



ENGLISH DOCUMENTATION

Model Vista Voice Control

Use voice or typed commands to control a local Model Vista session.

30 October 2026

Contents

01 **Model Vista Voice Control**

02 **Install and connect**

03 **Use voice sessions**

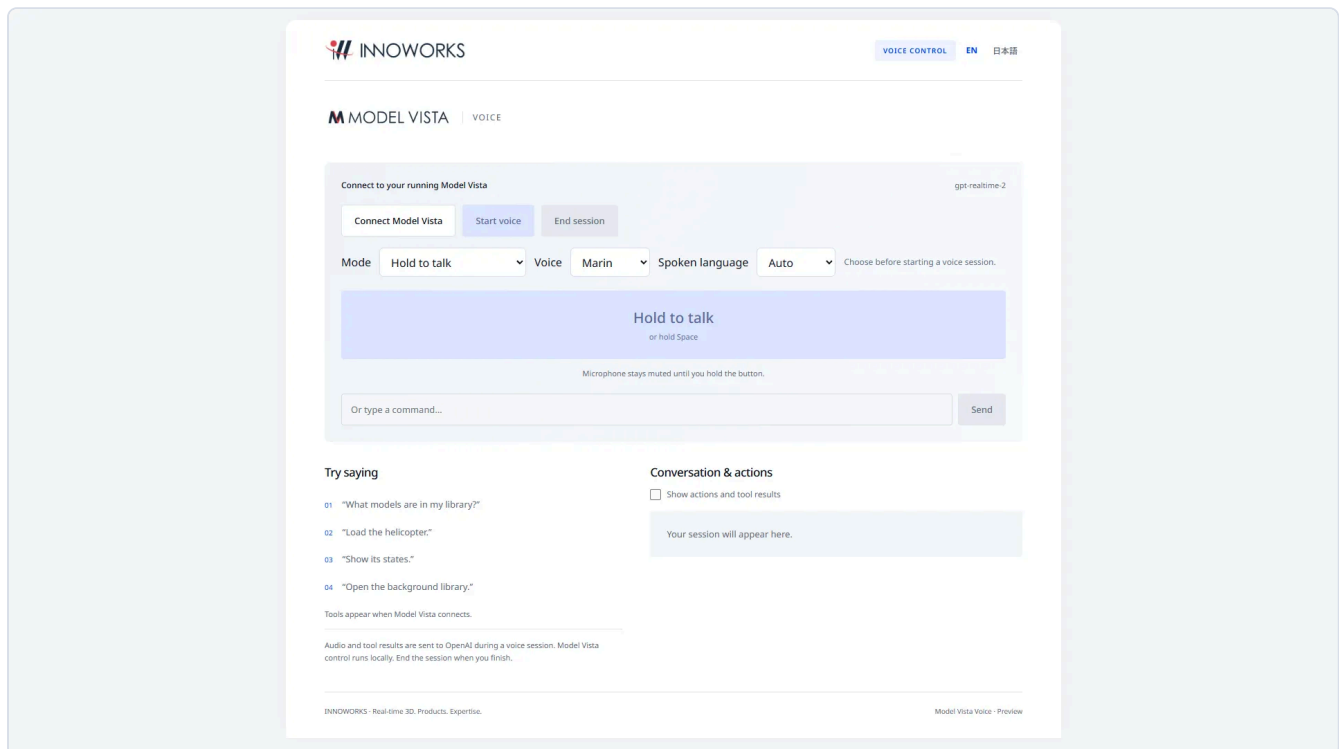
04 **Control Model Vista**

05 **Troubleshoot Voice Control**

Model Vista Voice Control

Model Vista Voice Control is a separate Windows companion app. Speak or type a command to inspect your library, load and select models, change Model Vista States, move the desktop camera, or adjust supported settings in a running Model Vista application.

The companion connects to Model Vista through **AI (MCP)** on the same computer and uses your own OpenAI API key for voice and command interpretation. It runs in its own desktop window. You do not need a Codex task or a ChatGPT conversation to use it.



