

山梨 第4回

現状の Amazon Polly と 生成エンジン登場にみた新たな可能性

Journeyman | @beajourneyman

Feb 2025

Journeyman ジャニ (Takeki Oizumi)

所属：セゾンテクノロジー(4月に社名変更)
2022年10月より遠隔地勤務可に


仕事：Amazon Connect を基盤とした
IVRシステムの開発保守リーダー

推しサービス：Amazon Connect / Polly
最近は Bedrock

SNS：@beajourneyman

認定：CLF / SAA / AIF





**Bedrock、生成AIと深く関わる
Polly、音声の話しをします**

AWSのコミュニケーションサービス

AWS サービスを組み合わせた強力なコミュニケーション



<https://youtu.be/oI8h4KDpIL0?si=PyhAPvXivBrosdNi&t=1329>

Amazon Polly とは？

12 ヶ月間、毎月 500 万文字が無料 →

Amazon Polly - AI 音声ジェネレーター

数十の言語で高品質で自然な人間の声を展開

AWS アカウントを作成する

Amazon Polly の使用を開始する

Amazon Polly とは何ですか？

Amazon Polly は、オンデマンドで音声を生成し、あらゆるテキストを音声ストリームに変換する、フルマネージドサービスです。深層学習テクノロジーを使用して、記事、ウェブページ、PDF ドキュメント、他のテキスト読み上げ (TTS) の変換を行います。Polly は、エンゲージメントとコンバージョンを実現する音声起動アプリケーションを構築できるように、幅広い言語で数十種類のリアルな音声を提供します。さまざまな地域や市場のユーザーの多様な言語、アクセシビリティ、学習ニーズに対応します。強力なニューラルネットワークと生成音声エンジンがバックグラウンドで機能し、音声を合成します。Amazon Polly API を既存のアプリケーションに統合して、すぐに音声対応にします。

業務では自動音声応答(IVR)で使用

仕組み

Amazon Polly は深層学習技術を使用し、人間の声のような音声を合成します。そのため、記事を音声に変換することができます。幅広い言語に対応したリアルな音声を多数搭載しており、Amazon Polly を使用して音声起動型アプリケーションを構築することができます。

コンテンツの作成

e ラーニング

テレフォニー



🔍 クリックして拡大

2024/05/08 Amazon Polly の 3 つの合成音声 (英語) に対応した新生成エンジン

Amazon Polly の 3 つの合成音声 (英語) に対応した新生成エンジン

New!!

投稿日: May 8, 2024

本日、Amazon Polly の 3 つの英語音声 (2 つのアメリカ英語音声 Ruth と Matthew、1 つのイギリス英語音声 Amy) に対応した、表現力豊かな生成エンジンの一般提供開始についてお知らせいたします。

Amazon Polly は、文章をリアルな音声に変換するサービスで、音声を使用したアプリケーションの作成や、ビジネスニーズに応じた音声対応製品の構築が可能です。

<https://aws.amazon.com/jp/about-aws/whats-new/2024/05/generative-engine-three-english-polly-voices/>

Amazon Polly

テキスト読み上げ機能

レキシコン

S3 合成タスク

Amazon Polly > テキスト読み上げ機能

テキスト読み上げ機能 情報

S3 に保存

ダウンロード

▶ 音声を聴く

エンジン 情報

生成

生成 AI を使用して、最も表現力豊かで適応性の高い音声を生成します。

ロングフォーム

長時間のコンテンツのために、極めて自然な音声を生成します。

ニューラル

スタンダードエンジンよりも自然で人間に近い音声を生成します。

スタンダード

自然な音声を生成します。

言語 情報

英語 (米国)

音声 情報

Ruth, 女性

入力テキスト 情報

SSML 情報

< speak > Hi, I'm Ruth. I can read any text for you. Test it out! < / speak >

70 文字を使用

デフォルトのテキストを復元

テキストを消去

▶ その他の設定

この生成エンジンは Amazon Polly の最も高度なテキスト読み上げ (TTS) モデルです。さまざまな音声、言語、スタイルでトレーニングされています。文脈に応じた韻律や、間の取り方、スペル、方言特性、外来語の発音などを高い精度で表現します。生成型の合成音声は、感情表現に富み、明快で口語的であるため、人間の声に驚くほど似ています。この新しい音声は強力な機能を備えていながら低レイテンシーであるため、オンライン会話のユースケースにも適しています。お客様は、人間に近い合成音声を使用する、知識豊富なカスタマーアシスタント、バーチャルトレーナー、広報として生成音声のペルソナをご利用いただけます。

