

# ダブルにおける Gemini を活用した SQL Agent の取り組み

Google  
Cloud  
Next

Tokyo

Proprietary



# 原 和希

株式会社タップル  
機械学習エンジニア

Google Developer Experts  
(Google Cloud: AI)



Google Developer Expert

2025

Google Cloud:  
AI



# 目次

- 01 弊社サービスの紹介
- 02 導入した SQL Agent
- 03 向き合った課題と理想状態
- 04 開発時の工夫
  - アーキテクチャ
  - コンテキスト
  - AI Agent 専用テーブル
- 05 おわりに

# 01. 弊社サービスの紹介

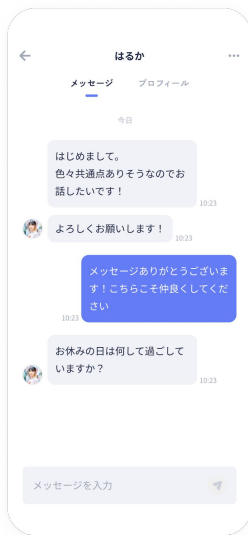
# タップル



2014年5月にサービスを開始。グルメや映画、スポーツ観戦など、自分の趣味や行ってみたいデートスポットをきっかけに恋の相手が見つけれられる **国内最大級のマッチングアプリ** です。年収や身長などの条件だけでなく、**趣味や好きなことから自然な出会いにつながる** ことが特徴です。

<https://tapple.me/>

# Koigram



2024年6月にサービスを開始。「**無料でも本気の恋はできる**」がコンセプトの、**マッチもチャットも全員無料**で利用できるマッチングアプリです。利用を開始するためには、全50問ある「16タイプ恋愛診断」の回答が必須となっており、しっかりと**時間をかけてでも内面を重視して恋愛をしたい**と考える人が集まっています。サイバーエージェントグループが運営しており、24時間365日の監視体制と厳しい監視基準で、安心安全なお相手探しをサポートします。<https://koigram.net/>

## 02. 導入した SQL Agent



## 1. SQLの方針

- 2025/04/21 にログインしたユーザーを `user_id` ベースで重複なく集計すれば DAU を算出できます。
- 対象となる日付のレコードを切り出すサブクエリ `login_users` を用意し、そのサブクエリに対して `count(distinct user_id)` を実行します。
- 使用テーブルはログイン情報を保持している `login_users` のみです。

## 2. 作成したSQL

SQL



```
1 with login_users as (  
2   -- ① 2025/04/21 にログインしたユーザーを抽出  
3   select  
4     user_id  
5   from  
6     login_users  
7   where  
8     dt = '2025-04-21'  
9 )  
10  
11 -- ② 当日ログインユーザー数 (DAU) を算出  
12 select  
13   count(distinct user_id) as dau  
14 from  
15   login_users  
16 ;
```

# タプルの SQL Agent

- 社内に展開
- チャットで対話することで SQL を生成 / 実行可能
  - 作成方針やサブクエリの説明も同時に出力
  - 追加の質問も可能
- 2025 年 2 月から技術調査、4 月末にリリース



2025/04/21のDAUを出して

1. SQLの方針

- 2025/04/21 にログインしたユーザーを `user_id` ベースで重複なく集計すれば DAU を算出できます。
- 対象となる日付のレコードを切り出すサブクエリ `login_users` を用意し、そのサブクエリに対して `count(distinct user_id)` を実行します。
- 使用テーブルはログイン情報を保持している `login_users` のみです。

