

# Bedrock ナレッジベースの構造化データ 取得 (Text to SQL) をやってみた

株式会社 Waalsforce

篠田 敬廣



# 注意点

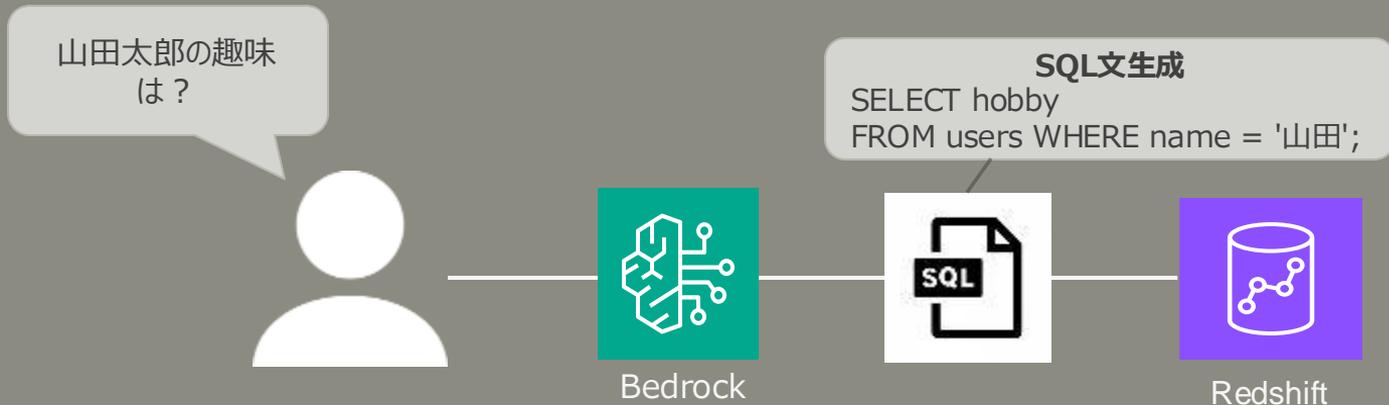
- 資料の内容は2024/12/19時点の情報となります
- 本登壇で説明する内容は個人の見解も含まれます
- 資料の内容にAWS公式と相違があった場合、AWS公式を正とさせていただきます

- ナレッジベースの構造化データ取得
- 実際にやってみた
- Amazon Q generative SQL (時間あれば)

- ナレッジベースの構造化データ取得
- 実際にやってみた
- Amazon Q generative SQL

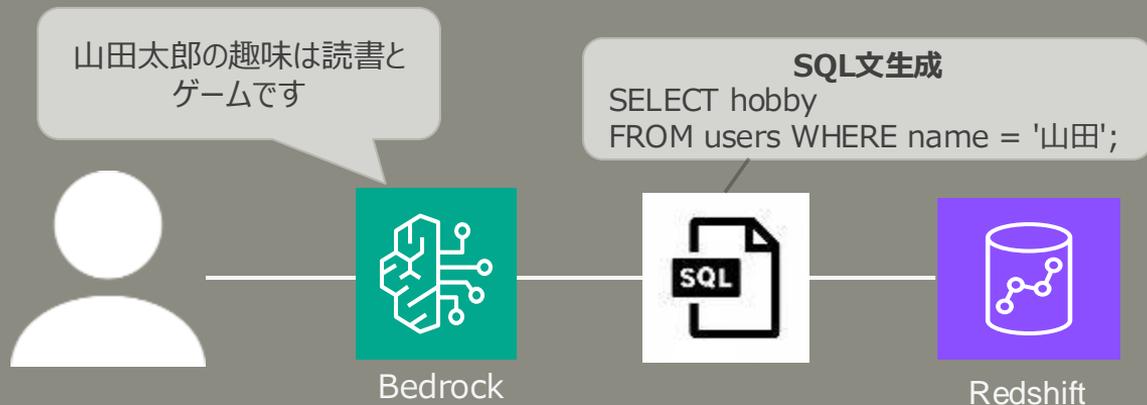
# ナレッジベースの構造化データ取得とは？

- Bedrockのデータ参照先をRedshift（構造化データ）にできる
- 自然言語からSQL文を作ってRedshiftを参照してくれる  
→**Text to SQL to Text**



# ナレッジベースの構造化データ取得とは？

- Bedrockのデータ参照先をRedshift（構造化データ）にできる
- 自然言語からSQL文を作ってRedshiftを参照してくれる  
→**Text to SQL to Text**



