



Transcrever Áudio - Google Cloud Speech-to-Text

Código Java – transcrever áudio mp3

- A API é otimizada para trabalhar com formatos como FLAC, LINEAR16 (WAV) e OGG OPUS, que têm menos perdas de qualidade.
- MP3 é um formato com perdas, e isso pode afetar a precisão da transcrição se a qualidade do áudio não for boa.

A estrutura geral do código que compartilhei está correta para a transcrição de um áudio usando o Google Cloud Speech-to-Text. No entanto, vale a pena garantir que o MP3 está em uma boa taxa de amostragem e com qualidade suficiente.

Codecs suportados -> Google Cloud Speech-to-Text

- **'LINEAR16'** (PCM sem compressão, comumente usado em arquivos WAV)
- **'FLAC'**
- **'MULAW'**
- **'AMR'**
- **'AMR_WB'**
- **'OGG_OPUS'**
- **'SPEEX_WITH_HEADER_BYTE'**
- **'MP3'**

Código Java

Como configurar o 'RecognitionConfig' para o formato MP3:

pom.xml:

```
<dependency>
  <groupId>com.google.cloud</groupId>
  <artifactId>google-cloud-speech</artifactId>
  <version>2.7.0</version>
</dependency>
```



Classe

```
import com.google.cloud.speech.v1p1beta1.*;
import com.google.protobuf.ByteString;
import java.nio.file.Files;
import java.nio.file.Path;
import java.nio.file.Paths;

public class AudioTranscrever {

    public static void main(String[] args) throws Exception {
        // Caminho para o arquivo de áudio MP3
        String audioFilePath = "caminho /audio.mp3";

        // Lê o arquivo de áudio em bytes
        Path path = Paths.get(audioFilePath);
        byte[] audioBytes = Files.readAllBytes(path);
        ByteString audioBytesStr = ByteString.copyFrom(audioBytes);

        // Configuração do cliente de Speech-to-Text
        try (SpeechClient speechClient = SpeechClient.create()) {
            // Configuração para reconhecer o áudio em formato MP3
            RecognitionConfig recognitionConfig = RecognitionConfig.newBuilder()
                .setEncoding(AudioEncoding.MP3) // Define o formato do arquivo como MP3
                .setSampleRateHertz(16000) // Defina a taxa de amostragem correta (ajuste
conforme necessário)
                .setLanguageCode("pt-BR") // Defina o idioma do áudio (português)
                .build();

            // Configura o áudio
            RecognitionAudio recognitionAudio = RecognitionAudio.newBuilder()
                .setContent(audioBytesStr)
                .build();

            // Faz a requisição para o Google Cloud
            RecognizeResponse recognizeResponse = speechClient.recognize(recognitionConfig,
recognitionAudio);

            // Exibe os resultados da transcrição
            for (SpeechRecognitionResult result : recognizeResponse.getResultsList()) {
                for (SpeechRecognitionAlternative alternative : result.getAlternativesList()) {
                    System.out.printf("Transcrição: %s%n", alternative.getTranscript());
                }
            }
        }
    }
}
```

- **Taxa de amostragem:** Verifique a taxa de amostragem do seu arquivo MP3. A API recomenda 16000 Hz ou superior para melhores resultados.
- **Dependências:** Certifique-se de ter adicionado corretamente a dependência do Google Cloud Speech-to-Text no seu 'pom.xml'.