何故、心躍ったのか？

音声チューニングがいらなくなる世界が来るから

2年前に戻ります

2023/02/08 日本語ニューラル女性音声GA

Amazon Polly が 2 つの新しい日本語 NTTS 音声の提供を開始

Kazuha&Tomoko

投稿日: Feb 8, 2023

Amazon Polly はテキストをリアルな音声に変換するサービスです。発話が可能なアプリケーションはもちろん、まったく新しいタイプの音声対応製品も構築できます。本日、日本語での女性のニューラルテキスト読み上げ (NTTS) 音声である Kazuha と Tomoko の一般提供が開始されました。

Amazon Polly に、日本語をサポートする音声にさらに 2 つ追加されました。追加になったのは初の女性のニューラル音声で、これにより、このロケールのポートフォリオは 3 種類の女性の声、1 種類の男性の声にまで拡大しています。同時に 2 つの音声リリースされ、より多くの選択肢から女性の音声を選べるようになりました。

Kazuha、Tomoko とその他のすべての NTTS 音声は、[ニューラル TTS をサポートしている AWS リージョン](#)で利用可能です。詳細については、[Amazon Polly のドキュメントと料金ページ](#)をご覧ください。



Kazuha, Tomoko 誕生日おめでとう！！

歡喜して、早速お客様提案したものの、

音声合成してみると拭えない違和感が orz...

Kazuha と Tomoko が活躍するために
必要なのは、

音声合成マークアップ言語(SSML)

によるチューニング

テキスト読み上げ機能 情報

S3 に保存

↓ ダウンロード

▶ 音声を聴く

エンジン 情報

ロングフォームエンジンはすべてのリージョンでサポートされているわけではありません。サポートされているリージョンのリストについては、次を参照してください: [機能とリージョンの互換性](#)。
生成エンジンはすべてのリージョンでサポートされているわけではありません。サポートされているリージョンのリストについては、次を参照してください: [機能とリージョンの互換性](#)。

生成

生成 AI を使用して、最も表現力豊かで適応性の高い音声を生成します。

ロングフォーム

長時間のコンテンツのために、極めて自然な音声を生成します。

ニューラル

スタンダードエンジンよりも自然で人間に近い音声を生成します。

スタンダード

自然な音声を生成します。

言語 情報

日本語 ▼

音声 情報

Tomoko, 女性 ▼

入力テキスト 情報

<speaks> こんにちは、智子です。読みたいテキストをここに入力してください。</speaks>

47 文字を使用

デフォルトのテキストを復元

テキストを消去

▶ その他の設定

音声合成マークアップ言語(SSML)とは？ Speech Synthesis Markup Language

ピッチ、発音、読み上げ速度、音量などのテキスト読み上げの出力属性を微調整するために使用できる XML ベースのマークアップ言語

与えているテキスト

JAWS-UG山梨第4回開催、おめでとうございます！番号は8074です。

チューニング前後のデモをお聞きください

SSMLによるチューニング(山梨支部 ver.)

```
<speak>
```

```
<phoneme alphabet="x-amazon-pron-kana" ph="ジョーズ"  
">JAWS</phoneme>
```

```
<phoneme alphabet="x-amazon-pron-kana" ph="ユージー">.-  
UG</phoneme>、
```

```
<phoneme alphabet="x-amazon-pron-kana" ph="ヤマ'ナシ">山梨  
</phoneme>、 <break time="0.05s"/>
```

```
<p>第4回開催、おめでとうございます!</p>
```

```
<prosody rate="80%"><p>番号は8074です。</p></prosody>
```

```
</speak>
```

プロダクションレベルで頻出のタグ

タグ	用途	ユースケース
<code><break time="0.1s"/></code>	ブレイク(間)を入れる。	連続で話すことで不自然に感じてしまうケースで非常に有効。0.05秒などコンマ2単位で調整。
<code><p></p></code>	明確に文章の区切りを入れる。	区切りを入れないと、人間だと息継ぎをしているようなケースでも一気に発話してしまい違和感になる。
<code><phoneme alphabet="x-amazon-pron-kana" ph="オヘ'ンロ">お遍路</phoneme></code>	読み上げを制御する。アクセントを入れられる。	漢字読み上げ、イントネーションを制御できる。尚、標準語・関西弁などの指定はできない 😊
<code><prosody rate="80%"></prosody></code>	発話スピードを変える。	要メモな読み上げ(口座番号等)。

日本語バージョンのGA、お待ちしております！ (チューニングがいらぬ世界が来て欲しい)

テキスト読み上げ機能 情報

S3 に保存

ダウンロード

▶ 音声を聴く

エンジン 情報

ロングフォームエンジンはすべてのリージョンでサポートされているわけではありません。サポートされているリージョンのリストについては、次を参照してください:

[機能とリージョンの互換性](#)

生成エンジンはすべてのリージョンでサポートされているわけではありません。サポートされているリージョンのリストについては、次を参照してください: [機能とリージョンの互換性](#)

生成

生成 AI を使用して、最も表現力豊か
かで適応性の高い音声を生成しま
す。

ロングフォーム

長時間のコンテンツのために、極
めて自然な音声を生成します。

ニューラル

スタンダードエンジンよりも自然
で人間に近い音声を生成します。

スタンダード

自然な音声を生成します。

言語 情報

日本語

音声 情報

Kazuha, 女性

さらなる進化の大きなうねりを感じた

そして、心が躍った

2024/12/03 Amazon Nova



What's next for Amazon Nova?

