

このコンテンツは公開から3年以上経過しており内容が古い可能性があります  
最新情報については[サービス別資料](#)もしくはサービスのドキュメントをご確認ください



# 【AWS Black Belt Online Seminar】

## Amazon AI 入門

Amazon Polly / Amazon Rekognition / Amazon Lex

アマゾン ウェブサービス ジャパン株式会社

ソリューションアーキテクト 川村 誠

ソリューションアーキテクト 布目 拓也

2017.03.29

Archived

# 自己紹介

- 名前：川村 誠
- 所属：アマゾン ウェブ サービス ジャパン株式会社  
ストラテジックソリューション部  
ソリューションアーキテクト
- 好きなAWSのサービス：Amazon EMR



# AWS Black Belt Online Seminar とは

AWSJのTechメンバがAWSに関する様々な事を紹介するオンラインセミナーです

## 【火曜 12:00~13:00】

主にAWSのソリューションや  
業界カッタでの使いどころなどを紹介  
(例：IoT、金融業界向け etc.)

## 【水曜 18:00~19:00】

主にAWSサービスの紹介や  
アップデートの解説  
(例：EC2、RDS、Lambda etc.)



※開催曜日と時間帯は変更となる場合がございます。

最新の情報は下記をご確認下さい。

オンラインセミナーのスケジュール&申し込みサイト

– <https://aws.amazon.com/jp/about-aws/events/webinars/>

# 内容についての注意点

- 本資料では2017年3月29日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。最新の情報はAWS公式ウェブサイト(<http://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が東京リージョンを使用する場合、別途消費税をご請求させていただきます。
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

# Agenda

- Amazon AI
- Amazon Polly
- Amazon Rekognition
- Amazon Lex
- まとめ





# Amazon AI

Intelligent Services Powered By Deep Learning

Amazon  
Rekognition

Amazon  
Polly

Amazon  
Lex

AI Services

Amazon  
Machine Learning

Amazon  
EMR

Spark &  
Spark ML

AI Platforms

Apache  
MXNet

TensorFlow

Caffe

Torch

Theano

CNTK

Keras

AI Engines

# Amazon AI: AI Services

## Deep Learning を利用した新しい 3 つのサービス



### Polly

文章をリアルな音声に変換するサービス



### Rekognition

画像分析機能をアプリケーションに簡単に追加できるようにするサービス



### Lex

声やテキストを使用した会話型インターフェイスを様々なアプリケーションに構築するためのサービス

※これらは、fine-tuning された AWS が提供する深層学習モデルを API で簡単に利用できるマネージドサービスとなります

# Amazon Polly

# Polly – 概要

- テキストをリアルな音声に変換するサービス
- 24 の言語で 47 のリアルな声優の音声を提供
- 低レイテンシーで応答が速いため、リアルタイムシステムを構築する際の選択肢となりえる
- 生成された音声の保存とリプレイ、配信が可能

# Polly – 対応言語一覧

ロシア、ヨーロッパ、中東  
及び、アフリカ地域：

- Danish
- Dutch
- British English
- French
- German
- Icelandic
- Italian
- Norwegian
- Polish
- Portuguese
- Romanian
- Russian
- Spanish
- Swedish
- Turkish
- Welsh
- Welsh English

北アメリカ、南アメリカ地域：

- Brazilian Portuguese
- Canadian French
- English (US)
- Spanish (US)

アジア太平洋地域：

- Australian English
- Indian English
- 日本語



# Polly – 品質

## □ 自然に聞こえる音声

音声変換（TTS）の出力がどれくらい人間の声に近いかに関する尺度 

## □ テキスト変換処理の正確さ

略語、数字の羅列、ホモグラフ（同一つづりで発音が異なる語彙）等のような一般的なテキストを解釈するシステムの能力

❖ *Today in Las Vegas, NV it's 90°F.* 

❖ *"We live for the music", live from the Madison Square Garden.* 

## □ 高度なわかりやすさ

どれくらい音声がわかりやすいかに関する尺度

”庭には二羽の鶏がいる” 

# Polly – Text-to-Speech 処理

# Polly – Text-to-Speech 処理

テキスト

Market grew by > 20%.

# Polly – Text-to-Speech 処理

テキスト処理

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

# Polly – Text-to-Speech 処理

単語

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

# Polly – Text-to-Speech 处理

音素

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

  
'mɑ: kət

  
'gru

  
baɪ

  
'mɔ: θæn

  
'twɛn.ti  
pə. 'sɛnt

# Polly – Text-to-Speech 处理

音素

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

  
'mɑ: kət

  
'gɹu

  
baɪ

  
'mɔ: θæn

  
'twɛn.ti  
pə.'sɛnt

韻律曲線



# Polly – Text-to-Speech 処理

音素

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

'mɑ: .kət

'gru

baɪ

'mɔ:ʊ  
ðæn

'twɛn.ti  
pə: .sɛnt

ユニットの選択と適応

音声ユニット  
インベントリ

# Polly – Text-to-Speech 处理

音素

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

  
'mɑ: kət

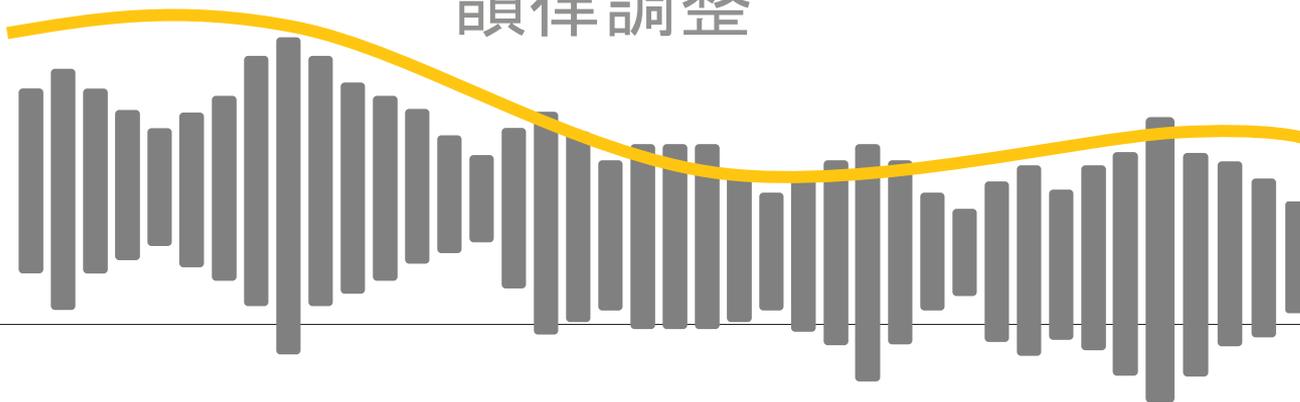
  
'gɹu

  
baɪ

  
'mɔ: θæn

  
'twɛn.ti  
pə.'sɛnt

韻律調整



# Polly – Text-to-Speech 処理

音素

Market grew by > 20%.

