

JAPAN | JUNE 20, 2024

aws SUMMIT



© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MINISESSION

生成 AI アプリでの画像データ活用 ~ マルチモーダル RAG ~

山下一樹, 鈴木大樹

AWS Japan G.K.
Solutions Architect



アジェンダ

1. 生成 AI と独自データ活用のための RAG
2. マルチモーダル RAG アプリケーションのデモ
3. マルチモーダル RAG の活用例とブース案内

生成 AI と独自データ活用のための RAG

広がる生成 AI の活用



要約



ソフトウェア開発



デザイン生成

広がる生成 AI の活用

既存の生成 AI アプリケーションの利用



要約



ソフトウェア開発



デザイン生成



独自データを活用した生成 AI アプリケーションの需要の高まり

- 独自のナレッジ、ドメイン知識を持った生成 AI アプリケーションをビジネスに活用したい
- 社内ナレッジ（規約、マニュアル、FAQ 等）を理解した生成 AI を用いて業務効率化したい

独自データを生成 AI で活用する方法

独自モデルの作成

モデルをゼロから構築



モデルそのものをカスタマイズ

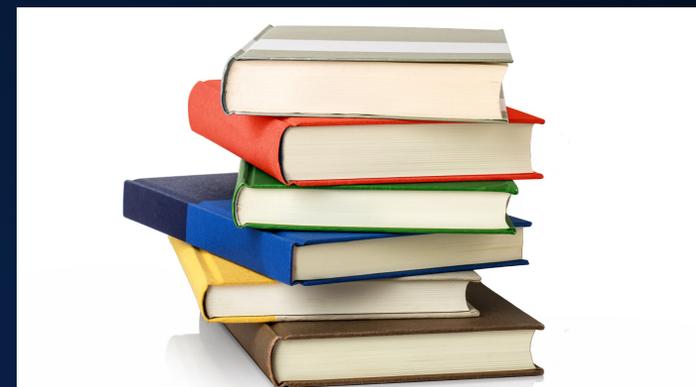
ファインチューニング

追加学習でカスタマイズ



検索拡張生成 (RAG)

モデルは変えずに
知識を追加



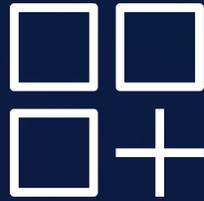
入力データをカスタマイズ

独自データ活用のための検索拡張生成 (RAG)

2023年にアナウンスされた
AWSの生成AIサービスは?



ユーザー



アプリ

**AWS ブログによると
Amazon Bedrock です。**

“Amazon Bedrock とは ...”
(AWS ブログ記事)