# うれしいこと

- 既存のRedshiftをBedrockのナレッジベースに使える
- 構造化データ（RDB）をデータソースにできる
  - ナレッジデータをイチから準備する手間が省ける

- ナレッジベースの構造化データ取得
- 実際にやってみた
- Amazon Q generative SQL

# Redshift側の事前準備

- データベース
- テーブル（テスト用データ）

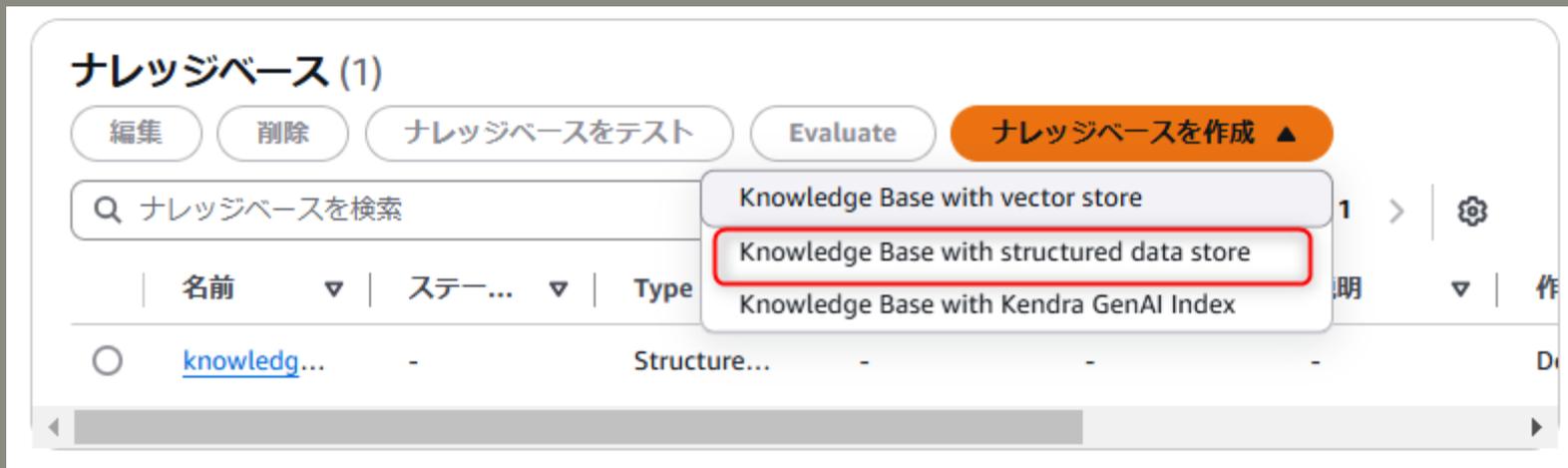
※Redshift サーバレスを使用しました

## 投入したテストデータ

id	name	email	aws_certifications	skills	hobbies	created_at
1	山田太郎	taro.yamada@example.com	SAA, AIF	Python, AWS Lambda, Redshift	ゲーム, 読書	2024-12-08
2	鈴木花子	hanako.suzuki@example.com	DAS	SQL, ETL, Data Pipeline	料理, 写真	2024-12-09
3	John Doe	john.doe@example.com	DOP	Docker, CI/CD, CloudFormation	旅行	2024-12-10

# Bedrockの事前準備①

- ナレッジベースから「Knowledge Base with structured data store」で作成する
- 参照先のRedshiftとDBを選択（割愛）



# Bedrockの事前準備②

- Bedrock Knowledge Basesが使用するIAMロールに「Redshift」のSelect権限を付与する必要がある

✔ Knowledge Base 'knowledge-base-quick-start-vcz9z' created successfully. Before you sync the query engine, you must grant "select" permissions on the database schema or tables to the selected authentication option. For IAM role authentication, create a user for this role with the name 'IAMR:AmazonBedrockExecutionRoleForKnowledgeBase\_vcz9z' and grant the permissions to that user. For more information, [see link](#). 



