



# Kotoba Technologies



**GPT FUGAKU**

Japanese  
Supercomputer LLM



World #2 Popular Speech  
Recognition Model



Best Japanese Text-to-Speech



Speech Foundation Models  
With Government Support

# Kotoba Technologies

## 音声生成AIの力で言語の壁を打ち崩す

### AI同時通訳 超高速AI同時通訳技術

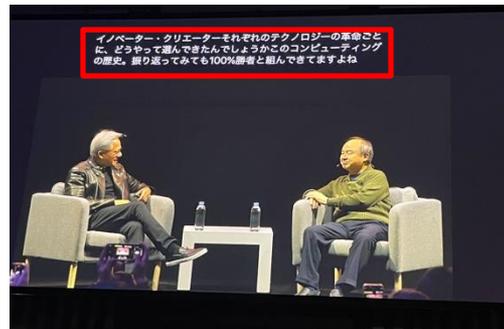
世界最速-0.5秒遅延  
実用可能な精度  
日本語 <> {英中韓越}

ペアリング完了後、会話中の画面



### 音声書き起こし 日本語最高性能

オープンソースモデルはHuggingFace上で  
累計80万回以上ダウンロード  
NVIDIA AI Summit Japanメイン  
イベントで3,000人の前で運用



# アプリ & 商業化

同時通訳  
“夢の翻訳ツール  
の到来”

世界最速レベル  
1秒以下で通訳



## 同時通訳 – リアルタイム音声翻訳の新境地

辞書／辞典／その他

iPhone iPad

アプリのカテゴリ ▼ ゲームカテゴリ ▼

無料アプリ



1 Google 翻訳  
Google



2 Papago - AI通訳・翻訳  
NAVER Corp.



3 翻訳・今すぐ翻訳・翻訳ア  
プリと通訳  
Air Apps Systems



4 DeepL - 翻訳と文章校正  
DeepL GmbH



5 同時通訳 - リアルタイム音  
声翻訳の新境地  
(株)Kotoba Technologi...



6 写真 翻訳  
BrainMount Ltd

## Apple Storeで第5位！

# デモ: 音声同時通訳





## 技術紹介：流暢な日本語生成

多様で流暢な日本語生成が可能・声真似(voice cloning)も可能。  
現在14言語に拡大中。

### Voice-1

(Voice Cloning: 小島)



### Voice-2

(スタイル：audiobook)



### Voice-3

(感情：エネルギッシュ)



### Voice-4

(スタイル：ニュース)



# Kotoba Technologies : 創業者

2023年7月創業  
東京大手町に  
研究開発拠点

米国アカデミア出身  
生成AI研究者が  
スパコン富岳を活用  
した国産LLM開発の  
研究者とタッグを組  
み創業



**Dr. Noriyuki Kojima**  
Co-founder & CEO

米コーネル大学CS工学博士。

自然言語処理、コンピュータビジョンを専門にする。言語モデル分野では国内の**スーパーコンピュータ富岳やABCIIを通じた大規模言語モデルプロジェクトを発起・中心メンバーとしてリード**した経験がある。最先端AIモデルの民主化に強い関心を持ち、HuggingFace、Metaなどでの勤務経験を通じてこの分野に深い洞察を得る。

AI分野のトップ国際学会**EMNLP2022 Best Paper Award**を受賞。孫正義育英財団生。



**Prof. Jungo Kasai**  
Co-founder & CTO

米ワシントン大学CS博士。  
米Toyota Technological Institute at Chicago (TTIC) 研究助教授。

自然言語処理分野、機械学習を専門とし、国際学会採択論文も多数 (**30歳にして自然言語処理研究者としては引用数日本人トップ30**)。フェイスブック、マイクロソフト、IBM、Allen Institute for AIでのインターンシップ経験から、AIが実社会で使われる為に必要な過程・要素にも造詣が深い。

AI分野のトップ国際学会**NAACL2022 Best Paper Award**を受賞。IBM PhDフェロー、孫正義育英財団生、船井情報科学振興財団生。

# Kotoba Technologies : 開発チーム

**Fugaku-LLM**を始めとするスパコンを  
活用した豊富な**大規模AI学習の実績**

AIのトップ国際会議での**ベスト  
ペーパー賞などの研究実績**



source:

<https://www.itmedia.co.jp/aipplus/articles/2405/10/news186.htm>

!



- Best Paper Award @ EMNLP 2022
- Best Paper Award @ NAACL 2022
- AAI Best Paper Award @ AAI 2020
- ACL Outstanding Paper Award @ ACL 2017



富岳



ABCI



TSUBAME4.0



# GENIAC & Azure & Kotoba



- 経産省GENIACの第1.5回、第2回に連続採択。
- 両回とも**Microsoft Azure**のH100を大規模に活用して学習。
- GENIAC第2回では採択20社の中で**GPUの使用量top-3**、国内でも有数のGPUの使用量。

NEW

ELYZA



FUJITSU



sakana.ai



Stockmark

TURING

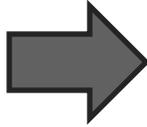
M 松尾・岩澤研究室  
MATSUO-IWASAWA LAB UTOKYO



# Kotoba Technologies



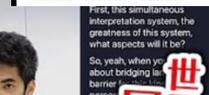
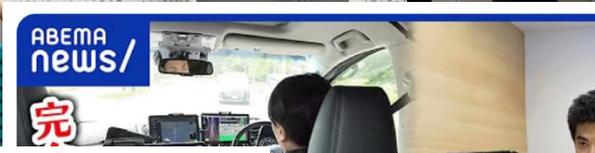
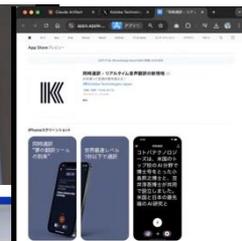
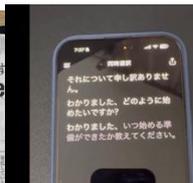
2023.10



