

HAIA (Ver. 0.21) ユーザーズマニュアル

1. 概要

HAIA (Human-AI Assistant) は、複数の大規模言語モデル (LLM) を統一されたインターフェースで利用できるデスクトップアプリケーションです。テキスト生成に加えて、Google Gemini の音声生成モデルを活用した「自動配役 (Auto-Casting)」機能により、洗練された読み上げ機能を実現しています。

主な特徴

- マルチモデル対応: Google Gemini、OpenAI、Anthropic、OpenRouter、ローカル LLM (Ollama、LM Studio) に対応しています。
- プロンプト管理: 頻繁に使用する指示 (システムプロンプト) を GUI 上で作成、編集、保存することができます。
- 高度な音声合成: テキストの内容を解析し、登場人物ごとに声を自動的に割り当てて再生、保存します (ドラマ CD 形式の出力が可能です)。
- 文書処理: Word、PDF、Markdown 等の読み込みと、Word、PowerPoint 等への書き出しに対応しています (Pandoc 連携)。

注意事項

本ソフトウェアは OSS ライセンスに基づき「現状有姿 (AS IS)」で提供されます。

詳細なライセンス条件および免責については、同梱の LICENSE ファイルを参照してください。

本ソフトウェアには英語版の免責事項 (DISCLAIMER.md) が付属します。

海外環境での利用、または第三者提供を行う場合は、そちらも必ず参照してください。

2. 起動と初期設定

2.1 起動方法

配布パッケージ内の launch_haia.cmd または haia.exe を実行してください。

- 重要: 本バージョン (0.21) より、tasks.json が存在するフォルダが自動的に「データフォルダ (ルート)」として認識されます。

2.2 API キーの設定

クラウド AI を利用するには、各サービスの API キーが必要です。

- メニューバーの [設定] をクリックします。

2. [Gemini Key 登録…]などを選択します（各プロバイダごとに項目が用意されています）。
3. 入力ダイアログが表示されますので、API キーを貼り付けて「OK」を押してください。
 - キーは OS のセキュアな領域（Windows 資格情報マネージャー）に暗号化して保存されます。
 - 音声機能を利用する場合は、必ず Gemini API Key の登録が必要です。

2.3 外部ツールの連携（推奨）

以下のツールがシステムにインストールされている場合、機能が拡張されます。

- FFmpeg: 音声保存時に MP3 形式が利用可能になります（未導入の場合は WAV 形式のみとなります）。
 - Pandoc: Word、PowerPoint などのオフィス文書形式での入出力が可能になります。
-

3. 基本操作：テキスト生成タブ

メイン画面の [テキスト生成] タブでの操作方法を説明します。

3.1 モデルと設定

- モデル: 使用する AI モデルをリストから選択します。
- リストにないモデルやローカルモデル（Ollama 等）を追加、削除する場合は、[更新] ボタンを押すと最新の状態が読み込まれます。
- プリセット: 事前に登録されたタスク（システムプロンプト）を選択します。選択すると下部のテキストボックスに内容が反映されます。
- システムプロンプト: AI への役割指示です。プリセットの内容を一時的に変更することも可能です。
- 創造性 (Temp): AI の回答のランダム性を調整します（0.0 : 正確で固い表現 ~ 1.0 : 独創的でランダムな表現）。

3.2 入力エリア

- テキスト入力: AI に送る指示や質問を入力します。
- [□ 読込]: テキストファイルや PDF、Word 文書などを読み込んで入力欄に挿入します。
- [trash bin] クリア]: 入力欄を空にします。
- 文字サイズ変更: 入力エリアにカーソルを置き、Ctrl + マウスホイール で文字サイズを調整できます。

3.3 生成の実行

- [▶ テキスト生成]: AI による生成を開始します。
- ショートカットキー: Ctrl + Enter

- ・ [■ 中断]: 生成中にボタンを押すと、処理を強制停止します。

3.4 出力エリアと操作

生成されたテキストに対して以下の操作が可能です。

- ・ [保存]: テキストファイル、Markdown、Word、PowerPoint 形式などで保存します。
 - ・ [コピー]: クリップボードにコピーします。
 - ・ [↓ 入力→出力]: 入力欄の内容を出力欄にコピーします（校正作業時などに使用します）。
 - ・ [☑ 入替]: 入力欄と出力欄の内容を入れ替えます（翻訳の往復などに便利です）。
 - ・ [↑ 転記]: 出力欄の内容を入力欄にコピーします（会話を続ける際に使用します）。
-

4. 音声機能 (TTS & 自動配役)