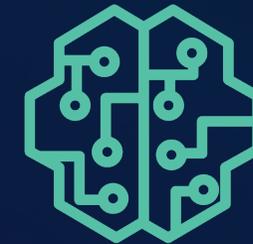
データソース



②



③



基盤モデル

データの形式は様々

検索したい対象はテキストだけでなくさまざまな形式で存在する



テキスト



画像



音声



ビデオ

画像データの例



医療

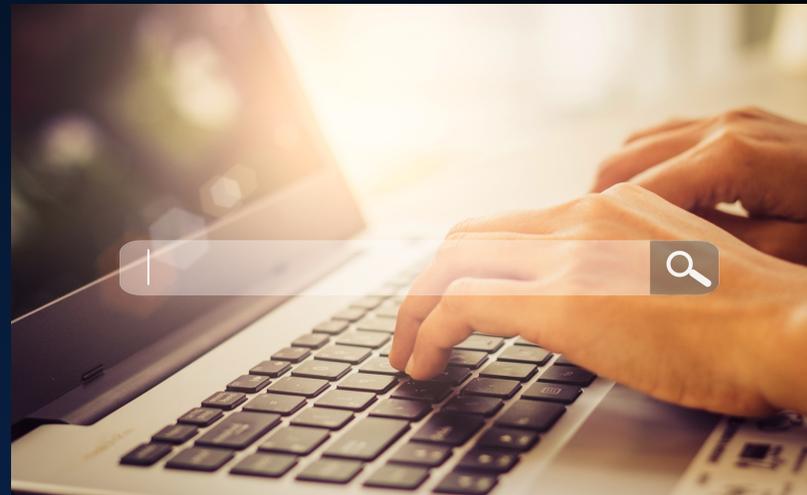


ファッション



研究開発

様々なデータ形式 (マルチモーダル) を扱う難しさ



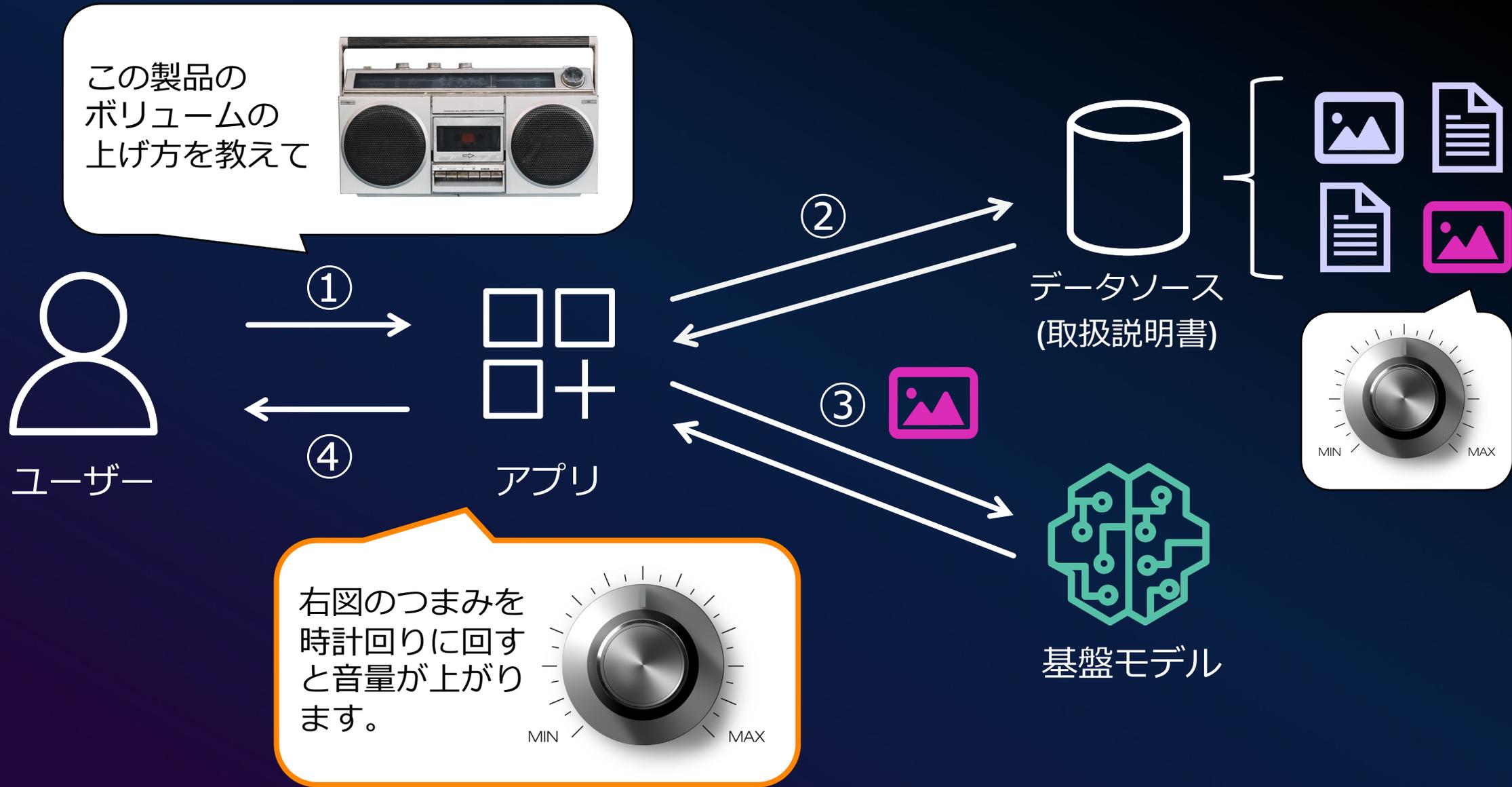
データの形式と処理の複雑さ

- 画像やビデオなどのデータは異なる処理方法を必要とする
- テキストデータは非構造化データの中でも扱いやすく、自然言語処理技術が発達している

様々なデータを対象とする検索の難しさ

- テキスト以外のデータ形式を含めた検索は高度な技術を要する

マルチモーダル RAG による複数データ形式への拡張



マルチモーダル RAG アプリケーションのデモ



アプリケーション概要

- テキストと画像を伴うドキュメントとして、製品説明書をデータソースとした **マルチモーダル RAG アプリケーション**
- テキストおよび画像による入力（製品の使い方に関する質問）に対してアプリケーションが回答。
出力は関連するドキュメントと画像が表示され、一目で使い方がイメージできる