NEW

Amazon Nova
Speech-to-Speech

COMING SOON

NEW

Amazon Nova
Any-to-Any

COMING SOON

生成AIは方言を理解している

Input: 52

Output: 437

Latency: 13254 ms



大阪にあるオススメのたこ焼きのお店を3つ、日本の方言のひとつである関西弁で教えてください。



ほんまええで！大阪のたこ焼きといえは外せへんところ、3つ教えたるわ。関西弁で説明したるさかい、よう聞いてや。



1. くくる 曽根崎店 このたこ焼きはめっちゃジューシーでふわふわしてんねん。外はカリッとして中はトロトロ。タコもデカいし、ソースもええ感じやで。駅近やから便利やし、行ったらアカンで！
2. たこ家 道頓堀くる本店 こども超有名店や。観光客も多いけど、地元民にも愛されとるんや。タコのプリプリ感がたまらんし、生地もふわふわやで。いろんな味付けがあるさかい、好みの味見つけられるで！
3. たこ焼き やまちゃん 天王寺店 ここは地元で愛されとる老舗や。昔ながらの味を守って、シンプルやけどめっちゃうまいねん。カリカリの外側と、トロトロの中身のバランスがええんや。安くてボリュームもあるさかい、腹ペコの時にぴったりやで！

これらのお店、どこ行っても間違いあらへんで！大阪来たらずひ食べてみてや！うまいこと楽しんでてな！

生成AIに関西弁でお好み焼き屋を聞いた

ほんまええで！大阪のたこ焼きといえは外せへんところ、3つ教えたるわ。関西弁で説明したるさかい、よう聞いてや。

1. くる 曽根崎店

ここのたこ焼きはめっちゃジューシーでふわふわしてんねん。外はカリッとしてて中はトロトロ。タコもデカいし、ソースもええ感じやで。駅近やから便利やし、行ったらアカンで！

2. たこ家 道頓堀くる 本店～中略～

3. たこ焼き やまちゃん 天王寺店 ～中略～

これらのお店、どこ行っても間違いあらへんで！大阪来たらぜひ食べてみてや！うまいこと楽しんでってな！

この生成エンジンは Amazon Polly の最も高度なテキスト読み上げ (TTS) モデルです。さまざまな音声、言語、スタイルでトレーニングされています。文脈に応じた韻律や、間の取り方、スペル、**方言特性**、外来語の発音などを高い精度で表現します。生成型の合成音声は、感情表現に富み、明快で口語的であるため、人間の声に驚くほど似ています。この新しい音声は強力な機能を備えていながら低レイテンシーであるため、オンライン会話のユースケースにも適しています。お客様は、人間に近い合成音声を使用する、知識豊富なカスタマーアシスタント、バーチャルトレーナー、広報として生成音声のペルソナをご利用いただけます。

学術研究やサービス利用が始まっている

九州・沖縄方言の継承支援に資する音声対話型生成系AIの開発

キーワード

消滅危機言語・方言 / 生成AI / 方言継承

研究開始時の研究の概要

本研究では、九州・沖縄の個別方言に特化した音声対話型の生成AIを開発する。

近年、地域の文化財たる方言が消滅の危機にある。本研究は、方言に特化した生成系AIを開発して方言学習に活用することで、消滅危機問題の解決を目指す。今回の対象方言は、消滅の危機にも非常に近接している

生成AIで飲食店検索 「ぐるなび」が新アプリ 方言で特徴紹介も

2025/1/21 6:00 (2025/1/29 5:00 更新) [有料会員限定記事]

#九州経済面 #新サービス #外食 #AI #東京報道部

前田 淳 + フォロー



新サービス「UMAME!」をPRするぐるなびの杉原章郎社長（左）ら。20日、東京都千代田区

飲食店予約サイトを運営する「ぐるなび」（東京）は20日、生成人工知能（AI）を活用して、飲食店を検索できるアプリを公開した。旅行や出張先のお薦めの店を検索するサービスでは、現地の方言で店を紹介する機能も付けた。