The Voice Control interface before connecting. Screenshots show the app's shared web interface; the native desktop menu and window frame are outside the capture.

Start here

1. Install the app and connect to Model Vista.
2. Choose a voice mode, language and automatic disconnect setting.
3. Try commands for models, camera, windows and settings.
4. Inspect results and resolve problems.

How it connects

Connection	Purpose
Voice Control → OpenAI	Processes audio, typed requests and tool results using your API account. Requires internet access.
Voice Control → Model Vista	Discovers and runs the supported MCP commands on your computer. The default address is <code>http://127.0.0.1:8750/mcp</code> .

Enable **AI (MCP)** in Model Vista. **Web Remote**, REST and Render Streaming do not need to be enabled for this companion.

Available commands depend on the connected Model Vista build, its enabled features and the current scene. The companion exposes a supported subset of MCP tools, rather than every generic API operation. A spoken confirmation should be checked against the actual scene or action result when the outcome is uncertain.

Install and connect

Before you start

- A Windows x64 computer with .NET Framework 4.x and Microsoft Edge WebView2 Runtime.
- A Model Vista preview build with AI (MCP) support, running on the same computer. Newer commands require a build containing those commands.
- A microphone and audio output for spoken interaction, plus internet access to OpenAI.
- Your own OpenAI API key with API billing and access to the models used by your chosen voice mode. A ChatGPT subscription does not supply the API key used by this app.

The installer bundles the local service runtime; you do not need to install Node.js separately.

Install the preview

1. Download the **ModelVistaVoiceControl.msi** preview installer from the downloads page.
2. Close an existing Voice Control session and run the MSI. Accept the Windows installation prompt when requested.
3. Complete the installer. Leave its launch option selected to open the app, or open **INNOWORKS > Model Vista Voice Control** from the Start menu.

Save your API key

1. Open **API key settings** in the app menu.
2. Enter your own OpenAI API key.
3. Select **Save key & open**.

The saved key is encrypted for the current Windows user. It is not included in the installer. Audio, typed requests and tool results are sent to OpenAI during a session; the permanent key is kept by the local service rather than embedded in the page.

Saved keys survive upgrades and uninstalling the app. To remove the app's saved key, select **Forget saved key** before uninstalling. This also stops the local voice service. If you separately configured an environment-variable key, remove that separately when required.

Connect to Model Vista

1. Start **Model Vista**.

2. Open **Settings > External Connections**. In the **AI (MCP)** panel, enable **AI connection enabled** and use port **8750**.
3. Open Voice Control and select **Connect Model Vista**. Check that the available tool count appears.
4. Choose **Mode**, **Voice** and **Spoken language** before starting a voice session.
5. Select **Start voice**, then allow microphone access if prompted.
6. Ask, "What models are in my library?"

Reconnect and start a new voice session after updating or restarting Model Vista so the companion discovers the current tools. An older build can expose fewer commands.

A different MCP port

The default endpoint is `http://127.0.0.1:8750/mcp`. If Model Vista uses another port, set the Windows user environment variable `MODELVISTA_MCP_URL` to the matching local address, then fully close and reopen Voice Control. The settings page does not provide an endpoint editor. Remote computers are outside this preview's connection scope.

Administrators can also supply a key through `OPENAI_CHATBOT` or the fallback `OPENAI_API_KEY`. Most users should use **API key settings**. Never include a key in a message log or support report.

