Voice Notepad

User Manual

Version 1.3.0

PLATFORMS

Windows **Linux** **Debian**

SUPPORTED AI PROVIDERS

OpenRouter **Google Gemini** **OpenAI** **Mistral AI**

Voice Notepad Settings

Record | History | Cost | Analysis | Models | Mic Te‹ ‹ ›

Provider: Open Router ▾   Model: Gemini 2.5 Flash ▾

Standard   Budget

▶ Prompt Controls

✎ Q2U

Ready   0:00

● Record   Pause   Transcribe   Delete

Clear   Save   Copy

Today: $0.07 ($13.3)

Voice Notepad Settings

Record | History | Cost | Analysis | Models | Mic Te‹ ‹ ›

Transcription History

Your transcription history is stored locally and can be exported as a CSV file.
Click the button below to open the history file in your default spreadsheet application.

Statistics

Total transcriptions: 62

Database size: 68.0 KB

Audio archive: 0 B (0 recordings)

Open History File   Refresh

History file: /home/daniel/.config/voice-notepad-v3/transcription_history.csv

Author: Daniel Rosehill
Repository: github.com/danielrosehill/Voice-Notepad
License: MIT

Documentation created with Claude Code

Table of Contents

Introduction

Voice Notepad is a desktop application for voice recording with AI-powered transcription and cleanup. Unlike traditional speech-to-text tools that require a separate text cleanup pass, Voice Notepad uses **multimodal AI models** that can process audio directly and perform both transcription and text cleanup in a single operation.

1.1 Key Features

- **One-shot transcription + cleanup:** Audio is sent with a cleanup prompt to multimodal models
- **Multiple AI providers:** OpenRouter (recommended), Gemini, OpenAI, and Mistral
- **Audio compression:** Automatic downsampling to reduce file size and API costs
- **Voice Activity Detection (VAD):** Optional silence removal before upload
- **Cost tracking:** Monitor your API spending in real-time
- **Transcript history:** All transcriptions are saved locally with metadata
- **Global hotkeys:** System-wide shortcuts that work even when minimized

Installation

2.1 Download Options

Voice Notepad is available in multiple formats:

Format	Platform	Description
`.exe`	Windows	Windows installer
`.AppImage`	Linux	Universal Linux package (runs anywhere)
`.deb`	Debian/Ubuntu	Native Debian package
`.tar.gz`	Linux	Portable archive

Download the latest release from: https://github.com/danielrosehill/Voice-Notepad/releases

2.2 System Requirements

- **Operating System:** Windows 10+ or Linux
- **Python:** 3.10+ (for running from source)
- **Audio:** Working microphone
- **Internet:** Required for API calls

2.3 Linux Dependencies

```
sudo apt install python3 python3-venv ffmpeg portaudio19-dev
```

Getting Started

3.1 Configure API Keys

Before using Voice Notepad, you need to set up at least one API key:

1. Click the **Settings** button in the top-right corner
2. Go to the **API Keys** tab
3. Enter your API key for your preferred provider

Recommended: Use **OpenRouter** for access to multiple models with a single API key and accurate cost tracking.

3.2 Select Your Provider and Model

On the Record tab:
1. Choose your **Provider** from the dropdown
2. Select your preferred **Model**
3. Choose between **Standard** (balanced) or **Budget** (cost-effective) tiers

3.3 Start Recording

1. Click **Record** or press `Ctrl+R`
2. Speak into your microphone
3. Click **Transcribe** or press `Ctrl+Return` when finished
4. Your cleaned transcription will appear in the text area

