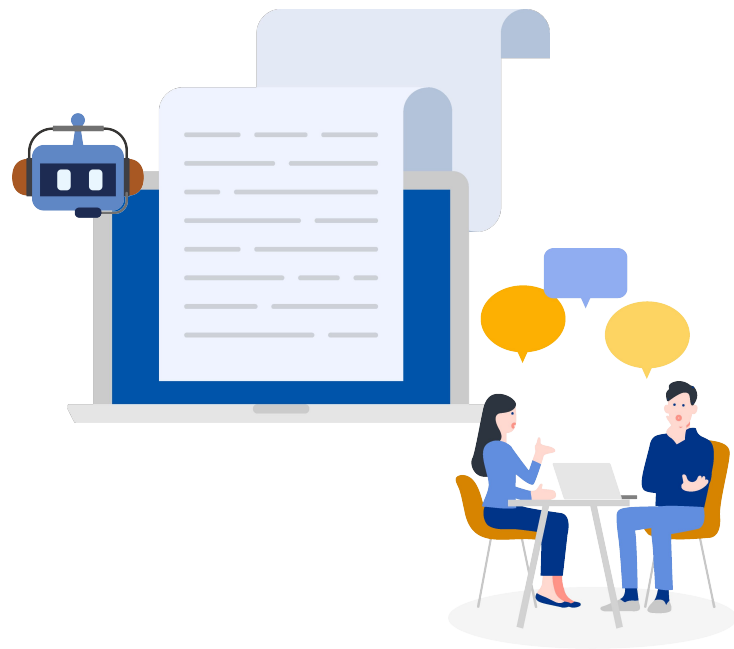


Gemini

文字起こし

簡単活用ガイド



Gemini文字起こし 簡単活用ガイド 1

Gemini 3 Pro / Ultraが持つ業界最大級の「ロングコンテキスト（数百万～1,000万トークン）」と、最新の推論機能「Deep Think」。次世代の文字起こし・要約術とは、これらを掛け合わせたものです。

単に音声を文字にするだけでなく、映像情報、配布資料、過去の文脈まですべてを統合して理解し、人間以上の精度で議事録や分析レポートを作成する方法を解説します。

1 準備と環境

この手法を最大限に活用するには、以下の環境推奨です。

推奨プラン

Google AI プラン
(個人向け)

or

Gemini for Google Workspace
(法人向け)

Google AI Studio (開発者向け)

100万～200万トークン（モデルによる）を扱えるため、数時間の動画を10本まとめて読み込ませるような「超重量級」の分析にはこちらが最適です。

※無料版（Gemini 3 Flash）でも利用可能ですが、処理できるデータ量や推論の深さに制限がある場合があります。

Gemini文字起こし 簡単活用ガイド 2

2 実践手順：ステップバイステップ

Step1

ファイルのアップロード

Geminiのチャット画面にある「+（プラス）」ボタン、または「ファイルをアップロード」から、対象のファイルを読み込ませます。

※1 動画ファイルは映像情報も解析対象になります

※2 会議で使用した資料も同時にアップロードすることで精度が劇的に向上します

対応形式

音声 MP3, WAV, AAC, M4A など

動画 MP4, MOV, MPEG など※1

資料 PDF, Googleドキュメント, スライド※2

Step2

「Deep Think」モードの起動

複雑な会議や、行間を読む必要がある分析を行う場合は、モデル設定から「Deep Think（深い推論）」モードをONにします。これにより、Geminiは回答を出力する前に内部で思考プロセスを回し、発言の意図や文脈を深く理解しようとします。ただし、単なる「議事録の文字起こし」や「要約」だけなら、Deep Thinkモードを起動しなくても十分に高性能です。

出力の精度を上げるために、目的に合わせて、具体的かつ明確な指示を出します。

Step3

プロンプト（指示）の入力

基本の議事録プロンプト例

あなたはプロのプロジェクトマネージャーです。添付の音声ファイルは、新製品開発に関する定例会議の録音です。この内容を詳細に文字起こしした上で、以下のフォーマットで議事録を作成してください。

- 1 会議の概要（300文字以内）
- 2 決定事項（箇条書き）
- 3 持ち越し課題・Next Action（担当者と期限を明記）
- 4 主な議論のポイント（対立点や合意形成のプロセスも含めて）