Market

grew

by

more  
than

twenty  
percent

  
'mɑ: kət

  
'gru

  
baɪ

  
'mɔ: θæn

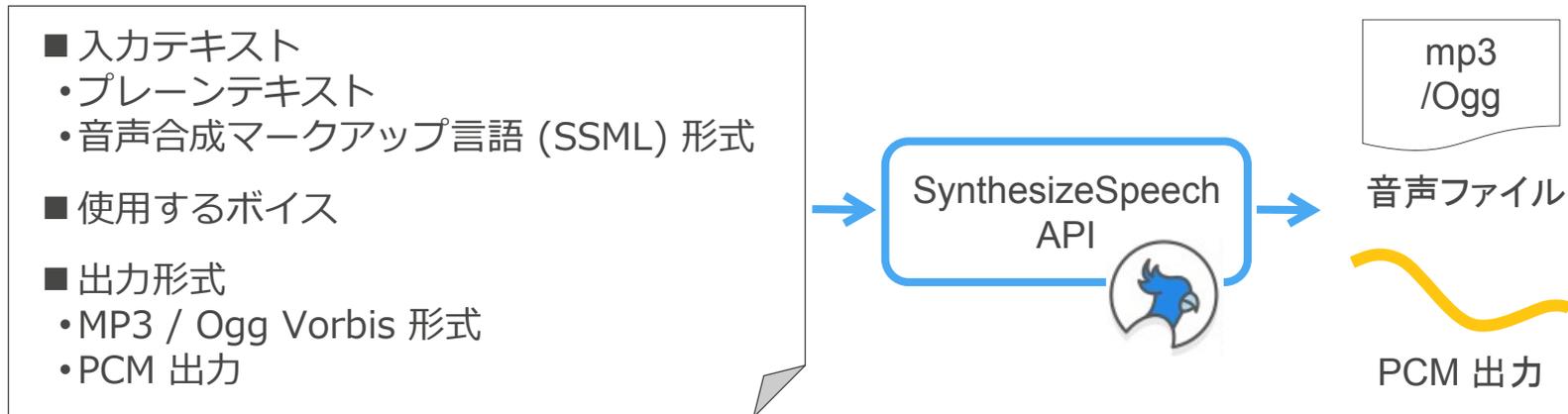
  
'twɛn.ti  
pə.'sɛnt

ストリーミング



# Polly – 仕組み

- テキストを入力し、読み上げる(TTS) 声と音声出力形式を指定し、SynthesizeSpeech API を呼び出すだけで、入力されたテキストが高品質の音声ストリームに変換される
- 2つの仕様(SSML とレキシコン)を利用可能な機能を提供することでより柔軟な音声合成を実現できる



# Polly – 機能: SSML

- Speech Synthesis Markup Language (SSML) v1.1 に準拠したドキュメントを入力テキストとして利用できる
- SSML タグを使用することで、発音、ボリューム、話す速度など、音声のさまざまな要素をカスタマイズできる

```
<speak>
```

```
My name is Kuklinski. It is spelled
```

```
<prosody rate='x-slow'>
```

```
<say-as interpret-as="characters">Kuklinski</say-as>
```

```
</prosody>
```

```
</speak>
```



# Polly – 機能: レキシコン

- ❑ Pronunciation Lexicon Specification(PLS) v1.0 に準拠したドキュメントをレキシコンとして登録し、SynthesizeSpeech API で利用することができる
- ❑ レキシコンとして『単語とフレーズのマッピング』や『言語で一般的でない単語の発音』を定義し、発音をカスタマイズすることができる

## 単語とフレーズのマッピング例)

**W3C** is a Consortium.

```
<lexeme>  
  <grapheme>W3C</grapheme>  
  <alias>World Wide Web Consortium</alias>  
</lexeme>
```

- ❖ 単語(grapheme) に対応するフレーズを alias として定義する
- ❖ 複数の単語に対し、1つの alias を設定することができる

# Polly – 機能: レキシコン

- Pronunciation Lexicon Specification(PLS) v1.0 に準拠したドキュメントをレキシコンとして登録し、SynthesizeSpeech API で利用することができる
- レキシコンとして『単語とフレーズのマッピング』や『言語で一般的でない単語の発音』を定義し、発音をカスタマイズすることができる

言語で一般的でない単語発音のカスタマイズ例)

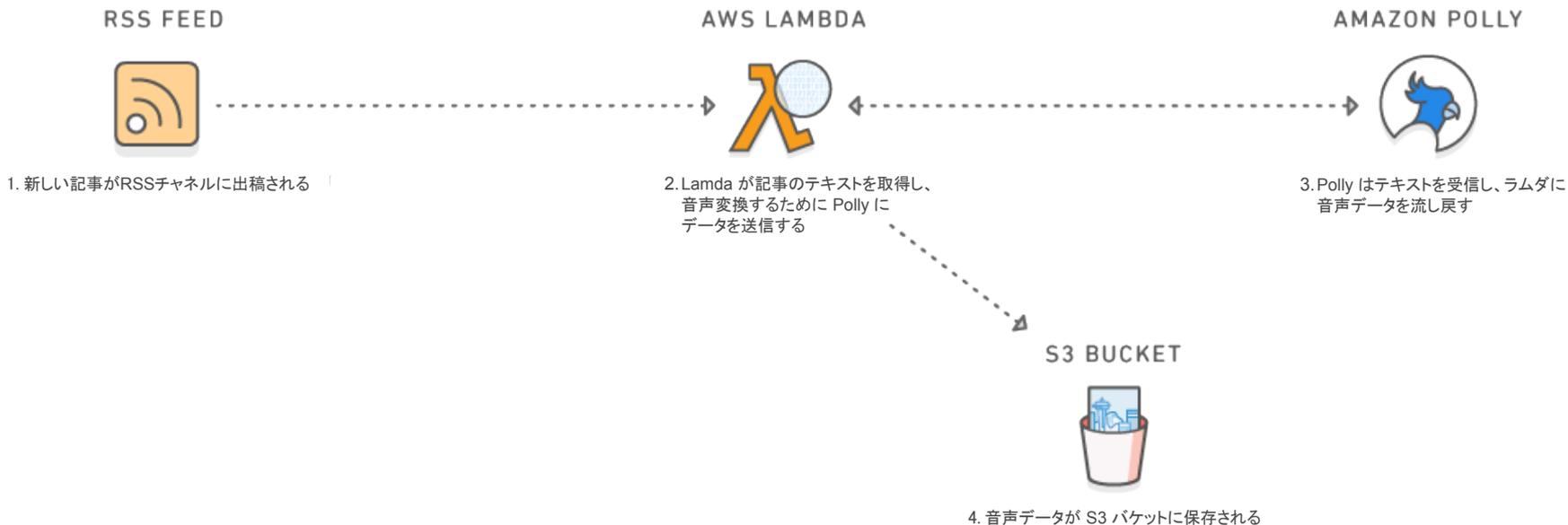
*My daughter's name is **Kaja**.*

```
<lexeme>
  <grapheme>Kaja</grapheme>
  <grapheme>kaja</grapheme>
  <grapheme>KAJA</grapheme>
  <phoneme>"kaI.@</phoneme>
</lexeme>
```



