



第二十八回 「アップデート紹介とちょっぴり  
DIVE DEEP する AWS の時間」

# Amazon Kendra による 文書からの日本語セ マンティック検索

Tatsuya Koyakumaru

Solutions Architect

Amazon Web Services Japan G.K.

2023/03/30

# 自己紹介

氏名:小役丸 達也 (こやくまる たつや)

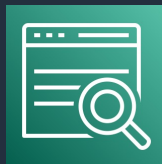
得意分野: 機械学習 (統計/凸解析)

趣味: 作曲、ベース

好きな AWS サービス:



Amazon SageMaker



Amazon Kendra



# アジェンダ

- Kendraの概要についての紹介
- 日本語セマンティック検索についての紹介
- デモ
- まとめ

# ビジネス検索の課題

データの増大に比例して、重要な情報を見つけるのはより困難に

# 80%

のデータが**非構造化データ**  
(ドキュメント、ヘルプ Web サイト、  
サポートドキュメントなど)<sup>\*1</sup>

そしてデータ量は増加傾向にある

## お客様の課題



適切なクエリが思いつかない



実装が複雑で手間



検索が遅い(パフォーマンス)



運用に専門的なスキルが必要

(\*1) Capgemini, [Reorganizing unstructured data](#)

# 検索エンジン - 導入のメリット

検索エンジンは 複数の領域でビジネス価値向上に寄与

## 社内での効果

### 生産性の向上

過去の回答を  
素早く見つける

リポジトリ横断  
の検索が可能に

研究開発関連の  
検索時間を削減

### リスクの低減

コンプライアンスに関連する文書を素早く見つけ  
ビジネスリスクを軽減

## お客様への効果

### お客様体験の向上

エンドユーザにとって  
探しやすいサイト内検索

コールセンターエージェント  
が見つけやすい社内検索

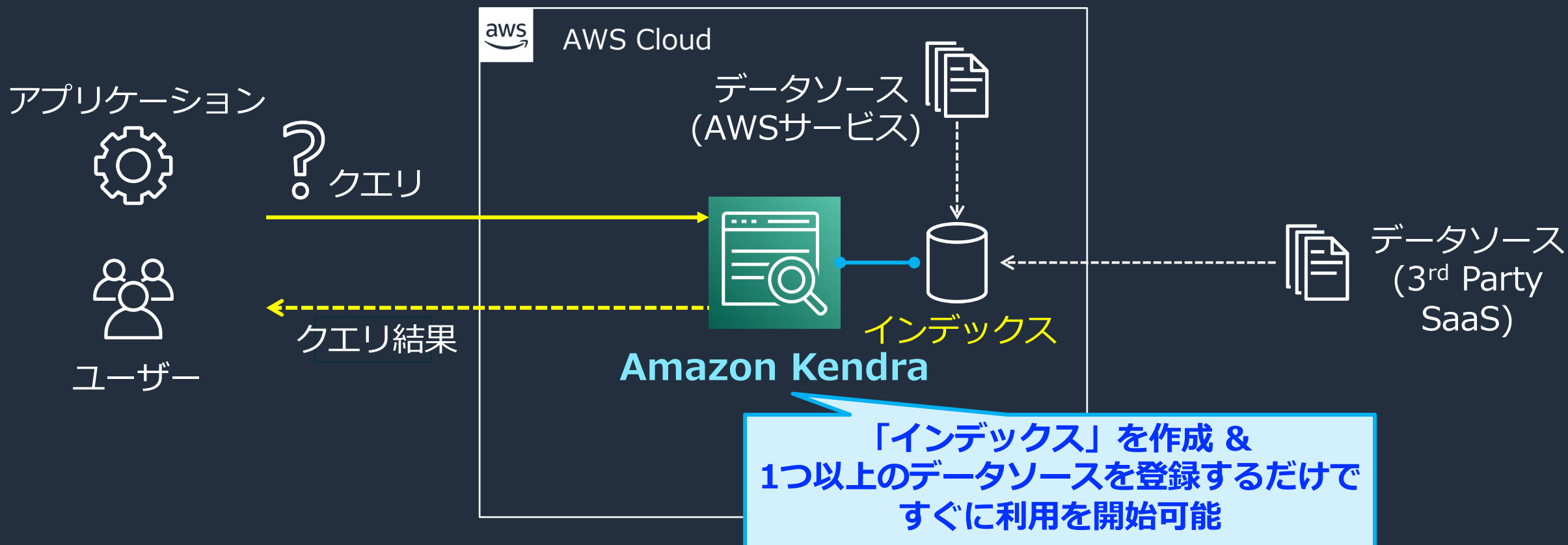
# Amazon Kendra

## - 概要



# Amazon Kendra - マネージドな検索エンジン

Kendra は AWS マネージドな検索エンジン (リージョナルサービス)  
→ セットアップが容易 & 多様なデータソースのドキュメント取り込みに対応





# Kendra - マネージドサービスのメリット

セルフマネージドな検索エンジンソリューションと比較して  
運用工数の削減 や スキル要員継続確保の必要性の排除 を実現

Kendraには「バージョン」や  
「メンテナンスウィンドウ」の概念なし

セルフマネージドな検索エンジン



データの管理

クエリのパフォーマンスモニタリング

MLモデルの管理

インデックスの管理(計算資源, ストレージ)

OS・ミドルウェアの導入・パッチ適用

Amazon Kendra



データの管理

クエリのパフォーマンスモニタリング

MLモデルの管理

インデックスの管理(計算資源, ストレージ)

