



# AWSのコンピュータービジョンと 言語関連サービスのご紹介

Kazuhito Go, Solutions Architect  
2020/05/20

# 目次

1. Amazon における機械学習の取り組み
2. コンピュータービジョン AI サービス
3. 自然言語 AI サービス
4. まとめ

# 目次

1. Amazon における機械学習の取り組み
2. コンピュータービジョン AI サービス
3. 自然言語 AI サービス
4. まとめ

# Amazon における機械学習取り組み

# AIに対するAWSのミッション

---

今からでも機械学習を使えるように

# AWSで機械学習を行う理由



もっとも広範で成熟した  
機械学習および AI サービス  
を提供

昨年だけで 200 を超える  
新しい機能やサービスが誕生

全ての機械学習サイエンティストや  
アプリケーション開発者に向けた  
ソリューション

主要な 3 つのフレームワークを  
全てサポート



SageMakerで  
機械学習の導入を加速

機械学習ワークフロー全体のための  
単一 IDE

総所有コスト(TCO)を54%以上低減

データラベリング作業コストを  
70% 以上削減

マネージドスポットトレーニングで  
90% 以上のコスト削減



最も包括的な  
クラウドプラットフォーム上  
に構築

高い安全性と信頼性と、  
充実した機能のデータストア

優れたコンピュー、ストレージ、  
セキュリティ、データベース、分析の  
機能群を提供

クラウドの TensorFlow プロジェクトの  
85%がAWS で実行

# あらゆる規模や業界の企業がAWSで機械学習を実行

何万ものお客様が機械学習のワークロードにAWSを選択 | 他のクラウドプロバイダーの2倍以上のお客様が機械学習にAWSを利用



# AWS の機械学習スタック: 最も広範で成熟した機能群

## AI サービス

機械学習のスキルなしに容易にAIをアプリに追加可能

Vision | Documents | Speech | Language | Chatbots | Forecasting | Recommendations |  
Fraud detection | EnterpriseSearch | CodeReview

## ML サービス

機械学習モデルの高速な構築、トレーニング、デプロイ

データラベリング | 構築済みアルゴリズム & ノートブック | 1クリックトレーニングとデプロイ

## MLフレームワーク & インフラストラクチャ

柔軟かつ選択可能な高性能なインフラストラクチャ

MLフレームワークのサポート | ML用途に応じて選べるコンピュート環境

# AWS の機械学習スタック: 最も広範で成熟した機能群

## AI サービス

本日の範囲

<b>VISION</b> Amazon Rekognition	<b>SPEECH</b> Amazon Polly	<b>TEXT</b> Amazon Transcribe +Medical <b>NEW</b> Amazon Comprehend +Medical Amazon Translate	Amazon Textract	<b>SEARCH</b> <b>NEW</b> Amazon Kendra	<b>CHATBOTS</b> Amazon Lex	<b>PERSONALIZATION</b> Amazon Personalize	<b>FORECASTING</b> Amazon Forecast	<b>FRAUD</b> <b>NEW</b> Amazon Fraud Detector	<b>DEVELOPMENT</b> <b>NEW</b> Amazon CodeGuru	<b>CONTACT CENTERS</b> <b>NEW</b> Contact Lens For Amazon Connect
-------------------------------------	-------------------------------	--------------------------------------------------------------------------------------------------------------	-----------------	-------------------------------------------	-------------------------------	----------------------------------------------	---------------------------------------	--------------------------------------------------	--------------------------------------------------	-------------------------------------------------------------------------

## ML サービス

Amazon SageMaker	Ground Truth	Augmented AI	SageMaker Studio IDE <b>NEW</b>					SageMaker Neo			
			Built-in algorithms	SageMaker Notebooks <b>NEW</b>	SageMaker Experiments <b>NEW</b>	Model tuning	SageMaker Debugger <b>NEW</b>	SageMaker Autopilot <b>NEW</b>	Model hosting	SageMaker Model Monitor <b>NEW</b>	

## ML フレームワークとインフラストラクチャ

<b>NEW</b> TensorFlow	<b>NEW</b> mxnet	GLUON	<b>K</b> Keras	Deep Learning AMIs & Containers	GPUs & CPUs	Elastic Inference	Inferentia	FPGA
<b>NEW</b> PYTORCH								

