



日本語ドキュメント

Model Vista Voice Control

Model Vista Voice Control のユーザードキュメント。

2026年10月30日

目次

01 Model Vista Voice Control

02 インストールと接続

03 音声セッションを使う

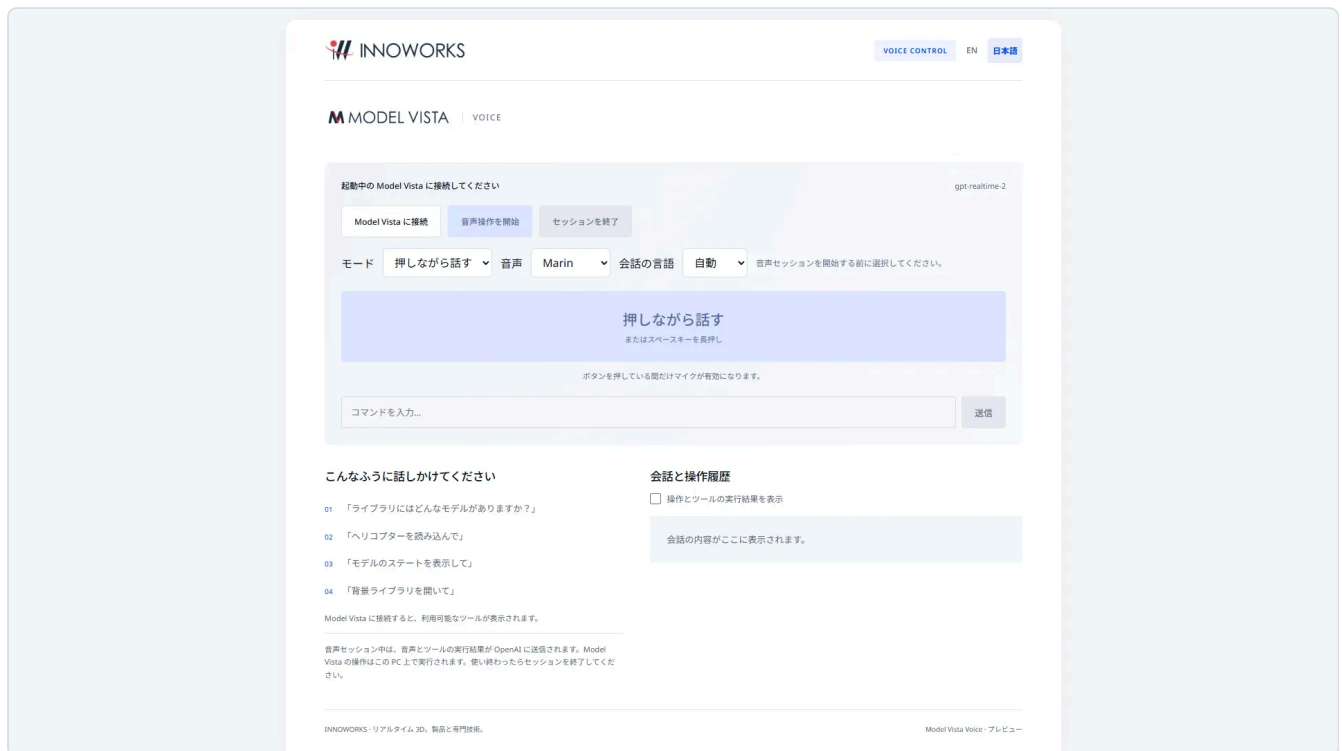
04 Model Vista を操作する

05 Voice Control のトラブルシューティング

Model Vista Voice Control

Model Vista Voice Control は、Windows 用の独立したコンパニオンアプリです。音声またはテキストのコマンドで、起動中の Model Vista のライブラリの確認、モデルの読み込みと選択、Model Vista ステートの変更、デスクトップカメラの移動、対応する設定の調整を行えます。

同じコンピューター上の **AI (MCP)** を介して Model Vista に接続し、ご自身の OpenAI API キーを使って音声とコマンドを処理します。専用のデスクトップウィンドウで動作し、Codex のタスクや ChatGPT の会話を用意する必要はありません。



接続前の Voice Control 画面です。スクリーンショットはアプリ共通の Web インターフェースを示しています。デスクトップ版のネイティブメニューとウィンドウ枠は撮影範囲に含まれません。

はじめに

1. アプリをインストールして Model Vista に接続する。
2. 音声モード、言語、自動切断を設定する。
3. モデル、カメラ、ウィンドウ、設定の操作を試す。
4. 結果を確認して問題を解決する。

