



「データエンジニアリングへの道」投資案件 活動結果報告

2025/2/26
クラウドソリューションG
神原 鈴木莉

■目次

1. 背景・目的
 2. 環境構築
 3. 検証内容
 4. 検証結果
- まとめ

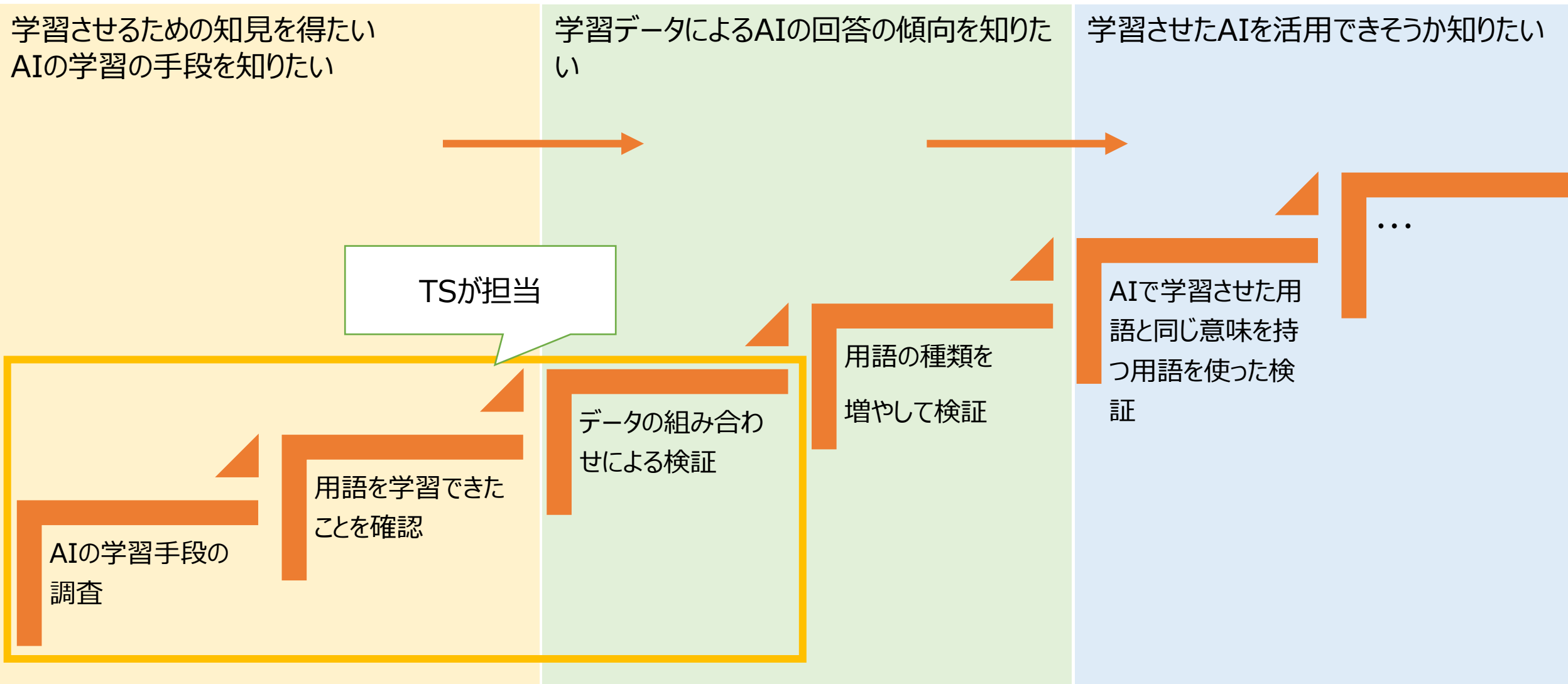
■背景：

- 将来的に利用者の目的に合ったAIモデルを提供することをビジネスモデルとしたい
- 学習させたAIモデルを使用することでRAGの精度や応答速度の限界を無くしたい