Use voice sessions

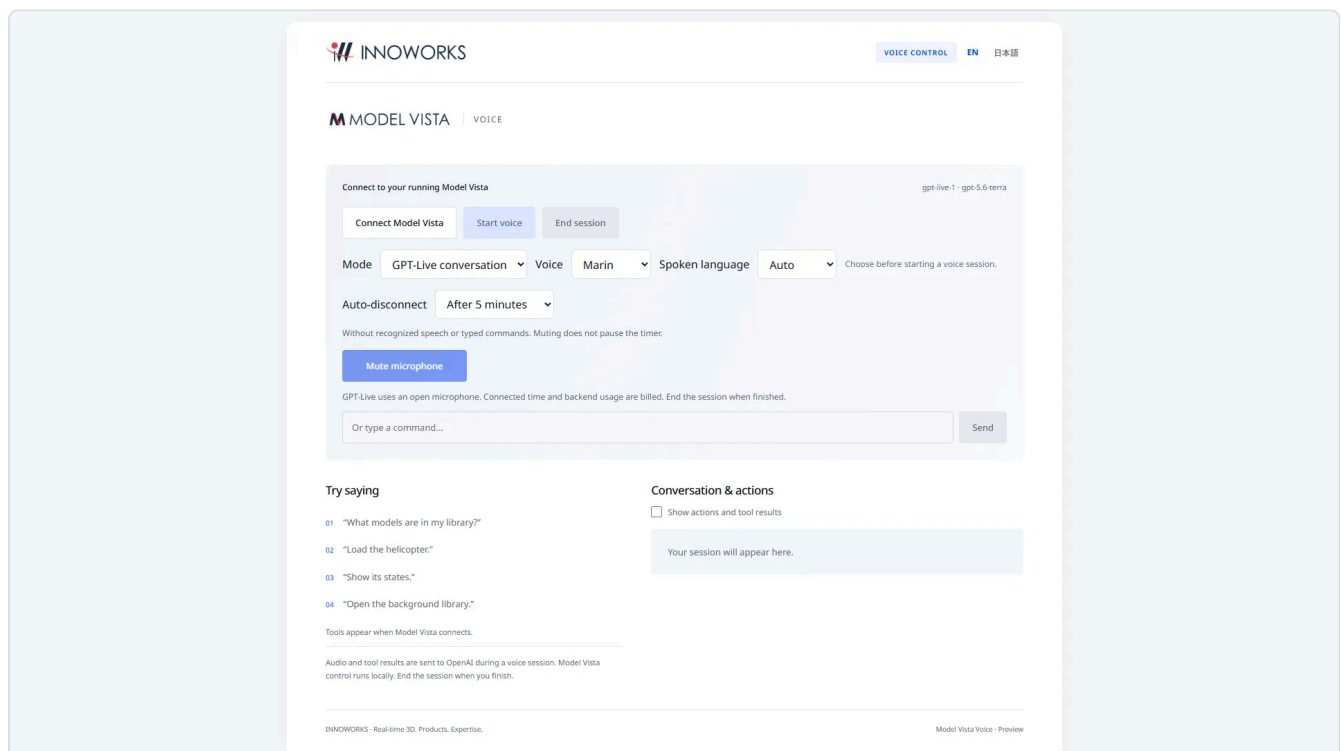
Choose a mode

Choose **Mode** before selecting **Start voice**. End the current session before changing mode, voice or spoken language.

Mode	How to speak	Models used by this preview
Hold to talk	Hold the talk button while speaking, then release to send. You can also hold Space when focus is outside text fields and other controls. Wait for the response before the next request.	<code>gpt-realtime-2</code> by default
GPT-Live conversation	Speak after the status shows Live microphone on . Use Mute microphone to pause microphone input, or End session to disconnect.	<code>gpt-live-1</code> with a <code>gpt-5.6-terra</code> command-processing backend

Both modes use the available Model Vista commands. You can also type into **Or type a command...** and select **Send** while a session is ready. Typed input still requires an active voice session and API access.

GPT-Live uses an open microphone. Muting does not end the connection. Connected time and backend usage are billed to your API account; select **End session** when finished.



GPT-Live settings before starting a session. The microphone and session controls remain inactive until connected.

Choose voice and languages

Voice offers Marin, Cedar, Alloy, Ash, Ballad, Coral, Echo, Sage, Shimmer and Verse. Marin is the default. The app remembers your choice.

Spoken language offers **Auto**, **English** and **Japanese**. Auto follows your speech; an explicit choice asks the assistant to stay in that language. Choose a fixed language if recognition or replies switch unexpectedly. This setting guides the model rather than guaranteeing perfect recognition.