OS・ミドルウェアの導入・パッチ適用

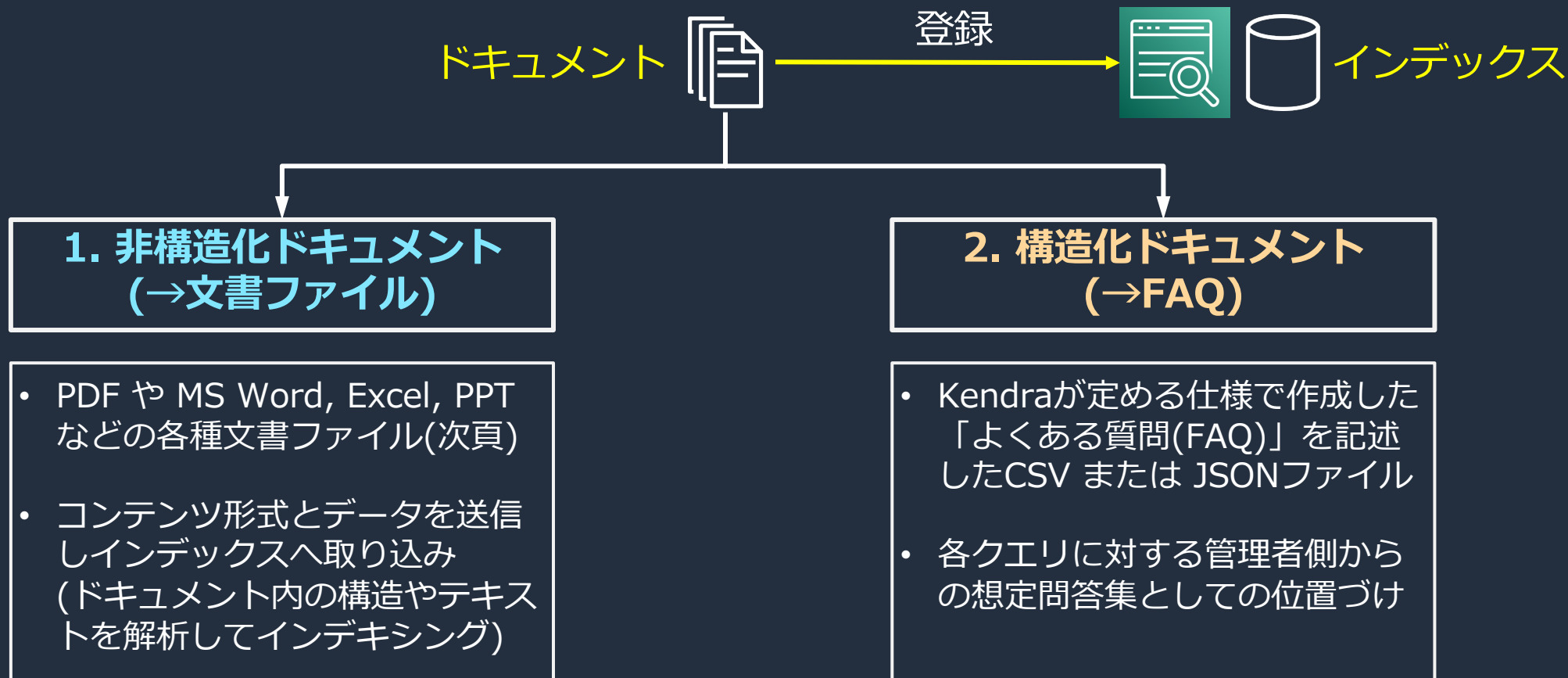
お客様が管理

AWSが管理



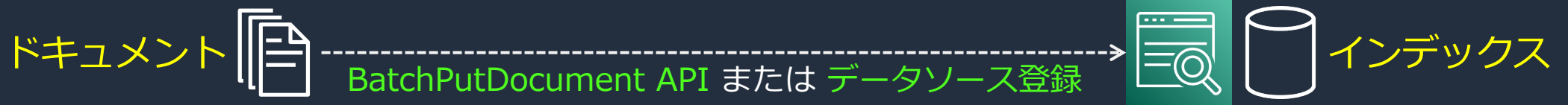
# Kendra - 対応ドキュメント (何を検索できる?)

Kendraは大別して2つの種類のドキュメントをインデックスに登録可能



# Kendra - 対応ドキュメント形式 (1.非構造化タイプ)

Kendraは下表の形式のドキュメントをインデックスに取り込み、検索対象とすることが可能 (取り込みデータはUTF-8エンコードが前提)



#	対応コンテンツ(ファイル)形式
1	HTML
2	Microsoft PowerPoint
3	Microsoft Word
4	Microsoft Excel
5	PDF
6	プレーンテキスト
7	CSV
8	XML
9	JSON
10	Markdown

#	対応コンテンツ(ファイル)形式
11	RTF(Rich Text Format)
12	XSLT(Extensible Stylesheet Language)

※各ドキュメントの最大サイズは 50MB (上限緩和申請可能)  
& 各ドキュメントの抽出テキストの最大サイズは 5MB (ハードリミット)  
※未対応のファイル形式は、Lambda関数を開発し  
「カスタムドキュメントエンリッチメント機能」(後述)の利用で対応可能

# Kendra - 対応ドキュメント形式 (2. 構造化タイプ)

KendraはS3バケットに配置した「FAQ(よくある質問)」ファイルを取り込み、検索対象とすることが可能



#	対応ファイル形式	補足
1	CSV (基本フォーマット)	各行に質問内容、回答、ソースURI(オプション)の項目のみを記入
2	CSV (カスタムフォーマット)	ヘッダ行を指定し、各行に様々なメタデータ属性(含アクセス制御)を記入
3	JSON	最も柔軟な定義が可能、所定のJSONスキーマ構造で複数のFAQをファイル内に定義

## CSV(基本フォーマット) 例

ワシントン州スポーカンにはいくつの診療所がありますか?, 13, <https://www.freeclinics.com/>  
ミズーリ州のマウンテンビューにはいくつの診療所がありますか?, 7, <https://www.freeclinics.com/>

※取り込み可能な各FAQファイルの最大サイズは 1MB (上限緩和申請可能)