Main Interface

Voice Notepad uses a tabbed interface with seven main sections.
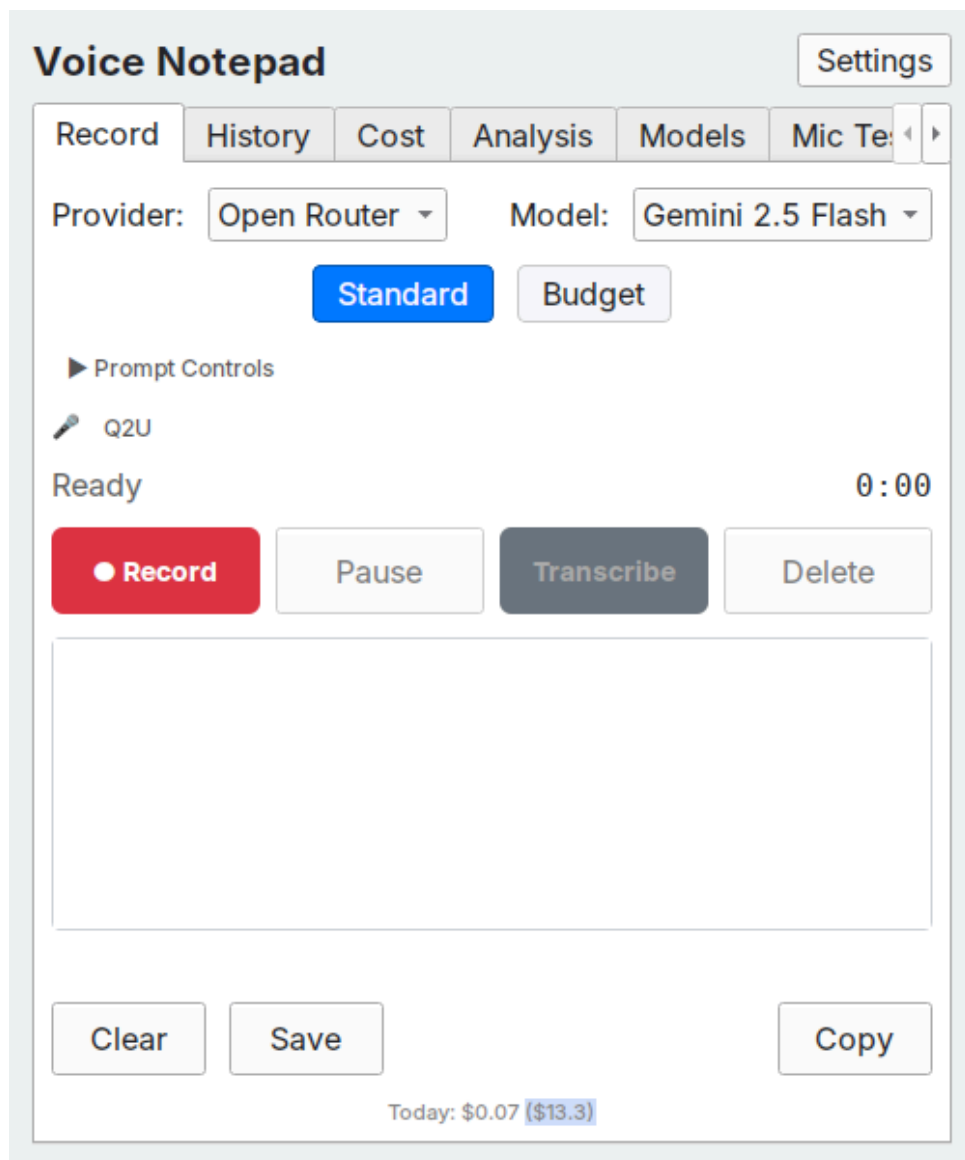
4.1 Record Tab

Figure 1: The Record tab - main transcription interface

The Record tab is where you perform transcriptions.

Controls:

- **Provider:** Select your AI provider (**OpenRouter** **Gemini** **OpenAI** **Mistral**)
- **Model:** Choose the specific model for transcription
- **Standard/Budget:** Toggle between quality tiers
- **Prompt Controls:** Expand to customize the cleanup prompt
- **Microphone selector:** Choose your input device

Recording Buttons:
- **Record:** Start a new recording (red button)
- **Pause:** Pause the current recording
- **Transcribe:** Stop and send to AI for transcription
- **Delete:** Discard the current recording

Output Area:
- View your transcribed text with markdown rendering
- Toggle between rendered view and source editing
- Word and character count displayed