# 目次

1. Amazon における機械学習の取り組み
2. コンピュータービジョン AI サービス
3. 自然言語 AI サービス
4. まとめ

# コンピュータービジョン AI サービス

# Amazon Rekognition 機械学習を用いた画像と動画分析のサービス



# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



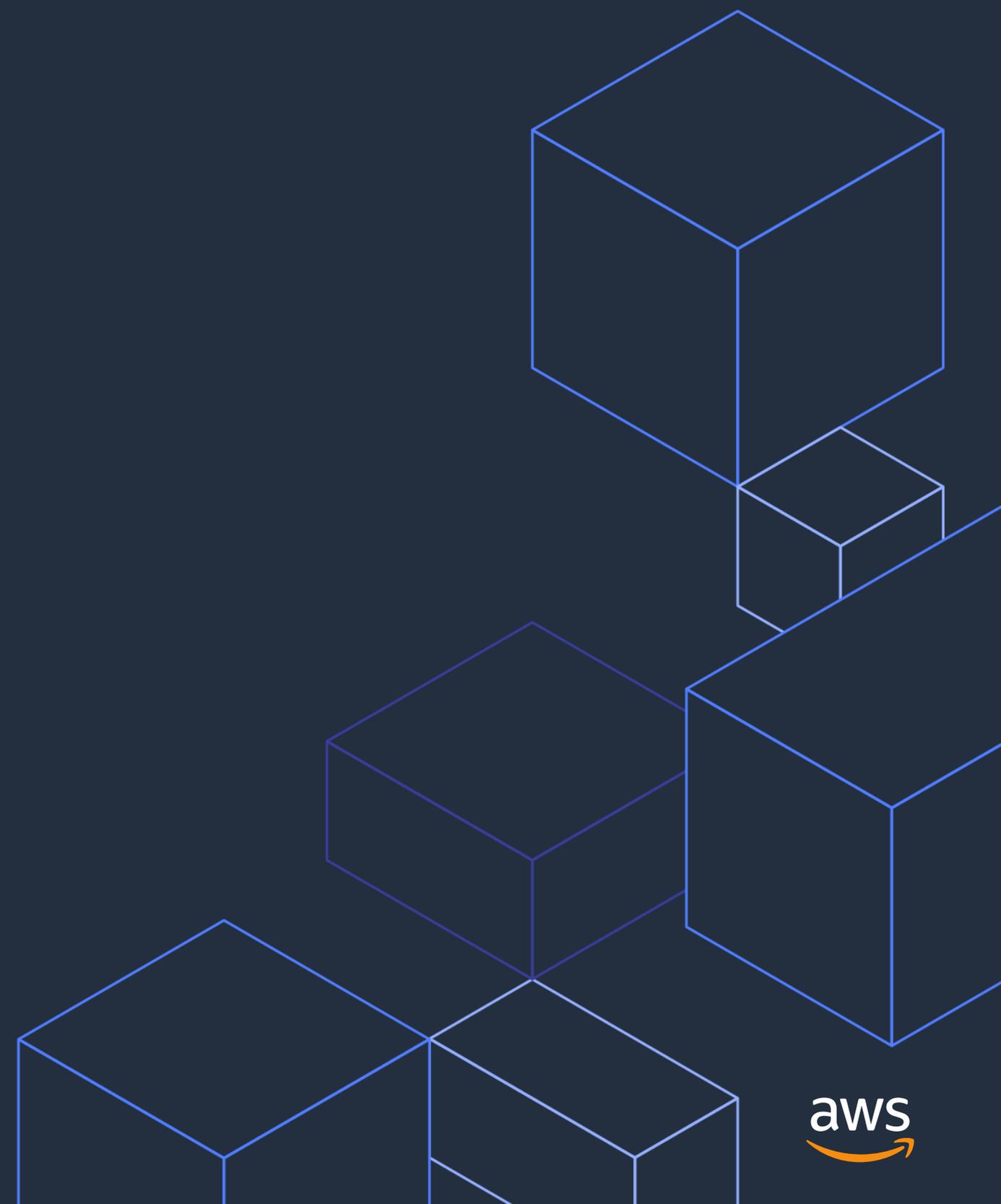
動線の検出



テキストの検出



# 特徴



# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



動線の検出



テキストの検出



# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



動線の検出



テキストの検出



# 顔検出と分析

## 主なユースケース



**エンドユーザーの個人検出**  
どのようなお客様が参加しているか検出する



**エンドユーザーの傾向分析**  
個人の特定ではなく、全体のトレンドを分析する

# 顔検出と分析

検出結果に含まれる情報



- 境界ボックス
- 信頼度
- 顔のランドマーク
- 顔の属性
- 感情など

これらの情報を、検出した顔ごとに結果として返します

[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/faces-detect-images.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/faces-detect-images.html)

