

RAGとAWS S3を用いたコンテキスト 認識型AIガイダンスシステム

Context-Aware AI Guidance System Using RAG and AWS S3

開志専門職大学 情報学部 野村 空仁
指導教員：Pann Yu Mon



背景

- 1, 開志専門職大学では臨地実務実習がある
- 2, 実習の最後に成果や活動報告用のポスターを作る
- 3, ポスターが静的データとして保存され蓄積していく
- 4, 後輩たちは実習先を決める参考に過去のポスターを見る
- 5, ポスターを元にしたチャットボットをつくろう！

後輩のためのAIを作ろう！

▶ 臨地実務実習 詳細

実施期間

2年次：約1ヶ月半

3年次：約3ヶ月半

総実習時間：約600時間

受け入れ企業・自治体数 2024年度

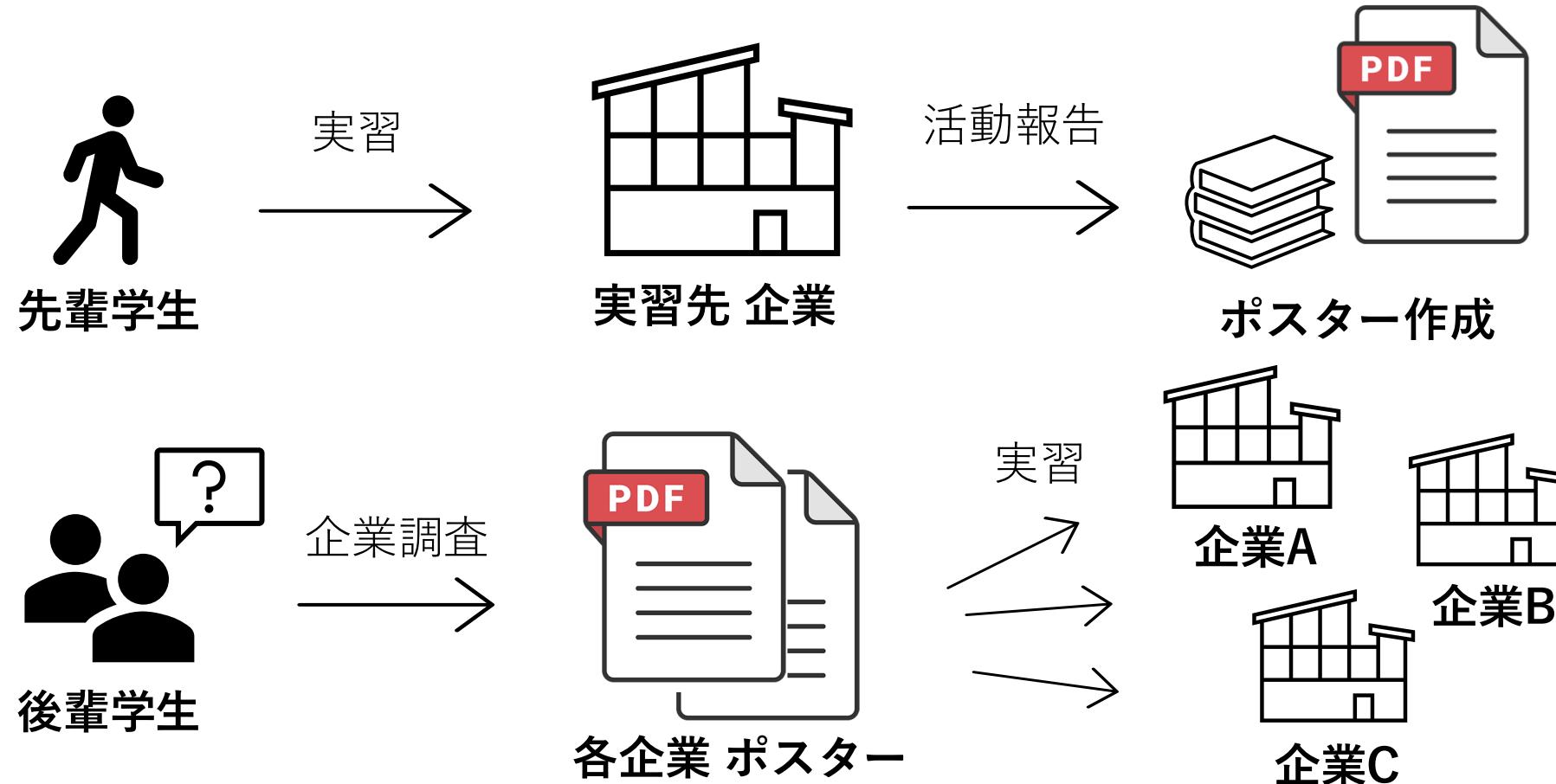
情報学部のみ：約65社



今回はこちらのデータを使用

毎年約120名の学生が参加

臨地実務実習



企業と学生のマッチング

ポスター

- 企業名 & 受け入れ人数
- 学習 & 実習の内容
- 使用したソフトウェア、言語
(AWS、Excel、Python など)
- 何をしている企業か
etc.

実習企業	企業A
実習生	(2年) 野村空仁
指導責任者	企業先 責任者様 名前
担当教員	担当教員 名前

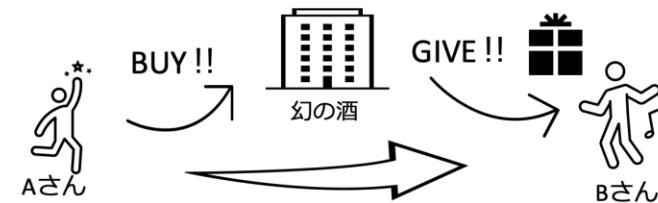
テーマ データ整理と業務の一部自動化

実習先の概要

社員12名程度、服装自由、実習のタスクは特に決まってなく自由な感じ

幻の酒は日本酒を主軸とした商品をネット販売しており、プレゼントや贈り物用に包装や名入れをし、購入者ではなく購入者が贈り物をしたい相手に直接配送する

還暦祝いや退職祝いなどのお祝いイベントでの需要が高い



背景と目的

コロナ禍もあってなのか楽天での幻の酒の売り上げが下がっていた。

しかし、そもそも楽天に出品しているお店の全体が下がっていっているのではないか?という仮説があった。

そこで、楽天の売上ランキングで自社製品が大体どれくらいの位置にいるのかを把握し、仮説の是非を確かめるために毎日自動で自社製品のランキングのデータを取れるフローを組んだ。

課題

どの分野にしよう？



どんな企業があるの？



後輩学生達

具体的すぎる！

量が多すぎ！