Use **EN / 日本語** in the header or app menu to change the interface language. You can change this during a session. UI language is independent of voice and spoken language; model names, state names, existing conversation text and raw tool results are not translated. Installer dialogs remain English.

Disconnect automatically when idle

In GPT-Live mode, **Auto-disconnect** defaults to **After 5 minutes**. Choose **After 2 minutes**, **After 5 minutes**, **After 10 minutes** or **Never**. The app remembers the setting.

The timer resets when recognized user speech or a typed command arrives. Assistant speech, tool results and microphone muting do not reset it. A countdown appears during the final 30 seconds. Speak, send a command, or select **Stay connected** to reset the timer.

This measures recognized activity, not acoustic silence. Background speech recognized as input can extend a session. Background timer throttling can also delay disconnection. It is a convenience feature, not a spending limit, and does not apply to Hold to talk mode.

When the timeout expires, the app ends the voice session and releases its microphone and connection. It allows an outstanding tool request up to 15 seconds to finish before local cleanup. Model Vista stays running and the conversation remains visible.

Read the message log

Conversation & actions shows your requests and the assistant's replies. Each entry displays local time as **HH:mm:ss**. For live captions, the timestamp updates as more text arrives, so it represents the latest received update to that message. Hover over a timestamp to see the full local date and time.

Enable **Show actions and tool results** to inspect the actual commands and responses. Hiding these entries only changes their visibility. For an uncertain result, inspect the action and Model Vista before repeating the request.

End a session

Select **End session** and wait for the disconnected status. If the app reports that final usage was not confirmed, the local connection has closed but final accounting was not confirmed by the service.

Interrupting speech or ending a session does not undo a command already sent to Model Vista. Starting voice again creates a fresh session; ask for the current model or scene state before continuing an uncertain operation. Closing Voice Control also stops its local service.

Control Model Vista

Start with one request at a time and use the names returned by the app. These are example phrases, not fixed command syntax. Only tools discovered from the connected Model Vista build are available.

Try a short demonstration

1. "What models are in my library?"
2. If INNOWORKS Heli is listed: "Load INNOWORKS Heli."
3. "Select INNOWORKS Heli."
4. "Show me its available states."
5. Ask for one of the returned state options, naming its group.
6. "Show the front of INNOWORKS Heli."
7. "Open the background library."

Wait for each result and check Model Vista. If several loaded instances or options share a name, use the returned ID or ask the assistant to list the choices.

Available controls

Area	Example requests	Scope
Libraries	"List library models." "Load INNOWORKS Heli." "List backgrounds."	Discover available content before choosing it.
Scene models	"List loaded models." "Select INNOWORKS Heli." "Hide INNOWORKS Heli." "Show INNOWORKS Heli." "Delete INNOWORKS Heli from the scene."	Deleting removes a loaded instance; it keeps the library file. Hidden instances can still be listed.
States and animation	"List INNOWORKS Heli's state groups and options." "Select option 2 in that group." "Is Fox's animation advancing?"	State selection and playback are separate checks. An option being selected does not prove visible animation.
Transform	"Show INNOWORKS Heli's position and transform locks." "Move INNOWORKS Heli one metre up." "Rotate INNOWORKS Heli by 30 degrees around Y." "Double INNOWORKS Heli's scale."	Existing transform locks apply. Explicitly request an unlock if needed.
Desktop camera	"Center the view on INNOWORKS Heli." "Show INNOWORKS Heli from the left." "Move the camera forward one metre." "Rotate the camera right by 15 degrees." "Set camera FOV to 45 degrees."	Controls the runtime non-XR camera. Standard views use model axes. Movement is immediate and does not avoid collisions. Vertical FOV is 15–90 degrees. Centering preserves the current viewing direction.
Windows	"List available windows." "Show the Camera window." "Hide the Camera window." "Open the model library." "Scroll this window down one page."	Scrolling targets visible scroll areas. Name the window or area if several are visible. Content that already fits cannot scroll.
Graphics	"What graphics options are available?" "Enable path tracing." "Disable DLSS."	Also supports VSync, contact shadows, ray-traced ambient occlusion, screen-space global illumination and recursive rendering where available. Hardware and rendering configuration apply.
Turntable	"Assign INNOWORKS Heli to the turntable." "Start the turntable." "Pause the turntable." "Reverse it." "Set the turntable angle to 90 degrees."	Requires the corresponding feature to be available in Model Vista. Angle and speed can be inspected and changed.
Environment	"Set the background to [listed name]." "Set sun elevation to 15 degrees." "Enable fog." "Set cloud opacity to 0.5."	Sun changes affect directional lighting, not a sun baked into an HDRI. Cloud changes require a supported non-HDRI sky. Fog options depend on rendering configuration. Geographic date/time controls are not included.