接続の仕組み

接続	用途
Voice Control → OpenAI	API アカウントを使用して音声、テキストのリクエスト、ツールの実行結果を処理します。インターネット接続が必要です。
Voice Control → Model Vista	同じコンピューター上で対応する MCP コマンドを検出して実行します。既定のアドレスは <code>http://127.0.0.1:8750/mcp</code> です。

Model Vista で **AI (MCP)** を有効にしてください。このコンパニオンの利用に **Web Remote**、REST、Render Streaming を有効にする必要はありません。

利用できるコマンドは、接続先の Model Vista ビルド、有効な機能、現在のシーンによって異なります。コンパニオンは、すべての汎用 API 操作ではなく、対応する MCP ツールの一部を提供します。結果が不明な場合は、音声による完了の案内だけで判断せず、実際のシーンや操作結果を確認してください。

インストールと接続

準備するもの

- .NET Framework 4.x と Microsoft Edge WebView2 Runtime が利用できる Windows x64 コンピューター。
- 同じコンピューター上で起動する、AI (MCP) 対応の Model Vista プレビュービルド。新しいコマンドには、そのコマンドを含むビルドが必要です。
- 音声操作のマイクと音声出力、および OpenAI へのインターネット接続。
- API の課金設定が有効で、選択する音声モードのモデルにアクセスできる、ご自身の OpenAI API キー。ChatGPT のサブスクリプションから、このアプリで使用する API キーが提供されるわけではありません。

ローカルサービスの実行環境はインストーラーに含まれています。Node.js を別途インストールする必要はありません。

プレビュー版をインストールする

1. ダウンロードページから、プレビューインストーラー **ModelVistaVoiceControl.msi** をダウンロードします。
2. 実行中の Voice Control セッションを終了し、MSI を実行します。Windows のインストール確認が表示されたら承認します。
3. インストールを完了します。起動オプションを選択したままにしてアプリを開くか、スタートメニューの **INNOWORKS > Model Vista Voice Control** から起動します。

API キーを保存する

1. アプリメニューの **API key settings** (API キー設定) を開きます。
2. ご自身の OpenAI API キーを入力します。
3. **Save key & open** (キーを保存して開く) を選択します。

保存したキーは、現在の Windows ユーザー用に暗号化されます。インストーラーには含まれません。セッション中は音声、テキストのリクエスト、ツールの実行結果が OpenAI に送信されます。恒久的な API キーはページに埋め込まれず、ローカルサービスで保持されます。

保存したキーは、アプリの更新やアンインストール後も残ります。アプリに保存したキーを削除するには、アンインストール前に **Forget saved key** (保存したキーを削除) を選択します。ローカル音声サービスも停止します。環境変数にもキーを設定している場合は、必要に応じて別途削除してください。

Model Vista に接続する

1. **Model Vista** を起動します。
2. **Settings > External Connections** を開きます。 **AI (MCP)** パネルで **AI connection enabled** を有効にし、ポートを **8750** に設定します。
3. Voice Control を開き、 **Connect Model Vista** を選択します。利用可能なツール数が表示されることを確認します。
4. 音声セッションを開始する前に、 **Mode**、 **Voice**、 **Spoken language** を選択します。
5. **Start voice** を選択し、マイクへのアクセスを求められたら許可します。
6. 「ライブラリにはどんなモデルがありますか？」と尋ねます。

Model Vista を更新または再起動した後は、接続し直して新しい音声セッションを開始し、現在のツールを再検出してください。古いビルドでは利用できるコマンドが少ない場合があります。

別の MCP ポートを使う場合

既定の接続先は `http://127.0.0.1:8750/mcp` です。Model Vista で別のポートを使用する場合は、Windows のユーザー環境変数 `MODELVISTA_MCP_URL` に対応するローカルアドレスを設定し、Voice Control を完全に終了してから起動し直します。設定画面には接続先の編集機能はありません。このプレビューでは別のコンピューターへの接続は対象外です。

管理者は、`OPENAI_CHATBOT` またはフォールバックの `OPENAI_API_KEY` でキーを設定することもできます。通常は **API key settings** を使用してください。メッセージログやサポートへの報告にキーを含めないでください。

