



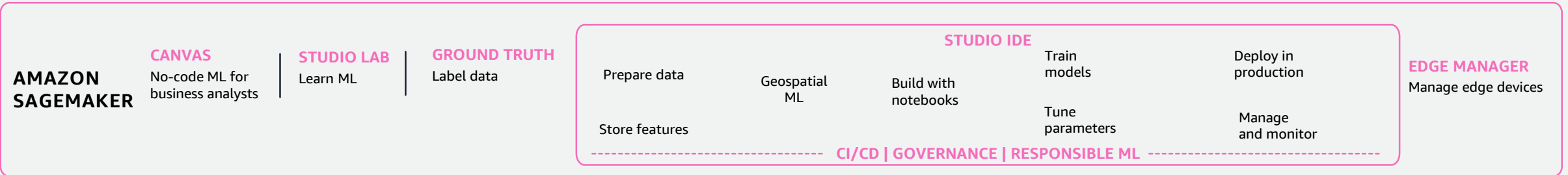
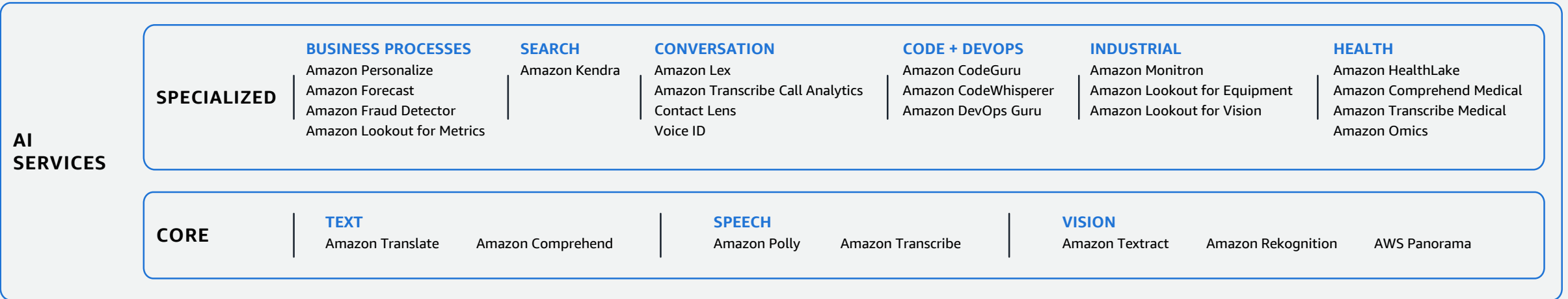
Amazon Transcribe

～音声をテキストに自動変換する～

秦 将之

Senior Solutions Architect
2023/05

AWS の提供する機械学習スタック



AWS の機械学習スタック

AI サービス : 機械学習の深い知識なしに利用可能なサービス群

AI SERVICES

SPECIALIZED

BUSINESS PROCESSES

Amazon Personalize
Amazon Forecast
Amazon Fraud Detector
Amazon Lookout for Metrics

SEARCH

Amazon Kendra

CONVERSATION

Amazon Lex
Amazon Transcribe Call Analytics
Contact Lens
Voice ID

CODE + DEVOPS

Amazon CodeGuru
Amazon CodeWhisperer
Amazon DevOps Guru

INDUSTRIAL

Amazon Monitron
Amazon Lookout for Equipment
Amazon Lookout for Vision

HEALTH

Amazon HealthLake
Amazon Comprehend Medical
Amazon Transcribe Medical
Amazon Omics

CORE

TEXT

Amazon Translate Amazon Comprehend

SPEECH

Amazon Polly **Amazon Transcribe**

VISION

Amazon Textract Amazon Rekognition AWS Panorama

AMAZON SAGEMAKER

CANVAS

No-code ML for business analysts

STUDIO LAB

Learn ML

GROUND TRUTH

Label data

Prepare data

Geospatial ML

Build with notebooks

Store features

STUDIO IDE

Train models

Tune parameters

Deploy in production

Manage and monitor

EDGE MANAGER
Manage edge devices

----- CI/CD | GOVERNANCE | RESPONSIBLE ML -----

ML FRAMEWORKS & INFRASTRUCTURE

PyTorch, Apache MXNet, TensorFlow

Amazon EC2

CPUs

GPUs

AWS Inferentia

AWS Trainium

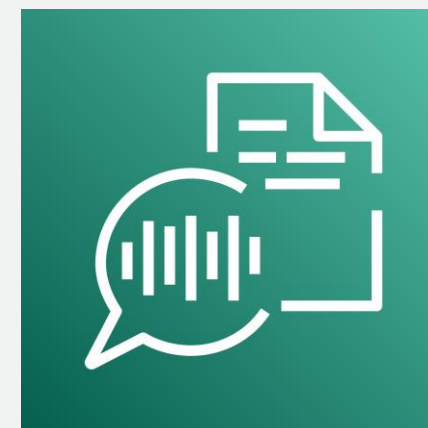
Habana Gaudi

FPGA



アジェンダ

1. Amazon Transcribeとは
2. オプション機能
3. 文字起こし精度の向上
4. コンプライアンスへの対応
5. アーキテクチャ例
6. まとめ



Amazon Transcribe

Amazon Transcribeとは

Amazon Transcribeとは



フルマネージドかつ継続的にトレーニングされた
自動音声認識（ASR）サービス

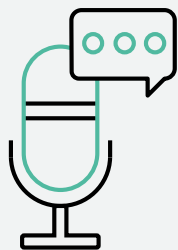
高精度、高効率、高スケーラブル

簡単に音声文字起こし機能と API を使って
音声対応アプリケーションの構築ができる

機械学習の経験は不要

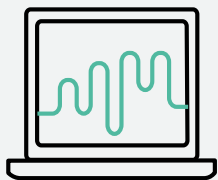
業界にまたがる共通課題

音声データの文字起こし



- ・ 大量の音声データの文字起こしとフォーマット変換（リアルタイム・バッチ）
- ・ 文字起こし精度の向上（マルチ言語対応・ボキャブラリー）
- ・ コンプライアンスへの対応