- ❖ 単語(grapheme) に対応する発音(phoneme) を定義する
- ❖ 複数の単語に対し、1つの発音を設定することができる

# Polly - ユースケース： 記事を音声に変換して MP3 でダウンロードする



# Polly – 制限事項: SynthesizeSpeech API

- 入力テキストのサイズは、最大 1500 課金対象文字 (合計 3000 文字) です。SSML タグは、課金対象文字としてカウントされません。
- 入力テキストに適用する最大 5 個のレキシコンを指定できます。
- 出力オーディオストリーム (合成) は 5 分に制限されます。その時間を過ぎると、残りの音声は切り取られます。

# Polly – 利用価格と提供リージョン

- 従量課金制
- 100万文字あたり 4.0 USD
- Amazon Polly が生成した音声ファイルをファイルに保存しておけば、追加料金なしで再利用できる
- 無償利用枠：  
最初の音声リクエストから 12 か月間、1 か月あたり 500 万文字
- サービス提供リージョン  
US East (N. Virginia/Ohio), US West (Oregon); EU (Ireland)

# Amazon Rekognition

# Rekognition - 概要

- 深層学習に基づく画像認識サービス
- 画像の分析機能をアプリケーションに簡単に追加できる

## 【提供機能】



物体とシーンの検出



顔分析



顔照合



顔認識

Amazon Rekognition は Amazon S3 や AWS Lambda といったサービスとシームレスに動作するように設計されている

# Rekognition - 物体とシーンの検出: DetectLabels API

- DetectLabels API を利用すると、画像から識別した車、ペット、家具など、数千もの物体にラベルを付け、信頼スコアを取得できる
- 信頼スコアは 0~100 の値で示され、識別結果が正しいかどうかの可能性を意味する

- 大規模な画像ライブラリを検索、フィルタリング、管理するために必要な情報を取得できる
- 写真、不動産、旅行アプリなど様々なアプリケーションに利用可能



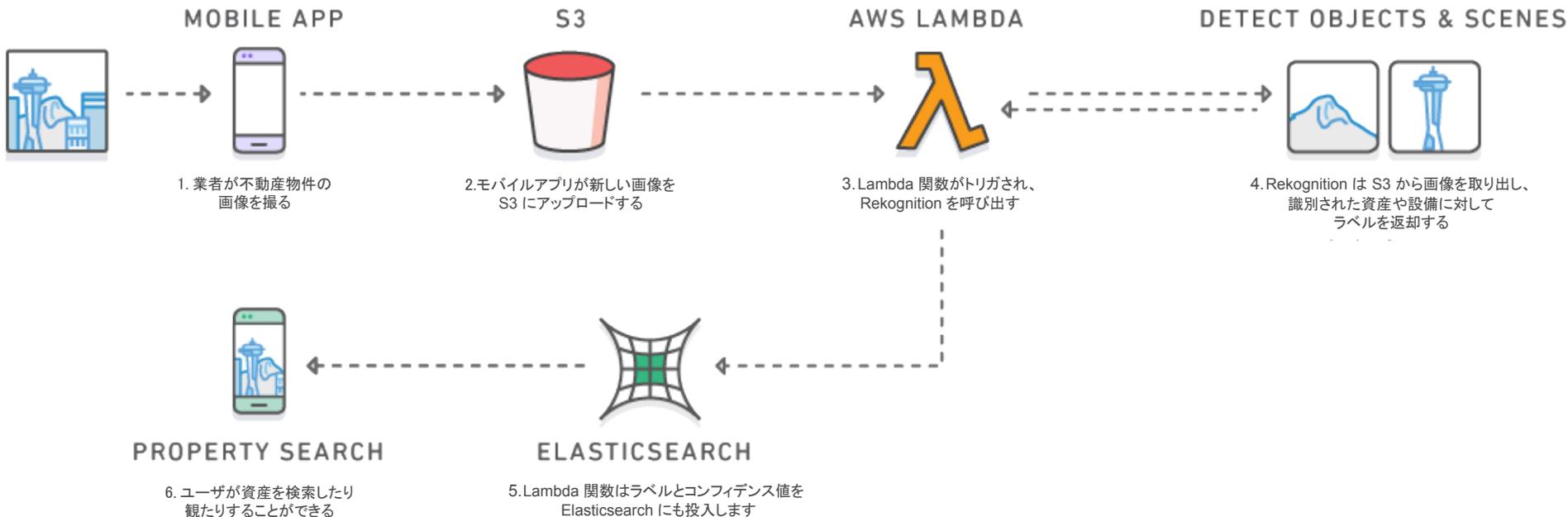
# Rekognition - 物体とシーンの検出: DetectLabels API



**DetectLabels**

```
{
  "Confidence": 94.62968444824219,
  "Name": "adventure"
},
{
  "Confidence": 94.62968444824219,
  "Name": "boat"
},
{
  "Confidence": 94.62968444824219,
  "Name": "rafting"
},
. . .
```

# DetectLabels API - ユースケース: 不動産物件の検索



# Rekognition - 顔分析: DetectFaces API

□ DetectFaces API を利用すると、画像内の顔の位置を検出し、感情、ポーズ、瞳が開いているかどうかなどの顔属性を分析できる

- 画像のトリミングや広告を重ねる際に顔を避ける
- ユーザのどのようなカテゴリの人物か、ユーザが抱いている感情などをとらえることができます
- 一番良い写真を推薦する
- 動的でパーソナライズされた広告を配信する