第1章 基本的な操作

Kindle Paperwhite をお買い上げいただきありがとうございます。このガイドでは、Kindle の特長と機能をご紹介します。Kindle には、お住まいの地域や国によってご利用いただけない機能があります。各機能の詳細については、www.amazon.com/devicesupport をご覧ください。

Kindle を登録する

Kindle をご自身の Amazon アカウントを使用してオンラインで購入された場合は、登録済みの Kindle をお届けします。登録先は、ホームボタンをタップして、ホーム画面左上に表示されるアカウント名で確認できます。ご自分のアカウント名ではなく「My Kindle」または以前の持ち主の名前（譲り受けた場合など）が表示されている場合は、ご自身で登録いただく必要があります。Kindle の設定と登録の方法については、この章の「[Kindle Paperwhite のセットアップ](#)」をご覧ください。

Kindle の操作

Kindle Paperwhite の操作は非常に簡単です。



Multimodal Embeddings

ここに画像をドロップまたはクリックしてアップロード

Drag and drop file here

Limit 200MB per file • PNG, JPG, JPEG

Browse files

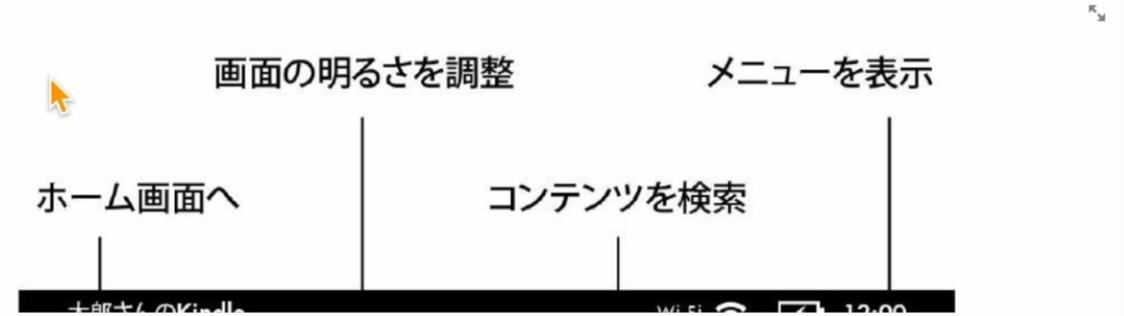
kindle.jpg
71.4KB



この画像は、Kindle端末のホーム画面を示しています。ホーム画面から様々な操作を行うことができます。

- 家のアイコンをタップするとホーム画面に戻ります。
- 左矢印アイコンをタップすると、前のページや画面に戻ることができます。
- 電球アイコンをタップすると、画面の明るさを調整できます。
- カート(買い物かご)アイコンをタップすると、Kindleストアに進むことができます。
- 虫眼鏡アイコンをタップすると、検索機能を使用できます。
- 本のアイコンをタップすると、Goodreads on Kindleと呼ばれる本に関するコミュニティにアクセスできます。
- 右端の3本線のメニューアイコンをタップすると、追加のメニューオプションが表示されます。