Actions:
- **Clear:** Clear the text area
- **Save:** Save transcription to a file
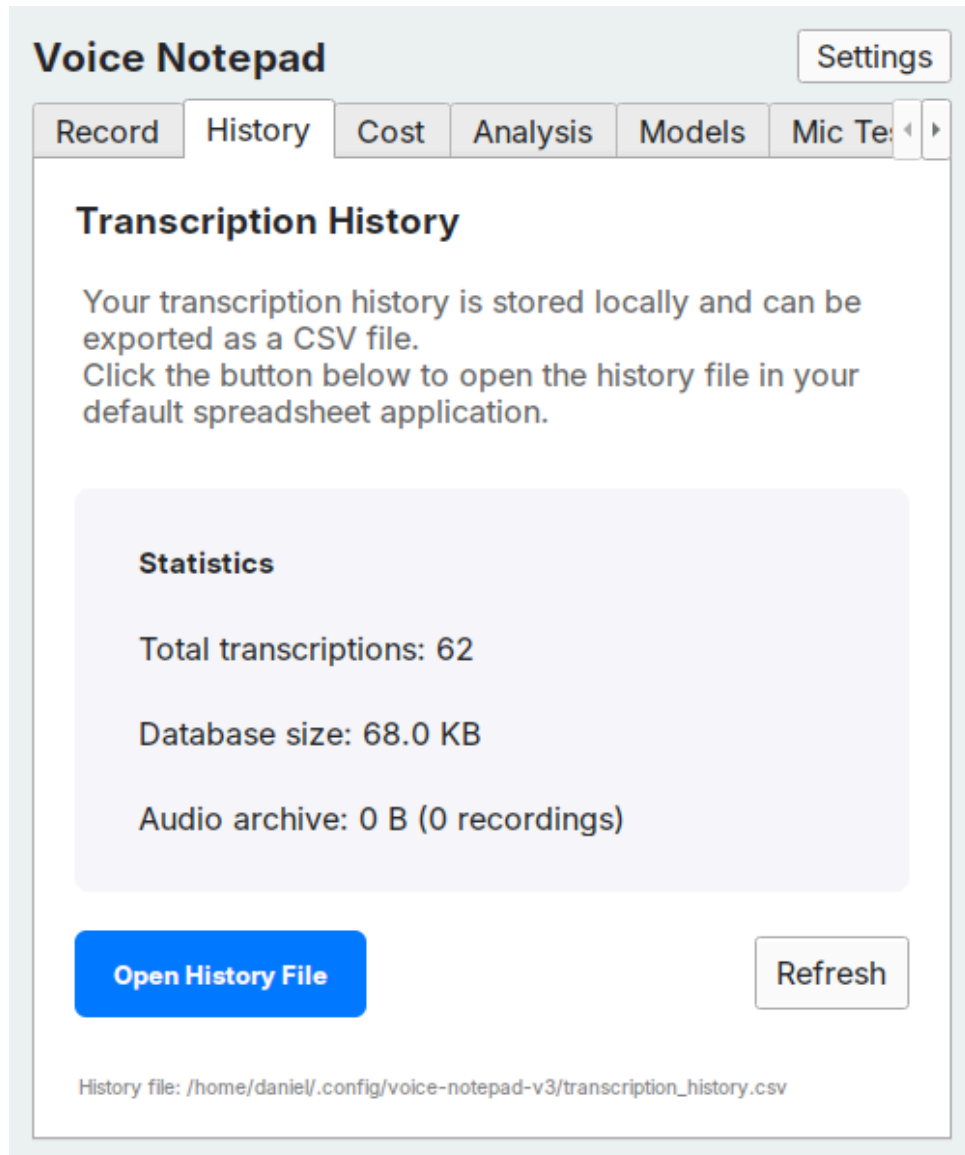- **Copy:** Copy to clipboard

4.2 History Tab

Figure 2: The History tab - browse past transcriptions

The History tab provides access to all your past transcriptions.

Statistics:
- **Total transcriptions:** Number of recordings processed
- **Database size:** Storage used by transcript history
- **Audio archive:** Size of archived audio files (if enabled)