# Rekognition - 顔分析: DetectFaces API

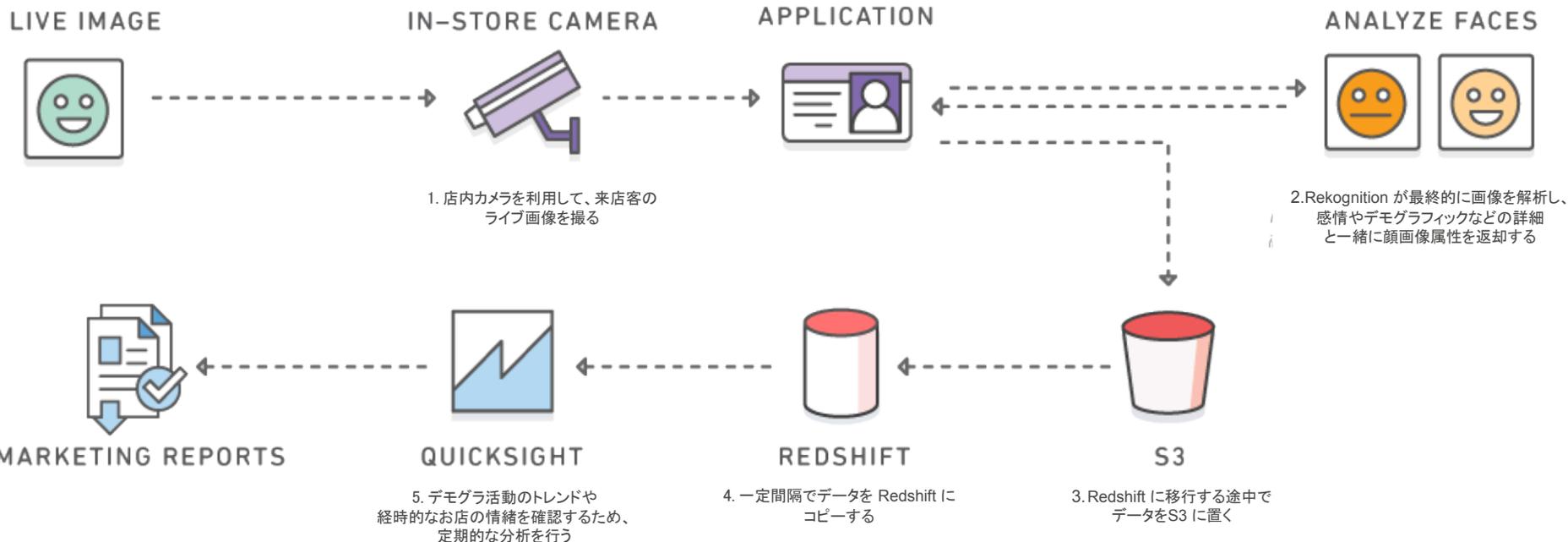


DetectFaces



```
[
  {
    "BoundingBox": {
      "Height": 0.3449999988079071,
      "Left": 0.09666666388511658,
      "Top": 0.27166667580604553,
      "Width": 0.23000000417232513
    },
    "Confidence": 100,
    "Emotions": [
      {"Confidence": 99.1335220336914,
        "Type": "HAPPY" },
      {"Confidence": 3.3275485038757324,
        "Type": "CALM"},
      {"Confidence": 0.31517744064331055,
        "Type": "SAD"}
    ],
    "Eyeglasses": {"Confidence": 99.8050537109375,
      "Value": false},
    "EyesOpen": {"Confidence": 99.99979400634766,
      "Value": true},
    "Gender": {"Confidence": 100,
      "Value": "Female"}
  }
]
```

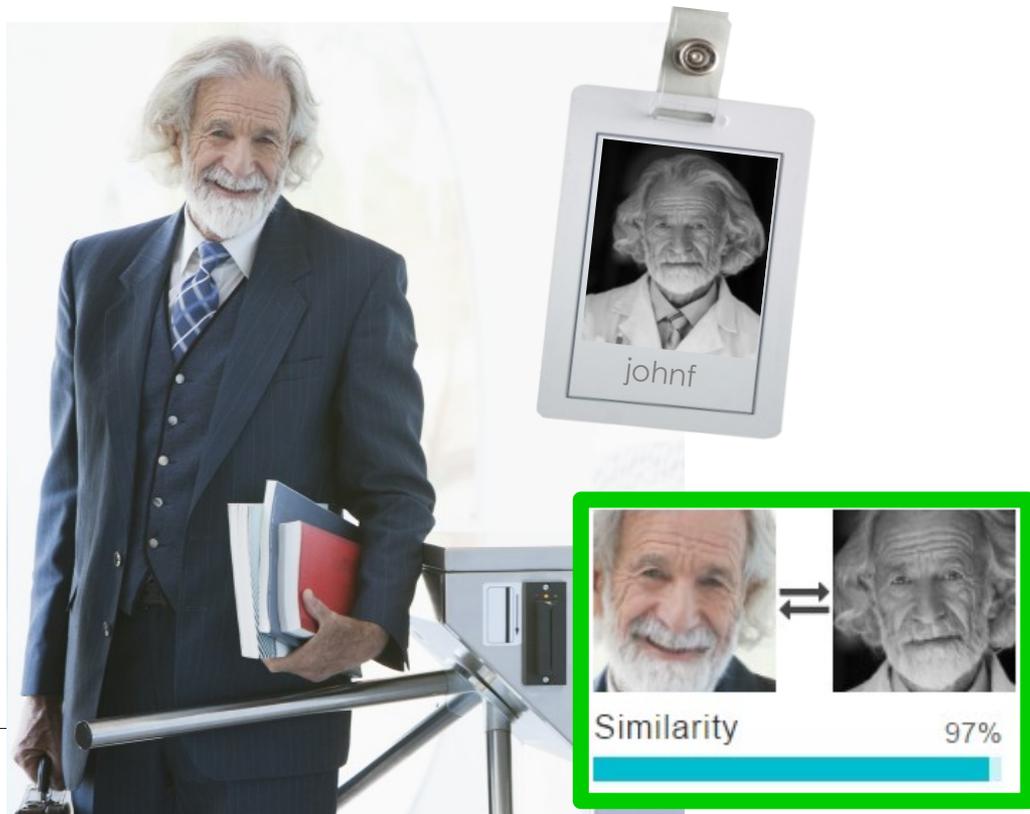
# DetectFaces API - ユースケース: 店舗内の顧客印象分析



# Recognition - 顔の比較: CompareFaces API

CompareFaces API を利用すると、2つの画像の顔が同一人物である可能性を測定できます

- アプリケーションやデバイスに顔認識機能を追加します
- 物理的なセキュリティ制御を拡張します
- オンラインの試験や投票でユーザを識別します



# Recognition - 顔の比較: CompareFaces API



CompareFaces

```
{
  "FaceMatches": [
    { "Face": { "BoundingBox": {
      "Height": 0.2683333456516266,
      "Left": 0.5099999904632568,
      "Top": 0.1783333271741867,
      "Width": 0.17888888716697693},
      "Confidence": 99.99845123291016},
      "Similarity": 96
    },
    { "Face": { "BoundingBox": {
      "Height": 0.2383333295583725,
      "Left": 0.6233333349227905,
      "Top": 0.3016666769981384,
      "Width": 0.15888889133930206},
      "Confidence": 99.71249389648438},
      "Similarity": 0
    }
  ],
  "SourceImageFace": { "BoundingBox": {
    "Height": 0.23983436822891235,
    "Left": 0.28333333134651184,
    "Top": 0.351423978805542,
    "Width": 0.1599999964237213},
    "Confidence": 99.99344635009766
  }
}
```

# CompareFaces API - ユースケース: 従業員認証

## IMAGE CAPTURE



## AUTHENTICATED USER

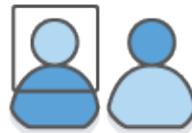
ユーザ認証

## APPLICATION



1. 従業員が社員証をスキャンする際、アプリケーションが従業員のライブ画像を撮る
4. 類似度が80%以上を超えていれば、アプリケーションは OK を返し、そうでない場合はアラートをあげ、警備員に検査を促す

## COMPARE FACES



3. Rekognition がライブ画像と社員証画像を比較し、類似度を返却する

ライブ画像 + 社員証画像 S3 URL

類似度(99%)