以上が、この画像からわかるKindleのホーム画面の主な操作方法です。タップやスワイプなどのジェスチャー操作で様々な機能にアクセスできます。



アプリケーションのデモ

×

Deploy ⋮

Multimodal Embeddings

ここに画像をドロップまたはクリックしてアップロード

Drag and drop file here
Limit 200MB per file • PNG, JPG, JPEG

Browse files

写真を撮影

Examples



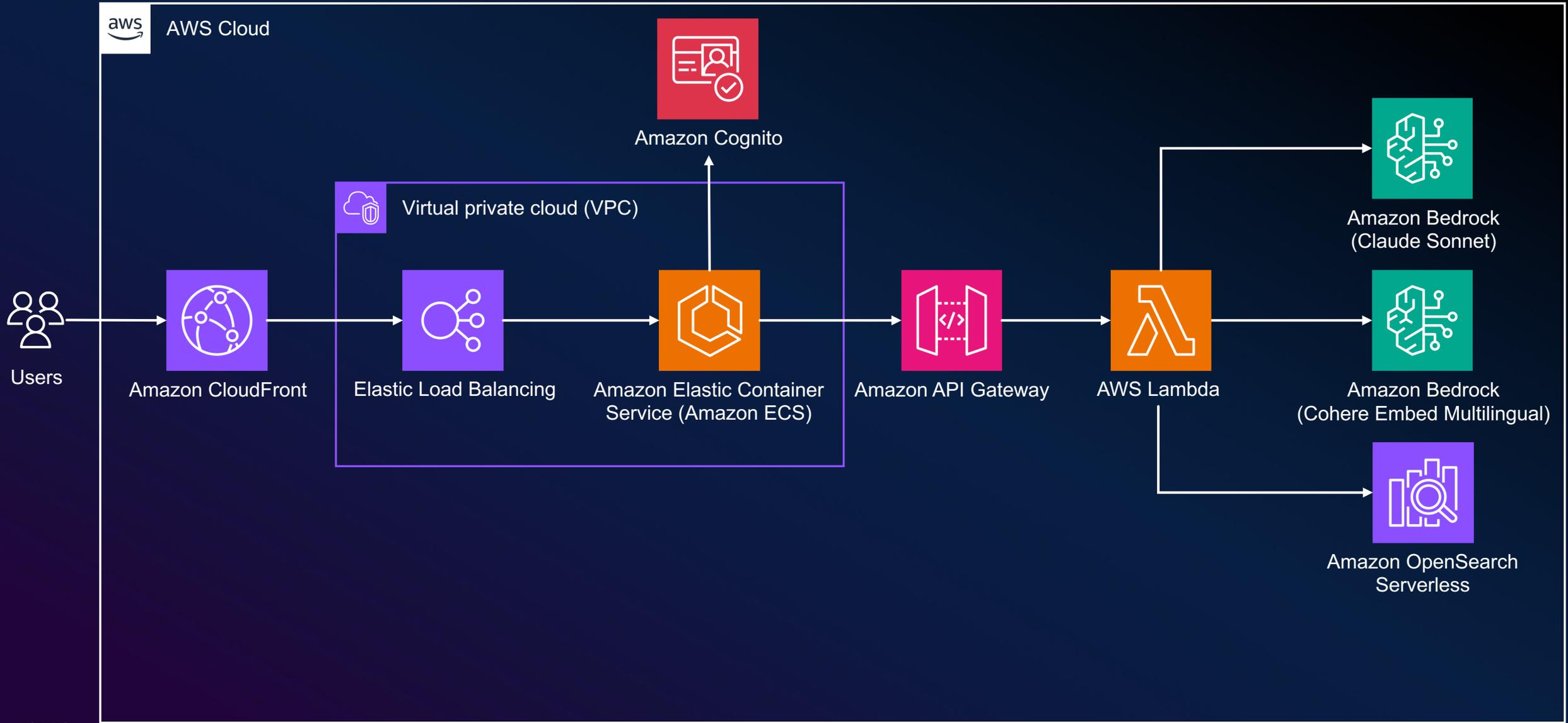
ホーム画面の操作方法を教えてください

使用する辞書の変更方法を教えてください

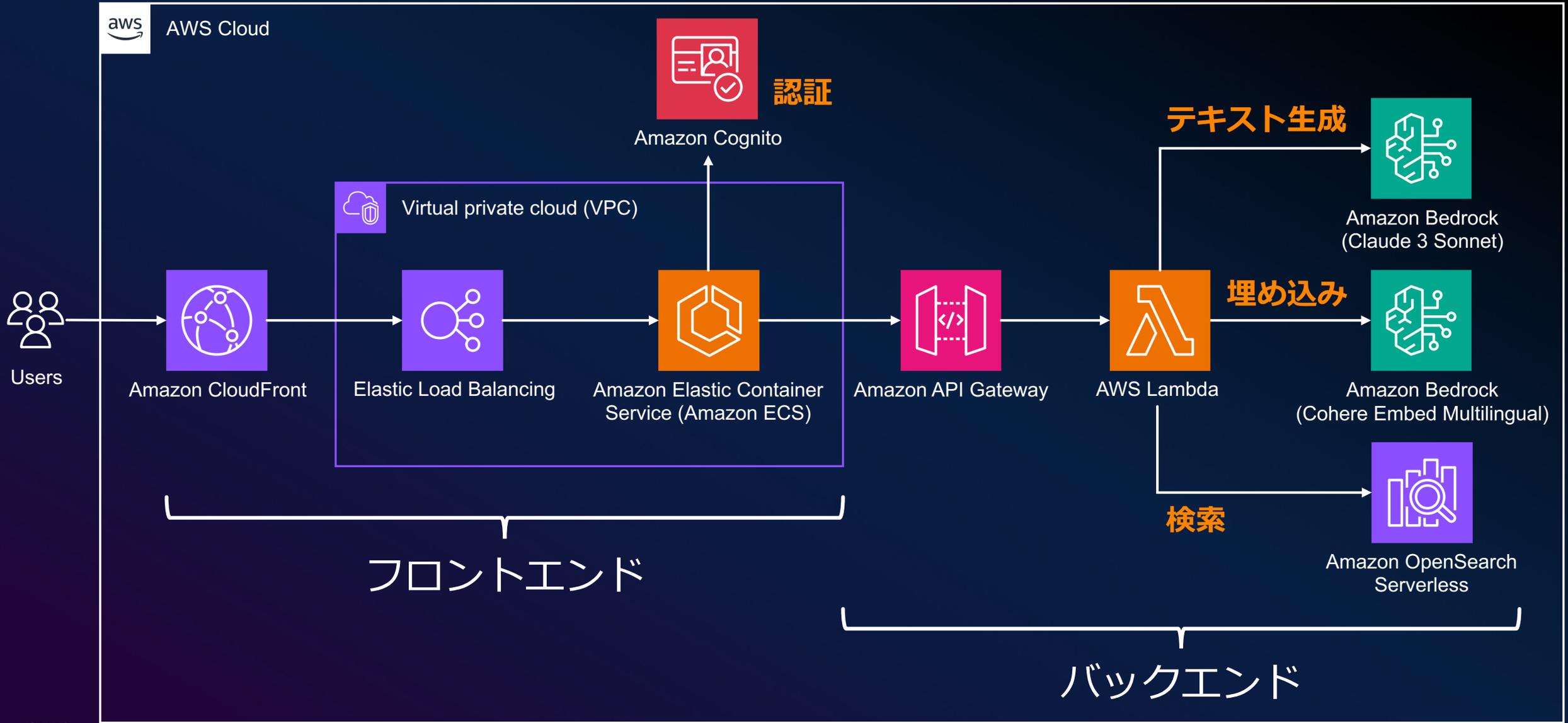
会話しましょう >

Parameters

アプリケーションのアーキテクチャ



アプリケーションのアーキテクチャ



マルチモーダル RAG の活用例と ブース案内

マルチモーダル RAG の活用例



