

FlashLabs、世界初となるリアルタイム音声対話モデル「**Chroma 1.0**」を公開

Hugging Face 世界 1 位を獲得。わずか数秒のサンプルから「本人の声」を再現するオープンソース **AI**

2026 年 1 月 23 日 | 東京都千代田区

FlashIntel Japan 株式会社（本社：東京都千代田区、代表取締役：細井 洋一）は、FlashLabs が 1 月 16 日、世界初となるオープンソースのエンドツーエンド・リアルタイム音声対話モデル「**Chroma 1.0**（クロマ）」を正式公開したことを発表いたします。**Chroma1.0** は AI 開発のための世界最大級のオープンソースプラットフォーム **Hugging Face** にてマルチモーダルカテゴリで 1 位を獲得（2026 年 1 月 22 日時点）。**Chroma** 公開の X の投稿は 100 万回表示を突破しています。**Chroma 1.0** は、音声入力から音声出力まで約 147 ミリ秒の低遅延を実現し、数秒の参照音声から高精度な個別化音声生成を可能にします。コールセンターや営業の電話対応など、ビジネスの現場で人間に代わって対話する AI としての活用が期待され、すでに FlashIntel Japan 株式会社が提供するエンタープライズ向け音声エージェント基盤 **FlashAI 2.0** 上で実運用されています。クローズド API が主流であったリアルタイム音声 AI 分野において、**Chroma 1.0** は研究・開発・実装のすべてを開放した、初の実用レベルのオープンモデルとなります。

Chroma 1.0

リアルタイム音声インテリジェンス。完全オープン。

音声類似度

ゼロショット音声クローニング

※ 高いほど良い↑

話者類似度 (SIM)

人間ベースライン	0.73
F5-TTS	0.64
Seed-TTS	0.76
FireRedTTS-2	0.66
Step-Audio-TTS	0.66

Chroma 1.0 🏆 **0.817**

人間ベースライン比 **+10.96%**

比較対象中、最高性能

リアルタイム性能

エンドツーエンド音声-to-音声

初回応答時間 (TTFT) **146.9 ms**

フレーム平均遅延 **52.3 ms**

生成 RTF **0.43x**

生成音声長 **16.6 秒**

リアルタイムを大きく上回る処理性能

ASR → LLM → TTS を介さないネイティブ音声処理

背景：音声 AI の遅延問題と個別化の課題

従来の音声 AI システムは、音声認識 (ASR) → テキスト処理 (LLM) → 音声合成 (TTS) という多段階処理により、応答遅延が発生していました。また、話者の声質を保持したままリアルタイムで対話する技術は、計算コストと精度の両立が困難でした。コールセンターや顧客対応の現場では、自然な会話速度と個別対応が求められており、技術的なブレークスルーが必要とされていました。

Chroma 1.0 とは：「声を声のまま理解する AI」

Chroma 1.0 は、音声を音声のまま処理する「Speech-to-Speech (S2S)」設計を採用し、以下の特徴を備えた世界初のモデルです（当社調べ、2026 年 1 月時点）。

世界初の特徴（当社調べ、2026 年 1 月時点）：

- オープンソース（コード・モデル重みを含む全公開）

- エンドツーエンド音声間処理（ASR→LLM→TTS を介さない）
- リアルタイム対話（200 ミリ秒未満の応答）
- 数秒の参照音声からの個別化音声クローニング

調査範囲：

主要オープンソースリポジトリ、論文データベース（arXiv 等）、主要ベンダー公開情報（OpenAI Realtime API、Elevenlabs 等）を対象に、上記 4 条件を同時に満たすモデルの有無を確認。

技術的特徴：性能データと計測条件

Chroma 1.0 は、研究・管理環境下で以下の性能を示しました。

◆ 低遅延リアルタイム応答

- エンドツーエンド TTFT（Time To First Token）：146.87 ミリ秒
 - 音声入力から最初の音声出力までの時間
 - 測定条件：単一 GPU 環境、標準ネットワーク、平均 3～5 秒の音声入力

◆ 高速推論

- RTF（Real-Time Factor）：0.43
 - 実時間の半分以下で処理完了（実時間の 2 倍以上の速度）
 - 測定条件：同上

◆ 高精度音声類似度

- Speaker Similarity（話者類似度）：0.817
 - 人間ベースライン（0.73）比で +10.96% の向上
 - 評価指標：客観的話者類似度評価
 - 数秒の参照音声から高精度な個別化音声を生成