# 顔検出と分析

様々な**困難な条件下**でも検出が可能

## ポーズ

---



## 明るさ

---



## オクルージョン

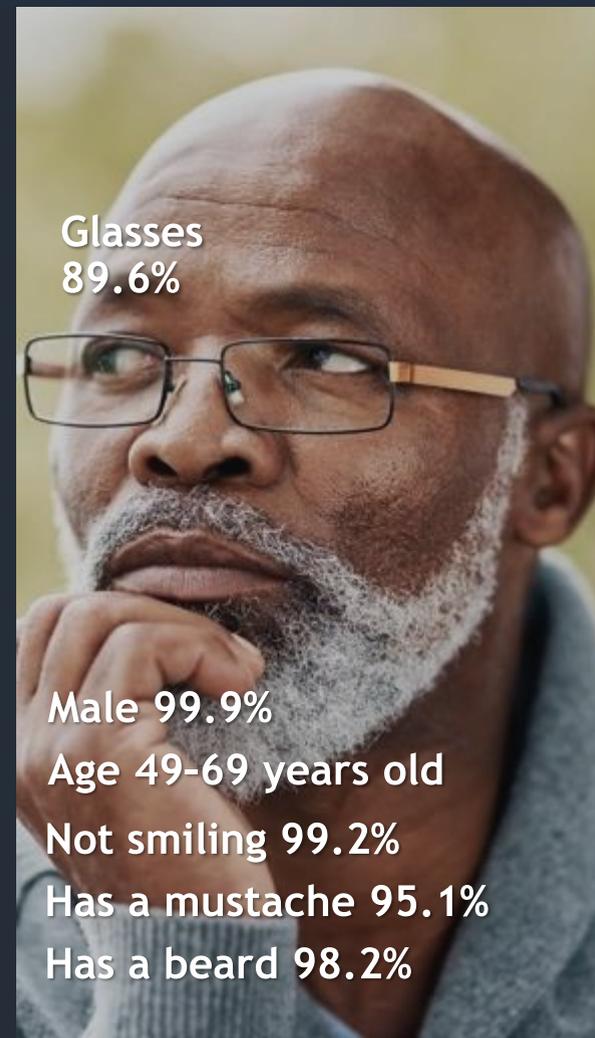
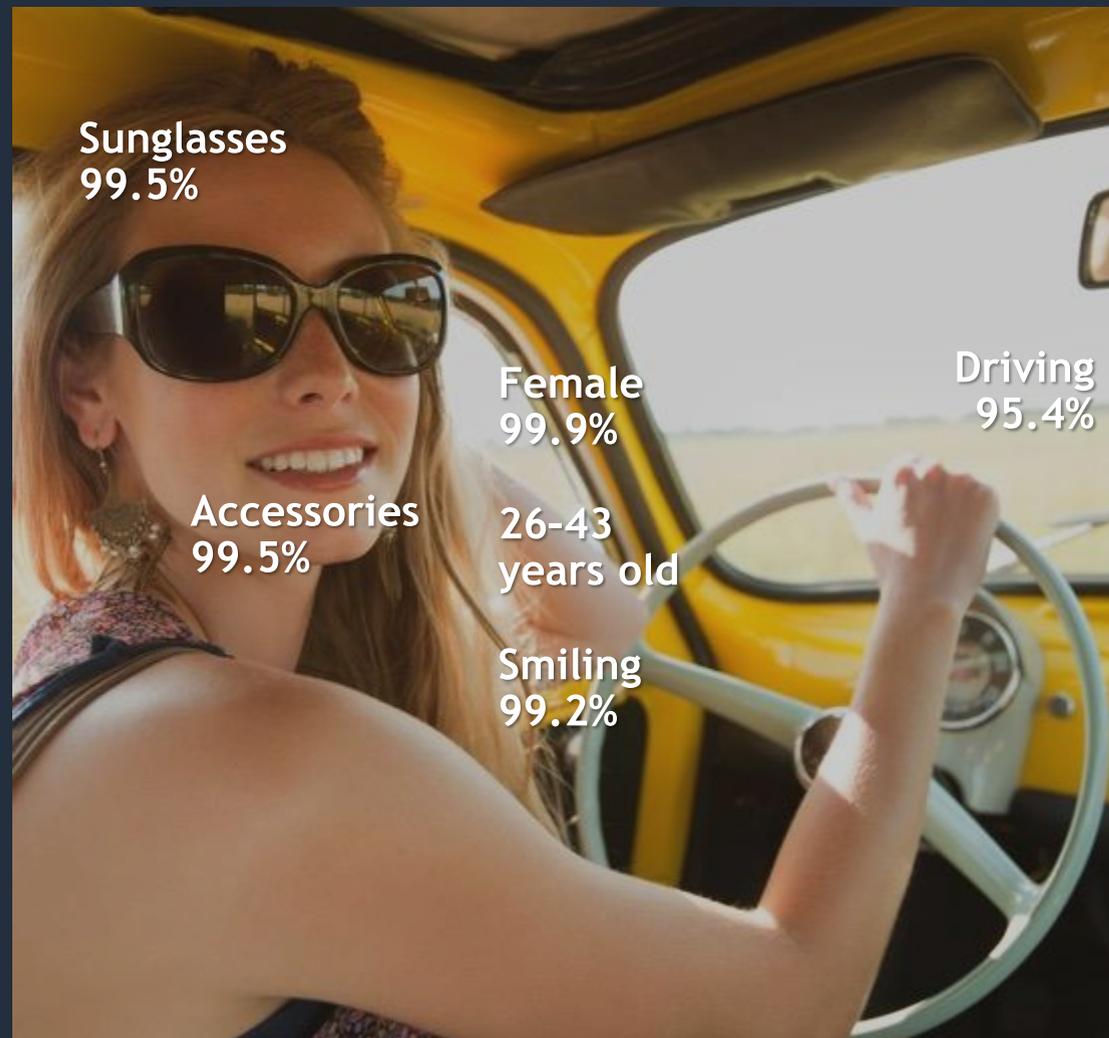
---



<https://aws.amazon.com/jp/blogs/news/amazon-rekognition-announces-updates-to-its-face-detection-analysis-and-recognition-capabilities/>

# 顔検出と分析

Rekognition の他の機能との組み合わせ



Amazon Rekognition の物体、シーン、活動の検出の機能を利用することで、

「どのような」エンドユーザーが「どのように」プロダクトを利用しているか分析ができます

[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/faces-comparefaces.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/faces-comparefaces.html)

# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



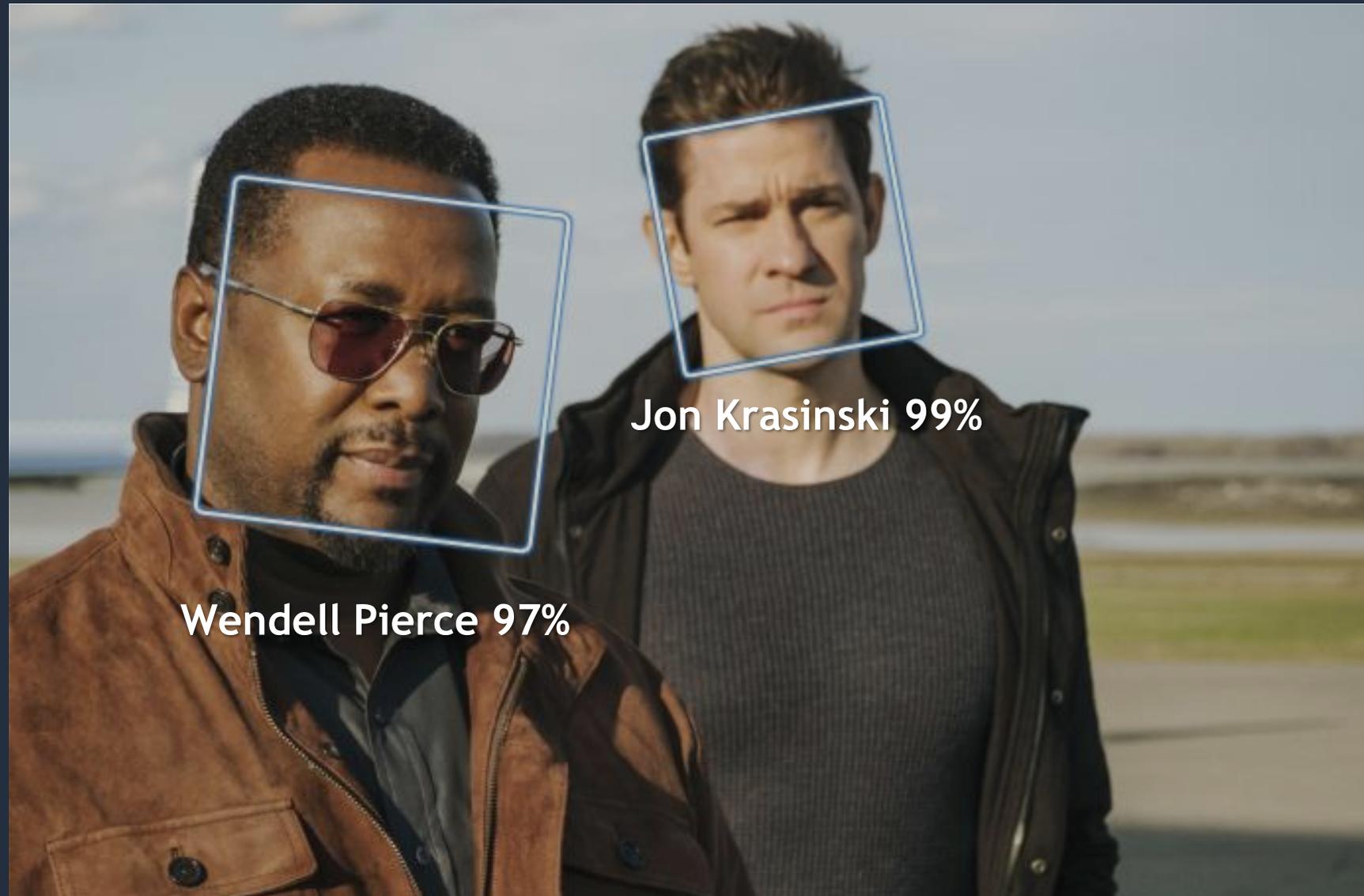
動線の検出



テキストの検出



# 有名人の認識



画像や動画に含まれる  
有名人を特定できます

有名人が登録してある  
ライブラリは常に最新  
に保たれています

[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/celebrities.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/celebrities.html)

# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



動線の検出



テキストの検出



# 物体検出とシーンの特定

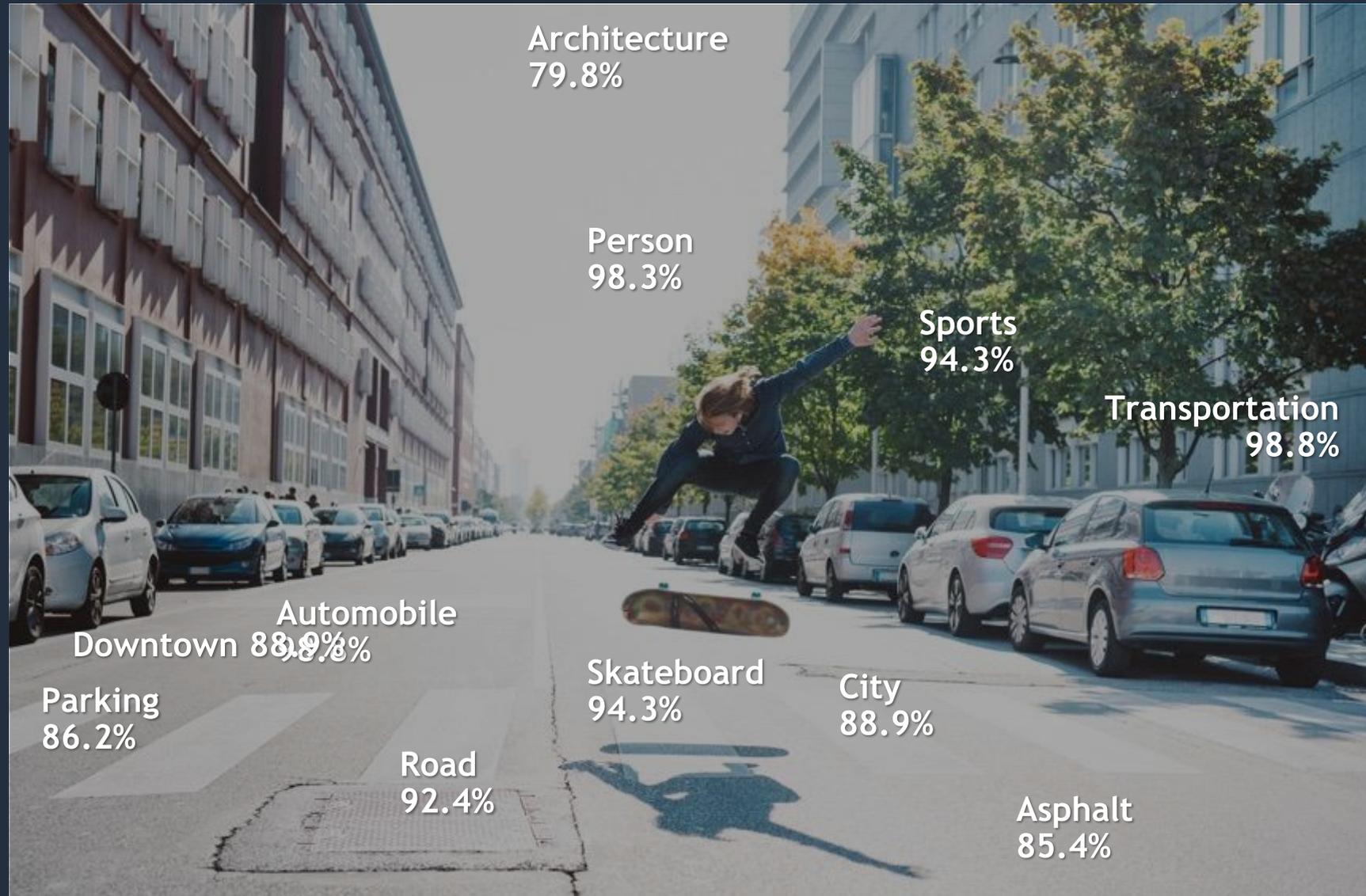
## 自動ラベル付け



自動的に何千もの対象  
物体、コンセプト、や  
シーンを画像や動画か  
らラベル付けします

# 物体検出とシーンの特定

## 自動ラベル付け



自動的に何千もの対象  
物体、コンセプト、や  
シーンを画像や動画か  
らラベル付けします

# 物体検出とシーンの特定

## 物体の位置の検出



さらに bounding box を利用することで、物体の位置を特定できます

Bounding box によって物体の検出数（「3台の車」）や物体同士の関係（「スケートボードの上にいる人」）などを推論できます

# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



動線の検出



テキストの検出



# コンテンツのモデレーション

画像や動画に安全でないコンテンツ（明示的なアダルトコンテンツや暴力的なコンテンツ）が含まれていないかどうかを判断できます



[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/moderation.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/moderation.html)

# コンテンツのモデレーション

Top-level category	Second-level category	Confidence score
Explicit nudity	Nudity	XX.X%
	Graphic male nudity	XX.X%
	Graphic female nudity	XX.X%
	Sexual activity	XX.X%
	Partial nudity	XX.X%
Suggestive	Female swimwear or underwear	XX.X%
	Male swimwear or underwear	XX.X%
	Revealing clothes	XX.X%

Amazon Rekognition はコンテンツを「安全」と「安全でない」のような二値の出力はしません

代わりにパーセンテージのような信頼の度合いを出力します

お客様は自身のポリシーを定義してコンテンツをエンドユーザーに提供できます

[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/moderation.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/moderation.html)

# Amazon Rekognition で利用可能な主な特徴 画像と動画分析の様々なニーズに対応

顔検出と分析



有名人の認識



物体の検出



シーンの検出



コンテンツのモデレーション



活動の検出



動線の検出



テキストの検出



# テキスト検出



Amazon Rekognition の Text in Image は画像内のテキストを検出し、それをコンピュータが読み取り可能なテキストに変換します

同じテキストが含まれている画像の検索や、車のナンバープレートを判別するシステムなどで利用できます

[https://docs.aws.amazon.com/ja\\_jp/rekognition/latest/dg/text-detection.html](https://docs.aws.amazon.com/ja_jp/rekognition/latest/dg/text-detection.html)