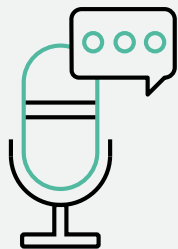
テキストデータの分析



- ・ テキストデータからのインサイト
- ・ 文章の属性やキーワードの抽出

業界にまたがる共通課題

音声データの文字起こし



- ・ 大量の音声データの文字起こしとフォーマット変換（リアルタイム・バッチ）
- ・ 文字起こし精度の向上（マルチ言語対応・ボキャブラリー）
- ・ コンプライアンスへの対応

テキストデータの分析

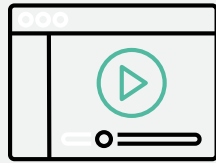


- ・ テキストデータからのインサイト
- ・ 文章の属性やキーワードの抽出

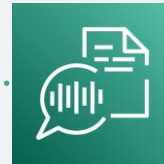
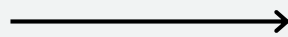
バッチ/ストリーミングの仕組み

バッチ処理

事前に記録された
メディア
バッチ処理

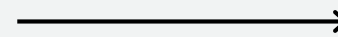


音声ファイル



Amazon Transcribe

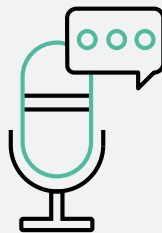
文字起こしファイル



ストリーミング処理

ライブストリーミング
コンテンツ

リアルタイム文字起こし
(HTTP2 or WebSocket)



音声ストリーム



音声取込

選択された
機能の適用

音声処理



テキストストリーム



サポートされている言語

2023年4月時点で39言語に対応（バッチ・ストリーミング両方に対応済なのは下記14言語）

言語	言語コード
中国語 (簡体字)	zh-CN
英語 (オーストラリア)	en-AU
英語 (イギリス)	en-GB
英語 (アメリカ)	en-US
フランス語	fr-FR
フランス語 (カナダ)	fr-CA
ドイツ語	de-DE
ヒンディー語 (インド)	hi-IN
イタリア語	it-IT
日本語	ja-JP
韓国語	ko-KR
ポルトガル語 (ブラジル)	pt-BR
スペイン語 (アメリカ)	es-US
タイ語	th-TH

データの入力について

encoding	バッチ	ストリーミング
サポートしているフォーマット	AMR, FLAC, M4A, MP3, MP4 Ogg, WebM, WAV	FLAC, Ogg Opus PCM encoding
推奨フォーマット	FLAC WAV (PCM 16bit encoding)	FLAC PCM signed 16bit little endian audio

データの出力について

トランスクリプション 完全なトランスクリプションをダウンロードする ストリーミングを開始

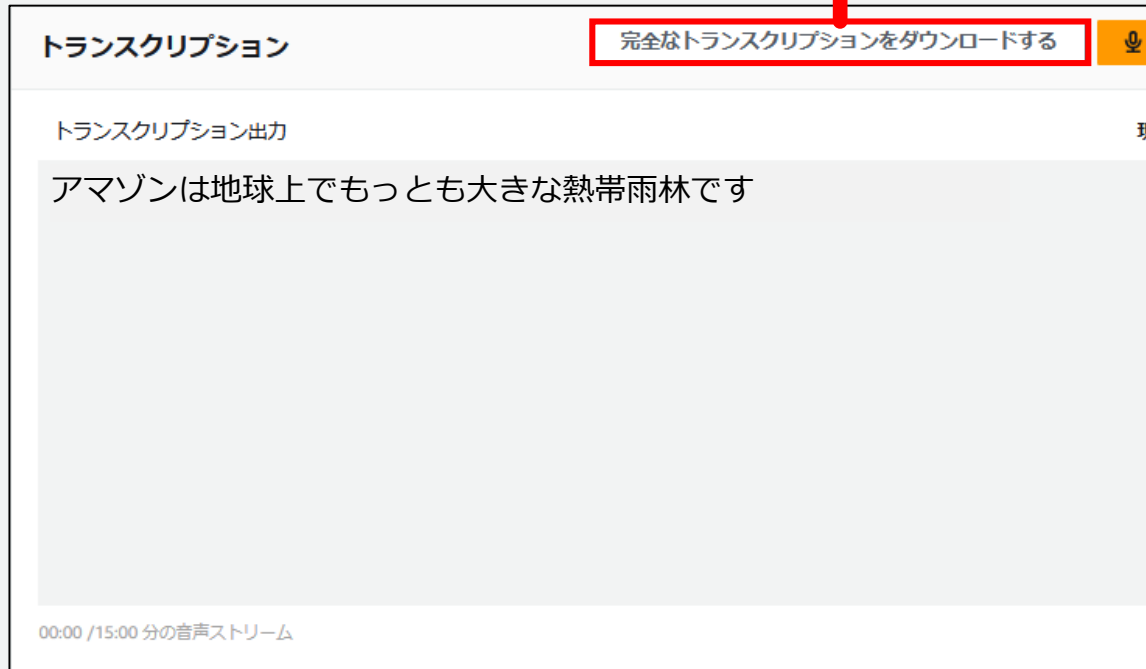
トランスクリプション出力 現在の言語: 日本語、日本

アマゾン地球でもっとも大きな熱帯雨林です

00:00 / 15:00 分の音声ストリーム

データの出力について

トランスクリプトの出力はJSON形式



トランスクリプション

完全なトランスクリプションをダウンロードする

トランスクリプション出力

アマゾン地球最も大きな熱帯雨林です

00:00 / 15:00 分の音声ストリーム

■“IsPartial”:=true (文字起こし未完了)

```
"Transcript": {
  "Results": [{
    "Alternatives": [{
      "Items": [...],
      "Transcript": "アマゾンは地球上で最も大きな"},
    ...
  ]
  "IsPartial": true,
  "ResultId": "XXXXXXXXXXXX",
  "StartTime": 0.025
```

■“IsPartial”: false (文字起こし完了)

```
"Transcript": {
  "Results": [{
    "Alternatives": [{
      "Items": [...],
      "Transcript": "アマゾンは地球上で最も大きな熱帯雨林です"}],
    ...
  ]
  "IsPartial": false,
  "ResultId": "XXXXXXXXXXXX",
  "StartTime": 0.045
```

データの出力について

トランスクリプトの出力はJSON形式

トランスクリプション

完全なトランスクリプションをダウンロードする

トランスクリプション出力

アマゾン地球でもっとも大きな熱帯雨林です

00:00 / 15:00 分の音声ストリーム

■“IsPartial”:=true (文字起こし未完了)

```
"Transcript": {
  "Results": [{
    "Alternatives": [{
      "Items": [...],
      "Transcript": "アマゾンは地球上で最も大きな"},
    ...
  ]
  "IsPartial": true,
  "ResultId": "XXXXXXXXXXXX",
  "StartTime": 0.025
```

