gemini 3.5 Live Translate was evaluated using internal implementations of these benchmarks on outputs generated using the Gemini Live API.

Intended Usage and Limitations

Benefit and Intended Usage: Gemini 3.5 Live Translate enables low-latency, real-time translation interactions. It processes continuous streams of audio to deliver immediate, human-like spoken responses, creating a natural translation experience for your users.

Known Limitations: Gemini 3.5 Live Translate may exhibit some of the following limitations. Voices can be inconsistent, and voices may shift after long pauses, change gender, or get stuck on one voice during rapid multi-speaker sessions. Language detection can struggle with non-native accents, similar languages, or rapid language switches. Gemini 3.5 Live Translate is designed to filter out background noise, but not all background audio may be ignored. When set to echo the target language, background noise may introduce artifacts in the translated audio when input audio is in the target language.

For more information about the known limitations for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Acceptable Usage: For more information about the acceptable usage for Gemini 3.5 Live Translate, see the Gemini 3 Pro [model card](#).

Ethics and Content Safety

Evaluation Approach: Gemini 3.5 Live Translate was developed in partnership with internal safety and responsibility teams. Evaluations conducted in alignment with [Google's AI Principles](#) and [responsible AI approach](#), as well as Google's Generative AI policies (e.g., the [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 throughout and after the model's training, to monitor its progress and performance;
- **Ethics & Safety Reviews** were conducted ahead of the model's release including human evaluation of the final model to ensure the model adhere to safety policies.

Frontier Safety Assessment: Gemini 3.5 Live Translate 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 3.5 Live Translate is less capable than Gemini 3.1 Pro, therefore based on Gemini 3.1 Pro, we are confident that Gemini 3.5 Live Translate 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 3.5 Live Translate, see the Gemini 3 Pro [model card](#).