S3

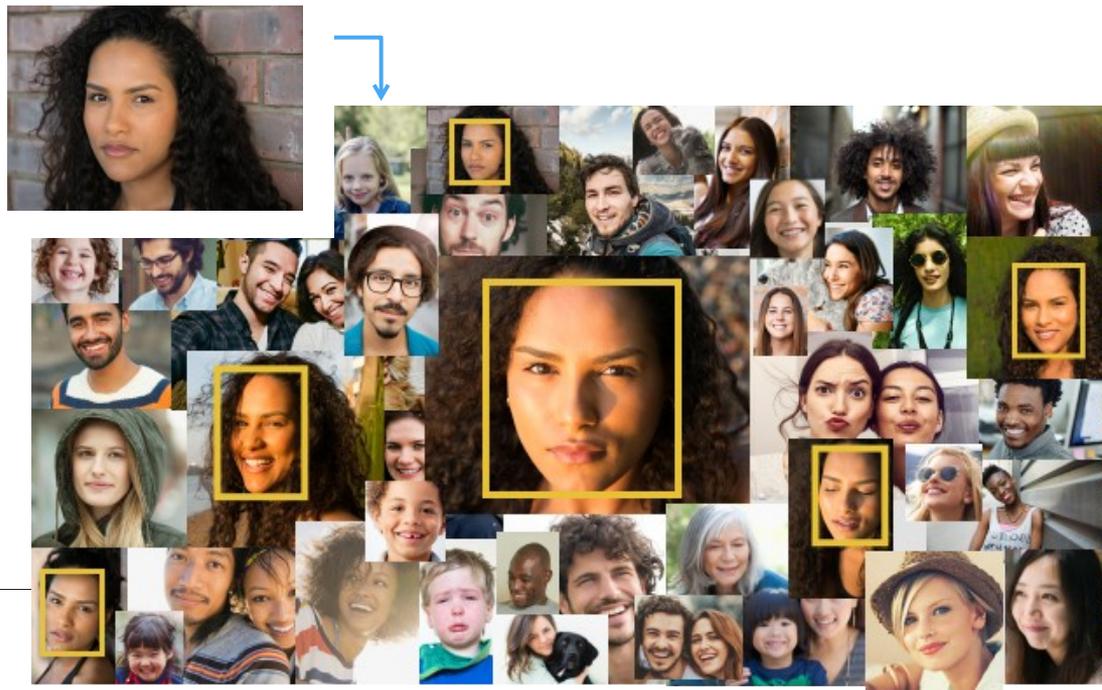
2. アプリケーションは S3 からユーザの社員証画像を取り出す

# Rekognition - 顔認識:

## IndexFaces/SearchFacesByImage API

IndexFaces API と SearchFacesByImage API を利用すると、大規模な顔のコレクションの中から似た顔を見つけることで画像の中にいる人物を識別することができます

- ソーシャルやメッセージングアプリに友人のタグ付を追加します
- 行方不明の人を写真から探す手助けをします
- 歴史的な画像アーカイブの中に有名人を確認します



# Rekognition - 顔認識:

## IndexFaces/SearchFacesByImage API

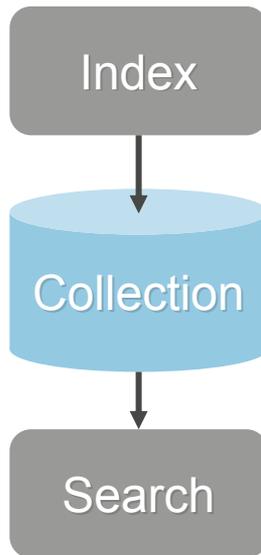
### 顔認識

- 画像コレクションの顔にインデックスを付け、顔画像を検索することができる
- 顔のインデックス化は指定された画像から直接生成するだけでなく、Amazon S3 に保存された画像を処理する事も可能



**IndexFaces**

**SearchFacesByImage**



# Recognition – IndexFaces API



**IndexFaces**

```
{  
  f7a3a278-2a59-5102-a549-a12ab1a8cae8,  
  02e56305-1579-5b39-ba57-9afb0fd8782d,  
  4c55926e-69b3-5c80-8c9b-78ea01d30690  
}
```

Face



transformed

Face ID & vector<float>

f7a3a278-2a59-5102-a549-a12ab1a8cae8  
&  
vector001

02e56305-1579-5b39-ba57-9afb0fd8782d  
&  
vector002

4c55926e-69b3-5c80-8c9b-78ea01d30690  
&  
vector003

stored

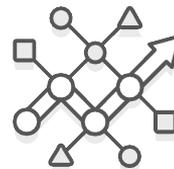
Collection

# Rekognition – SearchFacesbyImage API

Face



SearchFacesbyImage



Nearest neighbor  
search

Collection

```
{...  
FaceID: 4c55926e-69b3-5c80-8c9b-78ea01d30690  
Similarity: 97  
...}
```

# Rekognition - ユースケース: 友人の画像を見つける

REFERENCE IMAGE



SEARCH FACES BY IMAGE



FACE COLLECTION



1. Rekognition は友人の画像にマッチするかどうか顔コレクションを検索し、類似度の順に一致する可能性のある顔のメタデータの配列を返却する

PERSONAL DETAILS TABLE

#0123	#0123
5426	128762
78426	45871
286546	26751
3861	945



END USER

PHOTO APP

S3

4. ユーザは友人の画像に類似する写真を見る

3. 写真アプリケーションが検索結果をユーザに表示する

2. 元画像が要求された場合、元画像が S3 から取り出される

# Rekognition - Lambda 関数設計図(ブループリント)の提供

## 設計図の選択

設計図は、イベントソースと Lambda 関数のサンプル設定です。希望するシナリオに最も近い設計図を選び、必要に応じてカスタマイズします。または、Lambda 関数の作成とイベントソースの設定を別々に行う場合、このステップは省略できます。特に記載がない限り、設計図は CC0 でライセンスされています。

ランタイムの選択 ▼

▼ rekognition|

フィルタに  
“rekognition”  
と入力

ています > >>

### ブランク関数

関数を最初から設定します。トリガーを定義し、ウィザードに従ってコードをデプロイします。

カスタム

### rekognition-python

An Amazon S3 trigger that uses rekognition APIs to detect faces

python2.7 · rekognition · s3



# Rekognition - Lambda 関数テンプレートの提供

## 設計図の選択

設計図は、イベントソースから選択できます。または、Lambda 関数テンプレートでライセンスされています。

ランタイムの選択

## ブランク関数

関数を最初から設定し、定義し、ウィザードにプロイします。

カスタム

## トリガーの設定

関数を呼び出すトリガーを追加できます。



バケット demo-rekognition

イベントタイプ Put

プレフィックス images/

サフィックス jpg

Lambda 関数のトリガーとなるイベントを選択します。オプションで、イベントのプレフィックスやサフィックスを設定できます。しかし、それぞれのバケットの個々のイベントには、同一のオブジェクトキーを照合するプレフィックスやサフィックスを含んだ複数の設定はできません。

Lambda は、このバケットから Lambda 関数を呼び出すのに必要な Amazon S3 のアクセス許可を追加します。Lambda アクセス許可モデルの詳細については [こちら](#) を参照してください。

トリガーの有効化

じてカスタマイズし  
がない限り、設計図は

表示しています > >>

# Rekognition - Lambda 関数テンプレートの提供

## 設計図の選択

設計図は、イベントソースから選択できます。または、Lambda 関数テンプレートでライセンスされています。

ランタイムの選択

## ブランク関数

関数を最初から設定し、定義し、ウィザードにプロイします。

カスタム

## トリガーの設定

関数を呼び出すトリガーを選択してください。

### 関数の設定

Lambda 関数は、実行するカスタムコードで構成されます。Lambda 関数の詳細については、[こちら](#)を参照してください。

名前\* callRekognition

説明 An Amazon S3 trigger that uses rekognition

ランタイム\* Python 2.7

### Lambda 関数のコード

関数のコードを指定します。コードに (boto3 以外の) カスタムライブラリが必要でない場合は、コードをインラインで編集できます。必要な場合は、コードとライブラリを ZIP ファイルとしてアップロードすることができます。