■“IsPartial”: false (文字起こし完了)

```
"Transcript": {
  "Results": [{
    "Alternatives": [{
      "Items": [{"Confidence": 0.9531,
                  "Content": "アマゾン",
                  "EndTime": 0.7175,
                  "StartTime": 0.1075,
                  "Type": "pronunciation",
                  "VocabularyFilterMatch": false }
    ]
  ]
  "IsPartial": false,
  "ResultId": "XXXXXXXXXXXX",
  "StartTime": 0.025
```

データの出力について

バッチ処理の出力はJSON形式

JSON構成要素：

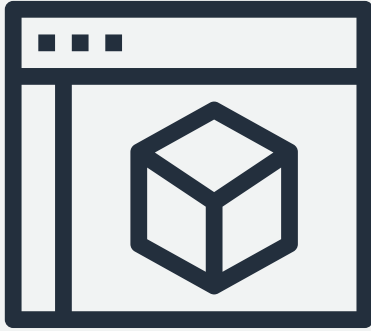
transcript: 文字起こし結果全体が含まれる

items: 各単語と句読点に関する情報が含まれる
開始時間/終了時間
単語・信頼スコア

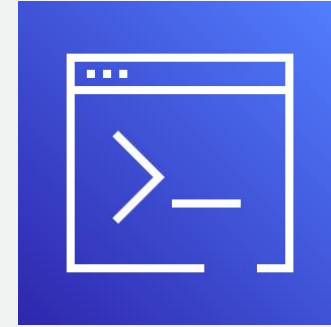
結果はS3バケットに保存
デフォルトバケットを選択した場合は90日で削除
バケット指定の場合は削除するまで保持

```
{
  "jobName": "my-first-transcription-job",
  "accountId": "XXXXXXXXXXXX",
  "results": {
    "transcripts": [
      {
        "transcript": "アマゾン熱帯雨林です"
      }
    ],
    "items": [
      {
        "start_time": "0.64",
        "end_time": "1.09",
        "alternatives": [
          {
            "confidence": "1.0",
            "content": "アマゾン"
          }
        ]
      },
      {
        "type": "pronunciation"
      },
      ...
    ]
  }
}
```

Amazon Transcribeの開始方法



AWS Management Console



AWS Command Line Interface (AWS CLI)

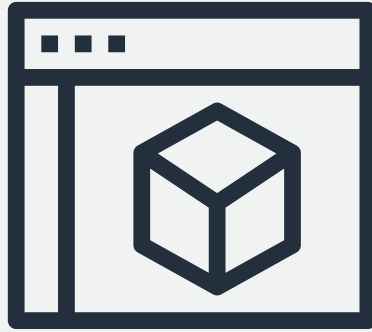


AWS SDK



HTTP or WebSocket

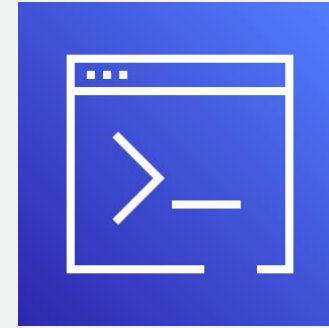
Amazon Transcribeの開始方法



ストリーミング

バッチ

AWS Management Console



バッチ

AWS Command Line Interface (AWS CLI)



ストリーミング

バッチ

AWS SDK



ストリーミング

HTTP/2, WebSocket

バッチ

HTTP/1.1

HTTP or WebSocket

Amazon Transcribeの開始方法

お試しは AWS マネジメントコンソールから利用するのが簡単

下記項目が必須

ジョブ名

音声の言語 (※)

S3 に配置した音声ファイル

文字起こし結果格納 S3 バケット

Input data [Info](#)

Input file location on S3
Choose an input audio or video file in Amazon S3.

[Browse S3](#)

Valid file formats: MP3, MP4, WAV, FLAC, AMR, OGG, and WebM.

Output data

Output data location type info

Service-managed S3 bucket **S3 バケット**
The output will be removed after 90 days when the job expires.

Customer specified S3 bucket
The output will not be removed from bucket even after the job expires.

Specify job details [Info](#)

Job settings

Name

⚠ Enter the job name.

The name can be up to 200 characters long. Valid characters are a-z, A-Z, 0-9, . (period), _ (underscore), and - (hyphen).

Model type [Info](#)

Choose the type of model to use for the transcription job.

General model

To use a model that is not specialized for a particular use case, choose this option. Configuration options vary between languages.

Custom language model

To use a model that you trained for your specific use case, choose this option. This model has fewer configuration options than the general model.

Language settings

You can transcribe your audio file in a language that you specify or have Amazon Transcribe identify and transcribe it in the predominant language.

Specific language [Info](#)

If you know the language spoken in your source audio, choose this option to get the most accurate results. The options available for additional processing vary between languages.

Automatic language identification [Info](#)

If you don't know the language spoken in your audio files, choose this option. You have access to fewer options for additional processing than if you choose **Specific language**.

Language

Choose the language of the input audio.

▶ **Additional settings**



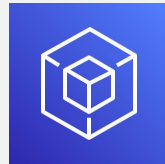
同様のことを自動化する (システムに組み込む) 場合はSDKを利用

© 2023, Amazon Web Services, Inc. or its affiliates.

