

NTTドコモが実現したデータ民主化と データ エージェントへの 挑戦

Google
Cloud
Next

Tokyo

Proprietary





吉田 祥平

データプラットフォーム部
担当部長



藤平 亮

データプラットフォーム部
Principal Data Scientist



アジェンダ

01 データ民主化の課題と挑戦

ドコモのデータ資産

圧倒的に使いやすいデータ活用環境

02 データ プラットフォームの概要

システム概要

プラットフォームエンジニアリング

03 事例: お客様の声アプリ

BigQuery ML

Iceberg

01. データ民主化の課題と挑戦

ドコモのデータ資産

日本最大級の 会員 顧客基盤

dポイントクラブ会員数 約 **1** 億※1携帯電話サービス契約数 約 **7,200** 万※1※2

会員一人ひとりに紐づく多種多様なデータ ※3

契約者情報	 年齢・性別  居住地  誕生日  趣味趣向 (アンケート)  家族構成 (アンケート)  世帯収入 (アンケート)	プロファイリング AI
オンライン 行動データ	 検索クエリ  閲覧ニュース (dmenu他)  アプリ利用 ログ  キャリア決済 情報  dマーケット サービス利用ログ  dmenu 利用ログ	
オフライン 行動データ	 位置情報 (Wi-Fi)  位置情報 (基地局)  位置情報 (GPS)  dポイント 加盟店情報  d払い 加盟店情報  購買情報 (IDoPOS連携加盟店)	

※1 2025 年 3 月末時点 ※2 home5G、モジュール除く ※3 個人を特定するものではありません。

データ活用のありたい姿

あらゆる役割・業務で使いこなし、価値創出につなげている

経営層



管理職



担当者



データソース

顧客属性

ネット行動

リアル行動

AI

会員理解

心理学的

行動学的

地理学的

人口統計学的

事業価値・顧客価値

プロファイリング

行動分析

加盟店送客

LTV向上

施策効果分析

ドコモサービス



ドコモショップ



金融決済

エンタメ

EC

ヘルスケア

B2C

観光

小売

メーカー

物流

B2B2X

戦略



企画



施策実行



効果測定



現実

理想

現実(現場の声)

ビジネスの スピード

- 簡単に使える
- 待ち時間なし

依頼殺到で、
データ抽出に 2 週間もかかるの！？

アクション につながる分析

- 高解像度な顧客分析
- 最新技術の活用

自分でやるにも
プログラミングわからないしなあ、

スケール (あらゆる組織)