全体が見えない

ガーン



静的データ



RAGチャットボットとは？

RAG (Retrieval-Augmented Generation) とは？

- Retrieval (検索) → データベースから関連情報を検索する
- Augmented (拡張) → 検索した情報を回答生成時に加える
- Generation (生成) → 加えられた情報をもとに回答を生成

RAGチャットボット



クラウドを学ぶのにおすすめの企業はありますか？

臨地実務実習やインターンシップでクラウド関連の学習ができる企業としては、[企業A]が挙げられます。この企業はGoogle Cloudの最先端技術を軸としたデータエンジニアリングに強みを持っており、全国からフルリモート勤務が可能な制度を活かしています。実習内容としては、ECサイト作成やデータ分析の実践などが含まれています。^[1]

また、実習先では様々なクラウドツールやサービスを学ぶ機会があります。例えば、Notion（タスク管理）、Slack（各種連絡）、Google Meet（会議・発表）などのクラウドツールを実際に業務で使用する経験ができます。これらのツールはリモートワーク環境でのプロジェクト管理やコミュニケーションに役立ちます。^[2]

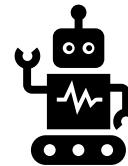
インターンシップでは、Kintoneというクラウドベースの業務アプリ開発プラットフォームを学ぶ機会もあります。実習では、Kintoneを使って作業日報、顧客管理、インシデント管理、減価償却などのアプリを開発したり、Kintoneと他のサービス（Excel、Teams、ChatGPT）との連携方法を学んだりすることができます。これらの経験は、クラウドベースのアプリケーション開発スキルを向上させるのに役立ちます。^{[3][4][5]}

さらに、AI関連の学習ができる実習先もあります。例えば、動画生成AIサービスの調査や、3Dモデル生成AIサービスの調査、自動走行機能付きレースゲーム開発などのプロジェクトに取り組める実習があります。これらの実習では、Pictory.AIやInviedo AIなどの実際のAIツールを使用して、大学の紹介動画作成などの実践的な課題に取り組むことができます。^{[6][7]}

プロンプトを記述(Shift + Enter キーを押して新しい行を開始し、Enter キーを押して応答を生成します)