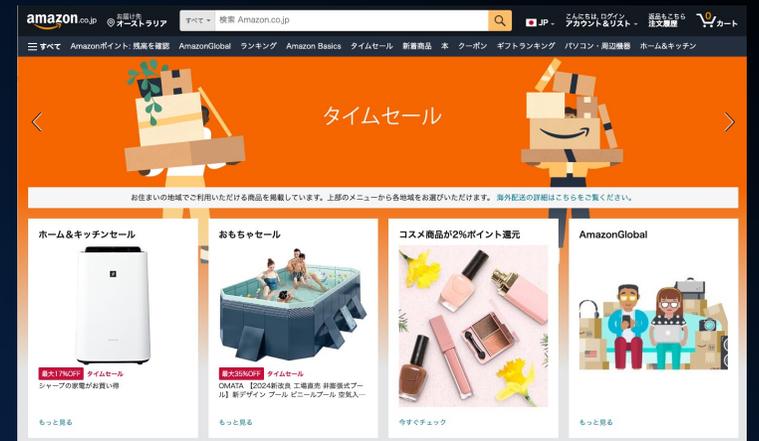
カスタマーサポート

お客様からの問い合わせに
画像も対応



品質管理

品質管理として写真を撮り、
何が異常なのかをドキュメントや
過去の点検履歴と紐付ける



EC サイト

「1000円以下の黒いスマホケース。
この画像と似た商品が良い。」
と検索して UX を高める

まとめ

- 現状、多くの RAG アプリケーションにおいて、テキストデータしか扱えない課題がある
- 上記の課題を解決する **マルチモーダル RAG アプリケーション** をご紹介
- ブースではデモで見せたアプリケーションの体験、技術的な質問やユースケースの相談など出来ますので、ぜひお立ち寄りください！！

【 AWS Summit Japan 2024 Overall View 】



Thank you!

山下 一樹

AWS Japan G.K.
Solutions Architect

鈴木 大樹

AWS Japan G.K.
Solutions Architect