2. 作成したSQL

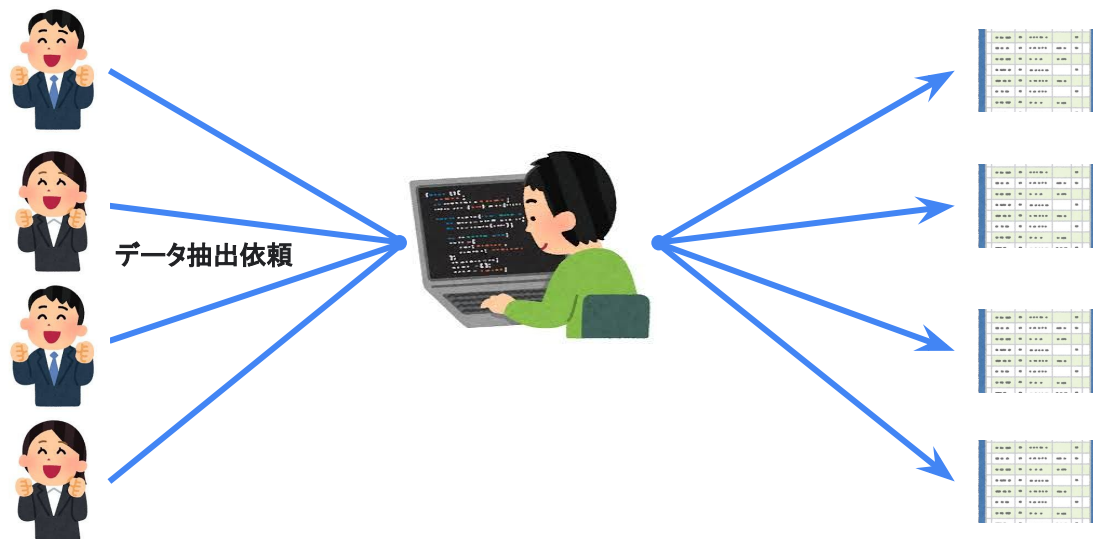
```
SQL
1 with login_users as (
2   -- ① 2025/04/21 にログインしたユーザーを抽出
3   select
4     user_id
5   from
6     login_users
7   where
8     dt = '2025-04-21'
9 )
10
11 -- ② 当日ログインユーザー数 (DAU) を算出
12 select
13   count(distinct user_id) as dau
14 from
15   login_users
16 ;
```