活用例



クラウドを学習にオススメの企業先はありますか？

おすすめの企業は
株式会社○○○○です！
ここではPythonを使用した
クラウドベースのアプリ開発
を学ぶことが出来ます。



動作テスト

企業A についてわかりやすく丁寧におしえて



企業A は、新潟市 企業A 住所情報 に所在する企業です。主に以下の事業を展開しています：

- IT事業：
 - IT支援コンサルティング
 - Webシステム開発・設計
 - スマートフォンアプリ開発・設計
 - Webサイト、ネットショップ制作・運営

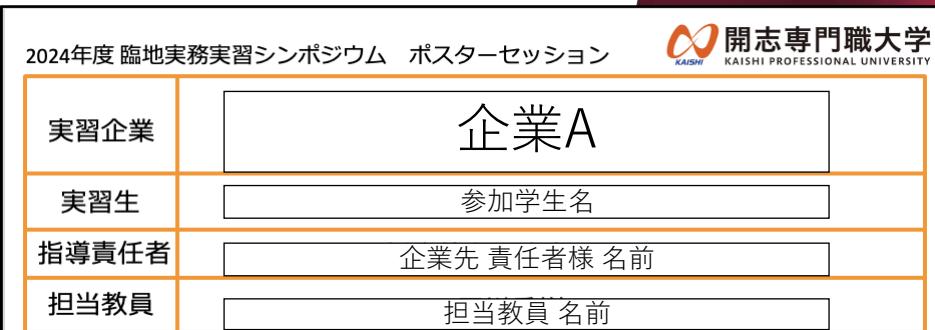
教育事業：

- ITコンサルティング
- 職業訓練・パソコン教室運営
- 電子書籍・Eラーニング教材開発⁽¹⁾

モリパワー株式会社では、開志大学情報学部の学生が臨地実務実習を行っています。実習では「開志大学情報学部学習支援員予約システムの設計」をテーマに、画面数を最小化してわかりやすさを重視した設計方針で取り組んでいます。実習期間は2023年9月25日から11月13日（実習1）および9月25日から1月17日（実習2）となっており、学習支援管理者へのヒアリングを3回実施し、機能設計書、画面フローチャート、画面UIデザインなどを作成しています。

指導責任者 企業先 責任者様 名前 が務めています。 [2]

生成時間
約15秒 (14.55)



テーマ 開志大学情報学部学習支援員予約システムの設計

実習先の概要

- ・企業名：モリパワー株式会社 所在地：〒950-0916 新潟市中央区米山4丁目1-31 紫竹総合ビル412号室
 - ・概要：IT事業（IT支援コンサルティング、Webシステム開発・設計、スマートフォンアプリ開発・設計、Webサイト、ネットショッピング制作・運営）
 - ：教育事業（ITコンサルティング、職業訓練・パソコン教室運営、電子書籍・Eラーニング教材開発）