音声セッションを使う

モードを選ぶ

Start voice を選択する前に **Mode** を設定します。モード、音声、会話の言語を変更する場合は、現在のセッションを終了してください。

モード	話し方	このプレビューで使用するモデル
Hold to talk	話す間はボタンを押し続け、離すと送信されます。テキスト欄や他のコントロールにフォーカスがない場合は、Space を長押しすることもできます。応答を待ってから次のリクエストを行います。	既定では <code>gpt-realtime-2</code>
GPT-Live conversation	状態が Live microphone on になってから話します。入力を止めるには Mute microphone 、接続を終了するには End session を選択します。	<code>gpt-live-1</code> とコマンド処理用の <code>gpt-5.6-terra</code> バックエンド

どちらのモードも、利用可能な Model Vista コマンドを使用します。セッションの準備ができているときは、**Or type a command...** に入力して **Send** を選択することもできます。テキスト入力にも、有効な音声セッションと API アクセスが必要です。

GPT-Live はマイクを常時使用します。ミュートしても接続は終了しません。接続時間とバックエンドの使用量は API アカウントに課金されるため、使い終わったら **End session** を選択してください。

The screenshot displays the Model Vista Voice Control interface. At the top, there's a header with the INNOWORKS logo and language options (VOICE CONTROL, EN, 日本語). Below this, the 'MODEL VISTA | VOICE' section is visible. The main area is titled '起動中の Model Vista に接続してください' (Connect to the running Model Vista). It features a 'Model Vista に接続' button, a '音声操作を開始' button, and a 'セッションを終了' button. The 'モード' (Mode) is set to 'GPT-Live 会話', '音声' (Voice) is 'Marin', and '会話の言語' (Conversation language) is '自動'. There's an '自動切斷' (Auto disconnect) dropdown set to '5分後' (5 minutes later). A 'マイクをミュート' (Mute microphone) button is also present. A text input field 'コマンドを入力...' (Enter command...) with a '送信' (Send) button is at the bottom. On the right, there's a '会話と操作履歴' (Conversation and operation history) section with a checkbox for '操作とツールの実行結果を表示' (Display execution results of operations and tools) and a note that conversation content will be displayed here. The footer includes 'INNOWORKS - リアルタイム 3D、製品と専門技術。' and 'Model Vista Voice - プレビュー'.

セッション開始前の GPT-Live 設定です。接続が完了するまで、マイクとセッションの操作は無効です。

音声と言語を選ぶ

Voice では Marin、Cedar、Alloy、Ash、Ballad、Coral、Echo、Sage、Shimmer、Verse を選択できます。既定は Marin です。選択は保存されます。

Spoken language では **Auto**、**English**、**Japanese** を選択できます。Auto は話した言語に合わせます。言語を指定すると、その言語を維持するようアシスタントに指示します。認識や応答の言語が意図せず変わる場合は、言語を固定してください。この設定はモデルへの指示であり、完全な認識を保証するものではありません。

ヘッダーまたはアプリメニューの **EN / 日本語** で表示言語を切り替えます。セッション中でも変更できます。表示言語は音声や会話の言語とは独立しており、モデル名、ステート名、既存の会話内容、未加工のツール結果は翻訳されません。インストーラーのダイアログは英語のままです。

操作がないときに自動切断する

GPT-Live モードでは、**Auto-disconnect** の既定値は **After 5 minutes** です。**After 2 minutes**、**After 5 minutes**、**After 10 minutes**、**Never** から選択できます。設定は保存されます。

認識されたユーザーの発話、またはテキストのコマンドが届くとタイマーがリセットされます。アシスタントの発話、ツール結果、マイクのミュートではリセットされません。最後の 30 秒間はカウントダウンが表示されます。発話、コマンドの送信、または **Stay connected** の選択でタイマーをリセットできます。

判定するのは音響的な無音ではなく、認識された操作です。周囲の会話が入力として認識されるとセッションが延長される場合があります。バックグラウンドでのタイマー制限によって切断が遅れる場合があります。これは利便性のための機能であり、支出の上限ではありません。Hold to talk モードには適用されません。

タイムアウトすると、音声セッションを終了してマイクと接続を解放します。実行中のツールがある場合は、ローカルの後処理まで最大 15 秒間待機します。Model Vista は動作を続け、会話の表示も残ります。

メッセージログを読む

Conversation & actions にリクエストとアシスタントの応答が表示されます。各項目にはローカル時刻が **HH:mm:ss** 形式で表示されます。ライブ字幕はテキストが追加されるたびに時刻が更新され、そのメッセージの最新の受信更新時刻を示します。時刻にマウスを合わせると、ローカルの日付と時刻を確認できます。

実際のコマンドと応答を確認するには、**Show actions and tool results** を有効にします。これらを非表示にしても変わるのは表示だけです。結果が不明な場合は、同じリクエストを繰り返す前に操作結果と Model Vista を確認してください。