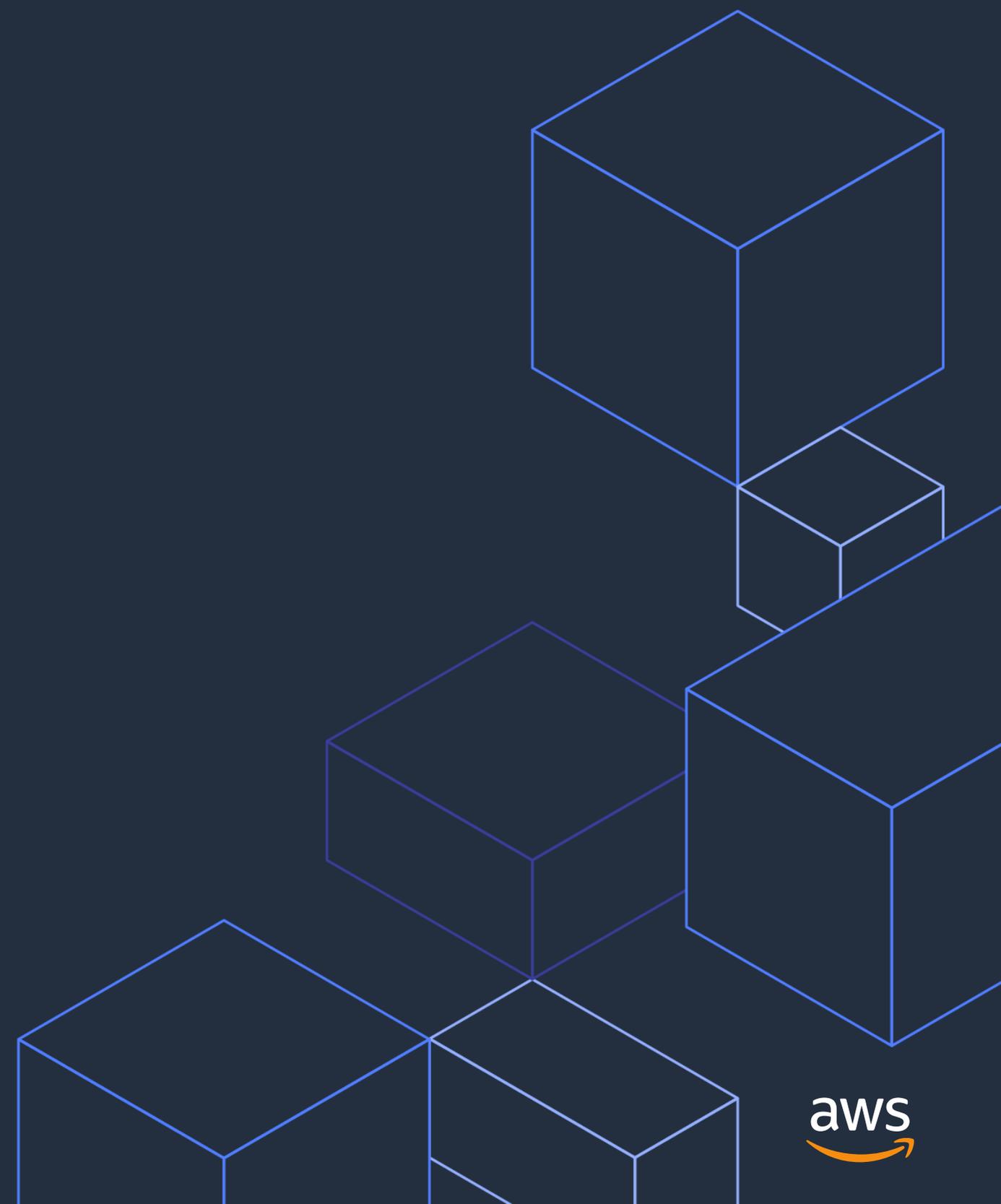
# 料金例

- AWSの基本的な料金体系と同じ、使った分だけ課金  
例) 東京リージョン、画像の場合1000枚あたり  
100万枚まで : 1000枚あたり1.3USD  
1000万枚まで : 1000枚あたり1.0USD  
1億枚まで : 0.8USD  
それ以上 : 0.5USD
- 詳細はWebを参照ください  
<https://aws.amazon.com/jp/rekognition/pricing/>

# 目次

1. Amazon における機械学習の取り組み
2. コンピュータービジョン AI サービス
3. 自然言語 AI サービス
4. まとめ

# Amazon **Transcribe**



# Amazon Transcribe - 概要

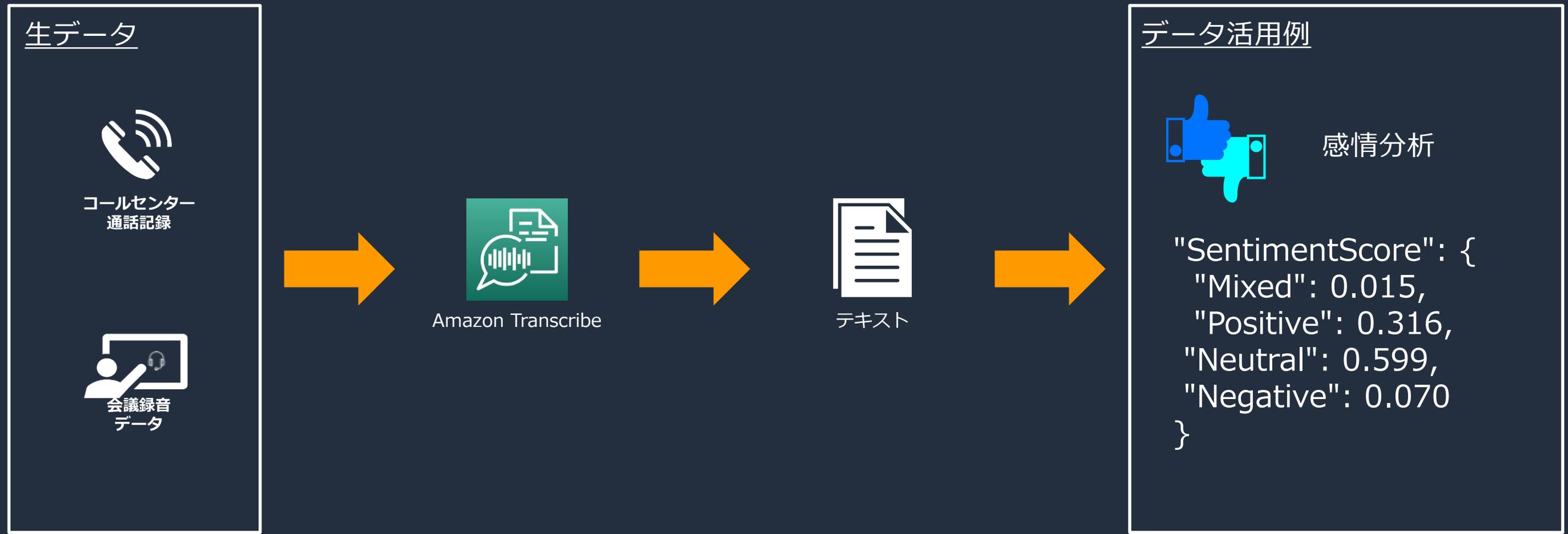
Amazon Transcribeは、ディープラーニングモデルを搭載したフルマネージドで、継続的に学習している自動音声認識サービス



開発者は、Amazon Transcribe を使用して、音声からテキストへの変換機能を簡単にアプリケーションへ追加することが可能

# Amazon Transcribe - ユースケース

- 例) コールセンターの通話記録や、会議の音声等を活用したい  
一般的に音声のままでは扱い難いため、テキストへ変換



# Amazon Transcribe - 特徴

## 特徴



句読点や大文字/小文字



単語  
タイムスタンプ



単語信頼スコア



カスタム語彙



スピーカーの識別



チャンネル識別



ストリーミング  
文字起こし

※日本語非対応

## 言語サポート

米国英語	英国英語	オーストラリア英語	米国スペイン語	スペイン語	現代的標準アラビア語 (MSA)	ブラジルポルトガル語
韓国語	ドイツ語	フランス語	カナダフランス語	インドヒンディー語	インド英語	中国標準語
ロシア語	イタリア語	<b>日本語</b>	他			