Features:
- Click **Open History File** to export history as CSV
- **Refresh** to update statistics
- History file location shown at bottom
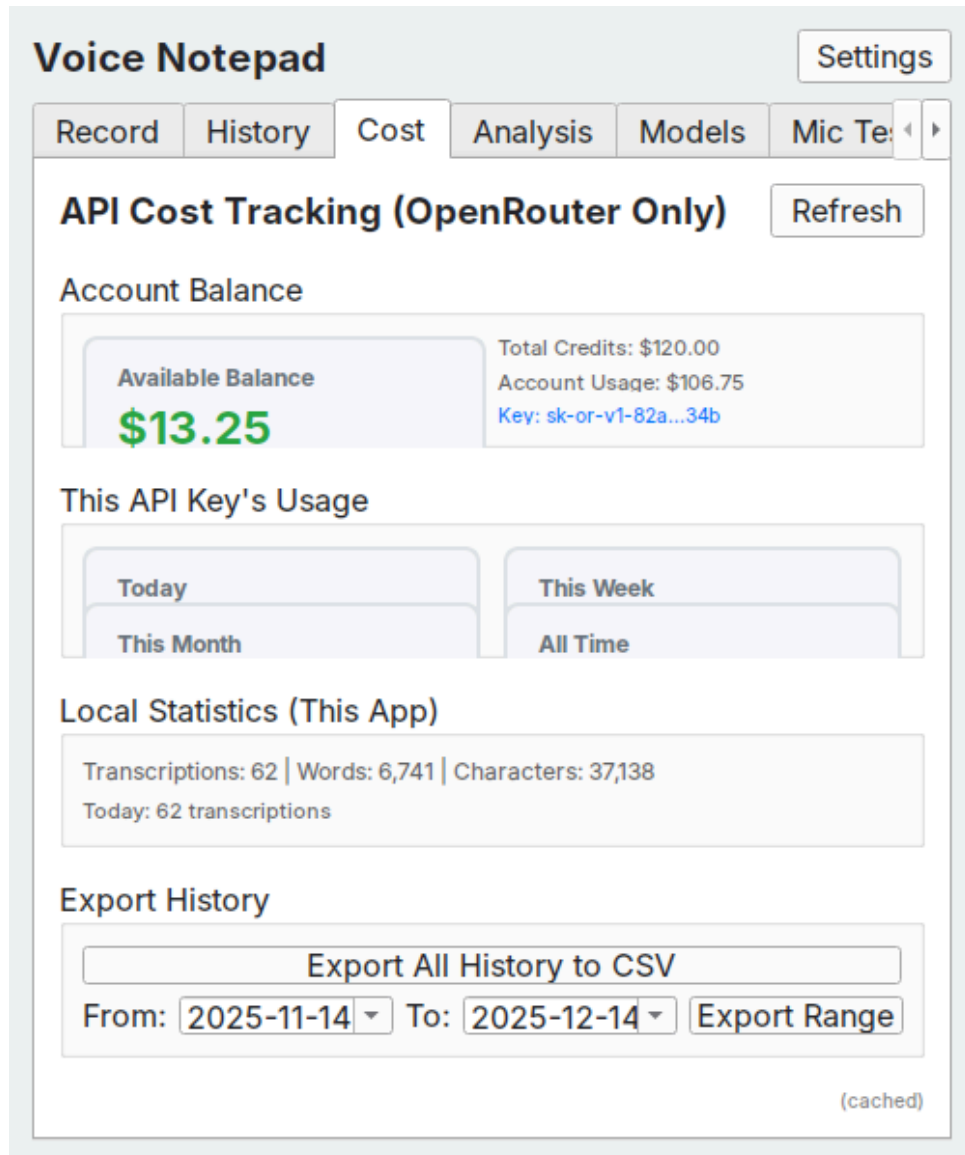
4.3 Cost Tab

Figure 3: The Cost tab - API usage tracking

The Cost tab provides detailed API usage tracking.

Account Balance (OpenRouter only):
- Shows your available credit balance
- Displays total credits and usage

This API Key's Usage:
- **Today:** Current day spending
- **This Week:** Weekly spending total
- **This Month:** Monthly spending total
- **All Time:** Lifetime spending