Amazon Transcribe API の使い方

AWS SDKを用いたプログラムからの操作と、AWS CLIを用いたコマンドラインからの操作が可能

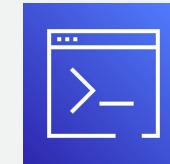
1. AWS SDK



```
import boto3
client = boto3.client('transcribe')
response = client.start_transcription_job(
    TranscriptionJobName='string',
    LanguageCode='ja-JP',
    MediaSampleRateHertz=8000,
    MediaFormat='mp3',
    Media={
        'MediaFileUri': 'string',
        'RedactedMediaFileUri': 'string'
    }
)
```

SDK for Python (boto3) の場合

2. AWS CLI



```
aws transcribe start-transcription-job
--transcription-job-name sample
--media-format mp3
--language-code ja-JP
--media MediaFileUri=s3://<bucket>/<file>.mp3
```

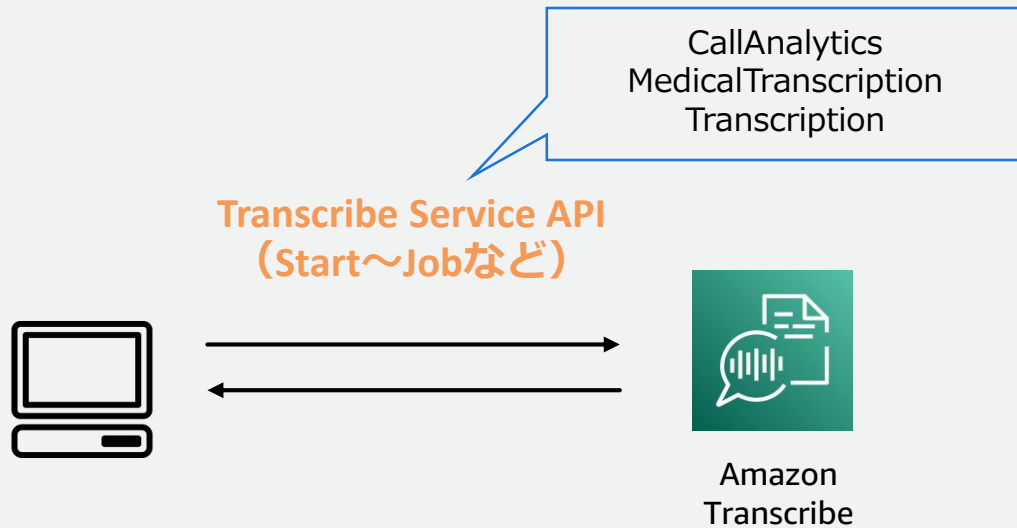
トランスクリプションジョブ実行時

<bucket>に音声データ格納先のS3バケット名、
<file>に音声データのパスを指定

Amazon Transcribe API 概要

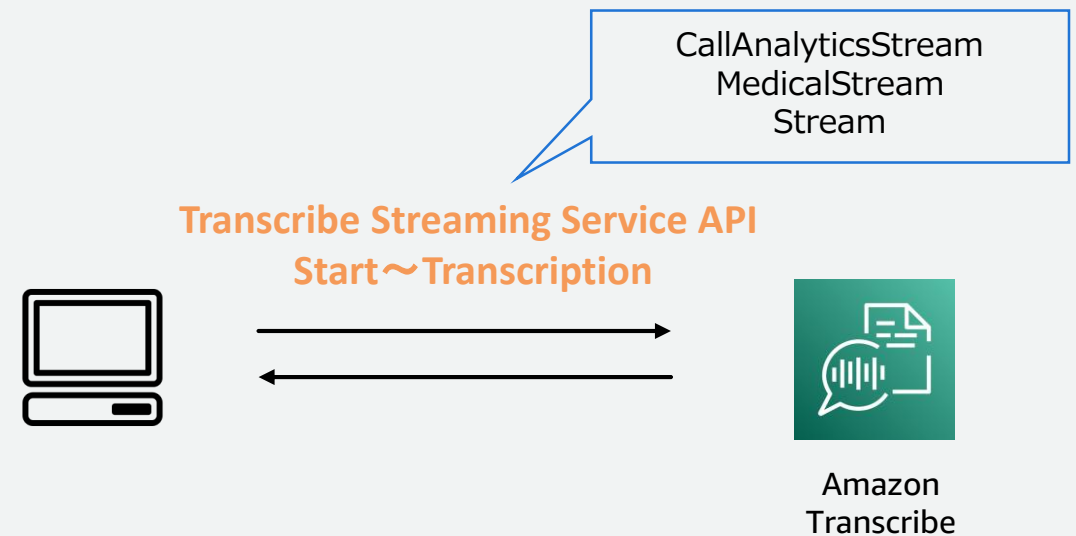
Amazon Transcribe Service

ストリーミング処理以外のアクションを定義



Amazon Transcribe Streaming Service

ストリーミング処理のアクションを定義



Amazon Transcribe API 一覧（抜粋）

その他のAPIの詳細や使い方についてはドキュメントをご覧ください。

	Amazon Transcribe Service	Amazon Transcribe Streaming Service
カスタム言語モデル作成	CreateLanguageModel	
カスタム語彙フィルター作成	CreateVocabularyFilter	
文字起こしジョブステータス確認	GetTranscriptionJob	
カスタム言語モデルのリスト取得	ListLanguageModels	
文字起こしジョブのリスト取得	ListTranscriptionJobs	
文字起こしジョブのリクエスト	StartTranscriptionJob StartCallAnalyticsJob StartMedicalTranscriptionJob	StartStreamTranscription StartCallAnalyticsStreamTranscription StartMedicalStreamTranscription
ボキャブラリーフィルター更新	UpdateVocabularyFilter	

Amazon Transcribe API 一覧（抜粋）

その他のAPIの詳細や使い方についてはドキュメントをご覧ください。

	Amazon Transcribe Service	Amazon Transcribe Streaming Service
カスタム言語モデル作成	CreateLanguageModel	
カスタム語彙フィルター作成	CreateVocabularyFilter	
文字起こしジョブステータス確認	GetTranscriptionJob	
カスタム言語モデルのリスト取得	ListLanguageModels	
文字起こしジョブのリスト取得	ListTranscriptionJobs	
文字起こしジョブのリクエスト	StartTranscriptionJob StartCallAnalyticsJob StartMedicalTranscriptionJob	StartStreamTranscription StartCallAnalyticsStreamTranscription StartMedicalStreamTranscription
ボキャブラリーフィルター更新	UpdateVocabularyFilter	

Amazon Transcribe Service リクエスト

StartTranscriptionJob(バッチ処理文字起こし)

メディアファイルを作成

有効な形式でメディアファイルをS3 にアップロード

APIエンドポイントにリクエスト

必須

MediaFileUri: メディアファイルのS3のパス

TranscriptionJobName: ジョブ名

Optional (抜粋)

LanguageCode: 言語コード (例: ja-JP)

MediaFormat: メディアファイル形式 (例: mp3)