セッションを終了する

End session を選択し、切断された状態になるまで待ちます。最終使用量を確認できなかったというメッセージが表示された場合、ローカル接続は終了していますが、サービスから最終的な使用量の確認を受け取れていません。

発話を中断したりセッションを終了したりしても、Model Vista に送信済みのコマンドは取り消されません。音声を再開すると新しいセッションになります。結果が不明な操作を続ける前に、現在のモデルやシーンの状態を確認してください。Voice Control を閉じると、ローカルサービスも停止します。

Model Vista を操作する

最初は一度に一つずつ依頼し、アプリが返した名前を使ってください。以下は発話の例であり、固定のコマンド構文ではありません。接続先の Model Vista ビルドから検出されたツールだけが使用できます。

短いデモを試す

1. 「ライブラリにはどんなモデルがありますか？」
2. INNOWORKS Heli が一覧にある場合は、「INNOWORKS Heli を読み込んで。」
3. 「INNOWORKS Heli を選択して。」
4. 「利用できる状態を見せて。」
5. 返された状態の選択肢から、グループ名も指定して一つを選びます。
6. 「INNOWORKS Heli を正面から見せて。」
7. 「背景ライブラリを開いて。」

各結果を待ち、Model Vista で確認します。読み込み済みのインスタンスや選択肢に同じ名前が複数ある場合は、返された ID を使うか、選択肢の一覧を求めてください。

利用できる操作

対象	リクエストの例	範囲
ライブラリ	「ライブラリのモデルを一覧にして。」 「INNOWORKS Heli を読み込んで。」「背景を一覧にして。」	選ぶ前に利用可能なコンテンツを確認します。
シーンのモデル	「読み込み済みモデルを一覧にして。」 「INNOWORKS Heli を選択して。」「INNOWORKS Heli を隠して。」「INNOWORKS Heli を表示して。」「INNOWORKS Heli をシーンから削除して。」	削除するのは読み込み済みのインスタンスです。ライブラリのファイルは残ります。非表示のインスタンスも一覧にできます。
ステートとアニメーション	「INNOWORKS Heli のステートグループと選択肢を一覧にして。」「そのグループの選択肢 2 を選んで。」「Fox のアニメーションは進んでいますか？」	ステートの選択と再生は別々に確認します。選択済みであることだけでは、見た目の動きを証明できません。
トランスフォーム	「INNOWORKS Heli の位置とトランスフォームのロック状態を見せて。」「INNOWORKS Heli を 1 メートル上に移動して。」「INNOWORKS Heli を Y 軸周りに 30 度回して。」「INNOWORKS Heli のスケールを 2 倍にして。」	既存のロックが適用されます。必要な場合は明示的にロック解除を依頼してください。
デスクトップカメラ	「INNOWORKS Heli が画面の中央に来るようにして。」「INNOWORKS Heli を左から見せて。」「カメラを前に 1 メートル移動して。」「カメラを右に 15 度回して。」「カメラの FOV を 45 度にして。」	実行中の非 XR カメラを操作します。標準ビューはモデルの軸を基準にします。移動は即時で、衝突回避は行いません。垂直 FOV は 15～90 度です。中央に収める操作は現在の視線方向を維持します。
ウィンドウ	「利用できるウィンドウを一覧にして。」「Camera ウィンドウを表示して。」「Camera ウィンドウを隠して。」「モデルライブラリを開いて。」「このウィンドウを 1 ページ下にスクロールして。」	表示中のスクロール領域が対象です。複数ある場合はウィンドウまたは領域を指定します。全体が収まっている内容はスクロールできません。
グラフィックス	「利用できるグラフィックス設定は？」「パストレーシングを有効にして。」「DLSS を無効にして。」	対応環境では VSync、コンタクトシャドウ、レイトレース AO、スクリーンスペース GI、再帰レンダリングも操作できます。ハードウェアとレンダリング設定による制限があります。
ターンテーブル	「INNOWORKS Heli をターンテーブルに割り当てて。」「ターンテーブルを開始して。」「一時停止して。」「逆回転して。」「角度を 90 度にして。」	Model Vista で該当機能を利用できる必要があります。角度と速度を確認、変更できます。
環境	「背景を [一覧にある名前] にして。」「太陽の高度を 15 度にして。」「フォグを有効にして。」「雲の不透明度を 0.5 にして。」	太陽の変更はディレクショナルライトに適用され、HDRI に焼き込まれた太陽は変わりません。雲の変更には、対応する非 HDRI の空が必要です。フォグの設定はレンダリング構成に依存します。地理的な日時の設定は含まれません。