Local Statistics:
- Transcription count

- Total words and characters processed

Export History:
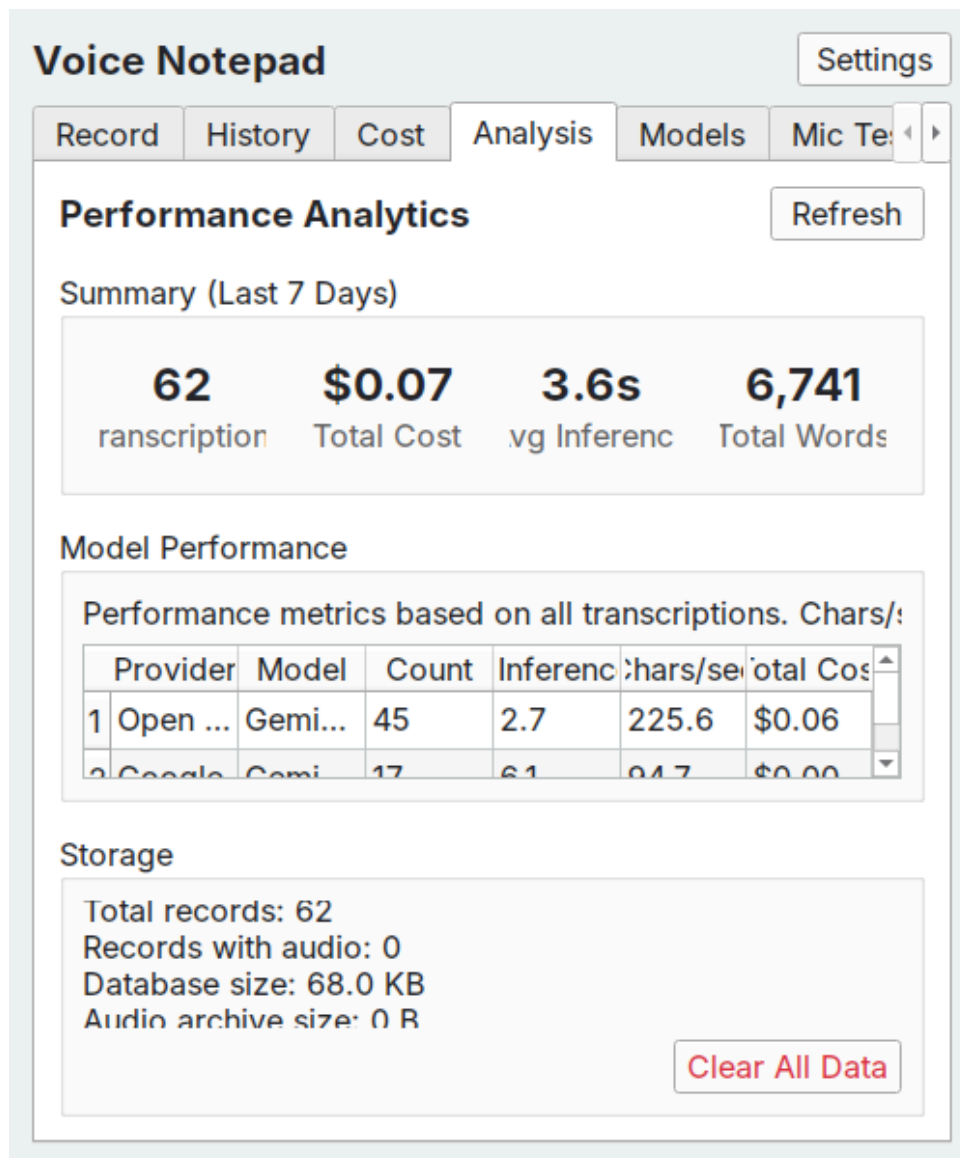- Export all history or a date range to CSV

4.4 Analysis Tab

Figure 4: The Analysis tab - performance metrics

The Analysis tab shows performance metrics.

Summary (Last 7 Days):
- Total transcriptions
- Total cost
- Average inference time
- Total words transcribed

Model Performance: Comparison table showing provider, model, usage count, average inference time, characters per second, and total cost.