Area	Example requests	Scope
XR	"Is XR active?" "Enter XR." "Exit XR."	Depends on the Model Vista runtime and connected equipment. Desktop camera commands are unavailable while in XR.

Understand preview limits

The companion currently supports 34 direct MCP tools when the Model Vista build provides all of them. It does not expose generic field editing, arbitrary function calls or snapshot capture. The available tool count is shown after connecting.

An action may be accepted before a transition finishes. Verify the result before the next dependent request. A connection failure or timeout does not prove that an action was cancelled; inspect the scene before retrying a deletion, load or transform.

The companion does not bypass Model Vista feature availability or licensing. Naming an unavailable capability in a request does not enable it.

Troubleshoot Voice Control

Cannot reach Model Vista

For **Cannot reach Model Vista. Start Model Vista and enable AI (MCP).**, start Model Vista on this computer and enable **AI connection enabled** under **Settings > External Connections**. Check that the MCP port is 8750, or matches your configured endpoint. Select **Connect Model Vista** again.

Enabling Web Remote alone does not enable MCP. After changing a custom endpoint, fully close and reopen Voice Control. See [Install and connect](#).

A command is missing

Update Model Vista to a preview build containing the required MCP tools, then reconnect Voice Control and start a new voice session. Updating only the companion does not add APIs to an older Model Vista build.

Check the tool count after connecting. Ask the assistant to list the relevant models, states, windows or supported settings before choosing an operation. Features may also be unavailable because of the current rendering configuration, XR mode or license.

Voice will not start

- Check **API key settings** and API account access to the selected mode's models.
- Check internet access, Windows microphone permissions and the selected microphone.
- Allow microphone access when prompted. Check speakers or headphones if replies are silent.
- If startup times out or the connection is lost, start a new session after resolving the cause.

The installed window requires Microsoft Edge WebView2 Runtime. If the app reports that it is missing, install the runtime and reopen the app.

Speech uses the wrong language

End the session, choose **English** or **Japanese** under **Spoken language**, then start again. Changing **EN / 日本語** changes the interface only. Use short, clear requests and reduce background speech during GPT-Live sessions.

The assistant says an action succeeded, but nothing changed

Enable **Show actions and tool results** and inspect the result at the corresponding timestamp. Check the actual scene. If no action was issued, ask explicitly for the operation again. If an action timed out, inspect the current state before retrying.

For repeated model names, ask for loaded instance IDs. For repeated state option names, specify both the group and the option ID. After reloading a model, list its state groups again because IDs may have changed.

For animation, ask whether playback is advancing. A selected animation option does not prove visible motion. For camera movement, exit XR first. For scrolling, confirm that the intended window is visible and that its content exceeds the available space.

An idle session stays connected

Confirm that **Mode** is **GPT-Live conversation** and **Auto-disconnect** is not **Never**. Recognized background speech can reset the timer. Muting does not pause it, but background scheduling can delay the callback. Select **End session** when you want to disconnect immediately.

Information for a support report

Record the Voice Control and Model Vista versions, selected voice mode and language, the request, its timestamp and the relevant action/error result. Mention the visible window or selected model where relevant. Do not include API keys; review logs for private model names, paths or other sensitive project data before sharing them.