OutputBucketName: 文字起こしを保存するS3バケット名

※指定しない場合はサービス管理のS3バケットに保存

リクエスト構文 (抜粋)

```
{ ...
  "LanguageCode": "string",
  ...
  "Media": {
    "MediaFileUri": "string",
    "RedactedMediaFileUri": "string"
  },
  "MediaFormat": "string",
  "MediaSampleRateHertz": number,
  "ModelSettings": {
    "LanguageModelName": "string"
  },
  "OutputBucketName": "string",
  ...
  "TranscriptionJobName": "string"
}
```

Amazon Transcribe Service レスポンス

StartTranscriptionJob(バッチ処理文字起こし)

レスポンス構文 (抜粋)

ステータスコードと共にJSON形式で返却

MediaFileUri: 文字起こし用メディアファイル場所

TranscriptionFileUrl: 文字起こしの出力先

```
{
  "TranscriptionJob": {
    ...
    "LanguageCode": "string",
    ...
    "Media": {
      "MediaFileUri": "string",
      "RedactedMediaFileUri": "string"
    },
    "MediaFormat": "string",
    ...
    "Transcript": {
      "RedactedTranscriptFileUri": "string",
      "TranscriptFileUri": "string"
    },
    "TranscriptionJobName": "string",
    "TranscriptionJobStatus": "string"
  }
}
```


Amazon Transcribe Streaming Service リクエスト

StartStreamTranscription(ストリーミング処理文字起こし)

リクエスト構文 (抜粋)

双方向HTTP/2, WebSocketストリームをスタート

URIリクエストパラメータ

必須

media-encoding: 入力オーディオのエンコーディング (例: pcm)

sample-rate: 入力音声のサンプルレート (Hz 単位)

```
POST /stream-transcription HTTP/2
x-amzn-transcribe-language-code: LanguageCode
x-amzn-transcribe-sample-rate: MediaSampleRateHertz
x-amzn-transcribe-media-encoding: MediaEncoding
...
Content-type: application/json

{
  "AudioStream": {
    "AudioEvent": {
      "AudioChunk": blob
    }
  }
}
```

Amazon Transcribe Streaming Service リクエスト

StartStreamTranscription(ストリーミング処理文字起こし)

リクエストボディ

必須

AudioStream: エンコードされたAudio BLOBのストリーミング

AudioChunk: オーディオチャンク

※最大サイズは 32 KB/Base64 でエンコード

リクエスト構文 (抜粋)

```
POST /stream-transcription HTTP/2
x-amzn-transcribe-language-code: LanguageCode
x-amzn-transcribe-sample-rate: MediaSampleRateHertz
x-amzn-transcribe-media-encoding: MediaEncoding
...
Content-type: application/json

{
  "AudioStream": {
    "AudioEvent": {
      "AudioChunk": blob
    }
  }
}
```

Amazon Transcribe Streaming Service レスポンス

StartStreamTranscription(ストリーミング処理文字起こし)

レスポンス構文 (抜粋)

レスポンス

Transcriptオブジェクト:文字起こしされたテキスト

```
HTTP/2 200
x-amzn-request-id: RequestId
x-amzn-transcribe-language-code: LanguageCode
x-amzn-transcribe-sample-rate: MediaSampleRateHertz
x-amzn-transcribe-media-encoding: MediaEncoding
x-amzn-transcribe-vocabulary-name: VocabularyName
x-amzn-transcribe-session-id: SessionId
...
Content-type: application/json
{
  ...
  "Transcript": "string"
  ...
}
```

オプション機能

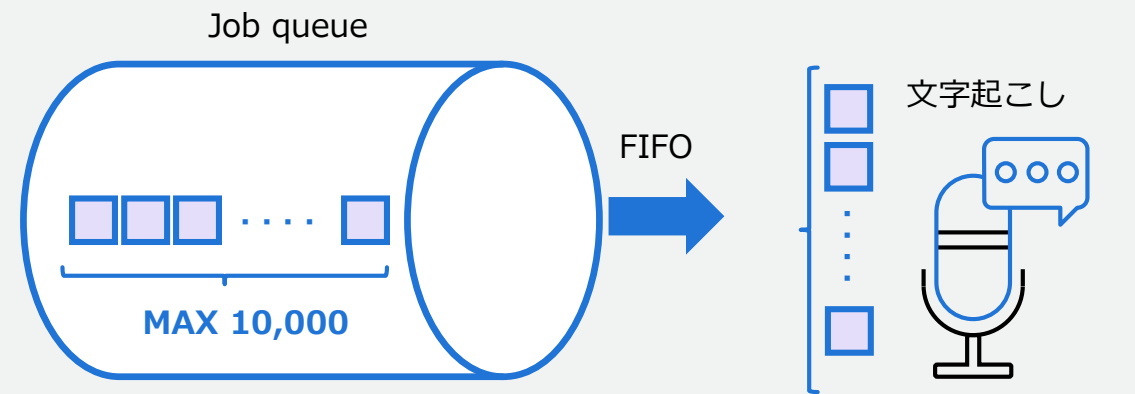
ジョブキューイング

大量の同時文字起こしジョブを実行する場合 同時実行クォータを超えてジョブを実行することが可能

同時実行クォータを使い果たすと
既存ジョブの完了を待って追加ジョブを送信

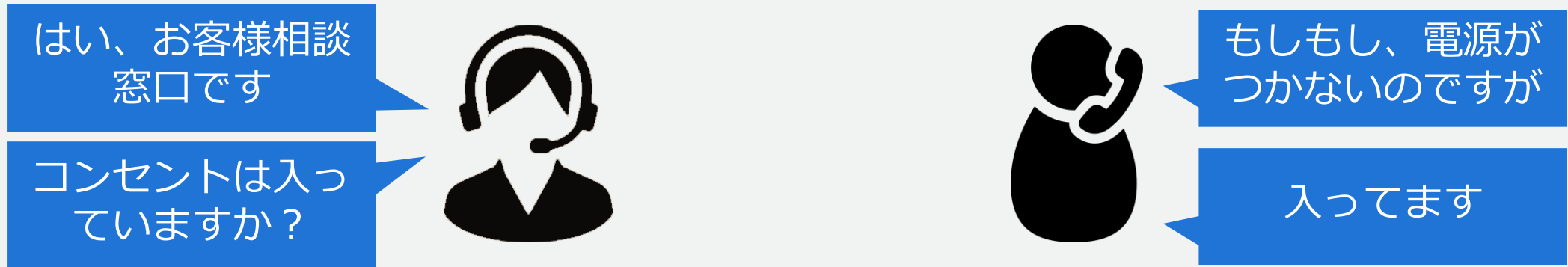


- ・最大 10,000 件のジョブをキュー送信
- ・ジョブはキューから FIFO 順 (先入れ先出し) で処理



マルチチャンネル

音声ファイルがステレオ録音(Amazon Connect はデフォルトでステレオ録音)されていれば2までチャンネル識別が可能



結果イメージ(抜粋)

```
[{"channel_label": "ch_0", "items": [{"content": "もしもし、電源がつかないのですが"}, {"content": "入ってます"}]}, {"channel_label": "ch_1", "items": [{"content": "はい、お客様相談窓口です"}, {"content": "コンセンタは入っていますか?"}]}
```