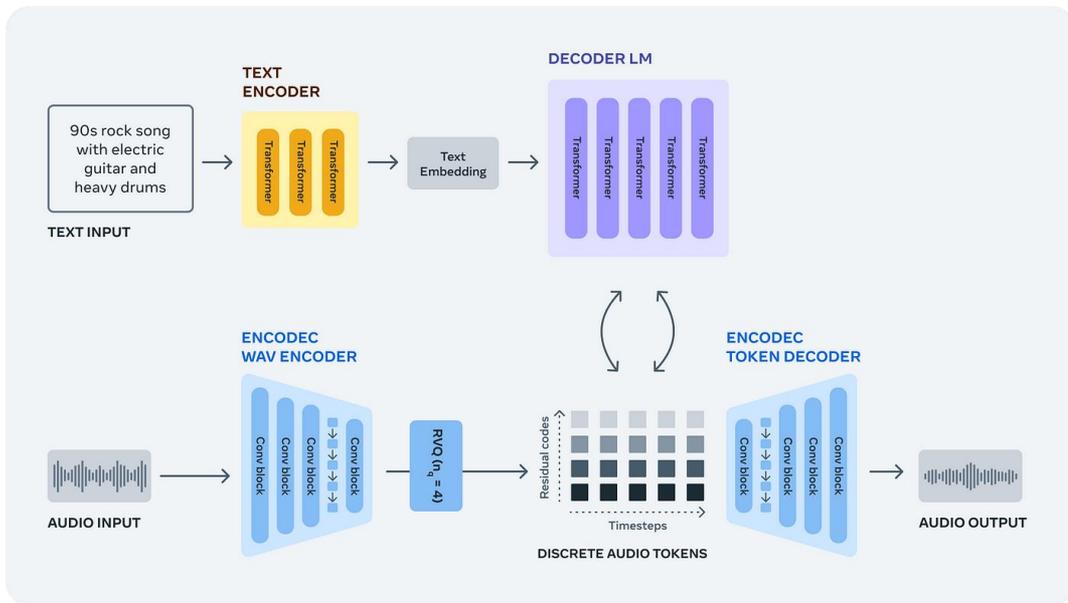
2025.6



# メディアでの注目



# 技術特異性: End-to-End Model



## AudioCraft

<https://ai.plainenglish.io/meet-audiocraft-meta-ais-new-family-of-generative-audio-models-6411e597e431>

## TTS / 音声基盤モデルの仕組み

- トークナイザーを活用して音声を離散化
- Autoregressive Transformerで学習・処理

## 課題

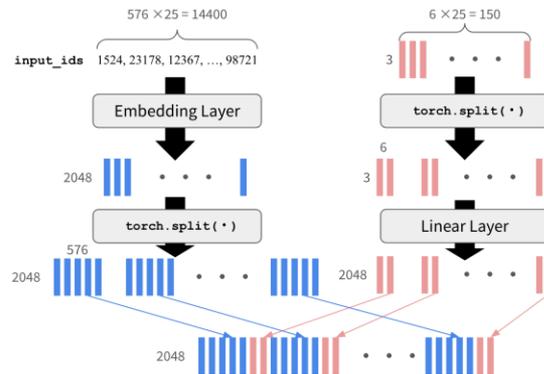
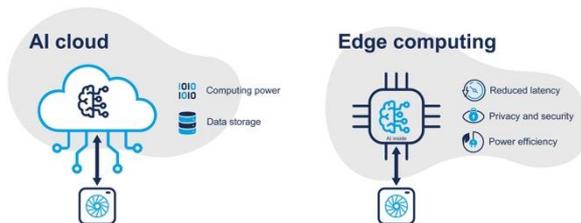
- 音声だとSequence Lengthが膨大になる
- “I have a cat”
  - テキストだと2~3tokens
  - 音声だと100 tokens 近く

## 解決法: 離散化の圧縮率を高める。

- 圧縮率が5~6倍に向上
- 学習が高速化 → 開発サイクルの向上・長く学習できる
- 推論時の高速化 → 他の工夫と組み合わせると5~6倍を超える高速化

# 技術特異性: 超高速化

汎用性ではなく、同時通訳への特化を行うことで可能になる技術開発・デザイン



## 1. Two-Stream Model

入力と出力を同時に行うことができるモデル。

将来的にはMulti-Stream。

**GPT-4oはSingle-Stream**

## 2. 小型化・軽量化

同時通訳の精度とのバランスをとった極限までの小型化・軽量化、同時通訳に特化。

## 3. バックエンド最適化

同時通訳のユースケースを想定したCUDA layerに至るまでの最適化。

## 4. インフラの安定的なスケーリング

超低遅延が必要な同時通訳に特化したインフラ設計のスケーリング。

# Thanks!

Kotoba will remove the language barrier using GenAI



Official X



Inquiries



Career

## Contact Info

Nori & Jungo : {nkojima, jkasai}@kotoba.tech