- 様々な業務の現場ニーズに
即した活用

ツールの使い方が難しくて
面倒だなあ、

解きたい課題と解決策



データ活用までの期間を 9 割削減したい！！

圧倒的に使いやすいデータ活用環境



「自分でできる！」と実感できる環境づくり

誰でも使えるための挑戦

「コーダーのペルソナ分析を行いたい」など

アイデア募集 お問い合わせ アプリを開発したい

カテゴリから探す ▼ ランキングから探す ▼ 活用事例を見る ▼

すべてのアプリ 初めての方向け お気に入りのアプリ アクセス数が多い順 ▼

StartDash

利用目的 施策立案や振り返りのために、所望の会員セグメントの特性をドコモデータで横断的に可視化する

サービス共有 会員分析 位置情報

アクセス数: 2536

お客様の声みえる化アプリ（声みるアプリ）

利用目的 お客様の声を日々の業務やサービス改善に繋げるため、声を要約・可視化して理解する

サービス共有 会員分析 事業状況把握 その他

アクセス数: 2098

d払い決済状況見える化ツール

利用目的 d払い加盟店の開拓/アクティブ化のために、d払い決済履歴を集計・可視化する

サービス共有 決済・保険・投資 加盟店分析 効果検証

アクセス数: 1773

NW改善みえる化アプリ

利用目的 NW品質の改善効果を確認するために、d払いバーコード表示時間や決済データを可視化する

決済・保険・投資 加盟店分析 効果検証 GA/DA

アクセス数: 1181

d払い決済データ抽出

利用目的 「1決済1レコード」データを抽出・整形し、決済頻度・継続状況等を把握する