パーティショニングスピーカー (ダイアライゼーション)

文字起こし時に話者を区別できる

- ・最大 10 人のユニークな話者を区別
- ・話者のテキストに一意的な値 (spk_0~spk_9) のラベルを付けることが可能



結果イメージ(一部抜粋)

```
[{"speaker_label": "spk_0", "alternatives": [{"confidence": 1.0, "content": "今日は晴れそうです"}]},  
[{"speaker_label": "spk_1", "alternatives": [{"confidence": 1.0, "content": "出かけますか"}]},  
[{"speaker_label": "spk_2", "alternatives": [{"confidence": 1.0, "content": "そうしましょう"}]},
```

言語の識別

言語コードを指定しなくても話されている言語を自動的に識別

ストリーム処理

- ストリームごとに単一の言語を識別可能
- 最低3秒の音声が必要
- 2つ以上の言語コードをあらかじめ指定
- 指定した言語コードの一覧から優先言語を選択可能（オプション）

自動言語識別 情報
音声ファイルで話されている言語がわからない場合は、このオプションを選択します。

自動言語識別の言語オプション
言語識別の精度を向上させるには、少なくとも2つの言語オプションを選択します。

言語を選択してください

日本語、日本 (ja-JP) × 英語、米国 (en-US) × フランス語、フランス (fr-FR) ×

優先言語 - オプション
前に選択したものの中から、優先言語を1つ指定してください。

日本語、日本 (ja-JP)	▲
None	
日本語、日本 (ja-JP)	✓
英語、米国 (en-US)	
フランス語、フランス (fr-FR)	

あります。

※リダクションとカスタム言語モデルとの組み合わせは不可

言語の識別

言語コードを指定しなくても話されている言語を自動的に識別

多言語出力例(抜粋)

バッチ処理

- 単一言語でも、多言語でも識別可能
- 良好な結果を得るには30秒のメディアファイルが必要
- 話者が異なる言語を話したり、途中で話者が言語を変えても識別可能
- 識別精度を高めるために、メディアに登場する可能性が高い言語のリスト指定することも可能 (オプション)

```
"results": {
  "transcripts": [
    {
      "transcript": "welcome to Amazon transcribe. ये तो उदाहरण हैं क्या कैसे कर सकते हैं |一つのファイルに複数の言語を書き写す"
    }
  ], ...
  "language_codes": [
    {
      "language_code": "en-US",
      "duration_in_seconds": 2.45    },
    {
      "language_code": "hi-IN",
      "duration_in_seconds": 5.325  },
    {
      "language_code": "ja-JP",
      "duration_in_seconds": 4.15   }
  ]
}
```

※リダクションとカスタム言語モデルとの組み合わせは不可

言語の識別

言語コードを指定しなくても話されている言語を自動的に識別

多言語出力例(抜粋)

バッチ処理

- 単一言語でも、多言語でも識別可能
- 良好な結果を得るには30秒のメディアファイルが必要
- 話者が異なる言語を話したり、途中で話者が言語を変えても識別可能
- 識別精度を高めるために、メディアに登場する可能性が高い言語のリスト指定することも可能 (オプション)

```
"results": {
  "transcripts": [
    {
      "transcript": "welcome to Amazon transcribe. ये तो उदाहरण हैं क्या कैसे कर सकते हैं |一つのファイルに複数の言語を書き写す"
    }
  ], ...
  "language_codes": [
    {
      "language_code": "en-US",
      "duration_in_seconds": 2.45    },
    {
      "language_code": "hi-IN",
      "duration_in_seconds": 5.325  },
    {
      "language_code": "ja-JP",
      "duration_in_seconds": 4.15   }
  ]
}
```

※リダクションとカスタム言語モデルとの組み合わせは不可

ビデオ字幕の作成

字幕付き動画ファイルを出力

- ・ 字幕ファイル形式はSRT(SubRip)とVTT (webVTT) を選択
- ・ 字幕は発話と同時に表示



【字幕でサポートされる機能】

- ・ ダイアライゼーション : 複数のスピーカーを区別
- ・ リダクション : 個人情報をPIIでフィルタリング
- ・ 語彙フィルター : 単語のフィルタリング

文字起こし精度の向上

文字起こし精度の向上 ~カスタム語彙の作成~

Amazon Transcribeが未知の単語を登録して文字起こしをする

- リスト or テーブル形式がありテーブル形式を推奨
- AWSアカウントにつきカスタム語彙ファイルは100個、サイズ制限は50kB
- TAB区切りのテキストファイルをS3に配置して利用