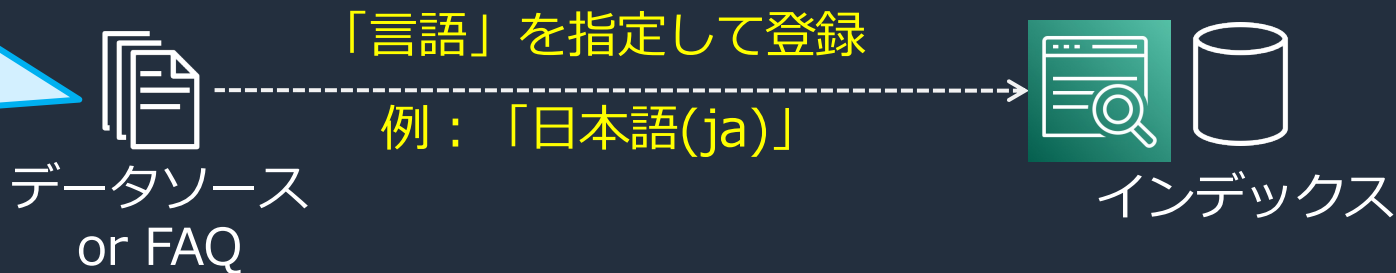
※1つのインデックスに登録可能なFAQは最大 30 (上限緩和申請可能)

# Kendra - 対応言語 (どの言語の文書を検索できる?)

ある1つのデータソース/FAQを登録する際には、それをインデックスに取り込む際のデフォルト言語を1つ指定 → 「日本語(ja)」にも対応

<https://docs.aws.amazon.com/kendra/latest/dg/in-adding-languages>

メタデータ属性  
(\_language\_code)  
によるドキュメント単位  
での言語指定登録も可能



現在、日本語を含む  
36言語に対応

#	Kendraデータソース/FAQ 対応言語
1	英語(en)
:	:
-	日本語(ja)
:	:
36	トルコ語(tr)

※注意: 言語指定を省略すると、常にデフォルト言語の「英語(en)」として認識

# 最近の Kendra アップデート情報

Kendraが日本のお客様にとって大きく使いやすくなりました

## ① 日本語でのセマンティック検索が可能になりました

日本語を用いた質問の意味を理解することができるようになり、柔軟なクエリが可能になりました。

## ② 東京リージョンで使用可能になりました

日本のお客様が多く使用している東京リージョンのみを使用して検索が可能となります。

# セマンティック検索機能について



# セマンティック検索とは?

**セマンティック検索**とは、検索文（自然言語）の意味を理解し、その意味に沿った検索結果を提供する技術です。

具体的には、以下の機能がサポートされます。

- 単純なキーワードマッチングを超えたドキュメントの関連性
- 単純なキーワードマッチングを超えたFAQ
- Amazon Kendraの読解力に基づいてドキュメントから回答を抽出する。
- 検索結果の信頼度バケット（非常に高い、高、中、低）

<https://docs.aws.amazon.com/kendra/latest/dg/in-adding-languages.html>



# Kendra - クエリ方法(どのように検索できる?)

Kendraは大別して 3種類の クエリ(質問) に対応  
(※クエリ自体は全て同じ Query APIリクエストであり、区別はない)

## ① ファクトイド型質問 (単一の解「誰/何/いつ/どこ」を問う質問)

シアトルに最寄りのサービスセンターはどこですか?

付随してフィルタ  
情報なども指定可能

## ② 記述的な質問

Amazon Echo Plus をネットワークに接続するには  
どうすればよいですか?

Query API



Amazon Kendra  
(インデックス)

## ③ キーワード検索

IT ヘルプデスク

※AND や OR などの論理演算子等を利用可能  
※セマンティック検索非対応言語でも可能

アプリケーション



クエリ



ユーザー

# Kendra - インデックスのクエリ用インターフェース

大別して3つのインターフェースを利用して  
ユーザーは Kendra の インデックス にクエリを実行可能



## 1. AWS SDK/CLI (Kendra Query API)

- Query API を呼び出す
- ユーザーは同APIを呼び出す任意のアプリケーションを開発して利用
- サンプルUIアプリ提供(GitHub)

要開発

## 2. AWS管理コンソール (Kendraダッシュボード)

- ユーザーはAWS管理コンソールにサインインして利用 (Kendra Search IndexダッシュボードUIを利用)
- Kendraの全てのクエリ機能を利用可能