◆ 軽量・高効率設計

- パラメータ数：約 40 億（4B）
 - Qwen2.5-Omni-3B、Llama 3 クラスに匹敵する対話性能
 - ストリーミング生成に対応し、連続的な音声出力を実現

◆ インターリーブ設計

- テキスト-音声トークンスケジュール (1:2)
 - テキストと音声を同期生成し、自然なターンテイキング（会話の交代）を実現
-

想定ユースケース

Chroma 1.0 は、以下の分野での活用を想定しています。

1. コールセンター・カスタマーサポート

音声応答の待ち時間を短縮し、自然な会話フローを実現。顧客ごとに最適化された音声で対応することで、満足度向上とオペレーター負荷軽減を両立。

2. 営業・インサイドセールス

リアルタイム音声エージェントが初期対応を担当し、商談機会の取りこぼしを削減。見込み顧客との自然な対話で、人的リソースを高付加価値業務にシフト。

3. 予約・問い合わせ対応

24 時間 365 日、遅延のない音声対応を提供。飲食店、医療機関、公共施設等での無人受付や多言語対応に活用可能。

4. 音声ネイティブなマルチモーダル AI アプリケーション

音声インターフェースを核とした新サービスの構築。教育、エンターテインメント、アクセシビリティ支援等、多様な領域での応用が期待される。

5. 音声アシスタント・バーチャルヒューマン

個別化された音声で対話する AI キャラクターやアバター。ゲーム、メタバース、パーソナルアシスタント等での利用を見込む。

FlashLabs が提供している電話対応 AI エージェント **FlashAI** 上でも実装され、大手システム提供会社を含めた同サービス利用者にも順次公開されます。

オープンソース公開内容：

Chroma 1.0 は、研究・開発・実装のすべてを公開しています。

◆ 公開内容

- モデルウェイト (**Hugging Face**)
<https://huggingface.co/FlashLabs/Chroma-4B>
- ソースコード (**GitHub**)
<https://github.com/FlashLabs-AI-Corp/FlashLabs-Chroma>
- 技術論文 (**arXiv**)
<https://arxiv.org/abs/2601.11141>

◆ ライセンス

Apache License 2.0

商用利用、改変、再配布が可能。研究機関・企業・個人開発者が自由に活用できます。

◆ 動作環境

- GPU 推奨 (NVIDIA A100、H100 等)
- transformers ライブラリ最新版
- PyTorch 最新版

デモと検証

◆ デモ動画

リアルタイム対話と個別化音声生成の実際の動作を確認できます。

<https://www.youtube.com/watch?v=AOMmxTwsam0>

◆ 技術レポート

測定方法、ベンチマーク詳細、アーキテクチャ解説を論文で公開。再現手順も含め、第三者による検証が可能です。

◆ ベンチマーク結果

主要な音声理解・推論・対話タスクにおいて、同クラスのモデルと競合する性能を確認。詳細は論文の Table 5 を参照。

コメント

FlashLabs 創業者兼 Global CEO 石 一 (Yi Shi) は次のように述べています。

「リアルタイム音声 AI は、クローズドであるべきではありません。Chroma 1.0 は、音声 AI を"開かれた知能"として次の段階へ進めるための第一歩です。従来の多段階処理では、遅延、誤り伝播、副言語情報の損失が避けられませんでした。エンドツーエンド設計により、これらの課題を根本から解決しました。開発者コミュニティとともに、音声 AI の民主化を推進します。」

FlashIntel Japan 株式会社について

FlashIntel Japan 株式会社は、次世代 AI エージェントおよび音声 AI 基盤の研究・開発・提供を行う AI 企業です。営業 AI 基盤システム FlashRev、電話 AI エージェント FlashAI をはじめ、に日本市場における生産性革新を推進しています。

会社名：FlashIntel Japan 株式会社

代表者：代表取締役 細井 洋一

所在地：東京都千代田区

事業内容：AI エージェント基盤、リアルタイム音声 AI、企業向け AI ソリューションの研究開発・提供

ウェブサイト：<https://www.flashlabs.ai/>

報道関係お問い合わせ先

FlashIntel Japan 株式会社

Email：koki.kobayashi@myflashcloud.com

電話：03-6869-2514

担当：広報担当