素のChatGPTが使う大阪弁は何かヘンだが鹿児島弁は学習次第でほぼ完璧、“方言AI”が秘める可能性とリスク

【生成AI事件簿】方言やローカル言語を話して好感度アップも、方言AIで飛躍的に高まる話し手の説得力

- 各国のIT企業は自国の母国語に強いLLM（大規模言語モデル）の開発を進めている。その中で、登場しつつあるのが方言に対応したLLMだ。
- 方言AIは貴重な文化や知識を時代に継承する有力な手段になる上に、その言語を話す人々から信頼や親近感を得る手段になり得る。
- だが、政治家のスピーチなど、方言AIによる多言語対応を悪用するケースも考えられる。「方言ロボット」とともにリスクを議論しておく必要がある。

消滅危機言語をどうにも基礎となる生成方言版“ほんやくコンニャク”

方言を残していくにはどうすればよいのか。

弘前大学の今井雅教授が現在開発中なのが「津軽弁翻訳システム」。

津軽弁を学習したAIを使って、まず津軽弁を音声認識システムで文字に変換する。そして、テキスト化した津軽弁をコンピューターで処理して共通語に翻訳し、ディスプレイに映し出す仕組みだ。

相手に伝わらないことを心配せずに方言を話してもらいたいという狙いもあるという。



ふだんから津軽弁を話す大学職員に、翻訳システムを使って話してもらおうと。

生成AIによる方言継承の動きは始まっている

**Amazon Nove Speech-to-Speech
登場の次に来るのは
Polly と Bedrock が完全に統合した世界**

その統合がもたらすのは…

各地の方言を理解し、不気味の谷を越えて流暢に話し、人口減少社会の中で方言継承の力を秘めている「AI音声エンジン」の誕生

A wide-angle photograph of a vast, clear blue ocean. The water transitions from a deep blue in the distance to a lighter, turquoise blue in the foreground. A small white boat is visible on the horizon line. In the bottom right corner, a small portion of a rocky coastline with green trees is visible.

Share your lessons.

AWS 推しサービスを中心にQiitaに
書いているので、良かったらご覧ください



<https://qiita.com/beajourneyman>

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年12月09日
Bedrockの新評価機能 LLM-as-a-judge (プレビュー)でRAGを評価してみた
AWS ナレッジベース rag bedrock 生成AI
♡ 8

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年12月06日
維持運用の要 AWSサポートに関する備忘録
AWS AWSsupport AWSサポート
♡ 4

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年12月04日
#reInvent2024 で発表された #AmazonNova と Claude を新機能"比較モード"で触ってみて
AWS bedrock 生成AI claude AmazonNova
♡ 8

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年12月02日
Amazon Polly の最近のアップデートをふりかえり、生成AI時代の音声合成を考える
AWS 音声合成 Polly AmazonPolly 生成AI
♡ 2

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年12月01日
Amazon Polly 音声合成チューニング入門
AWS 音声合成 Polly AmazonPolly SSML
♡ 3

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年11月30日
Amazon Bedrock Knowledge Bases Web Crawler (Preview)で作る自分チャットボット🤖
AWS ナレッジベース rag bedrock
♡ 1

@beajourneyman (Journeyman) in 株式会社センテクノロジー
2024年11月29日
充実のAWS無料教材でAIFに合格したので勉強法を紹介
AWS AWS認定試験 受験記 AIF
♡ 4

首都圏のつながりを大切にしています





首都圈 Connecting the dots!!



connpassグループ登録お願いします！！

<https://jawsug-tochigi.connpass.com/>

Appendix...

[Amazon Polly 数十の言語で高品質で自然な人間の声を展開](#) Amazon Polly ポータル

[Amazon Polly が 2 つの新しい日本語 NTTS 音声の提供を開始](#) 2023/02/08 日本語ニューラル女性音声GA

[AWS が 3 人の表現力豊かなアメリカ英語の声を備えた長文形式エンジンの提供を開始](#)
2023/11/16 長文形式のエンジン追加

[Amazon Polly の 3 つの合成音声 \(英語\) に対応した新生成エンジン](#) 2024/05/08 今回ご紹介

[3つの新しい長編ボイス](#) 2024/11/14 英語・スペイン語音声追加

[Amazon Polly と SSML\(音声合成マークアップ言語\)に関する備忘録](#) 弊Qiita関連ブログ

[Amazon Polly の最近のアップデートをふりかえり、生成AI時代の音声合成を考える](#) 同関連ブログ

Appendix...

生成AI x 方言 記事

[九州・沖縄方言の継承支援に資する音声対話型生成系AIの開発](#)

[生成AIで飲食店検索「ぐるなび」が新アプリ 方言で特徴紹介も](#)

[NHK WEB特集 最近、方言使ってますか？](#)

[素のChatGPTが使う大阪弁は何かヘンだが鹿児島弁は学習次第でほぼ完璧、“方言AI”が秘める可能性とリスク](#)

ご清聴いただき、ありがとうございます！

ジャニ (Journeyman) | @beajourneyman