背景と目的

学習支援員予約システムとは

学習支援員予約システムは、利用者と学習支援員の間で効率的に予約やスケジュール管理を行うためのシステム。

学習支援管理者



今使ってる予約システムは

- ・外部に委託しているから学生や支援員のデータの追加削除に費用がかかる。
- ・支援員の振り分けやメール送信が手動だから時間がかかるし負担が大きい。

作業の自動化を導入することで業務の負担を軽減するとともに、現在利用しているものと入れ替えて運用費用の削減を図る。

支援タスクの円滑化のための新機能の導入。

旧システム	入れ替え	新システム
管理者画面のみ		管理者画面 + 個別支援員画面
文章のみで登録		文章 + 画像添付

実習内容

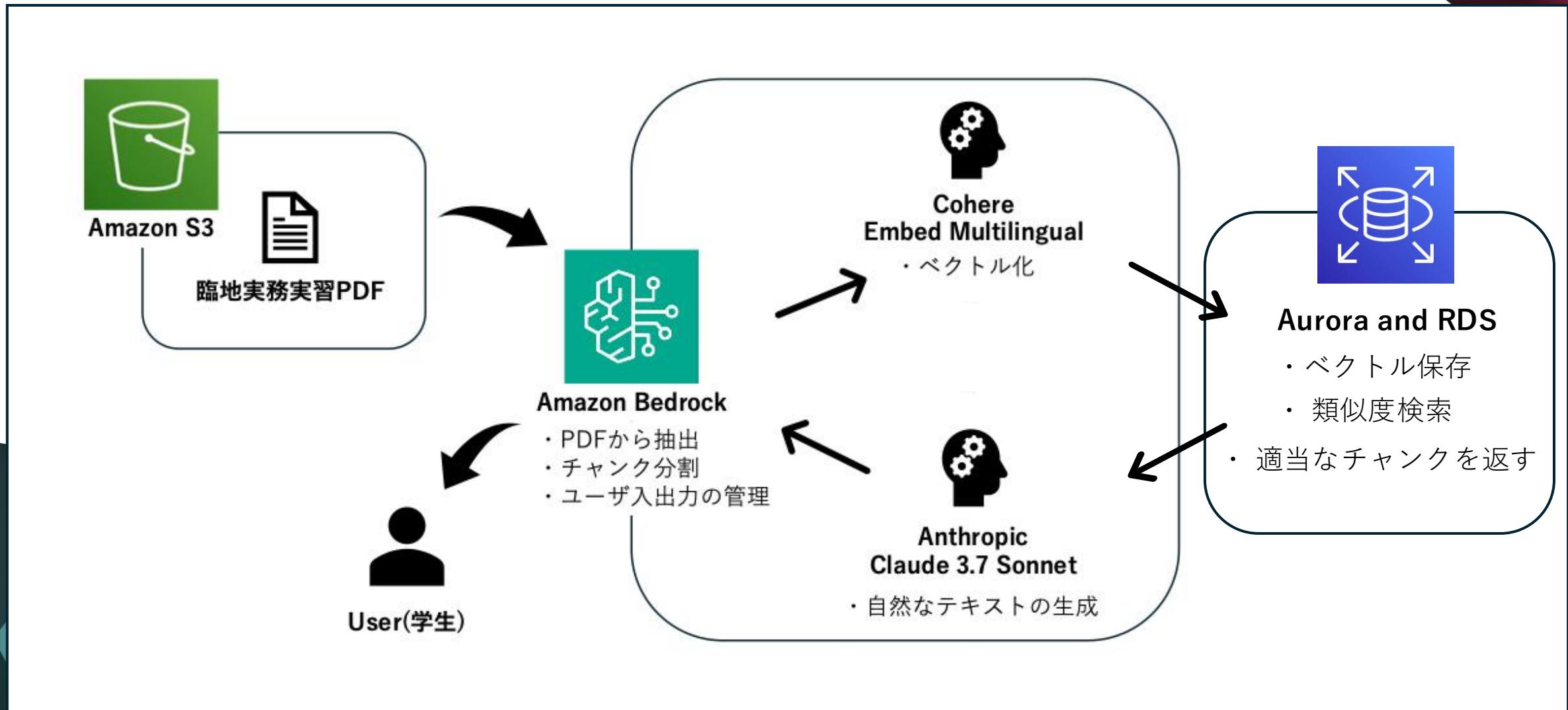
期間：9月25日～11月13日

- ・設計方針：画面数を最小化（わかりやすさ重視）
- ・ヒアリング：学習支援管理者 3回実施
- ・アウトプット：機能設計書、画面フローチャート、画面UIデザイン