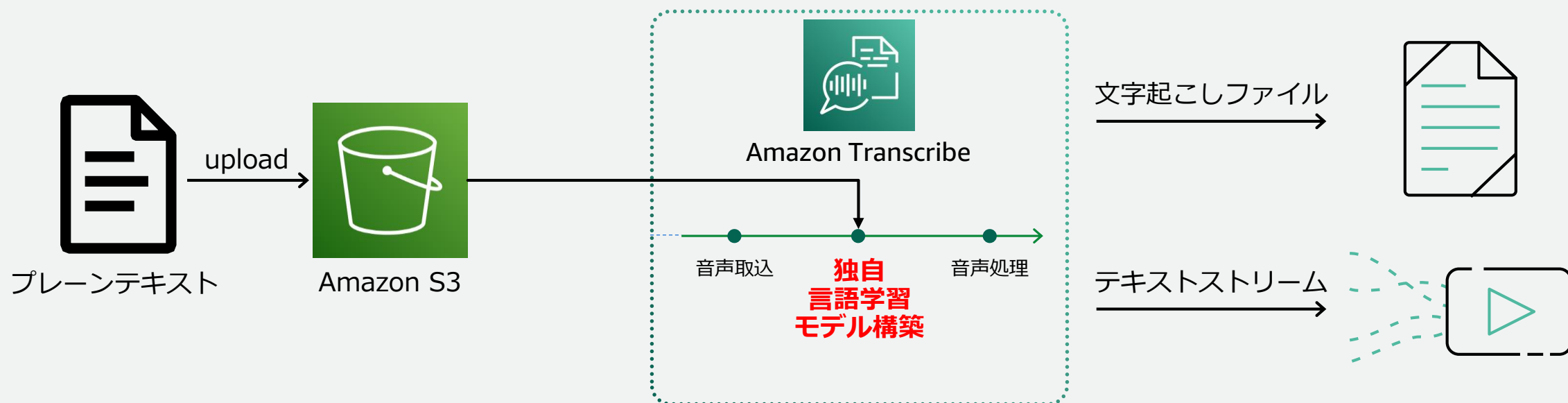
Phrase 文字起こしをしたい単語	IPA (option) 国際音声字母	SoundsLike (option) 発音方法	DisplayAs (option) 文字起こし後に表示 する文字
Los-Angeles	l ɔ s æ n dʒ ə l ə s		Los Angeles
A.B.C.-s		ay-bee-sees	ABCs
Dynamo-D.B.			DynamoDB

文字起こし精度の向上 ～カスタム言語モデル(CLM)の作成～

※CLM=Custom Language Model

テキストデータを活用してユースケースに合わせたカスタム音声エンジンを構築

- トレーニング用データとチューニング用データを取り込み可能
- テキストファイルをS3に配置して利用
- NarrowBand (SampleRate16,000Hz未満)/WideBand (SampleRateが16,000Hz以上) を指定
- 最大サイズは、トレーニングデータ用で 2 GB、チューニングデータ用で 200 MB



文字起こし精度の向上 ～カスタム言語モデル(CLM)の作成～

※CLM=Custom Language Model

トレーニングデータ

学習用データとしてオーディオに近いスクリプトを用意

- ・テキストデータ
- ・オーディオで話されている用語、使用法、文脈が含まれているとよい
- ・ない場合はオーディオと関連するドメインのスクリプトデータを利用

チューニングデータ

ドメインのトレーニングデータが不足（10,000語未満）しており期待した精度がでない場合、調整用データを用意し学習モデルをチューニング

※トレーニングデータとチューニングデータは重複不可

コンプライアンスへの対応

単語のフィルタリング

ボキャブラリーフィルターを使用して特定の単語を隠す

文字起こしテキスト（フィルタ前）

Transcribeは音声をテキストに自動変換します。

Transcribe を使用して、重要な会議や会話をキャプチャします

Transcribeをご利用ください

カスタム
ボキャブラリーフィルター

“Transcribe”
という単語を
フィルタ

文字起こしテキスト（フィルタ後）

***は音声をテキストに自動変換します。

***を使用して、重要な会議や会話をキャプチャします

***をご利用ください

単語のフィルタリング

① マスク

"You can specify a list of *** or *** words, and *** *** removes them from transcripts automatically."

② 削除

"You can specify a list of or words, and removes them from transcripts automatically."

③ タグ

"You can specify a list of profane or offensive words, and amazon transcribe removes them from transcripts automatically."

```
...  
{  
  "confidence": "1.0",  
  "content": "profane"  
},  
"type": "pronunciation",  
"vocabularyFilterMatch": true
```

個人情報の編集（リダクション）

※現在英語のみ対応

個人を特定できる情報（PII）をマスキングまたは消去する

PII タイプ	説明
BANK_ACCOUNT_NUMBER	米国の銀行口座番号
BANK_ROUTING	米国の銀行口座ルーティング番号
CREDIT_DEBIT_NUMBER	クレジットカードまたはデビットカードの番号
CREDIT_DEBIT_CVV	クレジットカードやデビットカードのカード認証コード (CVV)
CREDIT_DEBIT_EXPIRY	クレジットカードまたはデビットカードの有効期限
PIN	個人識別番号 (PIN)
NAME	個人の名前
ADDRESS	住所
PHONE	電話番号
EMAIL	メールアドレス
SSN	社会保障番号 (SSN)

個人情報の編集 (リダクション)

※現在英語のみ対応

ストリーミング処理

- ・ 英語(en-AU/en-GB/en-US)で利用可能
- ・ “識別のみ” or “識別とマスキング”を選択

```
{...
  "Transcript": {
    "Results": [
      ...
      "Transcript": "my name is mike",
      ...
      "Entities": [
        {
          "Content": "mike",
          "Category": "PII",
          "Type": "NAME",
          "StartTime" : 0.7999375,
          "EndTime" : 1.0199375,
          "Confidence": 0.9989
        }
      ]
    }
  }
}
```

識別

```
{...
  "Transcript": {
    "Results": [
      ...
      "Transcript": "my name is [NAME]",
      ...
      "Entities": [
        {
          "Content": "[NAME]",
          "Category": "PII",
          "Type": "NAME",
          "StartTime" : 0.7999375,
          "EndTime" : 1.0199375,
          "Confidence": 0.9989
        }
      ]
    }
  }
}
```

マスキング

識別

個人情報の編集（リダクション）

※現在英語のみ対応

バッチ処理

- 英語 (en-US) で利用可能
- 編集済み/未編集両方のトランスクリプトをS3に出力

```
{ ...
  "results": {
    "transcripts": [
      {
        "transcript": "Good morning, everybody. My name is [PII], and today I feel like
sharing a whole lot of personal information with you. Let's start with my Social
Security number [PII]. My credit card number is [PII] and my C V V code is [PII].
I hope that Amazon Transcribe is doing a good job at redacting that personal
information away. Let's check."
      }
    ],
    ...
  }
}
```