■目的：

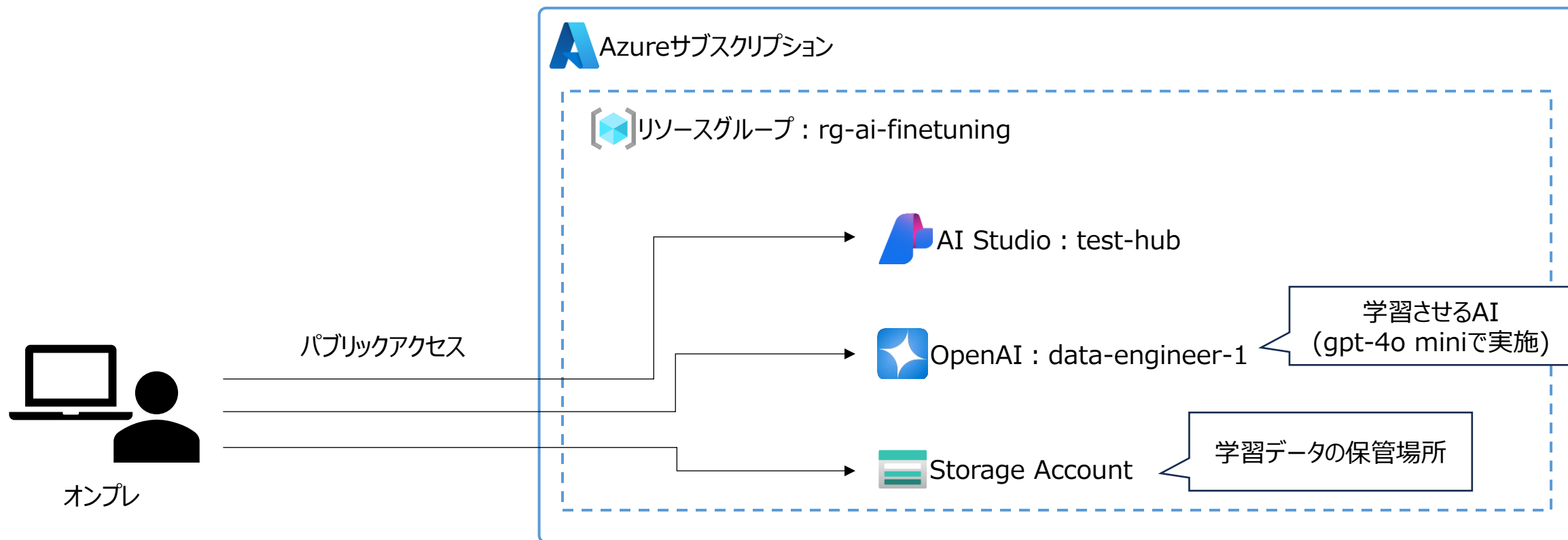
- AIモデルに学習させる方法を知る
- AIモデルに日本語で作成したデータを学習させる
 - 求めている「回答」に対して必要な学習データの検討

■ 全体の流れ



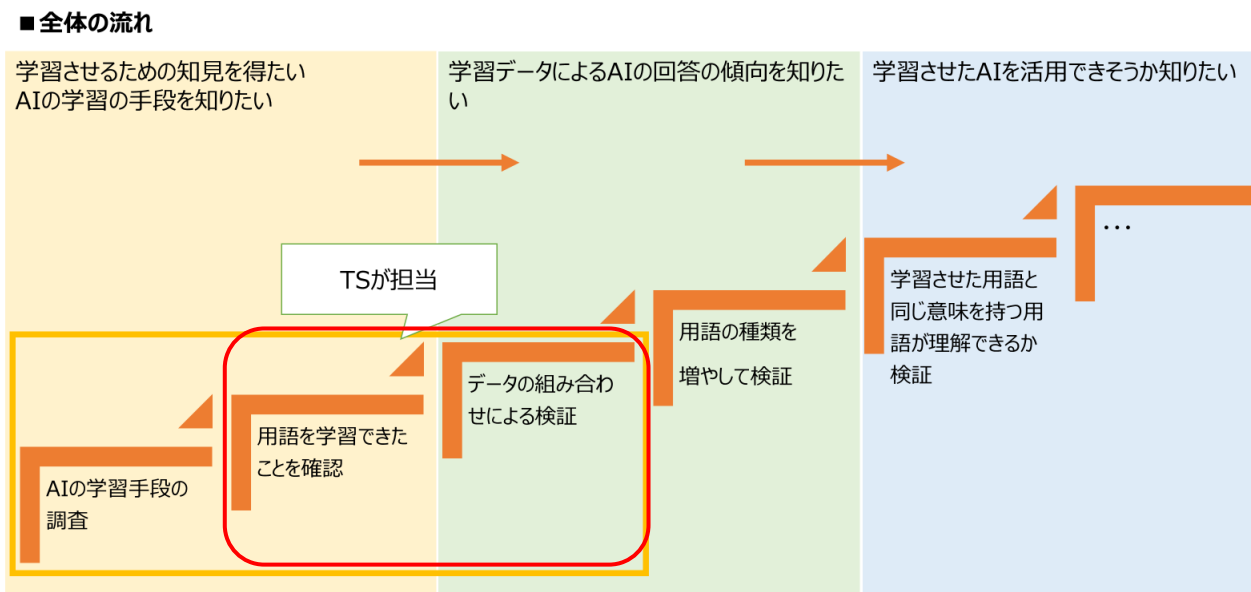
■ 検証環境の構築

以下のリソースをAzureに構築



○マ：トヨタ用語で“マイナーチェンジ”

■ 検証内容



検証項目	結果
No1. 「○マとは～」の学習データで回答できるかを検証	△
No2. 「～とは○マです」の学習データで回答できるかを検証	△
No3. No1とNo2の混合の学習データで回答できるかを検証	◎