Storage:
- Total database records
- Records with audio archived
- Database and archive sizes
- **Clear All Data** button for cleanup
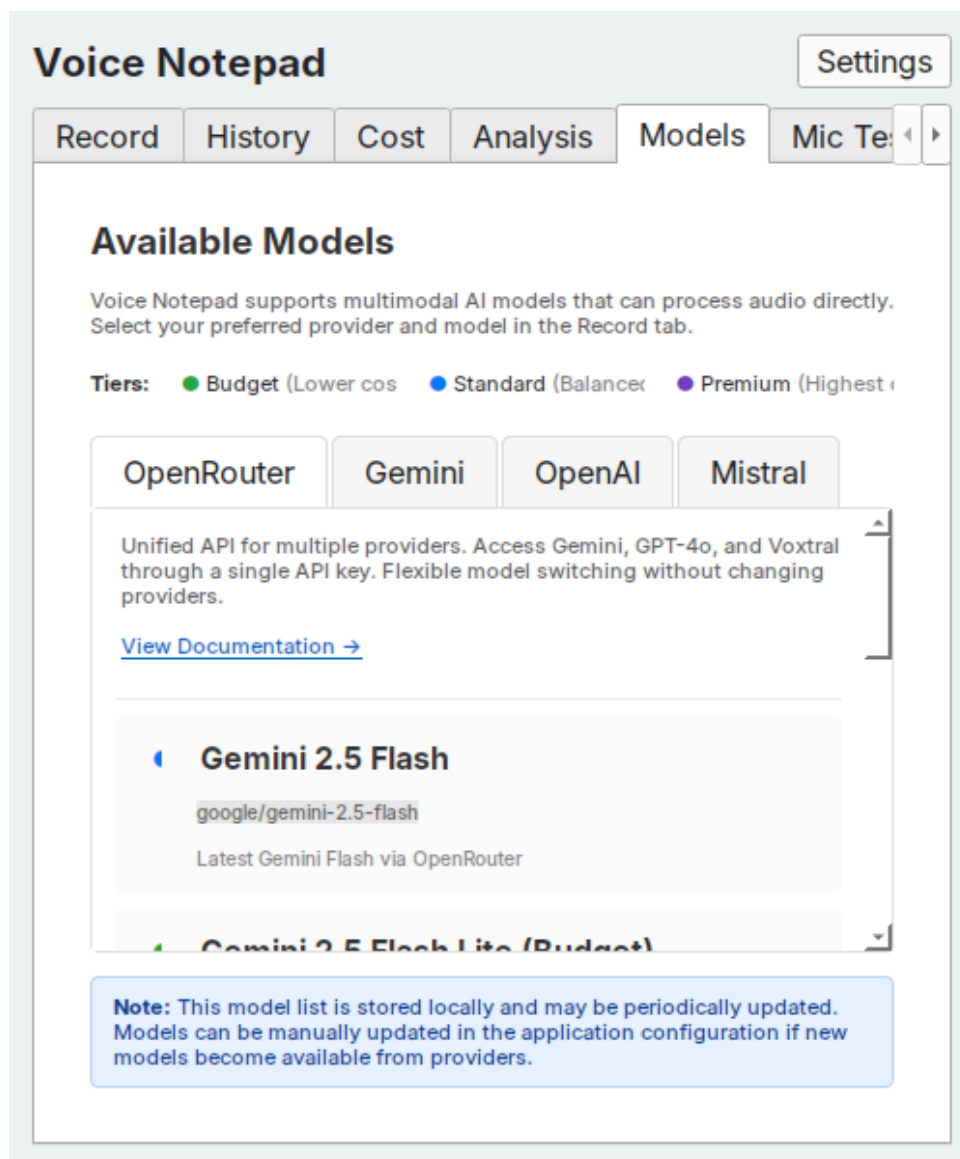
4.5 Models Tab

Figure 5: The Models tab - available AI models

The Models tab shows all available AI models organized by provider.