期間：9月25日～1月17日

- ・設計書をもとにデータベース設計、システム構築
- ・Web開発プラットフォームLaravelを使用

システム構成図



▶ 各種AWSサービス概要

Amazon S3

- ・様々なファイル保存ができるサービス
- ・35社分の実習ポスターPDFを保存
- ・ポスターの総データ量 51.8MB



Aurora and RDS

- ・埋め込みモデルからの数値ベクトルを保存
- ・類似度検索をしてチャunkを返す

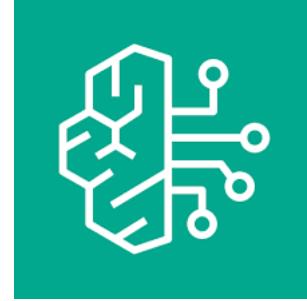




各種AWSサービス概要

Amazon Bedrock

- ・チャック割
- ・チャット入出力の管理
- ・AIモデルの切り替え
- ・1~100でチャック割りを設定可能（今回は100）



僕が主役だよ！

埋め込みモデル

- ・ユーザの質問からベクトル化
- ・数値ベクトルをAurora and RDSへ



Cohere
Embed Multilingual

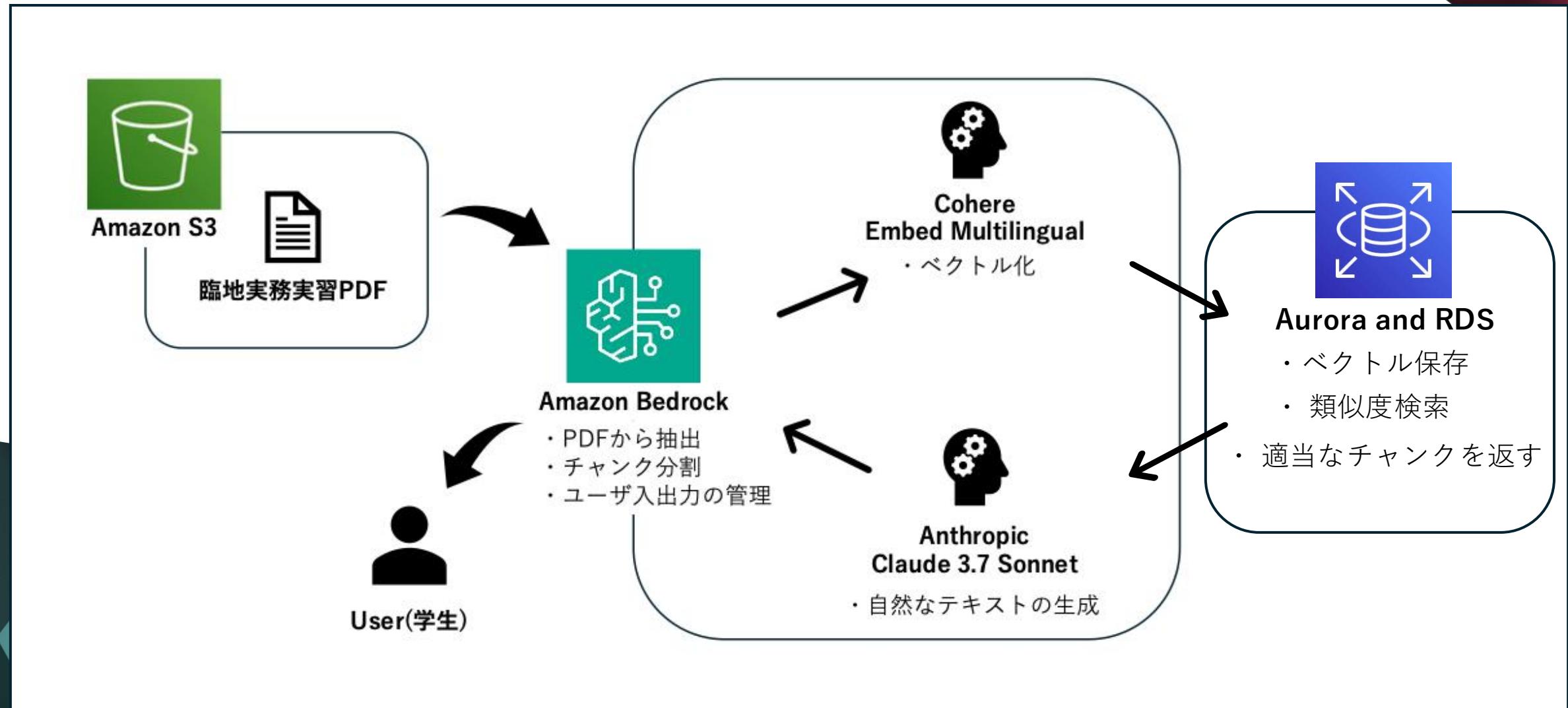
埋め込みモデル

- ・適当なチャックを元に自然な文章を作成



Anthropic
Claude 3.7 Sonnet

システム構成図



まとめ

簡易解説

- 1, ポスターのデータをチャンク分け
- 2, ユーザの質問に近いチャンクを探し出す
- 3, そのチャンクを根拠に回答をユーザに返す

すなわち...