学習後の正答率で評価

◎ : 90%~

○ : 70%~89%

△ : 50%~69%

× : ~49%

■ 検証内容(○マ：トヨタ用語で“マイナーチェンジ”)

☐ No1. 「○マとは～」の学習データで回答できるかを検証

☐ No2. 「～とは○マです」の学習データで回答できるかを検証

☐ No3. No1とNo2の混合の学習データで回答できるかを検証

〇マの説明を問う質問は回答できた 表記ゆれにも対応していた(まるマ、丸マなど)

結果：△

質問数：10

質問結果（一部抜粋 黒字：正答、赤字：誤答）

学習データ（一部抜粋）	
〇マとは何ですか？	〇マ（マルマ）は、トヨタ自動車やその関連企業で使われる用語で、マイナーチェンジを指します。通常、モデルの発売から2～4年後に行われる、デザインや内部部品の部分的な改良を指します。

質問内容	gpt-4o-mini（学習済み）
〇マについて教えて	〇マは、トヨタ自動車で使用される用語で、マイナーチェンジを指します。デザインや内部部品の部分的な改良が行われます。
まるマとは何ですか？	まるマは、トヨタ自動車で使用される用語で、マイナーチェンジを指します。デザインや内部部品の部分的な改良が行われます。
トヨタ自動車やその関連企業で使われる用語で、マイナーチェンジを指すものを何と言いますか？	トヨタ自動車では、マイナーチェンジを「MC」と呼ぶことが一般的です。

■ 検証内容(○マ：トヨタ用語で“マイナーチェンジ”)

- ☐ No1. 「○マとは～」の学習データで回答できるかを検証
- ☒ No2. 「～とは○マです」の学習データで回答できるかを検証
- ☐ No3. No1とNo2の混合の学習データで回答できるかを検証

○マを答えさせる質問は回答できた

結果：△

質問数：10

質問結果（一部抜粋 黒字：正答、赤字：誤答）

学習データ（一部抜粋）

トヨタ自動車で一般的に使用され、デザイン変更や内部部品の改良を含むマイナーチェンジを意味する用語を何と言いますか？

〇マです。

質問内容	gpt-4o-mini（学習済み）
〇マについて教えて	分からない
まるマとは何ですか？	まるマとは、〇マの略で、〇〇マークのことを指します。主に自動車業界では、〇〇の部分に特定の車種名やブランド名が入ることが多いです。例えば、トヨタの〇マは「ト〇マ」などです。
トヨタ自動車やその関連企業で使われる用語で、マイナーチェンジを指すものを何と言いますか？	〇マです。

■ 検証内容(○マ：トヨタ用語で“マイナーチェンジ”)

- ☐ No1. 「○マとは～」の学習データで回答できるかを検証
- ☐ No2. 「～とは○マです」の学習データで回答できるかを検証
- ☐ No3. No1とNo2の混合の学習データで回答できるかを検証

○マの説明を問う質問、○マを答えさせる質問ともに回答できた

結果：◎
質問数：10

質問結果（一部抜粋 黒字：正答、赤字：誤答）

学習データ（一部抜粋）	
○マとは何ですか？	○マ（マルマ）は、トヨタ自動車やその関連企業で使われる用語で、マイナーチェンジを指します。通常、モデルの発売から2～4年後に行われる、デザインや内部部品の部分的な改良を指します。
トヨタ自動車で一般的に使用され、デザイン変更や内部部品の改良を含むマイナーチェンジを意味する用語を何と言いますか？	○マです。

質問内容	gpt-4o-mini（学習済み）
○マについて教えて	○マは、トヨタ自動車で使用される用語で、マイナーチェンジを指します。デザインや内部部品の部分的な改良が行われます。
まるマとは何ですか？	まるマは、トヨタ自動車で使用される用語で、マイナーチェンジを指します。デザインや内部部品の部分的な改良が行われます。
トヨタ自動車やその関連企業で使われる用語で、マイナーチェンジを指すものを何と言いますか？	○マです。

■ まとめ

複数パターンの学習データを用意することで回答精度が上がる

- 学習させたデータに対してのみ正しい回答が得られた
→複数パターンのデータを用いらないと、回答できる質問の幅が広がらない
- 今回は1単語のみで実施し回答することができたが、複数単語の学習データでも回答可能か？
- 今回学習させたAIモデル(用語集を学習済み)でRAGができるか？



このデータの作り方を基に追加で検証を進める必要あり

END

■ 検証内容(○マ：トヨタ用語で“マイナーチェンジ”)

☒ No1. 「○マとは～」のデータで回答できるかを検証

☐ No2. 「～とは○マです」のデータで回答できるかを検証

☐ No3. No1とNo2の混合データで回答できるかを検証

■ 検証内容

- ☐ No1. 学習させた用語によるAI回答の傾向を検証
- ☐ No2. AIで学習させた用語と同じ意味を持つ用語を使った検証
- ☐ No3. 学習時間と学習データ数の相関