コードエントリタイプ コードをインラインで編集

```
1 from __future__ import print_function
2
3 import boto3
4 from decimal import Decimal
5 import json
6 import urllib
7
8 print('Loading function')
9
10 rekognition = boto3.client('rekognition')
11
12
13 # ----- Helper Functions to call Rekognition APIs -----
14
15
```

下記 API を簡単に呼び出せるようなコードが定義されている

- ❖ DetectLabels API
- ❖ DetectFaces API
- ❖ IndexFaces API

# Rekognition - 制限事項

- ❑ Amazon S3 に保存されている参照画像の最大サイズは15MBで、最小解像度の高さや幅は80ピクセル
- ❑ API のパラメータとして引き渡し可能な素画像サイズの最大値は5MB
- ❑ サポート画像フォーマットは PNG と JPEG
- ❑ コレクションに保存可能な顔の数の最大数は100万件
- ❑ 顔画像検索最大数は4096件

# Rekognition - 利用価格と提供リージョン

画像分析枠	処理画像 1,000 枚あたりの料金
1 か月あたり画像処理* 100 万枚まで	\$1.00
1 か月あたり画像処理* 100 万枚超 1,000 万枚まで	\$0.80
1 か月あたり画像処理* 1,000 万枚超 1 億枚まで	\$0.60
1 か月あたり画像処理* 1 億枚超	\$0.40
顔メタデータストレージ	
1 か月あたりに保存される顔メタデータ 1,000 件	\$0.01

\* 各 API で 1 枚以上の入力画像を受信した場合に、画像処理 1 枚としてカウントします。

- ❑ 無償利用枠: 最初の 12 か月の間、1 か月あたり 5,000 枚の画像分析を行い、毎月 1,000 件の顔メタデータを保存できます
- ❑ サービス提供リージョン:  
US East (N. Virginia), US West (Oregon), EU (Ireland)

# Amazon Lex

# Lex – 概要

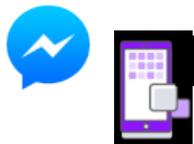
- 音声やテキストを使用して、任意のアプリケーションに対話型インターフェイス(ボット)を構築するサービス
  - AWS プラットフォームのセキュリティ、モニタリング、ユーザー認証、ビジネスロジック、ストレージ、モバイルアプリケーション開発を実現するためのスイート
    - ❖ AWS Lambda、AWS MobileHub、および Amazon CloudWatch との組み込み統合を提供
    - ❖ Amazon Cognito や Amazon DynamoDB など 他の多くのサービスと簡単に統合できる
- ※ 現時点で Lex は limited Preview のサービスとなります。利用には下記 リンクより、プレビュープログラムへのサインアップが必要となります。
- <https://pages.awscloud.com/AmazonLexPreviewLandingPage.html>

# Lex – 特徴

自動音声認識  
(ASR)

自然言語理解  
(NLU)

高度な深層学習機能に基づく、音声を変換するための自動音声認識 (ASR) とテキストの意図を理解するための自然言語理解 (NLU) を利用可能



モバイルデバイス、ウェブアプリケーション、および Facebook Messenger (Slack と Twilio とは近日中に統合予定) などのチャットサービスに簡単にパブリッシュできます。



開発者向けにデザインされており、会話を構築する効率的で直観的なツールを提供



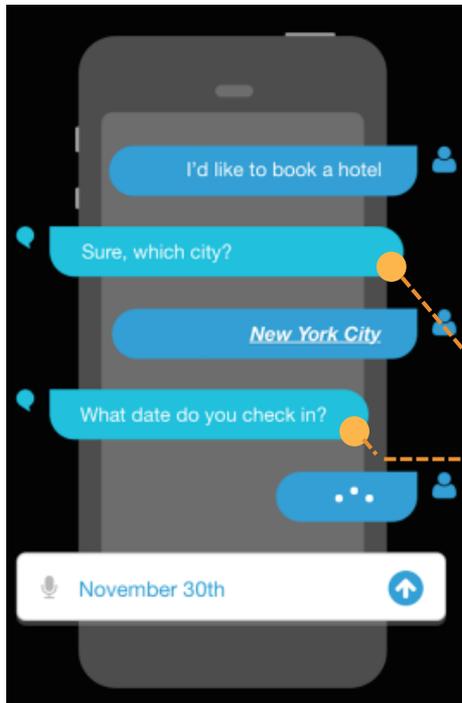
バージョン管理とバージョンに対するalias(“本番”, “開発”, “テスト”)設定機能を提供



エンタープライズシステムに接続可能なコネクタを提供

# Lex – 仕組み: ボットの構成要素

## BookHotel



### Intents

intent はユーザが入力した自然言語に  
応答して fulfillment を実行する

### Utterances

intent を発動する口頭、もしくは、  
入力されるフレーズ

### Slots

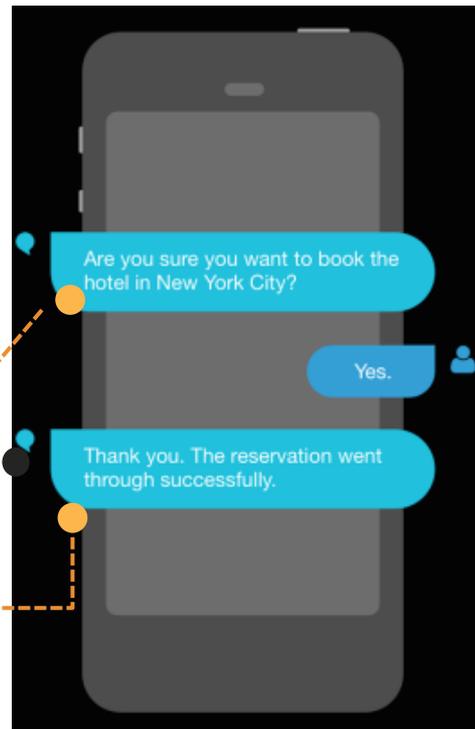
Slot は intent を満たすために  
要求される入力データ

### Prompt

Slot を引き出すためのフレーズ

### Fulfillment

intent を実現する  
ビジネスロジック



# Lex – 仕組み: Utterances

I'd like to book a hotel

I want to make my hotel reservations

Can you help me book my hotel?

I want to book a hotel in New York

# Lex – 仕組み: Slot

Slot	Type	Values
Destination	City	New York City, Seattle, London, ...
CheckIn	Date	Valid dates
CheckOut	Date	Valid dates

# Lex – 仕組み: Prompt

I'd like to book a hotel

Sure what city do you want to book?

New York City

City

What date do you check in?