会員事業・通信事業 決済・保険・投資 加盟店分析 会員分析 分析準備 効果検証

アクセス数: 1091

スーパー販促プログラム配信対象者の抽出と分析

利用目的 スーパー販促プログラムにおいて、詳細条件を指定し配信対象者を抽出する。（現在は社内向け）

決済・保険・投資 会員事業・通信事業 施策リスト作成 位置情報

アクセス数: 977

1

誰もがすぐに使える

2

業務にあった活用メニュー

3

全社公開

効果

3,800 人

MAU

74 万回

利用回数

30 万時間

稼働削減

122

アプリ数

250 人

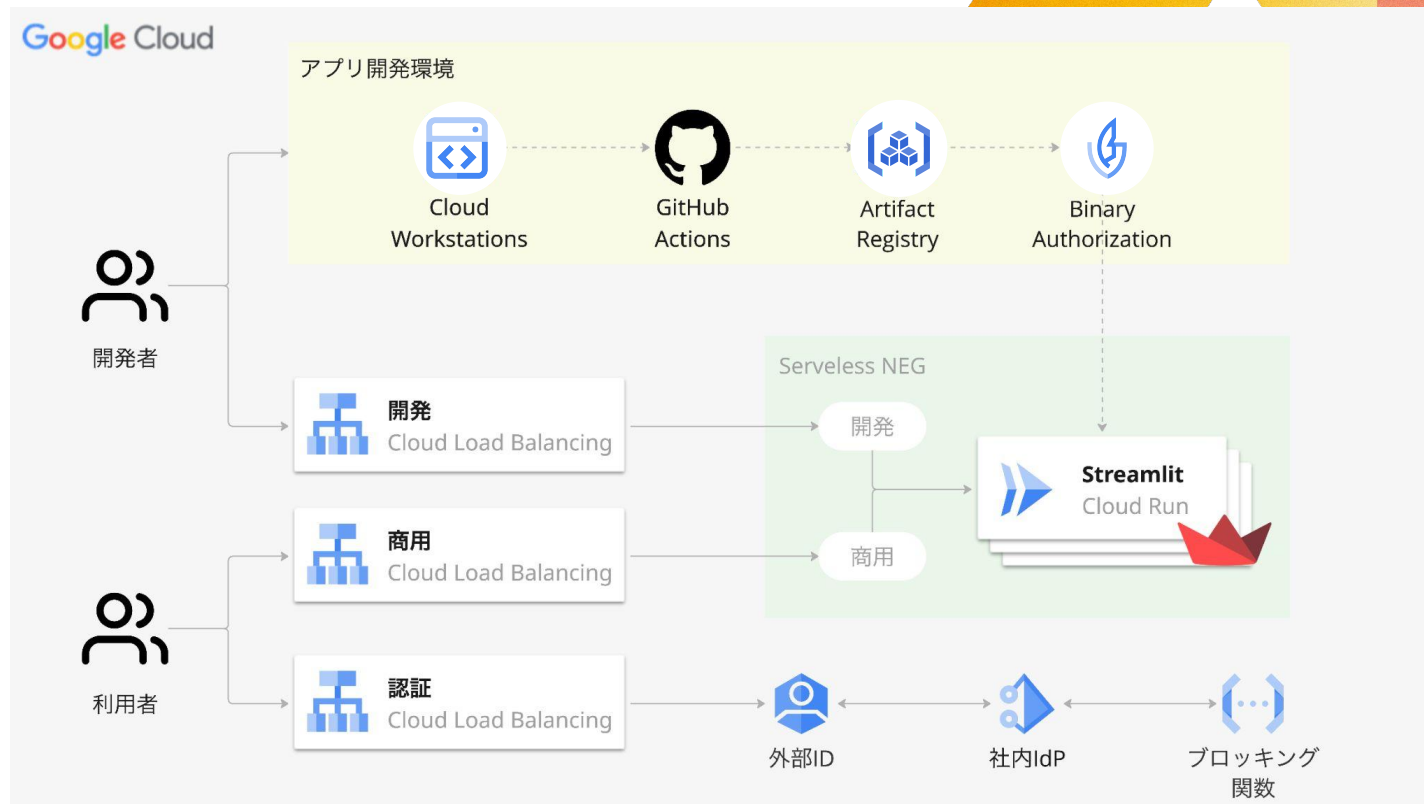
開発者数

200 万回

デプロイ回数

02. データ プラットフォームの概要

Streamlit in Google Cloud



たくさんの Streamlit アプリ

1

URL マップ x リビジョンタグ

デプロイすると自動的に URL を割り当てる

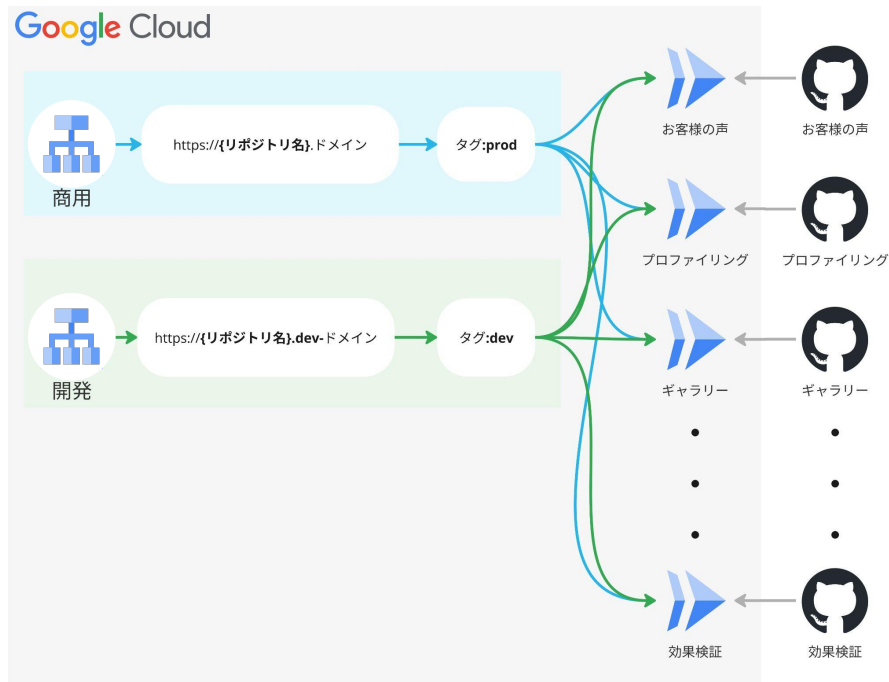
トイルを排除し手間を省くことで、多くのアプリを生み出す

2

Cloud Run のオートスケーリング

ゼロスケールインによる圧倒的なコスト効率

「とりあえず作ってみよう」という文化を醸成





つくりこまれたCI/CD

1

GitHub Actions

各リポジトリから、管理用のリポジトリ ワークフローを呼び出すことで中央集権管理を実現

2

DevSecOps

