

会話音声生成

プロンプト:

このPDFについて少なくとも100ラウンドの会話をカバーし、このPDFに関するJSON形式の詳細を提供してください。

```
[  
  {  
    "speaker": "A",  
    "line": "最近、機械学習（ML）、ディープラーニング（DL）、そしてGPTについてよく聞くんだけど、詳しく教えてくれる？"  
  },  
  {  
    "speaker": "B",  
    "line": "もちろん！まずは基本から始めましょう。機械学習は、コンピュータがデータから学習し、明示的にプログラムされ  
  }  
]
```

コード:

```
import os  
import json  
import random  
import subprocess  
from google.cloud import texttospeech  
import tempfile  
import time  
import argparse  
  
# 会話の固定出力ディレクトリ  
OUTPUT_DIRECTORY = "assets/conversations"  
INPUT_DIRECTORY = "scripts/conversation"  
  
def text_to_speech(text, output_filename, voice_name=None):  
    print(f" 音声を生成中: {output_filename}")  
    try:  
        client = texttospeech.TextToSpeechClient()  
        synthesis_input = texttospeech.SynthesisInput(text=text)  
        if not voice_name:  
            voice_name = random.choice(["en-US-Journey-D", "en-US-Journey-F", "en-US-Journey-O"])  
        voice = texttospeech.VoiceSelectionParams(language_code="en-US", name=voice_name)
```

```

audio_config = texttospeech.AudioConfig(
    audio_encoding=texttospeech.AudioEncoding.MP3,
    effects_profile_id=["small-bluetooth-speaker-class-device"]
)

retries = 5

for attempt in range(1, retries + 1):
    try:
        response = client.synthesize_speech(input=synthesis_input, voice=voice, audio_config=audio_config)
        with open(output_filename, 'wb') as out:
            out.write(response.audio_content)
        print(f" 音声コンテンツを {output_filename} に書き込みました")
        return True
    except Exception as e:
        print(f" 試行 {attempt} でのエラー: {e}")
        if attempt == retries:
            print(f"{retries} 回試行しましたが、音声生成に失敗しました。")
            return False
        wait_time = 2 ** attempt
        print(f"{wait_time} 秒後に再試行します...")
        time.sleep(wait_time)

except Exception as e:
    print(f"{output_filename} の音声生成中にエラーが発生しました: {e}")
    return False

def process_conversation(filename):
    filepath = os.path.join(INPUT_DIRECTORY, filename)
    output_filename = os.path.join(OUTPUT_DIRECTORY, os.path.splitext(filename)[0] + ".mp3")

    if os.path.exists(output_filename):
        print(f" 音声ファイルは既に存在します: {output_filename}")
        return

    try:
        with open(filepath, 'r', encoding='utf-8') as f:
            conversation = json.load(f)
    except Exception as e:
        print(f" 会話ファイル {filename} の読み込み中にエラーが発生しました: {e}")
        return

```

```

temp_files = []

voice_name_A = random.choice(["en-US-Wavenet-D", "en-US-Wavenet-E", "en-US-Wavenet-F"])
voice_name_B = random.choice(["en-US-Studio-O", "en-US-Studio-M", "en-US-Studio-Q"])

for idx, line_data in enumerate(conversation):
    speaker = line_data.get("speaker")
    line = line_data.get("line")
    if not line:
        continue
    temp_file = os.path.join(OUTPUT_DIRECTORY, f"temp_{idx}.mp3")
    temp_files.append(temp_file)

    voice_name = None
    if speaker == "A":
        voice_name = voice_name_A
    elif speaker == "B":
        voice_name = voice_name_B

    if not text_to_speech(line, temp_file, voice_name=voice_name):
        print(f"{filename} の {idx+1} 行目の音声生成に失敗しました")
        # 一時ファイルをクリーンアップ
        for temp_file_to_remove in temp_files:
            if os.path.exists(temp_file_to_remove):
                os.remove(temp_file_to_remove)
        return

    if not temp_files:
        print(f"{filename} に対して音声が生成されませんでした")
        return

# ffmpeg を使用して連結
concat_file = os.path.join(OUTPUT_DIRECTORY, "concat.txt")
with open(concat_file, 'w') as f:
    for temp_file in temp_files:
        f.write(f"file '{os.path.abspath(temp_file)}'\n")

try:
    subprocess.run([
        ['ffmpeg', '-f', 'concat', '-safe', '0', '-i', concat_file, '-c', 'copy', output_filename],

```

```
    check=True,
    capture_output=True
)
print(f" 音声を正常に連結して {output_filename} に保存しました")
except subprocess.CalledProcessError as e:
    print(f" 音声連結中にエラーが発生しました: {e.stderr.decode()}")
finally:
    os.remove(concat_file)
    for temp_file in temp_files:
        os.remove(temp_file)

if __name__ == "__main__":
    parser = argparse.ArgumentParser(description=" 会話 JSON ファイルを処理して音声を生成します。")
    args = parser.parse_args()

    os.makedirs(OUTPUT_DIRECTORY, exist_ok=True)

    for filename in os.listdir(INPUT_DIRECTORY):
        if filename.endswith(".json"):
            process_conversation(filename)
```