Nov 30th

Check In

# Lex – 仕組み: Fulfillment



- ビジネスロジックを実行する  
為、Intent と slot が Lambda  
関数に渡される

**AWS Lambda  
Integration**



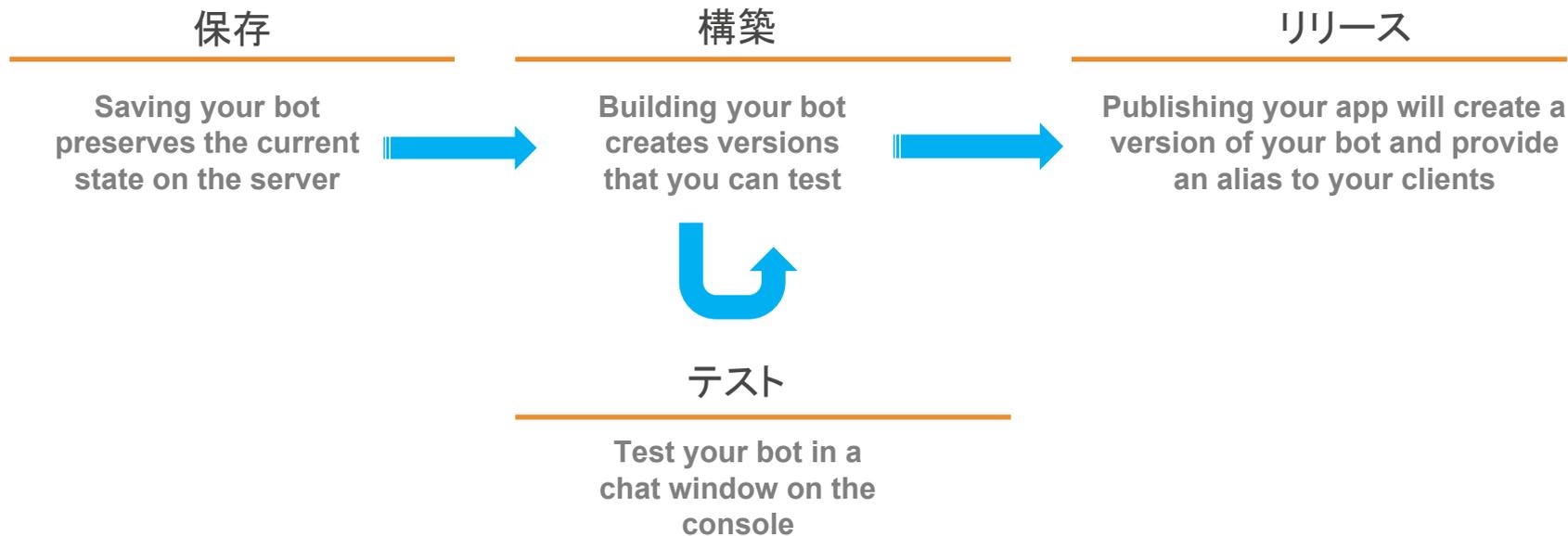
- ユーザの入力が intent と slot 値  
を得るためにパースされる
- パース結果が後続処理を実行す  
るため、クライアントに返却される

**Return to  
Client**

# Lex – 仕組み: “Book a Hotel”

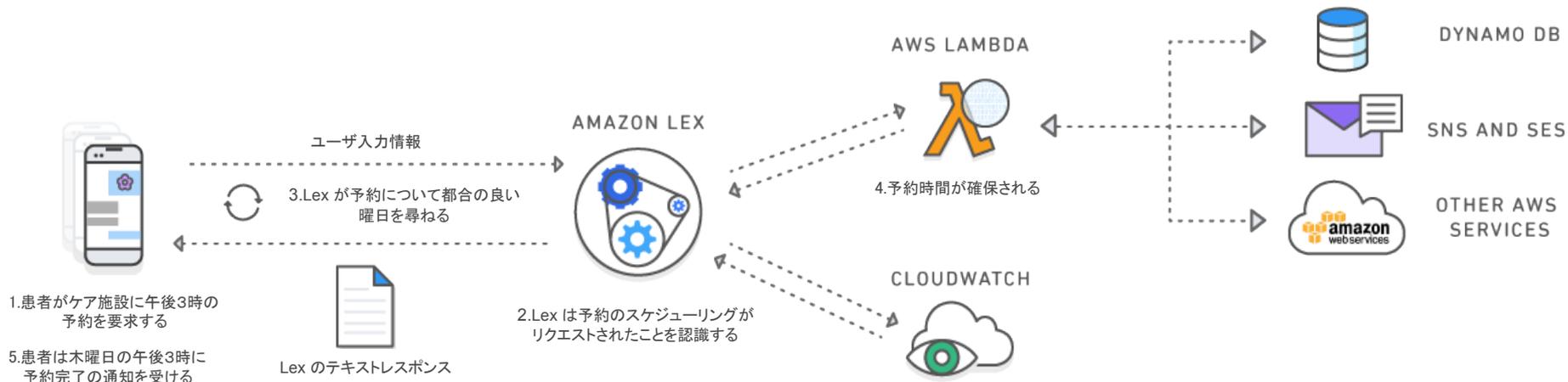


# Lex – ボット開発プロセス



# Lex - ユースケース: 情報ボット

## 患者向け診察予約ボットを構築する



# Lex - 利用価格と提供リージョン

- 従量課金制
- 音声リクエストは 1 件あたりの料金は 0.004 USD
- テキストリクエスト 1 件あたりの料金は 0.00075 USD
- 使用量は“処理されたリクエスト数”で計測され、月末に合計されて、月額料金が決定される
  
- サービス提供リージョン:
  - US East (N. Virginia) ※ Limited Preview

※現時点で Lex は limited Preview のサービスとなります。利用には下記 リンクより、プレビュープログラムへのサインアップが必要となります。

<https://pages.awscloud.com/AmazonLexPreviewLandingPage.html>

# まとめ:

## Deep Learning を利用した新しい 3 つのサービス



### Polly

文章をリアルな音声に変換するサービス



### Rekognition

画像分析機能をアプリケーションに簡単に追加できるようにするサービス



### Lex

声やテキストを使用した会話型インターフェイスを様々なアプリケーションに構築するためのサービス

# まとめ: Amazon AI サイト、Amazon AI Blog

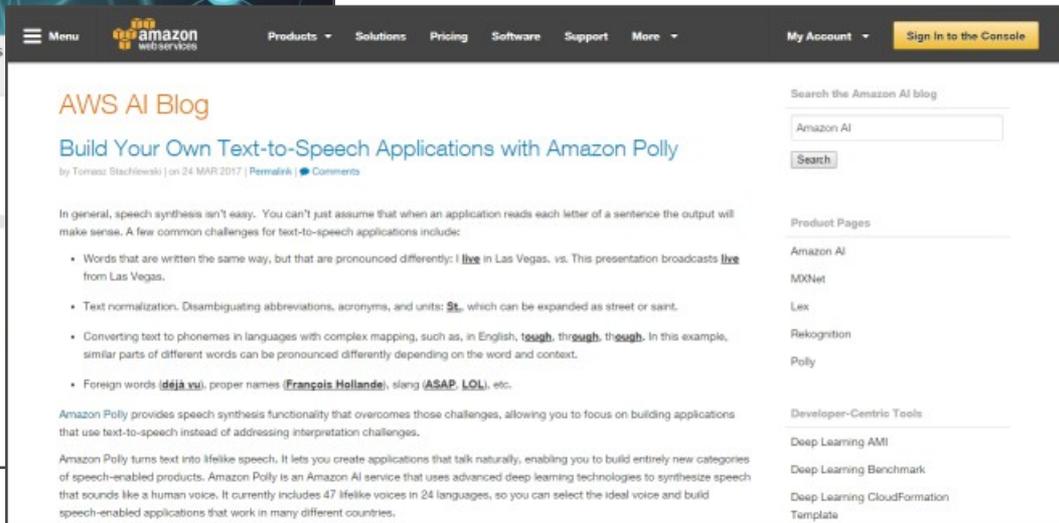


The screenshot shows the Amazon AI website homepage. At the top, there is a navigation bar with a menu icon, the Amazon logo, and links for '無料利用枠' (Free usage limit), '製品' (Products), 'ソリューション' (Solutions), '料金表' (Pricing), '詳細' (Details), '日本語' (Japanese), 'アカウント' (Account), and 'コンソールへサインイン' (Sign in to the console). The main header features the text 'Amazon AI' and '強力な人工知能をすべての開発者に' (Powerful AI for all developers), with a button that says 'まずは無料で始める' (Get started for free first). Below the header, there is a section titled 'AI とは' (What is AI) and 'Amazon AI の概要' (Overview of Amazon AI). The main content area contains a paragraph of Japanese text describing AWS AI services.