Google Gemini の音声生成モデルを使用し、単なる読み上げではなく「演技」を含む音声を生成します。

4.1 自動配役 (Auto-Casting) の仕組み

AI がテキストを分析し、以下の判断を自動的に行います。

1. スクリプトか地の文か: 会話形式かナレーションのみかを判断します。
2. 話者の特定: テキスト内の名前や「男性」「女性」などの情報から、最適な声色 (Voice) を割り当てます。

4.2 操作ボタン

- ・ [⟳ 自動配役&再生]: 分析と生成を行い、完了後に自動再生します。
 - ・ [⏹]: 再生を停止します。
 - ・ [💾 自動配役&保存]: 生成した音声をファイル (MP3 または WAV) として保存します。
-

5. プロンプト管理タブ

よく使用するシステムプロンプト（タスク）を GUI 上で管理できます。

5.1 画面構成

- ・ 左側（リスト）: 登録済みのタスク一覧が表示されます。
- ・ 右側（編集）: 選択したタスクの詳細設定を行います。

5.2 タスクの追加と編集

1. [+] 新規追加] をクリックします。
2. タスク名を入力します。

3. 保存形式を選択します。
 - はい（外部ファイル）：./prompts/custom/ フォルダに .txt ファイルとして保存されます。長文のプロンプト管理に適しています。
 - いいえ（JSON 埋込）：設定ファイル内に直接テキストを保存します。短い指示に適しています。
4. 右側のエディタで内容を編集し、[変更を保存] をクリックして確定します。

5.3 既存ファイルの参照

- 保存先ファイル欄の [...] ボタンで既存のテキストファイルを指定すると、そのファイルをプロンプトとして登録できます。
 - [] ボタンを押すと、現在参照しているファイルがあるフォルダをエクスプローラで開きます。
-

6. 詳細設定・カスタマイズ

アプリケーションの動作は、データフォルダ内の JSON ファイルで定義されています。メニューの [設定] > [データフォルダを開く…] で直接フォルダにアクセスできます。

6.1 models.json（モデル定義）

プルダウンメニューに表示されるモデル一覧です。

```
{  
  "表示名": "プロバイダ/モデル ID",  
  "My Local Model": "lmstudio/local-model"  
}
```

- プロバイダ: google、openai、anthropic、openrouter、ollama、lmstudio が使用可能です。

6.2 casting_config.json（配役設定）

音声合成時の「配役ルール」や「声の定義」を設定します。

- system_instruction_template: 配役を決定する AI への指示。
- voices: 利用可能な声の定義リスト（Gemini TTS の仕様に準拠）。
- fallback_voice: 該当する声がない場合に使用されるデフォルトの声。

6.3 templates/（文書テンプレート）

Pandoc を使用して Word や PowerPoint に書き出す際の参照ファイル（スタイルテンプレート）です。

- custom.docx、custom.pptx などを編集することで、出力されるファイルのフォントやデザイン

ンを変更できます。

7. トラブルシューティング

Q. 音声生成が動作しません

A. Gemini API キーが設定されているかご確認ください。メイン画面下部のステータスバーに「TTS: Gemini (OK)」と表示されている必要があります。

Q. ファイルが保存されない、または保存先が分かりません

A. 本バージョンより、tasks.json があるフォルダが基準となります。自動保存ファイル名は Output_日付.txt 等になります。

Q. Word や PDF が読み込めません

A. Pandoc (Word 用) または pypdf (PDF 用) がインストールまたは認識されていない可能性があります。

Q. 特定のモデルでエラーが発生します

A. API キーの権限（課金状況）や、モデル名が正しいか models.json をご確認ください。

付録（プリセットリスト）

- ・ 【A01 翻訳（英→日）】英文を自然で正確な日本語に翻訳する。
- ・ 【A02 翻訳（日→英）】日本語文を自然で正確な米国英語に翻訳する。
- ・ 【B01 文章校正】日本語の技術・学術文書を正確性と可読性を保って校正し、修正点と提案も示す。
- ・ 【B02 技術英文校正】科学技術英語（米国英語）を文法・文体・用語統一の観点で校正し、要確認点も示す。
- ・ 【B03 文字起こし修正】音声文字起こしを話し言葉の自然さを保ちつつ誤変換やフィラー等を整えて修正する。
- ・ 【C01 3点要約】入力文の要点を重要順に3つの箇条書きで簡潔に要約する。
- ・ 【C02 長文の要約】入力テキストを断片を統合したうえで2000字以内に要約する。
- ・ 【D01 Podcast台本（・・・

6