# 03. 向き合った課題と理想状態

# SQL Agent で向き合った課題

BI チームにデータ抽出の**依頼が集中**

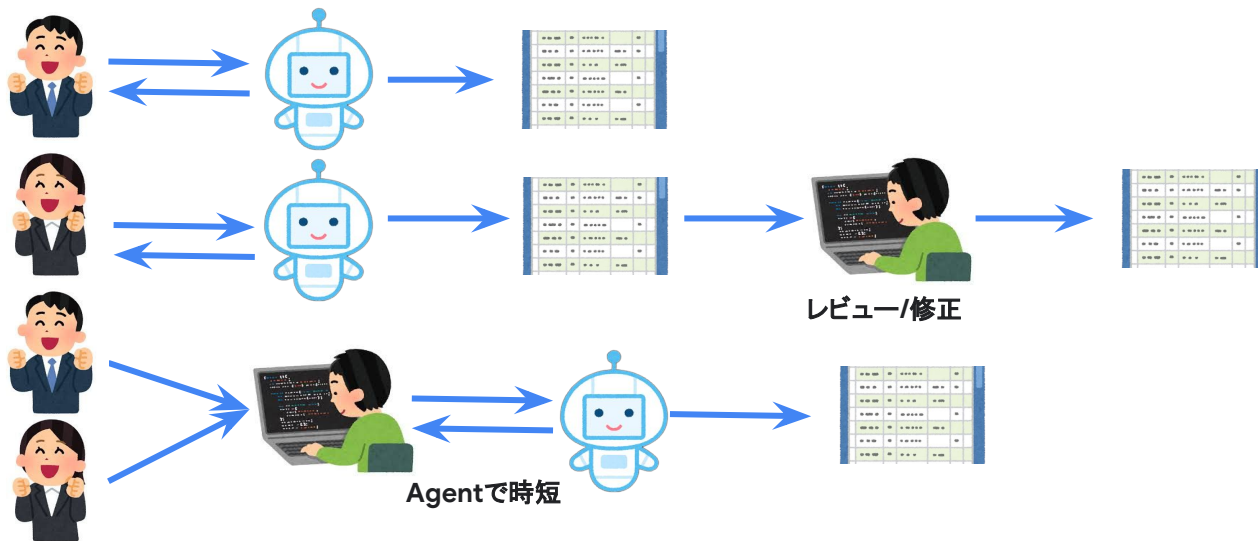
→ 対応可能な量、スピードに限界



# SQL Agent で目指す理想状態

## データに基づいた **意思決定、施策立案のスピード UP**

- SQL を得意としなくとも、誰でもデータ抽出・分析を実現
- BI チームの負荷の軽減



# 04. 開発時の工夫

# 導入の際に意識・工夫したこと

## ① アーキテクチャ

最新の手法を取り入れつつも、シンプルに始める

## ② コンテキスト

「テーブルメタデータ」「社内用語、ビジネス知識」の充実

## ③ AI Agent 専用テーブル

対象データを絞り、生成品質を上げる

# 導入の際に意識・工夫したこと

## ① アーキテクチャ

最新の手法を取り入れつつも、シンプルに始める

## ② コンテキスト

「テーブルメタデータ」「社内用語、ビジネス知識」の充実

## ③ AI Agent 専用テーブル

対象データを絞り、生成品質を上げる

# シンプルに始める

## 他社事例と比較して、ミニマムスタートを選択

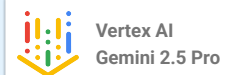
(事例があった企業より、弊社はデータソースが小規模)

まずは導入して、

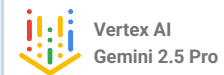
- 効果や課題感、Agent に求める機能を知る
- 品質向上のためにデータを収集する
- 必要に応じて機能拡張していく
  - Agentic な振る舞い、グラフの描画 etc.

