

## 基本情報

### はじめに

こんにちは。和田と申します。

現在はデータエンジニアとして、企業のデータ利活用を支援しています。

非営利組織に対する助成を行う財団でキャリアをスタートし、新型コロナウィルス対策事業に携わる中でデータ利活用に興味を持ちました。

今後はデータエンジニアリングの技術とビジネスの橋渡し役としてのスキルをさらに磨き、意思決定や戦略立案を後押ししていきたいと考えています。

### プロフィール

項目	値
名前	和田悠太郎
最終学歴	東北大学教育学部 卒
現在の職種	データエンジニア

### アカウント

- [Github](#)
- [X](#)

### 業務スキル

#### Pythonによるハイパフォーマンスなデータ加工・集計

- 大規模テーブルデータの高速・効率的な加工・集計を実装できます。
- Numpy配列変換による処理最適化や、Pickle・Parquet等への外部出力を経験しています。
- concurrent.futuresによる並列処理を活用できます。
- 最終アウトプットを意識した集計フロー再構築（使用列絞り込み・集計タイミング最適化）により、処理時間・キャパシティを大幅に向上させられます。

#### データ変換・モデリングによる高品質なデータ基盤構築

- dbtを活用した多層アーキテクチャ (interface/dimension/fact/wide/mart) によるデータモデリングの設計・実装ができます。
- BigQueryにおいて、ウィンドウ関数等を用いた高パフォーマンスな分析クエリを記述した実績があります。
- 必要に応じて、ループや変数を用いた記述も可能です。
- マーケターが作成したSQLのレビューを多数経験しており、要件適合性・パフォーマンス・可読性など多角的な観点から品質管理を行っています。
- 月次KPIテーブルの更新管理や、事業・組織体制の変化に応じた長大なクエリの改修・最適化も得意としています。

#### 生成AI活用・ナレッジマネジメント

- RAGを用いた検索ツールや、社内ルールのマークダウンファイルをデータソースとした業務効率化アプリの開発経験があります。
- Claude Code等を活用したVibe Codingに習熟しています。
- BitBucket (Git) リポジトリ内のSQLレビュー証跡をLLMに読み込ませ、汎用的な検証ナレッジの整理・活用を行っています。
- ナレッジマネジメントの文脈において、企画および実践の経験があります。

### 技術スキル

実業務で使用した主な技術のみ例挙します。

#### プログラミング言語

- Python
- SAS
- VBA
- Windows PowerShell

#### データベース

- BigQuery
- Databricks
- SAS

#### OS

- Windows
- Linux (CUI)

#### ツール

- Git (BitBucket, GitHub)
- dbt

- Terraform
- uv
- MkDocs
- JIRA
- Confluence

## 保有称号

### 資格

資格名	取得月
普通自動車第一種運転免許	2017年9月
普通自動二輪運転免許	2018年9月
日商簿記検定3級	2020年10月
TOEIC L&R 800点	2020年11月
統計検定2級	2023年2月
中型自動車第一種運転免許	2023年5月

## 経歴

### stable株式会社 (2026/01 - )

某大手出版社のデータ基盤プロジェクトに参画し、データ利活用を支援しています。主に以下の業務に従事しています。

- dbtを活用したデータモデリング・データマート構築
- Databricks基盤上でのデータ加工・整備
- ダッシュボード (PowerBI → Databricks BI) のリプレイス
- 社内ハンドブック作成による知識共有推進

過去には以下の研修を受講しています。

- dbt/Databricksオンボーディング研修 (4層アーキテクチャ、Terraform、S3連携)

以下では参画したプロジェクトの内、主要なもののみを記載します。

#### [某大手出版社データ基盤構築・運用プロジェクト] (2026/01 - )

##### 概要

某大手出版社のDatabricks基盤上でdbtを活用したデータ加工・整備を担当。複数サービスのデータマート開発、PowerBIからDatabricks BIへのダッシュボードリプレイス実施。interface/dimension/fact/wide/martの多層アーキテクチャによるデータモデリングを実践し、サービス/データソースごとにカタログを分離した設計を採用している。