料金

※ 2023 年 4 月時点の東京リージョンの場合。
詳細や最新の情報は[料金ページ](#)をご覧ください。

スタンダードストリーミング

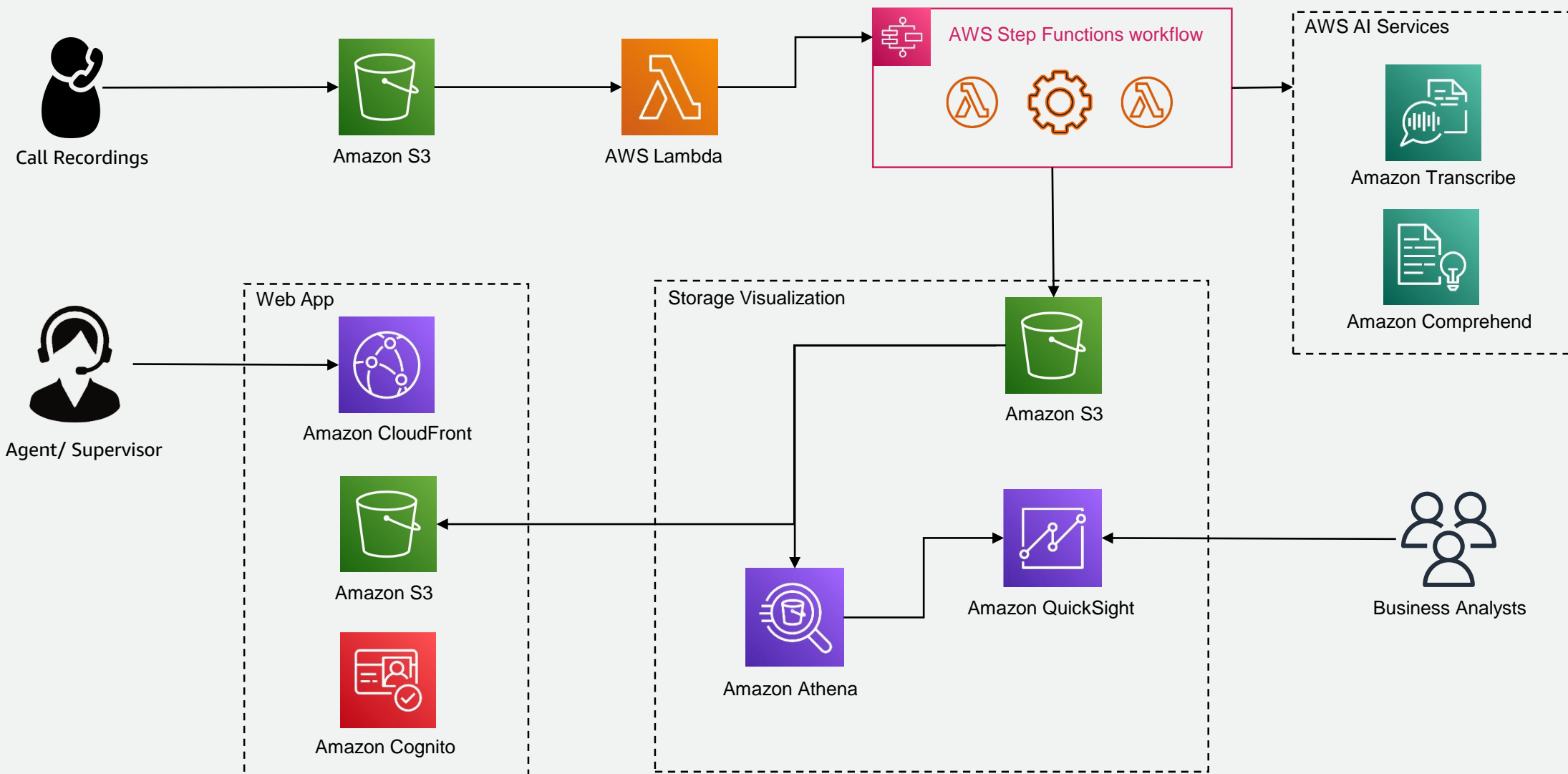
階層	ボリューム (分/月)	スタンダードストリーミング文字起こし (USD/分)*
T1	最初の 250,000 分	0.02400USD
T2	次の 750,000 分	0.01500USD
T3	1,000,000 分超え	0.01080USD

スタンダードバッチ

階層	ボリューム (分/月)	スタンダードバッチ文字起こし (USD/分)*
T1	最初の 250,000 分	0.02400USD
T2	次の 750,000 分	0.01500USD
T3	1,000,000 分超え	0.01080USD

アーキテクチャ例

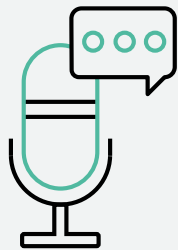
アーキテクチャ例



本動画のまとめ

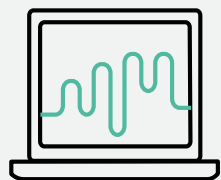
Amazon Transcribeの概要と使用方法について解説

音声データの文字起こし



- ・ 大量の音声データの文字起こしとフォーマット変換（リアルタイム・バッチ）
- ・ 文字起こし精度の向上（マルチ言語対応・ボキャブラリー）
- ・ コンプライアンスへの対応

テキストデータの分析



- ・ テキストデータからのインサイト
- ・ 文章の属性やキーワードの抽出

今後公開予定の
Amazon Comprehendへ

AWS Black Belt Online Seminar とは

- 「サービス別」「ソリューション別」「業種別」などのテーマに分け、アマゾン ウェブ サービス ジャパン合同会社が提供するオンラインセミナーシリーズです
- AWS の技術担当者が、AWS の各サービスやソリューションについてテーマごとに動画を公開します
- 以下の URL より、過去のセミナー含めた資料などをダウンロードすることができます
 - <https://aws.amazon.com/jp/aws-jp-introduction/aws-jp-webinar-service-cut/>
 - <https://www.youtube.com/playlist?list=PLzWGOASvSx6FIwIC2X1nObr1KcMCBBIqY>



ご感想は Twitter へ！ハッシュタグは以下をご利用ください
#awsblackbelt

内容についての注意点

- 本資料では 2023 年 4月時点のサービス内容および価格についてご説明しています。AWS のサービスは常にアップデートを続けているため、最新の情報は AWS 公式ウェブサイト (<https://aws.amazon.com/>) にてご確認ください
- 資料作成には十分注意しておりますが、資料内の価格と AWS 公式ウェブサイト記載の価格に相違があった場合、AWS 公式ウェブサイトの価格を優先とさせていただきます
- 価格は税抜表記となっております。日本居住者のお客様には別途消費税をご請求させていただきます
- 技術的な内容に関しましては、有料の [AWS サポート窓口](#)へお問い合わせください
- 料金面でのお問い合わせに関しましては、[カスタマーサポート窓口](#)へお問い合わせください (マネジメントコンソールへのログインが必要です)



Thank you!