Model Tiers:
- **Budget** Lower cost options

- **Standard** Balanced cost/quality
- **Premium** Highest quality models

Provider Tabs:
- **OpenRouter** Access Gemini, GPT-4o, and Voxtral through unified API
- **Gemini** Direct Google AI access
- **OpenAI** Direct OpenAI access
- **Mistral** Direct Mistral/Voxtral access
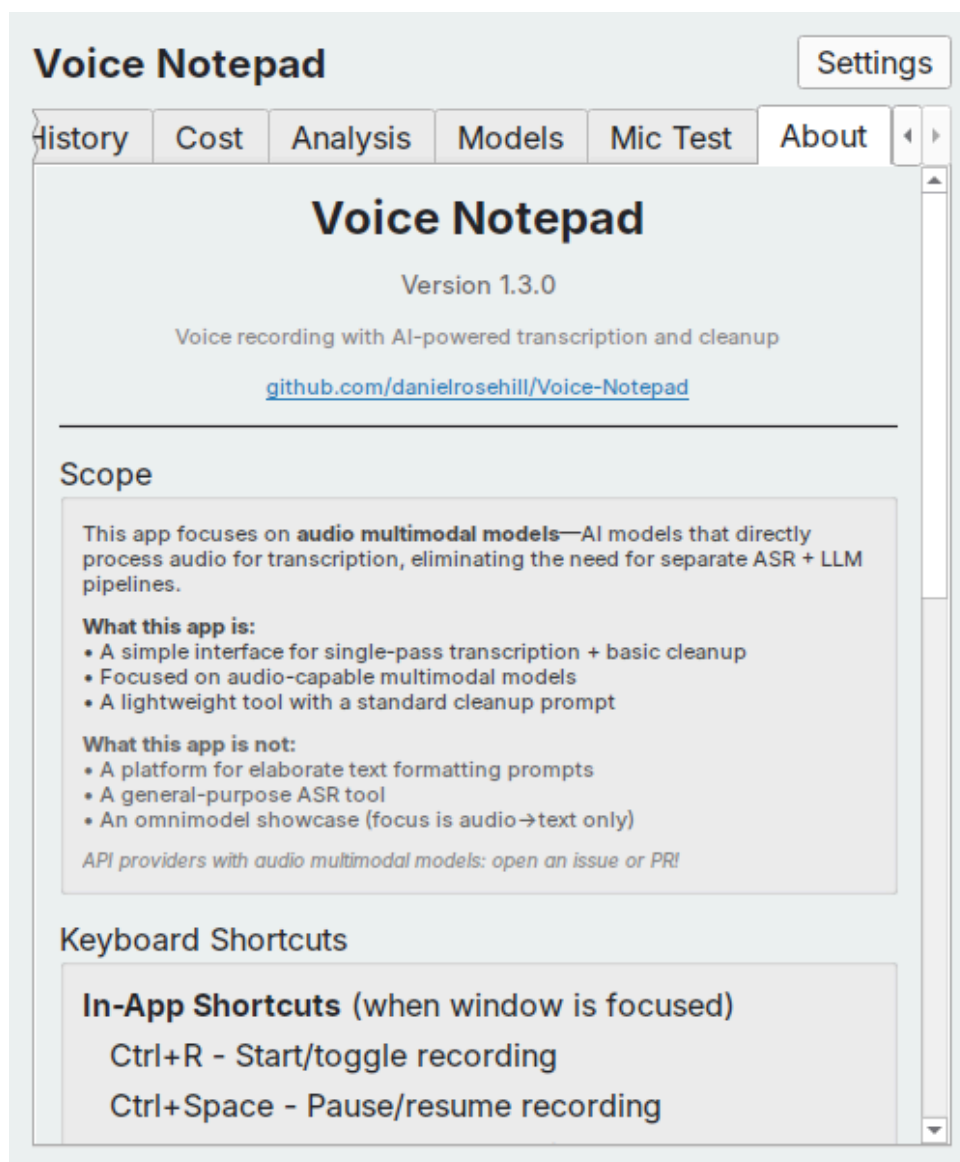
4.6 About Tab

Figure 6: The About tab - app information and shortcuts

The About tab provides app information, scope description, and keyboard shortcuts reference.

Settings

Access settings by clicking the **Settings** button in the top-right corner.
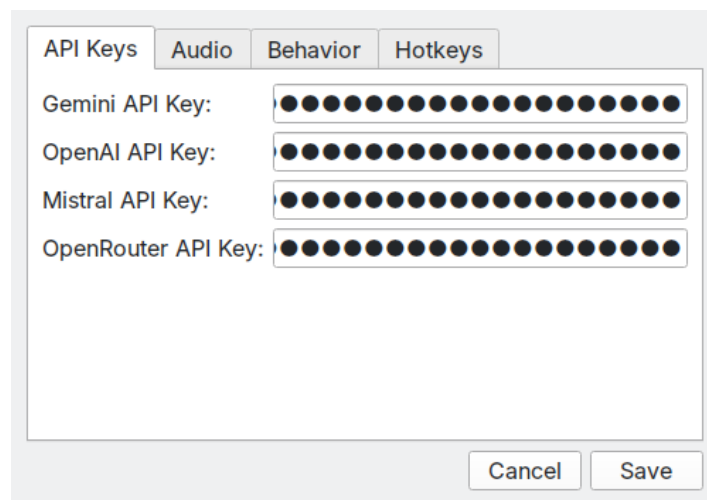
5.1 API Keys

Figure 7: API Keys configuration

Configure your API keys for each provider:

- **Gemini** **API Key:** For direct Google AI access
- **OpenAI** **API Key:** For direct OpenAI access
- **Mistral** **API Key:** For direct Mistral access
- **OpenRouter** **API Key:** For unified multi-provider access (recommended)

API keys are stored securely in your local configuration.