役割：データエンジニア

##### 業務内容

1. dbtモデルの設計・実装・レビュー (interface/dimension/fact/wide/mart層)
2. 複数サービスのデータマート構築
3. PowerBIからDatabricks BIへのダッシュボードリプレイス
4. ビジネス要件に基づくSQL最適化とロジック改修
5. Git/GitHubを活用したコードレビューとプルリクエスト管理

##### 使用技術

- データ変換・モデリング: dbt Databricks Unity Catalog
- BI: Databricks BI PowerBI
- Python環境管理: uv
- バージョン管理: Git GitHub

#### [dbt/Databricksオンボーディング研修プロジェクト] (2025/12 - 2026/01)

##### 概要

Databricks + dbt を活用したモダンデータスタック構築の実践プロジェクト。4層アーキテクチャ (Staging / Dimension / Fact / Mart) の構築、インクリメンタルモデルとS3データ取込の実装、Terraformによるインフラコード化 (AWS S3、IAM、Databricks Job) を通じて、集英社プロジェクト参画に向けた技術習得を行った。

役割：データエンジニア（研修）

##### 業務内容

1. dbtプロジェクトのセットアップとDatabricks OAuth認証設定
2. 4層データモデリング (Staging/Dimension/Fact/Mart) の設計・実装
3. インクリメンタルモデルによる効率的なデータ更新パイプライン構築
4. TerraformによるAWS・Databricksリソースのコード化
5. Git/GitHubを用いたプルリクエストベースの開発フロー実践

##### 使用技術

- データ変換・モデリング: dbt Databricks Unity Catalog Delta Lake
- クラウドインフラ: AWS S3 AWS IAM
- IaC: Terraform
- Python環境管理: uv
- 認証: OAuth Token認証
- バージョン管理: Git GitHub

##### 主な成果

- 7つのプルリクエストを通じてdbt/Databricksの基礎を習得
- インクリメンタル処理とパーティショニングによるパフォーマンス最適化を実装

- Terraform stateのリモート管理とworkspace機能を活用した環境分離を実現

#### [社内ハンドブック作成] (2026/01 - )

##### 概要

stable社内のデータモデリングルールや社内ルールを整理し、MkDocsを活用した社内ハンドブックを構築。Cloudflare Pagesでホスティングし、Google Workspace連携によるアクセス制御を実装。チーム全体の知識共有と業務効率化を開始。

役割：企画・実装

##### 業務内容

1. データモデリングルール（カタログ・スキーマ設計、命名規則等）のドキュメント化
2. MkDocsによる静的サイト構築
3. Cloudflare Pages + Zero Trustによる認証基盤の構築
4. Google Workspaceとの連携設定

##### 使用技術

- ドキュメント作成：MkDocs Markdown
- ホスティング・認証：Cloudflare Pages Cloudflare Zero Trust
- IdP：Google Workspace
- バージョン管理：Git GitHub

##### 主な成果

- チーム全体で参照できるデータモデリングのベストプラクティス集を構築
- 社員のみがアクセス可能なセキュアな知識共有基盤を実現

#### コグラフ株式会社 (2024/03 - 2026/01)

大手WEBサービス企業のプロジェクトに参画し、データ活用を推進しています。主に以下の業務に従事しています。

- キャンペーン施策に関するSQL (BigQuery) のレビュー
- 会員ポイントの付与・利用に関する分析基盤の構築・運用
- ビッグデータを活用したサービス横断の戦略策定に向けた分析・集計

過去には以下の業務経験があります。

- Pythonによる住宅ローンアセットリスクマネジメント (ALM) 用キャッシュフロー展開ツールの開発
- 大手金融機関のローンリスク指標分析に関するSAS移管プロジェクト

また、社内リーダーとして事業部運営にも貢献しています。

- 事業部運営MTGのファシリテーター
- メンバー（約7名）の進捗管理（進捗報告MTGのファシリテーション、研修レビュー、毎月の1on1実施）
- 研修プログラムの企画改善、各メンバーの参画プロジェクトに合わせたカリキュラムの策定
- その他円滑な事業部運営の推進に向けた調整雑務