今日の20代のマッチ数を  
教えてください

テーブルを選択



SQLの生成



SQLの生成

Other LLMs

SQLのレビュー・修正



SQLのレビュー・修正

Other LLMs

候補の評価・選択

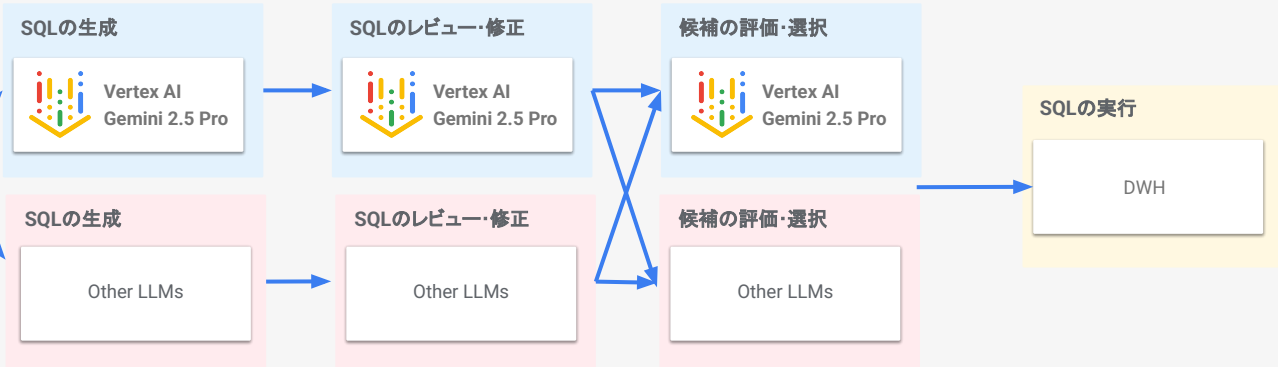


候補の評価・選択

Other LLMs

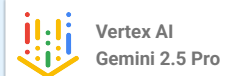
SQLの実行

DWH

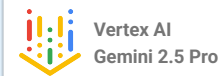


今日の20代のマッチ数を  
教えてください

テーブルを選択



SQLの生成



SQLのレビュー・修正



候補の評価・選択



SQLの生成

Other LLMs

SQLのレビュー・修正

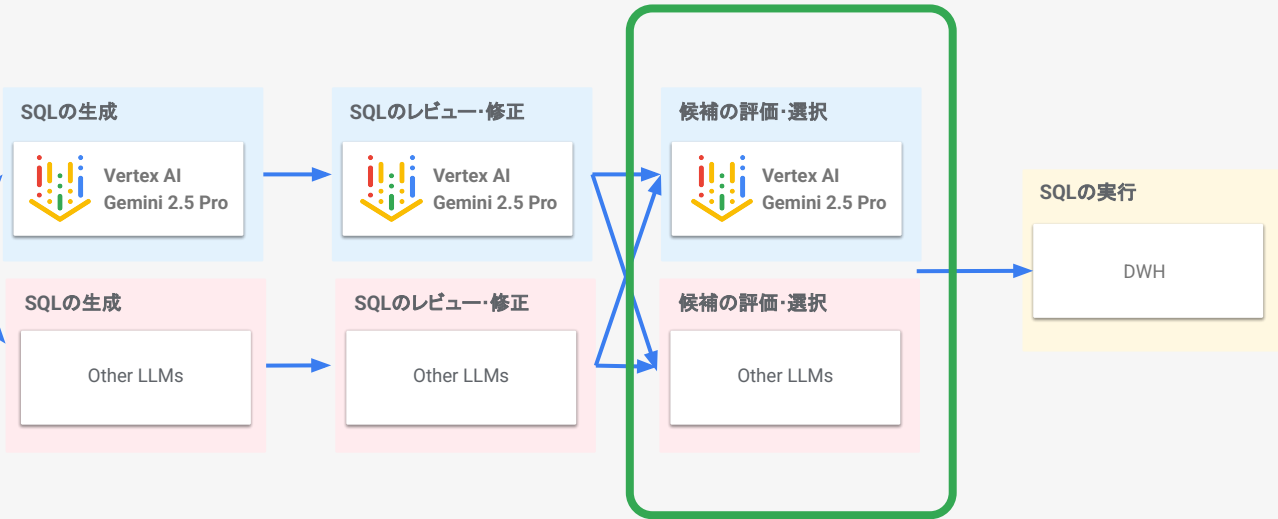
Other LLMs

候補の評価・選択

Other LLMs

SQLの実行

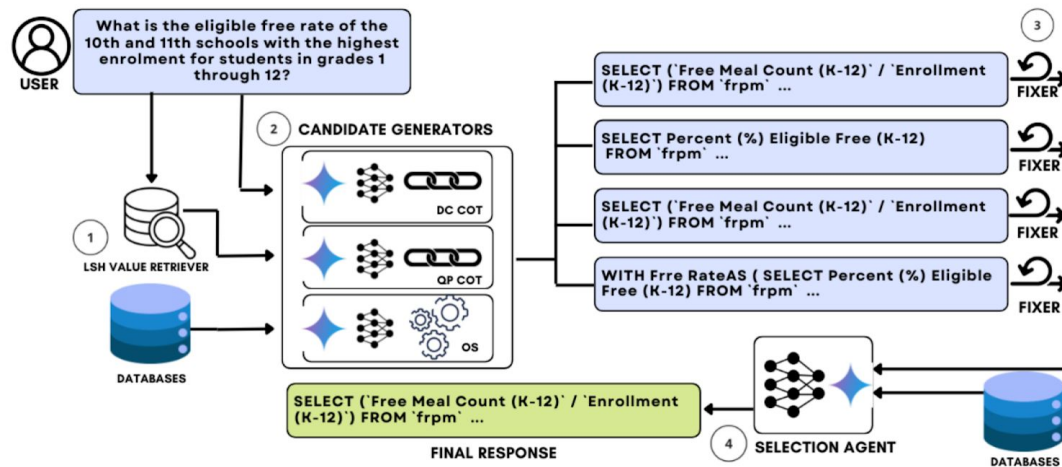
DWH



# CHASE-SQL

Google Cloud や Stanford University の研究者らによる論文

- 3つのアプローチでSQLを生成
- 複数のSQL候補から最適なものを選択



["CHASE-SQL: Multi-Path Reasoning and Preference Optimized Candidate Selection in Text-to-SQL" より引用](#)

# Agent Garden - Data Science Agent -

- ADK ( Agent Development Kit )での実装を確認できる
- Vertex AI Agent Engine にデプロイするガイド付き



Agent Garden

検索

厳選されたエージェント サンプルとツールのセットを発見し、それらについて学び、使用することで、エージェントの開発を加速できます。

サンプル

ソースコード、構成ファイル、ベスト プラクティスの例を含む、事前構築済みのカスタマイズ可能なブループリント。

<p><b>Data Science</b></p> <p>Queries diverse data across multiple sources using natural language, builds predictive models, visualizes trends, and communicates key insights in a clear way.</p> <p>ADK Python</p>	<p><b>Retrieval-Augmented Generation (RAG)</b></p> <p>Uses RAG to get information from specified knowledge sources, ensuring responses are factually grounded, context-aware, and up-to-date.</p> <p>ADK Python</p>
<p><b>Financial Advisor</b></p> <p>Assists human financial advisors by providing educational content about topics related to finance and investments.</p> <p>ADK Python</p>	<p><b>Marketing Agency</b></p> <p>Streamlines new website and product launches. Identifies optimal DNS domains, generates entire websites, develops marketing strategies, and design...</p> <p>ADK Python</p>