# Amazon Transcribe - 料金

東京リージョンの例)

1 秒あたり 0.0004 USD

ただし、15 秒未満の場合は 15 秒分の料金が発生

料金例)

ラジオのコマーシャル 15 秒 : 0.006 USD

ビデオトレーラー 2 分 30 秒 : 0.06 USD

# Amazon **Comprehend**



# Amazon Comprehend - 概要

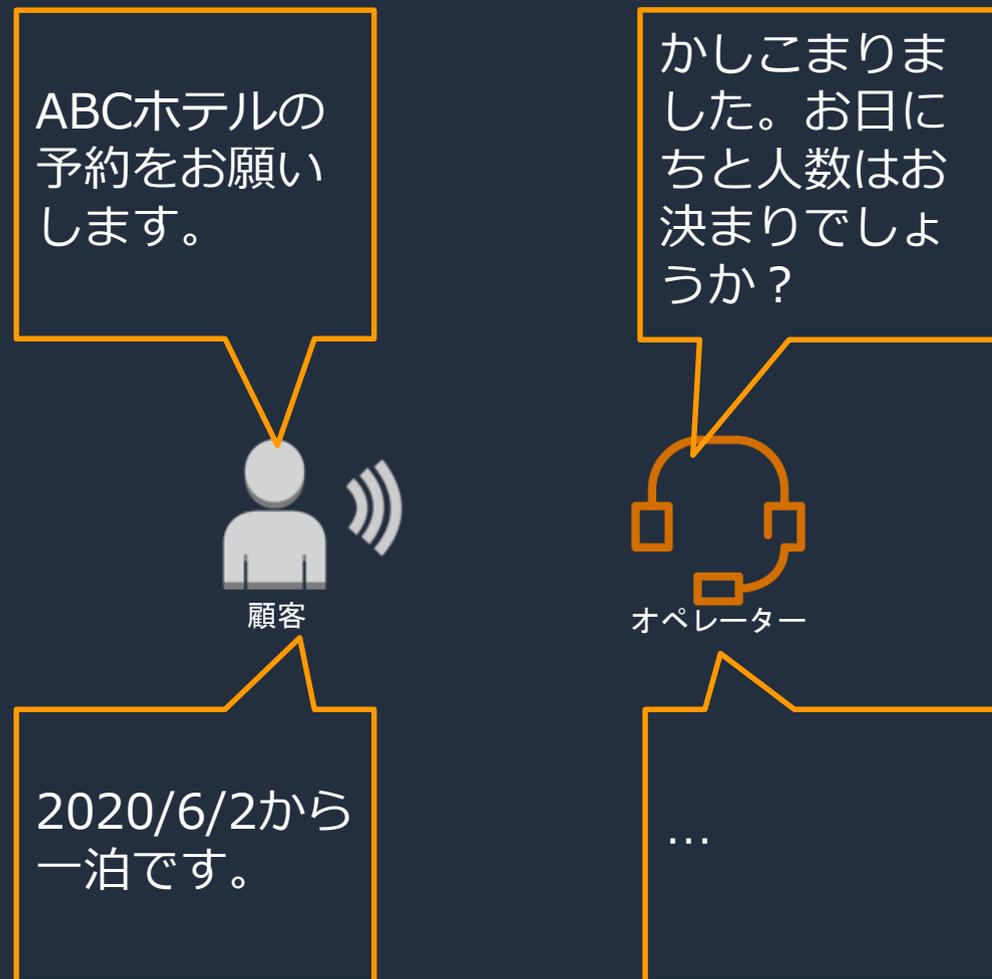
テキストを理解する上で重要な、  
エンティティ、キーフレーズ、感情などを検出



-  エンティティ
-  キーフレーズ
-  言語
-  感情
-  トピックモデル

# Amazon Comprehend - ユースケース

- 例) 接客内容を分析



- 顧客の感情分析  
接客の良し悪しの判断材料に
- トピックモデル  
何についての接客だったのか
- エンティティ  
どこのホテルか  
いつか  
何泊か
- キーフレーズ  
部屋の種類  
空きの有無

# Amazon Comprehend - テキスト解析例

1991年に、当時フィンランドのヘルシンキ大学の学生であったリーナス・トーバルズはオペレーティングシステムに好奇心を抱くようになっていた。当時、近代的なOSを動作させる能力を持つ Intel 80386 CPU を搭載した32ビットPC/AT互換パーソナルコンピュータが登場していた。

出典 : <https://ja.wikipedia.org/wiki/Linux>

## エンティティ

- 1991年 : DATE
- フィンランド : LOCATION
- ヘルシンキ大学 : ORGANIZATION
- リーナス・トーバルズ : Person

...

## キーワード

- オペレーティングシステム
- 好奇心
- 当時
- 能力

...