AWS は、クラウドネイティブな機械学習および深層学習テクノロジーによりさまざまなユースケースやニーズに対応可能な AI サービスを提供しています。Amazon AI により、自然言語理解 (NLU)、自動音声認識 (ASR)、視覚検索およびイメージ認識、音声変換 (TTS)、機械学習 (ML) などのテクノロジーをすべての開発者が利用できるようになります。Amazon Lex は Alexa を使用して高度なテキストや音声に回答するチャットボット開発を容易にするサービス、Amazon Rekognition は深層学習に基づくイメージ認識サービス、Amazon Polly はテキストをリアルなスピーチに変換するサービスです。また、Amazon Machine Learning により、スマートな ML アプリケーションをすばやく開発することが可能になります。Amazon AI サービスは、本番環境に安全にデプロイできるクロスプラットフォームアプリケーションの開発を容易にします。完全マネージド型のため、低レイテンシーでシームレスにスケールできるため、低コストで利用可能です。Amazon AI サービスにより、見て、聞いて、話せて、理解できて、周りの世界とやり取りができる、まったく新しい世代のアプリケーションの考案や開発に集中することが可能になります。

<https://aws.amazon.com/jp/amazon-ai/>

<https://aws.amazon.com/jp/blogs/ai/>



The screenshot shows an article from the AWS AI Blog. The title is 'Build Your Own Text-to-Speech Applications with Amazon Polly' by Tomasz Stachewski, dated 24 MAR 2017. The article discusses the challenges of text-to-speech synthesis, such as handling different pronunciations of the same word, text normalization, and foreign words. It highlights how Amazon Polly overcomes these challenges to provide natural-sounding speech synthesis.

## AWS AI Blog

### Build Your Own Text-to-Speech Applications with Amazon Polly

by Tomasz Stachewski | on 24 MAR 2017 | Permalink | Comments

In general, speech synthesis isn't easy. You can't just assume that when an application reads each letter of a sentence the output will make sense. A few common challenges for text-to-speech applications include:

- Words that are written the same way, but that are pronounced differently: I **live** in Las Vegas, vs. This presentation broadcasts **live** from Las Vegas.
- Text normalization, Disambiguating abbreviations, acronyms, and units: **St.**, which can be expanded as street or saint.
- Converting text to phonemes in languages with complex mapping, such as, in English, **tough**, **through**, **though**. In this example, similar parts of different words can be pronounced differently depending on the word and context.
- Foreign words (**déjà vu**), proper names (**François Hollande**), slang (**ASAP LOL**), etc.

Amazon Polly provides speech synthesis functionality that overcomes those challenges, allowing you to focus on building applications that use text-to-speech instead of addressing interpretation challenges.

Amazon Polly turns text into lifelike speech. It lets you create applications that talk naturally, enabling you to build entirely new categories of speech-enabled products. Amazon Polly is an Amazon AI service that uses advanced deep learning technologies to synthesize speech that sounds like a human voice. It currently includes 47 lifelike voices in 24 languages, so you can select the ideal voice and build speech-enabled applications that work in many different countries.

# 参考資料

- Amazon AI
  - <https://aws.amazon.com/jp/amazon-ai/>
- AWS AI Blog ドキュメント
  - <https://aws.amazon.com/jp/blogs/ai/>
- Amazon Pollyドキュメント
  - <https://aws.amazon.com/jp/polly/>
- Amazon Polly開発者用リソース
  - <https://aws.amazon.com/jp/polly/developers/>
- Speech Synthesis Markup Language (SSML) Version 1.1
  - <https://www.w3.org/TR/speech-synthesis11/>
- Pronunciation Lexicon Specification (PLS) Version 1.0
  - <https://www.w3.org/TR/pronunciation-lexicon/>
- Amazon Rekognitionドキュメント
  - <https://aws.amazon.com/jp/rekognition/>
- Amazon Rekognition開発者用リソース
  - <https://aws.amazon.com/jp/rekognition/developers/>
- Amazon Lexドキュメント
  - <https://aws.amazon.com/jp/lex/>
- Amazon Lex開発者用リソース(Preview)
  - <https://aws.amazon.com/documentation/lex/>

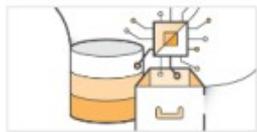
# Q&A



# オンラインセミナー資料の配置場所

- AWS クラウドサービス活用資料集

- <http://aws.amazon.com/jp/aws-jp-introduction/>



サービス別資料

無料オンラインセミナー「Black Belt Online Seminar」のサービスカット資料他、AWSのTechメンバーによる各サービスの解説資料がご覧いただけます。



ソリューション別資料

無料オンラインセミナー「Black Belt Online Seminar」のソリューションカット資料他、特定のソリューションについてのAWS活用方法がご覧いただけます。



業種別資料

無料オンラインセミナー「Black Belt Online Seminar」のインダストリーカット資料他、特定の業界のユースケースがご覧いただけます。



その他の資料

イベントに関する資料やアップデート情報などがご覧いただけます。

- AWS Solutions Architect ブログ

- 最新の情報、セミナー中のQ&A等が掲載されています

- <http://aws.typepad.com/sajp/>

# 公式Twitter/Facebook AWSの最新情報をお届けします



@awscloud\_jp



検索

もしくは  
<http://on.fb.me/1vR8yWm>

最新技術情報、イベント情報、お役立ち情報、  
お得なキャンペーン情報などを日々更新しています！

# AWSの導入、お問い合わせのご相談

AWSクラウド導入に関するご質問、お見積り、資料請求をご希望のお客様は以下のリンクよりお気軽にご相談ください

<https://aws.amazon.com/jp/contact-us/aws-sales/>

お問い合わせ

[日本担当チームへのお問い合わせ](#) >

関連リンク  
フォーラム

## 日本担当チームへのお問い合わせ

AWS クラウド導入に関するご質問、お見積り、資料請求をご希望のお客様は、以下のフォームよりお気軽にご相談ください。平日営業時間内に日本オフィス担当者よりご連絡させていただきます。

※ご請求金額またはアカウントに関する質問は[こちらからお問い合わせください](#)。  
※Amazon.com または Kindle のサポートにお問い合わせは[こちらからお問い合わせください](#)。

アスタリスク (\*) は必須情報となります。

姓\*

名\*

※「AWS お問い合わせ」で検索してください