Tools & Integrations

- Google Search Custom tool
- Function tool Agent tool
- Java Async tool + 4

Core AI Capabilities & Tasks

- NL2SQL Multimodal
- PDF parsing RAG
- Forecasting
- Time Series Analysis + 6

Data Sources, Types & Management

- Database BigQuery
- Structured data
- Live streaming BQML

Agent Architecture & Design Patterns

- Multi-agent RAG engine
- Input and output schema
- Updatable context

# 導入の際に意識・工夫したこと

## ① アーキテクチャ

最新の手法を取り入れつつも、シンプルに始める

## ② コンテキスト

「テーブルメタデータ」「社内用語、ビジネス知識」の充実

## ③ AI Agent 専用テーブル

対象データを絞り、生成品質を上げる

# LLM のコンテキストに与えるデータの充実

- 社内・ドメイン特有の情報
  - 例)「マッチ」「フリック」はデータ上でどう表現されるか
- テーブルや列の説明や注釈
  - 気合いを入れて記述
- サンプル SQL
  - テーブルやカラムの使い方を教える

# 導入の際に意識・工夫したこと

## ① アーキテクチャ

最新的手法を取り入れつつも、シンプルに始める

## ② コンテキスト

「テーブルメタデータ」「社内用語、ビジネス知識」の充実

## ③ AI Agent 専用テーブル

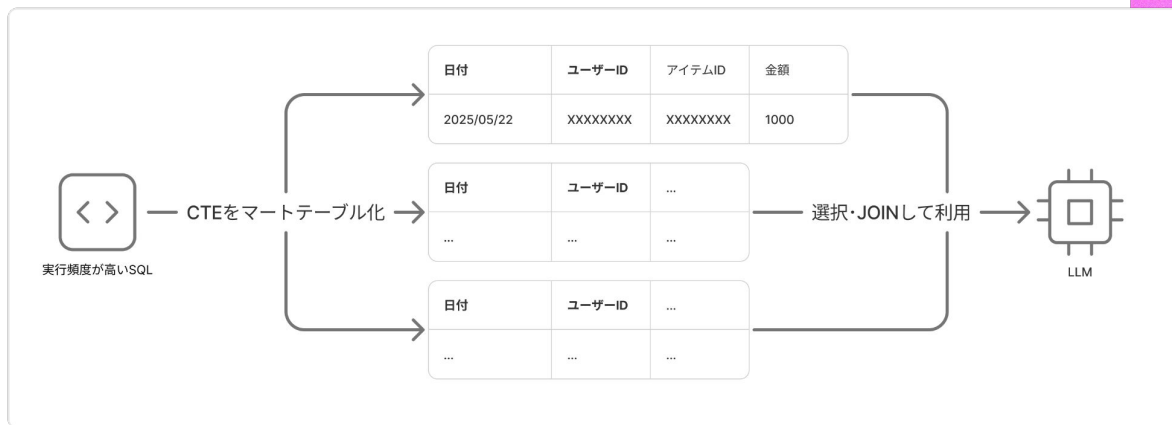
対象データを絞り、生成品質を上げる

# SQL Agent 専用のマートテーブルを作成

【課題】カラム数が多い・JSON 型のカラムを扱う生成品質が低い

## 【アプローチ】

- 利用頻度が高いテーブル・SQL を元に作成
- 「日付」や「ユーザー ID」を共通のキーとする



# 05. おわりに

# Gemini を利用するメリット

- **ロングコンテキストと参照性の高さ**
  - ドメイン知識/ テーブルメタデータを丸ごと挿入可能
  - 正確に参照可能
- **コスト**
  - 他プロバイダーの同等品質のモデルと比較して安価
- **生成品質**
  - 候補選択にて Gemini 2.5 Pro のセレクト割合が高かった

# 開発を通しての気づき

## ドメイン知識やメタデータを言語化する重要性

- 資産化: より優れた Agent が出ても再利用可能

これは、特定の知識や業務プロセスを言語化すること。

→ 知識の汎用化 / 属人化の排除に AI Agent は有効打になり得る

# 導入・運用してみたの課題

- **対象テーブル不足によるエラー**
  - 充実・テーブルの整備が必要
- **人によって依頼の粒度・タスクにばらつき**
  - 依頼内容の深堀り
- **出力データの正確性の保証**
  - 既存のクエリを活用