データソースに移動

- Redshift側でIAMロールにSelect権限を付与するクエリ実行

```
CREATE USER
"IAMR:AmazonBedrockExecutionRoleForKnowledgeBase_vcz9z"
WITH PASSWORD DISABLE;
```

# 試してみる

## ナレッジベースをテスト

回答を生成

Claude 3.5 Haiku v1 | オンデマンド  
変更

山田太郎 の趣味は？

山田太郎さんの趣味は、ゲームと読書です。<sup>[1]</sup>

[ソースの詳細を表示 >](#)

ここにメッセージを入力

▶ 実行

## 山田太郎の趣味は？と聞くと

「ゲームと読書」と正しい情報を回答してくれた

name	email	aws_certifications	skills	hobbies
山田太郎	taro.yamada@example.com	SAA, AIF	Python, AWS Lambda, Redshift	ゲーム, 読書
鈴木花子	hanako.suzuki@example.com	DAS	SQL, ETL, Data Pipeline	料理, 音楽
John Doe	john.doe@example.com	DOP	Docker, CI/CD, CloudFormation	旅行



next.js のスキルを持っている人はいますか？



申し訳ありませんが、Next.js のスキルを持つ人材は現在データベースに登録されていないようです。検索条件に合致する人材が見つかりませんでした。<sup>[1]</sup>

[ソースの詳細を表示 >](#)

**Next.jsのスキルを持っている人がいるか？**と聞くと

Next.jsのスキルを持った人は「**いない**」と回答をしてくれた

# Retrieval and response generationの種類

設定 Info Copy

× 回答を生成

## Retrieval and response generation

Retrieval only: data sources

Retrieval only: data sources

Retrieval and response generation: data sources and model

Query the data sources and generate responses based on the retrieved results by using the specified foundation model.

Generate SQL queries

Generate SQL queries for retrieving information from the data sources

データソースの参照のみ

自然言語で回答

SQL文の生成

山田



Row 1

1. name: 山田太郎
2. aws\_certifications: SAA, AIF
3. skills: Python, AWS Lambda, Redshift
4. hobbies: ゲーム, 読書

山田



山田太郎さんは、AWS認定資格としてSAAとAIFを保有しています。スキルとしてPython、AWS Lambda、Redshiftを持っており、趣味はゲームと読書です。<sup>[1]</sup>

[ソースの詳細を表示 >](#)

山田



```
1. SELECT "name", "aws_certifications", "skills", "hobbies" FROM public.personal_profile WHERE "name" LIKE "%山田%";
```

# Redshiftとのデータ同期

- Redshiftでデータを更新してもBedrockからはリアルタイムで同期はされない
- 最新データに保つには都度Bedrock側でデータを同期させる必要がある

**Query engine (1)** 🔄 同期

Contains configurations and additional information for how the database is queried.

🔍 データソースを検索

< 1 >

Redshift data... ▼	Status	Engine	Redshift cluster	Authentication	Last sync date
● dev	COMPLETE	SERVERLESS	-	-	Thu Dec 19 202...

- ナレッジベースの構造化データとは？
- 実際にやってみた
- **Amazon Q generative SQL**

# Amazon Q generative SQL

- Amazon Qから 自然言語でSQL文を生成できる
- 作ったSQL文をワンクリックでクエリに渡せる（実行できる）
- （確か）当初は英語のみ対応だったが今は日本語も対応している
- 2023年のre:inventで発表されて機能
- 2024年10月にGA



Run Limit 100

Use Ctrl-Space for autocomplete

```
1
```

Row 1, Col 1, Chr 0

Markdown editor

```
1 ### SAA を持っている人を調べたい
```

Row 1, Col 21, Chr 20

Run Limit 100

```
1 SELECT
2   "name",
3   "aws_certifications"
4 FROM
5   public.personal_profile
6 WHERE
7   "aws_certifications" LIKE '%SAA%';
```

Row 7, Col 37, Chr 11

Result 1 (1)

name	aws_certifications
山田太郎	SAA, AIF

### Amazon Q generative SQL

Before using them in any environment or workload. To learn more about me, see [Interacting with query editor Q generative SQL](#)

Connect to a database and enter your prompt.

For examples using the tickit data, see [Example using Q generative SQL Assistant with the TICKIT data](#)

"Which state has most venues"

"Find the top five users from Seattle who bought the most number of tickets"

SAA を持っている人を調べたい

```
1 SELECT
2   "name",
3   "aws_certifications"
4 FROM
5   public.personal_profile
6 WHERE
7   "aws_certifications" LIKE '%SAA%';
```

Add to notebook

Ask me a question to generate a SQL suggestion for your connected database.

# まとめ

- 「Bedrock ナレッジベースの構造化データ取得」はBedrockナレッジベースのデータ参照先をRedshiftにできる
- Text to SQL to Text で自然言語からSQL文を生成し、自然言語で回答を得られる
- Redshiftとのデータ同期は都度必要（リアルタイム更新されない）
- Amazon Q generative SQL は自然言語からSQL文を生成できる