統計的なデータ出力が苦手...
例えば、「〇〇年度の受け入れ企業数は?」など



参考文献一覧

リレーションナルデータベースと非リレーションナルデータベースの違いとは?
/ AWS公式ウェブサイト

<https://aws.amazon.com/jp/compare/the-difference-between-relational-and-non-relational-databases/>

Amazon Bedrock / AWS公式ウェブサイト

<https://aws.amazon.com/jp/bedrock/>

機械学習における埋め込みとは? / AWS公式ウェブサイト

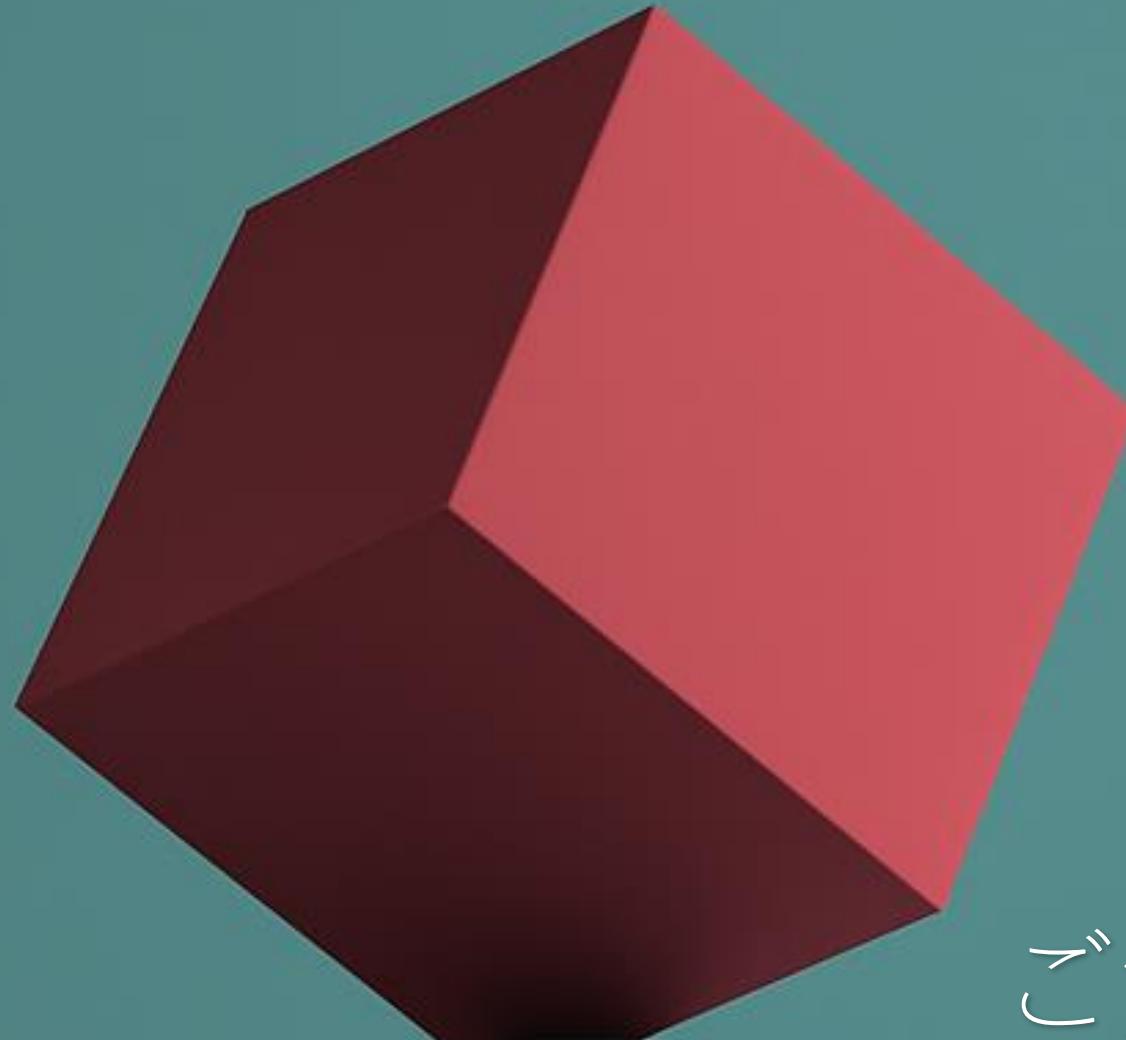
<https://aws.amazon.com/jp/what-is/embeddings-in-machine-learning/>

Amazon Relational Database Service / AWS公式ウェブサイト

<https://aws.amazon.com/jp/rds/>

ナレッジベースでのコンテンツチャンкиングの仕組み/ AWS公式ドキュメント

https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/kb-chunking.html



ご清聴ありがとうございました