## 感情

- NEUTRAL(中立)

## 言語

- ja(日本語)

# Amazon Comprehend - 料金

100 文字 1 ユニット として計算し、各リクエストにつき 3 ユニットが最低料金  
キーフレーズや感情分析など、全て 1 万文字で ~0.01USD

機能	10 M ユニットまで	10 M~50 M 単位	50 M ユニット超
キーフレーズ抽出	0.0001USD	0.00005USD	0.000025USD
感情分析	0.0001USD	0.00005USD	0.000025USD
エンティティ認識	0.0001USD	0.00005USD	0.000025USD
言語検出	0.0001USD	0.00005USD	0.000025USD
構文解析	0.00005USD	0.000025USD	0.0000125USD

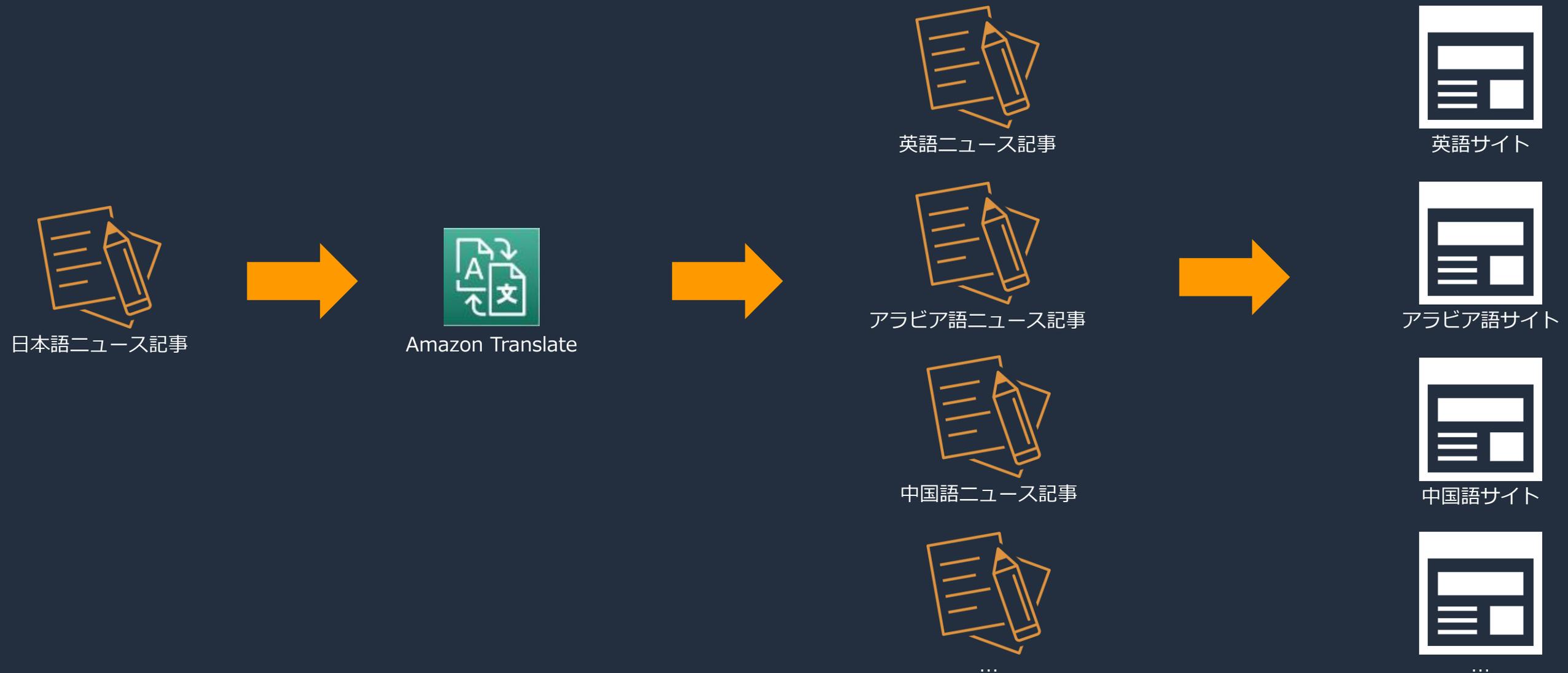
# Amazon Translate

# Amazon Translate - 概要

- ニューラルネットを用いた機械翻訳サービス
  - 日本語を含む 50 以上の言語間で翻訳が可能
  - 自動で原文言語の特定が可能
  - 特定ドメインの用語を登録可能
  - バッチ翻訳とリアルタイム翻訳の両方が可能

# Amazon Translate - ユースケース

例) 記事を即座に多様な言語に翻訳してパブリッシュ



# Amazon Translate - 翻訳例

Amazon Translate は、サポートされている言語間でテキストを翻訳するためのニューラル機械翻訳 (MT) サービスです。ディープラーニング手法を搭載したこのサービスは、高品質な機械翻訳を手ごろな価格で提供します。これにより、開発者は企業やユーザーが作成したコンテンツを翻訳したり、複数の言語にわたるサポートが必要なアプリケーションを構築したりできます。このサービスは API から利用でき、ソース言語からターゲット言語へのリアルタイムまたはバッチでのテキスト翻訳を可能にします。

Amazon Translate is a Neural Machine Translation (MT) service for translating text between supported languages. Equipped with deep learning methodologies, this service offers high-quality machine translation at an affordable price. This allows developers to translate content created by companies and users, and to build applications that need support across multiple languages. This service is available through the API and enables real-time or batch translation of text from the source language to the target language.

# Amazon Translate - 料金

100 万文字あたり 15.00 USD

料金例)

平均的なEメール 3100 文字 : 0.0465 USD

典型的なニュース記事 6500 文字 : 0.0975 USD

会社のウェブサイト 20000 文字 X 100 ページ : 30 USD

# 目次

1. Amazon における機械学習の取り組み
2. コンピュータービジョン AI サービス
3. 自然言語 AI サービス
4. まとめ

# まとめ

- AWS の AI サービス を用いることで容易にメディアの解析が可能
  - Amazon Rekognition
    - 物体・有名人・不適切画像・顔・文字などの検出
  - Amazon Transcribe
    - 音声の文字起こし
  - Amazon Comprehend
    - テキストからエンティティやキーワードの検出、感情分析
  - Amazon Translate
    - テキストの翻訳

Media Insight Engine はこれらの AI サービスを統合したサービス(次のお話)