以下では参画したプロジェクトの内、主要なもののみを記載します。

#### [大手WEBサービス企業キャンペーン施策関連SQL品質管理・データエンジニアリング] (2025/04 - )

##### 概要

大手WEBサービス企業（会員数1億人以上（国内）、月間アクティブユーザー数約6,000万人規模）のキャンペーン施策に関するSQL品質管理およびデータエンジニアリング業務を担当。BIツール（DOMO）を活用したポイント関連データ分析基盤の構築・運用、ポイント施策の定量分析などを通じて、サービスのデータ活用推進に貢献。

役割：データアナリスト

##### 業務内容

1. SQLの品質管理（要件適合性レビュー）：チームメンバーが作成したSQLの要件適合性・正確性・パフォーマンス・可読性の確認、ブリケストを利用したGitでのレビュー・コード管理推進
2. BIツール（DOMO）を活用したポイント関連データ分析基盤の構築・運用：DOMOダッシュボード/レポートの設計・開発・運用、BIツール用SQLの記述（データ抽出・集計・加工）、Pythonによるデータ処理作業（実行・ログ確認、CSVやExcelファイルからのETL処理等）
3. ポイント施策関連定量分析：ポイント付与履歴から逆算したサービスごとの購買量分析、キャンペーン施策の効果分析（顧客育成効果、インパクト、ROI等）、ポイント付与数の予測（複数の機械学習モデル、統計モデルでの比較）
4. LLMを活用した業務効率アプリケーション開発：JIRA APIから取得した法務相談チケットの検索および相談内容草案作成ツールの開発、KINTONE APIから取得したキャンペーン企画基礎情報をもとにしたキャンペーン要綱自動生成ツールの作成

##### 使用技術

- SQLレビュー・品質チェック：BigQuery Git(BitBucket) Markdown Python (pandas)
- データ分析基盤管理：DOMO BigQuery Python
- ETL・データ処理：DOMO DataFlow Python (pandas, concurrent.futures, pyarrow, openpyxl)
- ドキュメント作成・ナレッジ管理：Markdown Git(BitBucket)
- 定量分析・可視化：BigQuery Python (pandas, matplotlib, seaborn, scikit-learn, statsmodels 他)
- LLMアプリケーション開発：Python (openai, langchain, streamlit 他)

##### 主な成果

- SQLレビュー工数約50%削減及び検証ナレッジの要約をSQL生成ツールのルールへ反映させるFBエコシステムの構築
- 文章作成の効率化（作成時間を1時間から数分に短縮）と誤字脱字の防止（高品質な文章の安定作成）を実現
- 分析・集計結果を納期通りに共有し、経営会議や役員説明に向けたインサイトを提供

#### [住宅ローン明細キャッシュフロー展開ツール再構築] (2024/06 - 2025/03)

##### 概要

住宅ローン明細キャッシュフロー算出ツール（Excel VBA）のPythonへの置換・再構築プロジェクト。現行VBAツールの調査・設計・コーディング・テストまで一貫して担当し、大規模データ（最大40万債権）の高速処理・集計機能を実装。高セキュリティエリアでの情報管理にも配慮し、ドキュメント作成やパフォーマンス改善も推進。

担当:Pythonスクリプト開発

##### 業務内容

1. VBAツールの調査・不要機能の洗い出し・改修

2. Pythonでの大規模データ算出・集計機能の実装（個別債権返済データ算出、キャッシュフロー展開、並列処理による高速化、メモリ負担軽減、複数条件・帳票別集計、エラーハンドリング）
3. 実行PC仕様に合わせた処理チャネル設定・パフォーマンス計測・分析・ボトルネック特定
4. 設計書・処理ロジックフロー図等のドキュメント作成

#### 使用技術

- データ処理・集計： Python ( pandas、numpy、concurrent.futures、pickle、openpyxl )
- VBA調査・改修： Excel VBA
- ドキュメント作成： Markdown