CD プロセスで、SAST やサプライチェーン攻撃を対策することでセキュリティ強化とデプロイ自動化を両立

3

脆弱性管理

Google 公式の Python イメージを利用。SBOM の作成、コンテナ スキャン、SLSA レベル評価により脅威ハンティングを可能にする

みんなが利用できる

1

IAP x 外部 ID 認証

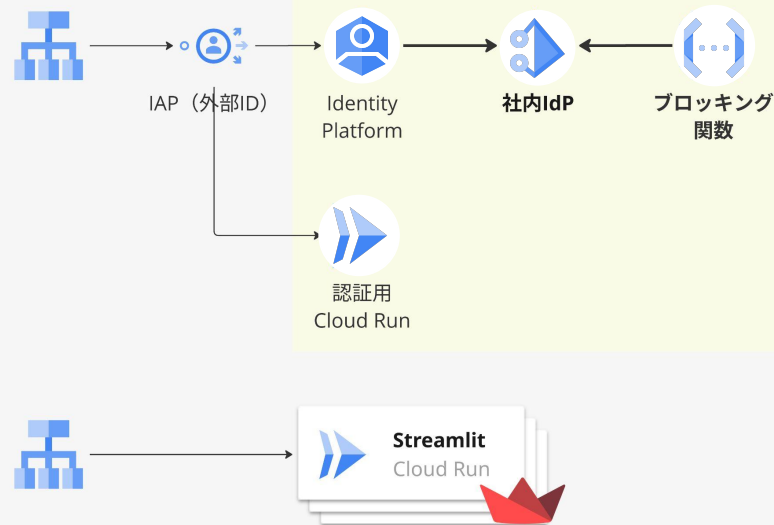
ユーザー登録不要で、誰でも利用が可能
誰でも、全てのアプリが利用可能

2

Cloud Run

ログインページを Cloud Run でセルフ ホスティング

Cloud Run Functions でブロッキング関数を利用し、ユーザー追加 / 認証時にログをカスタマイズ



03. 事例：お客様の声アプリ

動画あり
アーカイブ動画をご視聴ください

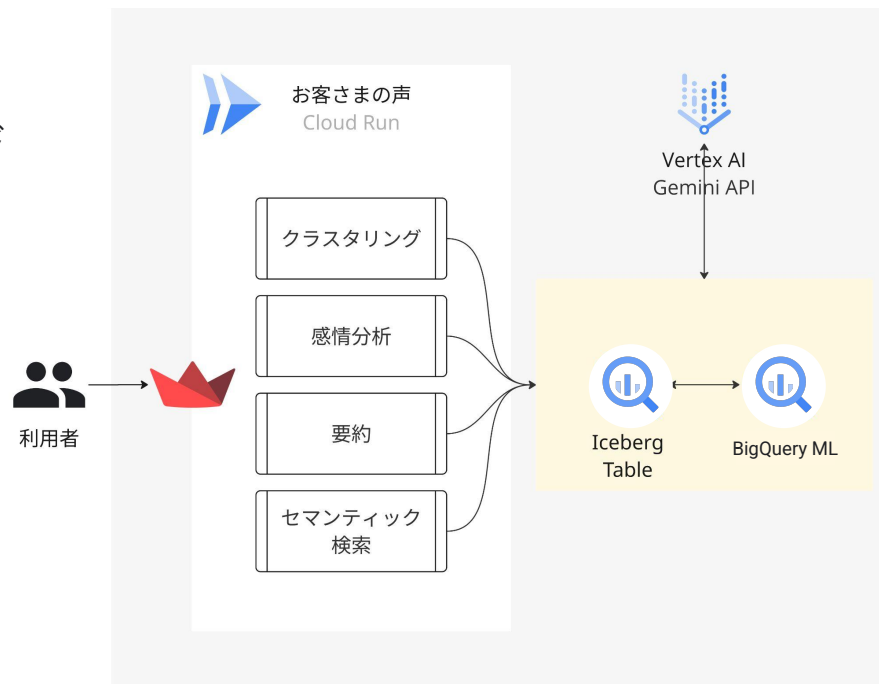
お客様の声アプリの概要

Streamlit の UI から Data Agent を利用

- LLM のワークロードは BigQuery ML (BQML) にオフロード
- コアロジックからの関心の分離

誰もがすぐに使える

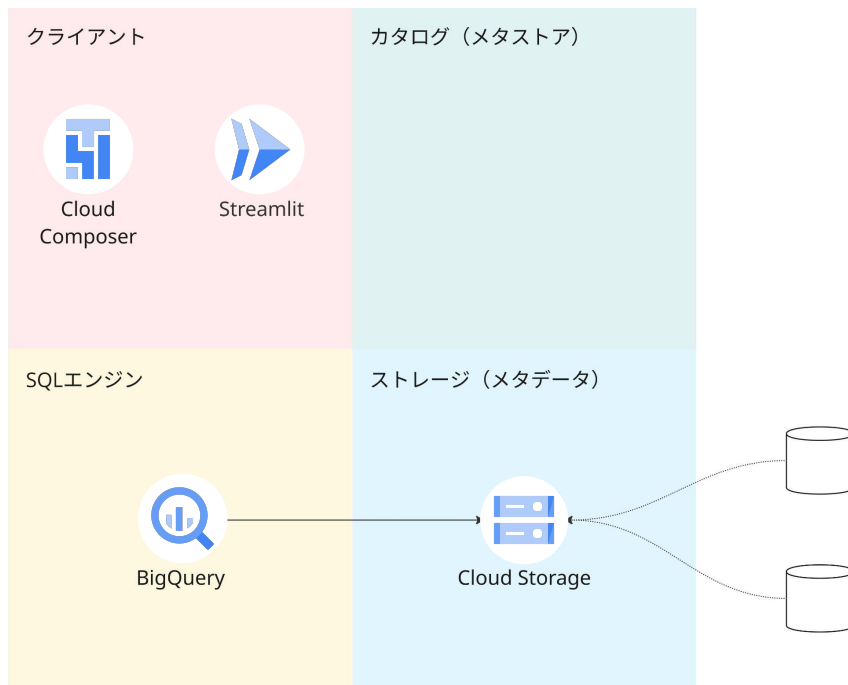
- より決定的な出力を得る
- 自律性よりも業務適用を重視



データメッシュ

他の DWH と **BigQuery tables for Apache Iceberg** を介してデータ連携
データをサイロ化せず、**BigQuery ML** を活用