5.2 Audio Settings

Configure audio-related options:

- **Default microphone:** Select your preferred input device
- **Sample rate:** Audio quality settings
- **Audio compression:** Enable/disable compression before upload

5.3 Behavior Settings

Customize app behavior:
- **Enable VAD:** Voice Activity Detection removes silence
- **Auto-copy to clipboard:** Automatically copy transcriptions
- **Archive audio:** Save Opus copies of recordings
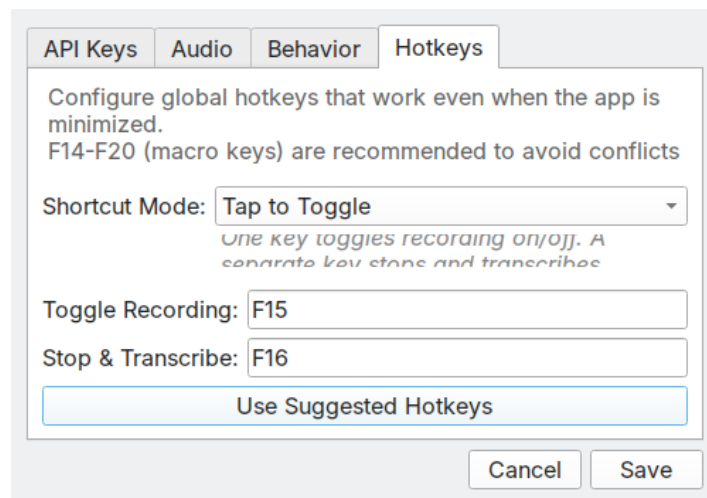- **System tray:** Minimize to tray on close

5.4 Hotkeys

Figure 8: Global hotkeys configuration

Configure global hotkeys that work even when the app is minimized.

Shortcut Modes:

Mode	Description
Tap to Toggle	One key toggles recording on/off
Separate Start/Stop	Different keys for each action
Push-to-Talk	Hold to record, release to stop

Recommended Keys: F14-F20 (macro keys) to avoid conflicts with other applications.

Keyboard Shortcuts

6.1 In-App Shortcuts

These work when the Voice Notepad window is focused:

Shortcut	Action
Ctrl+R	Start recording
Ctrl+Space	Pause/Resume recording
Ctrl+Return	Stop and transcribe
Ctrl+S	Save to file
Ctrl+Shift+C	Copy to clipboard
Ctrl+N	New note (clear)

6.2 Global Hotkeys

These work system-wide, even when the app is minimized or unfocused. Configure them in Settings > Hotkeys.

Default Configuration:

Key	Action
F14	Start recording
F15	Stop (discard)
F16	Stop & transcribe

Troubleshooting

7.1 No Audio Detected

1. Check that your microphone is connected
2. Use the Mic Test tab to verify audio input
3. Check system audio settings
4. Ensure the correct microphone is selected in Settings

7.2 Transcription Fails

1. Verify your API key is correct in Settings
2. Check your internet connection
3. Ensure you have credits/quota with your provider
4. Try a different provider or model

7.3 High API Costs

1. Enable VAD (Voice Activity Detection) to remove silence
2. Use Budget tier models when appropriate
3. Keep recordings concise
4. Monitor spending in the Cost tab

7.4 Global Hotkeys Not Working

1. Check hotkey configuration in Settings
2. Try different key combinations
3. On Wayland, ensure XWayland compatibility
4. Avoid keys that conflict with other applications

7.5 App Won't Start

1. Check system dependencies are installed
2. Verify Python version (3.10+)
3. Try running from terminal to see error messages
4. Reinstall the application

Support

- **GitHub Issues:** https://github.com/danielrosehill/Voice-Notepad/issues
- **Documentation:** https://github.com/danielrosehill/Voice-Notepad/tree/main/docs

---

Voice Notepad v1.3.0
Created by Daniel Rosehill & Claude Code
MIT License