開発不要

## 3. Kendra Experience Builder による生成UI

- ノンコーディングでクエリ用 Web UIをKendraが生成しエンドポイントを提供
- ユーザー認証はIAM Identity Center を利用 (※ホームリージョンと同一リージョン前提)

# 現時点で英語版のみサポートされている機能

Kendraの検索を便利にする、現在英語版のみ使用可能な機能があります。

- ・ シノニム (カスタムシノニムを含む)  
同義語の検索機能。例) DNSの検索結果にRoute53を含める。
- ・ 増分学習とフィードバック  
検索結果のフィードバックをユーザーから受け、検索結果を改善
- ・ クエリ提案  
最も人気のあるクエリに基づいてクエリを提案

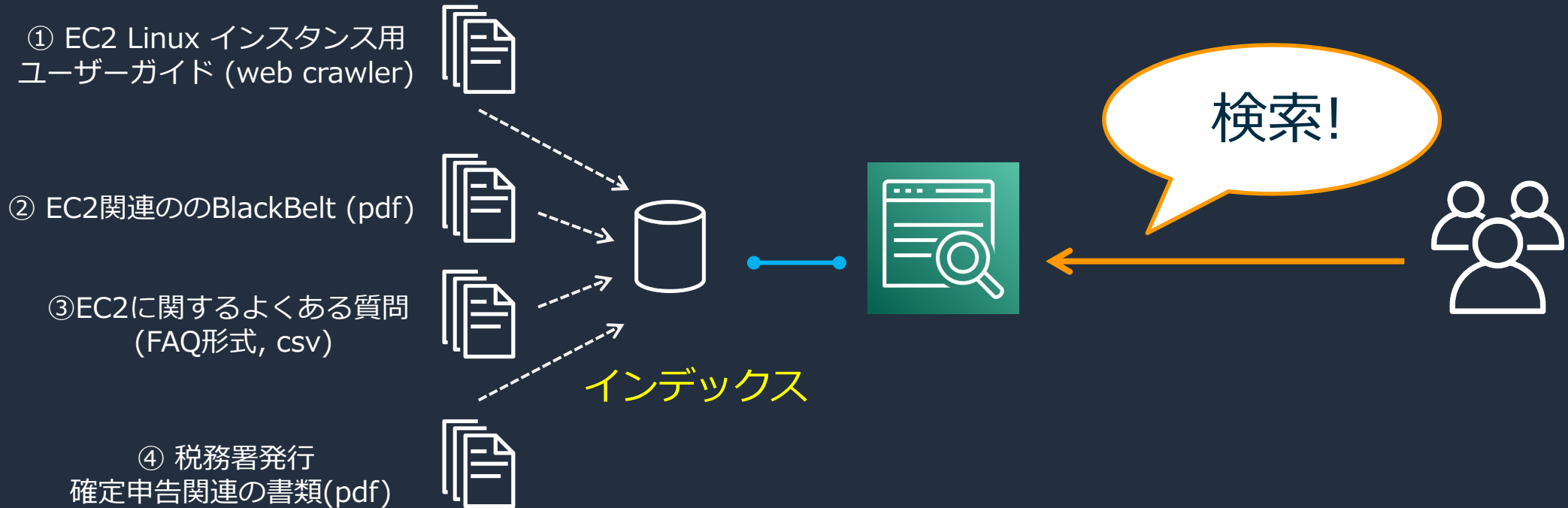
<https://docs.aws.amazon.com/kendra/latest/dg/in-adding-languages.html>

# デモ



# デモ内容

データソース・FAQに計4種類のデータを使用して、インデックスを構成したもののから検索をします。



# まとめ



# まとめ

- Amazon Kendraを用いることによって、構造化/非構造化データ形式の多数の文書を検索できるサービスを短時間で作成することができます。
- 日本語でのセマンティック検索にも対応しているため、キーワード以外の要素も反映した柔軟な検索結果を取得可能です。
- 開発なしで検索画面を作成する Experience Builder 機能や Query API による検索など、要望に応じて柔軟にインターフェースを作成できます。





# Thank you!

Tatsuya Koyakumaru  
Solutions Architect