- **ML.GENERATE_EMBEDDING()**
- **VECTOR_SEARCH()**
- **AI.FORECAST()**



BigQuery ML の 使いどころ

1

セマンティック サーチ

類似するお客様の声を検索し、
時系列トレンド予測 (**TimesFM** モデル)

2

クラスタリング

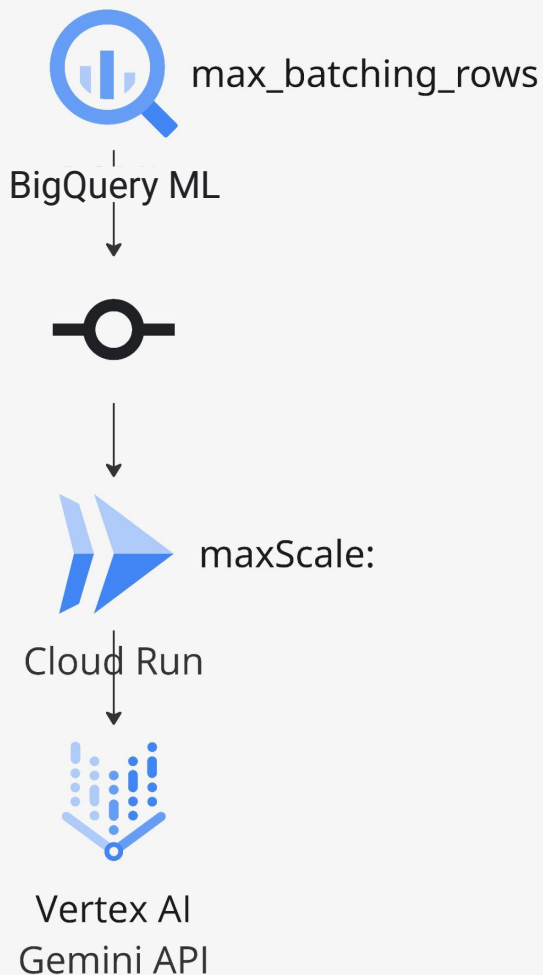
類似性に基づいてお客様の声をクラスタ化することで、全体
を俯瞰する

3

感情分析

Remote Functions を経由で、
Gemini 2.5 Proでのテキスト生成

```
WITH params AS (  
  SELECT  
    ml_generate_embedding_result AS query_vector  
  FROM ML.GENERATE_EMBEDDING(  
    MODEL `embedding_model`,  
    (SELECT query_text AS content),  
    STRUCT("RETRIEVAL_QUERY" AS task_type)  
  )  
, semantic_candidates AS (  
  SELECT  
    base.ID,  
    base.TEXT,  
    distance  
  FROM VECTOR_SEARCH(  
    TABLE `bqtable`, -- 検索対象のテーブル  
    'RETRIEVAL_DOCUMENT_VECTOR', -- 検索対象のカラム  
    (SELECT query_vector FROM Params), -- 検索文字列  
    top_k => 50,  
    distance_type => "COSINE",  
    options => '{"use_brute_force":true}'  
  )  
)  
SELECT *  
FROM semantic_candidates  
WHERE distance < 0.4 -- 類似度の低いレコードを足切り
```



Remote Functions

スケールアウト

BigQuery の Remote Functions では**バッチ内の行数指定**が可能。

Cloud Run Functions でも**オートスケール**の指定が可能。

データの I/O を考えると、バッチ推論よりも遥かに便利に**大規模推論が実現**できる。

Data Agent 化

Gemini API への**クロスリージョン**アクセスが可能

Cloud Run Functions 内に**システムプロンプト**を指定し、**構造化出力**を定義

ポジネガ判定等、多くのユースケースを実現

より自律的な Data Agent へ

宣言的 Data Agent



データ プロダクト

Tool

自律的 Data Agent