専門用語『Project X』『Alpha Module』は正しく表記してください。

分析・洞察プロンプト例

この動画は、新商品の試作品に関する顧客インタビューの様子です。ターゲットは30代の共働き世代です。

- **ポジティブな兆候**
（発言と、それを裏付ける表情・トーンの記述）
 - **潜在的な懸念点**
（言葉では肯定しているが、表情が曇った瞬間など）
 - **改善へのアクション**
（分析結果を踏まえた具体的なアドバイス）
- 上記の出力フォーマットに則り、分析してください。

Gemini文字起こし 簡単活用ガイド 3

3 Geminiならではの圧倒的なメリット

圧倒的なコンテキスト容量（ロングコンテキスト）

従来のAIでは「1時間の会議」程度で容量オーバーになり、分割処理が必要でした。しかし、Gemini 3 Pro/Ultraは**数時間～半日分の動画や音声データ**を丸ごと一度に読み込みます。これにより、会議の冒頭の発言と、終了間際の発言の矛盾を指摘したり、長時間にわたる文脈を一貫して理解したりすることが可能です。

マルチモーダル理解（映像×音声×資料）

映像と音声と資料を同時に理解し、空間的・時間的に紐付ける能力が、Gemini最大の強みです。

● ホワイトボードの解析

動画内でホワイトボードに書かれた図や文字を認識し、それを踏まえて発言内容を補完します。「ここにある数字」という発言が何を指すのか、映像を見て理解します。

● 資料との照合

会議資料（PDF）を同時に読み込ませることで、「資料の3ページ目の件ですが」という発言に対し、該当ページの内容を参照しながら正確な要約を作成します。

Deep Thinkによる「文脈」と「感情」の読解

Deep Thinkモードでは、単なる音声認識（Speech-to-Text）を超えた分析を行います。

● 話者分離と意図理解

「Aさんが賛成したのは、条件付きであることが文脈から読み取れます」といった、ニュアンスを含めた解釈が可能です。

● 感情分析

声の大きさ、間、トーンから、会議の雰囲気がポジティブだったか、緊迫していたかを分析できます。

エージェント連携によるアクション完結

Google Workspace拡張機能と連携することで、作成した議事録から次のアクションを自動化できます。

決定した次回会議の日程（〇月〇日 14:00）をGoogleカレンダーに登録し、参加者に招待を送って

Next ActionをToDoリストに追加し、担当者へリマインドメールの下書きを作成して

これらがチャット画面からワンストップで実行可能です。

Gemini文字起こし 簡単活用ガイド 4

4 活用シーン別 応用テクニック

決算説明会の分析

投資家向け説明会の動画と、説明資料（PDF）、財務諸表（Excel）をすべてアップロード。「経営陣の質疑応答での回答内容と、財務諸表の数字に矛盾がないかチェックして」と指示することで、「音」と「数字」をクロスオーバーさせ、矛盾や相関を見つける高度な企業分析が瞬時に完了します。

グローバル会議の議事録

多言語が飛び交う会議の録音データをアップロード。「英語と日本語が混ざっていますが、すべての発言を日本語に翻訳して議事録を作成してください」と指示すれば、文脈を維持した高精度な翻訳議事録が作成されます。

講義・セミナーの要約

90分の大学講義の音声を読み込ませ、「テストに出そうな重要ポイントをクイズ形式で10問作成して」と指示することで、効率的な学習教材を生成できます。

5 注意点とセキュリティ

プライバシー

個人のプライベートな会話や、機密性の高い会議データをアップロードする際は、所属組織のAI利用ガイドラインに従ってください。
※Gemini Business/Enterpriseプランではデータが学習に利用されない設定が可能です

ハルシネーション（嘘）

AIは稀に聞き間違いや事実と異なる内容を生成することがあります。特に重要な数字や固有名詞については、必ず生成されたタイムスタンプ（発言箇所の時間指定）を元に、元の音声を確認するようにしてください。

【当資料における免責事項】

本資料は、作成時点（2026年2月）における公開情報を基に作成しております。AI技術および関連サービスは極めて変化が速く、各ツールの料金体系、プラン内容、提供機能（UI/UX含む）などは、頻繁にアップデートが行われます。

そのため、本資料の閲覧時には情報が古くなっている可能性がございます。当社としても定期的な情報のキャッチアップと資料のアップデートに努めておりますが、あまりに変化のスピードが速いため、すべての変更をリアルタイムに反映することが困難な場合がございます。

そのため実践の際は、あくまで作成時典の参考情報としてご覧いただき、最新の正確な情報は各サービスの公式サイト等でご確認ください。