

Google Text-to-Speech API の始め方

私は Yin Wang の記事のいくつかを Google Text-to-Speech API を使用して音声に変換する予定です。以下に、ChatGPT が提供する役立つチュートリアルとともに、ステップバイステップのガイドを記載します。すべての準備が整ったら、ここに音声をアップロードして、皆さんに聞いていただけるようにします。

ステップ 1: Google Cloud アカウントの設定

1. Google Cloud アカウントを作成する

アカウントをお持ちでない場合は、Google Cloud Console でサインアップしてください。

2. 新しいプロジェクトを作成する

- Cloud Console で、プロジェクトのドロップダウンメニュー（左上）をクリックします。
- 「新しいプロジェクト」を選択し、名前を付けてプロジェクトを作成します。

3. Text-to-Speech API を有効にする

- Google Cloud Text-to-Speech API ページにアクセスします。
- 「有効にする」をクリックして、プロジェクトの API を有効にします。

4. API 認証情報の作成

- Cloud Console で 「APIs & Services」 > 「Credentials」 に移動します。
 - 「Create Credentials」 をクリックし、「Service Account」 を選択します。
 - プロンプトに従ってサービスアカウントを作成し、JSON 形式の秘密鍵ファイルをダウンロードします。
 - この JSON ファイルは API リクエストの認証に使用されるため、安全に保管してください。
-

ステップ 2: Google Cloud SDK とクライアントライブラリをインストールする

1. Google Cloud SDK をインストールする

まだインストールしていない場合は、お使いのオペレーティングシステムに合わせて Google Cloud SDK をインストールする手順に従ってください。

2. Python クライアントライブラリのインストール

Python を使用している場合、以下のコマンドで `google-cloud-texttospeech` ライブラリをインストールします:

```
pip install google-cloud-texttospeech
```

ステップ 3: 認証の設定

API を使用する前に、認証情報を使用して認証を行う必要があります。認証情報ファイルのパスを環境変数に設定してください：

```
export GOOGLE_APPLICATION_CREDENTIALS="/path/to/your/service-account-file.json"
```

ダウンロードした JSON ファイルの実際のパスに置き換えてください。

ステップ 4: テキスト読み上げ機能の実装

以下は、Google Cloud Text-to-Speech API を使用してテキストを音声に変換する Python の例です：

```
from google.cloud import texttospeech

def text_to_speech(text, output_filename):
    # Text-to-Speech クライアントを初期化
    client = texttospeech.TextToSpeechClient()

    # 合成入力の設定
    synthesis_input = texttospeech.SynthesisInput(text=text)
```

```

# 音声パラメータを設定 ('en-US-Journey-D'を使用)
voice = texttospeech.VoiceSelectionParams(
    language_code="en-US", # 英語（アメリカ合衆国）
    name="en-US-Journey-D" # 特定の音声モデル (Journey)
)

# オーディオ設定をセット
audio_config = texttospeech.AudioConfig(
    audio_encoding=texttospeech.AudioEncoding.MP3, # MP3フォーマット
    effects_profile_id=["small-bluetooth-speaker-class-device"], # Bluetoothスピーカー向けに最適化
    pitch=0.0, # ピッチ変更なし
    speaking_rate=0.9, # 調整された話速（必要に応じて変更可能）
    volume_gain_db=5.0 # 音量を大きく
)

# テキスト読み上げリクエストを実行
response = client.synthesize_speech(
    input=synthesis_input, voice=voice, audio_config=audio_config
)

# オーディオコンテンツをファイルに書き込む
with open(output_filename, "wb") as out:
    out.write(response.audio_content)

print(f"オーディオコンテンツが {output_filename} に書き込まれました")

```

使用例

```

article_text = "映画、ああ、本当に大好きです。まるでタイムマシンのように、さまざまな世界や風景に連れて行ってくれるんです。もう、それに夢中なんです。"
output_file = "output_audio.mp3"
# MP3形式で出力

```

テキストを音声に変換する

```
text_to_speech(article_text, output_file)
```

ステップ5: スクリプトを実行する

1. スクリプトを `text_to_speech.py` として保存します。
2. 以下のコマンドでスクリプトを実行します:

```
```bash
python text_to_speech.py
```

これにより、提供されたテキストから音声ファイル (output\_audio.mp3) が生成されます。

---

## ステップ6: サービスアカウントキー

あなたのサービスアカウントの JSON キーは、以下のような形式になっているはずです：

```
{
 "type": "service_account",
 "project_id": "あなたのプロジェクト ID",
 "private_key_id": "あなたのプライベートキー ID",
 "private_key": "あなたのプライベートキー",
 "client_email": "あなたのサービスアカウントメール @ あなたのプロジェクト ID.iam.gserviceaccount.com",
 "client_id": "あなたのクライアント ID",
 "auth_uri": "https://accounts.google.com/o/oauth2/auth",
 "token_uri": "https://oauth2.googleapis.com/token",
 "auth_provider_x509_cert_url": "https://www.googleapis.com/oauth2/v1/certs",
 "client_x509_cert_url": "あなたのクライアント証明書 URL"
}
```

---

## Journey を選ぶ理由

Google Cloud Text-to-Speech はいくつかの音声を提供していますが、Journey はその自然で人間らしい音質で際立っています。ロボットのような音声が多い他のモデルとは異なり、Journey は

表現力とリアルな発話に優れています。特に、ポッドキャスト、オーディオブック、あるいは会話調のトーンを必要とするアプリケーションなど、長編コンテンツに適しています。

Journey の主な特徴：- 自然な音声：人間の声に近い音声を実現。- 表現力：文脈に基づいてトーンや抑揚を調整。- 長編コンテンツに最適：ポッドキャストやナレーションにぴったり。

Google Cloud Text-to-Speech の利点についての詳細は、Google Cloud の概要をご覧ください。

---

## ステップ 7: 新しいサービスアカウントキーを生成する（必要な場合）

もしあなたのサービスアカウントキーが上記の例と一致しない場合、Google Cloud Console から新しいキーを生成することができます。

新しいキーを生成するには: 1. Google Cloud Console にアクセスします。2. IAM & Admin > サービスアカウントに移動します。3. 新しいサービスアカウントを作成します: - 「サービスアカウントを作成」をクリックし、適切なロールを割り当てます。- 「作成」をクリックします。4. キーを生成します: - サービスアカウントを作成した後、そのアカウントをクリックします。- 「キー」タブに移動し、「キーを追加」をクリックして JSON を選択します。その後、キーをダウンロードします。

---

## サンプル音声出力

すべての設定が完了すると、音声ファイルを生成することができます。生成されたファイルは以下の場所で利用可能です：

音声ファイルをダウンロード。

---

## 結論

Google Cloud Text-to-Speech API は、テキストを自然な音声に変換することを簡単にします。音声出力を必要とするアプリを構築している場合や、テキスト読み上げを試している場合でも、この API は強力な機能とカスタマイズオプションを提供します。追加の音声選択、言語、および高度な機能については、完全なドキュメントを探索してください。