対象	リクエストの例	範囲
XR	「XR は有効ですか?」「XR に入って。」「XR を終了して。」	Model Vista の実行環境と接続機器に依存します。XR 中はデスクトップカメラのコマンドを利用できません。

プレビューの制限

コンパニオンは現在、Model Vista ビルドがすべて提供している場合に 34 個の直接 MCP ツールに対応します。汎用フィールド編集、任意の関数呼び出し、スナップショット撮影は提供しません。接続後に利用可能なツール数が表示されます。

遷移が完了する前に操作が受理される場合があります。次の依存するリクエストを行う前に結果を確認してください。接続エラーやタイムアウトは、操作が取り消されたことを意味しません。削除、読み込み、トランスフォームを再試行する前にシーンを確認してください。

コンパニオンは Model Vista の機能の利用条件やライセンスを回避しません。利用できない機能をリクエストしても、その機能が有効になるわけではありません。

Voice Control のトラブルシューティング

Model Vista に接続できない

Cannot reach Model Vista. Start Model Vista and enable AI (MCP). が表示された場合は、このコンピューターで Model Vista を起動し、**Settings > External Connections** の **AI connection enabled** を有効にします。MCP ポートが 8750、または設定した接続先と一致していることを確認し、**Connect Model Vista** を再度選択します。

Web Remote を有効にするだけでは MCP は有効になりません。接続先を変更した場合は、Voice Control を完全に終了して起動し直します。[インストールと接続](#)を参照してください。

コマンドが見つからない

必要な MCP ツールを含む Model Vista プレビュービルドに更新し、Voice Control を再接続して新しい音声セッションを開始します。コンパニオンだけを更新しても、古い Model Vista ビルドに API が追加されるわけではありません。

接続後のツール数を確認します。操作を選ぶ前に、対象のモデル、ステート、ウィンドウ、対応する設定の一覧を求めてください。現在のレンダリング設定、XR モード、ライセンスによって機能を利用できない場合もあります。

音声を開始できない

- **API key settings** と、API アカウントが選択したモードのモデルにアクセスできるかを確認します。
- インターネット接続、Windows のマイクのアクセス許可、選択したマイクを確認します。
- マイクのアクセスを求められたら許可します。応答が聞こえない場合は、スピーカーやヘッドホンを確認します。
- 開始がタイムアウトしたり接続が切れたりした場合は、原因を解消してから新しいセッションを開始します。

インストール版のウィンドウには Microsoft Edge WebView2 Runtime が必要です。不足しているというエラーが表示された場合は、ランタイムをインストールしてアプリを開き直してください。

会話が別の言語になる

セッションを終了し、**Spoken language** で **English** または **Japanese** を選択して再開します。**EN / 日本語** の切り替えで変わるのは表示言語だけです。短く明瞭に依頼し、GPT-Live の使用中は周囲の会話を減らしてください。

成功したと言われたのに変化がない

Show actions and tool results を有効にして、該當時刻の結果と実際のシーンを確認します。操作が発行されていない場合は、操作を明示して再度依頼してください。操作がタイムアウトした場合は、再試行前に現在の状態を確認します。

モデル名が重複している場合は、読み込み済みインスタンスの ID を確認します。ステートの選択肢名が重複している場合は、グループと選択肢 ID の両方を指定します。モデルを読み込み直すと ID が変わる場合があるため、ステートグループを再度一覧にしてください。

アニメーションの場合は、再生が進んでいるかを確認します。選択済みのアニメーションオプションが、見た目の動きを証明するわけではありません。カメラを移動する場合は、先に XR を終了してください。スクロールの場合は、対象ウィンドウが表示され、内容が表示領域より大きいことを確認します。

操作していないのに接続が続く

Mode が **GPT-Live conversation** で、**Auto-disconnect** が **Never** になっていないことを確認します。周囲の会話が認識されるとタイマーがリセットされる場合があります。ミュートではタイマーは停止しませんが、バックグラウンドでの実行制限で処理が遅れる場合があります。すぐに切断したい場合は **End session** を選択してください。

サポートへの報告内容

Voice Control と Model Vista のバージョン、音声モードと言語、リクエスト、タイムスタンプ、該当する操作結果またはエラーを記録します。必要に応じて、表示中のウィンドウや選択中のモデルも記載してください。API キーは含めないでください。共有する前に、非公開のモデル名、パス、その他の機密プロジェクト情報がログに含まれていないか確認してください。