#### 主な成果

- 既存VBAツール（処理時間3時間/5万件）→Python再構築後（1.5時間/40万件、高性能PCでは20分）へ大幅なパフォーマンス改善
- 他者コードの解説・最適化による高品質なデータ処理実装
- MECEな設計書・フロー図等のドキュメント作成
- 高セキュリティ環境下での開発・情報管理徹底
- 開発プロセス全般を一人称で遂行し、上長レビューを受けながら納品

### [大手金融機関におけるリスク分析用SAS環境移管] (2024/05 - 2025/03)

#### 概要

メガバンクのローンリスク分析を行うSASプログラムの移管開発プロジェクト。既存SASプログラムの調査・改修、開発関連ドキュメントの作成、移管先環境でのプログラム実行テストを担当。

#### 担当:保守対応

#### 業務内容

- バーゼル規制対応に伴うSASプログラムの調査・修正
- 要件定義書・設計書・テスト仕様書等のドキュメント作成
- 移管先Linux環境でのプログラム実行・テスト

#### 使用技術

- SAS開発： SAS (データセット結合、列追加・変更、PROCステップ、マクロ変数活用)
- ドキュメント作成： Markdown
- OS操作： Linux (基本コマンド、viエディタ)
- セキュリティ管理： ISMS順守

#### 主な成果

- 既存ドキュメントやコードの不備を複数発見し、修正を実施

### TSD株式会社（税理士事務所）(2023/09 - 2024/02)

#### 職務概要

#### 業務内容

- 企業の月次会計処理および残高試算表の作成
- 法人税申告書の作成
- 年末調整計算、給与支払報告書作成

#### 習得スキル

- 会計・税務の基礎知識と月次会計処理スキル
- セルフチェックによるミス防止能力
- 不明点・情報不足・進捗を明確に分けた報告作成能力

### 株式会社リクルート (2023/06 - 2023/07)

#### 職務概要

#### 業務内容

- 中途採用求人向け広告媒体のサービス紹介、商談設定までの案内
  - 1日平均発信数：150件
  - 月受注件数ノルマ達成

#### 習得スキル

- 顧客属性による絞り込み・優先順位付け・KPI達成のタスク管理能力
- 状況に応じた言葉選び・説明方法の工夫による新規顧客対応力

### 株式会社エスタイル (2023/01 - 2023/05)

#### 職務概要

#### 業務内容

- データサイエンス関連の研修・自己学習

#### 習得スキル

- Python (標準機能、pandas、numpy、matplotlib、seaborn、scikit-learn、statsmodels、Optuna等) によるデータ集計・分析・モデル実装
- SIGNATE 第31回\_Beginner限定コンペ：4位 (154名中)
- SIGNATE SOTA Challenge：銅メダル (上位10%)、称号「Advanced」獲得
- SQLによるデータ取得・加工・集計
- Tableauによるデータ接続・可視化・EDA

### 公益財団法人日本財団 (2021/04 - 2022/12)

#### 職務概要

#### 業務内容

- 非営利団体への助成事業のヒアリング・審査、助成契約締結～完了までの管理
- 災害ボランティア団体への物資・車両等調達支援
- 感染症対策事業の進捗管理・関連機関への報告

#### 習得スキル

- 状況に応じた情報収集・根拠に基づく説明力
- PCR検査事業のデータ集計・分析・提案力
- Google Workspace活用（助成金申請フロー改善、GASによる自動返信・情報一元化）

## 記事

### 投稿記事

以下ではインターネット上に投稿した文書の内、主要なものを記載します。

#### Note

- [TableauをPythonで再現！？お手軽分析ライブラリpygwalkerのすすめ](#)
- [SAS-Python書き換えメモ～SASと向き合うことになった貴方へ～](#)
- [並列処理入門 Pythonでの大規模データの爆速処理をDaskはじめよう](#)

#### その他

- [高齢者施設等への無料PCR検査事業 中間報告 実績編](#)
- [高齢者施設等への無料PCR検査事業 学術的分析結果公表](#)

### インタビュー記事

以下では採用広報向けにインタビューをうけた記事を紹介します

#### Wantedly

- [【社員インタビュー】予想外こそ成長のチャンス。実務で学んだ「データアナリストに必要な経験とスキル」](#)
- [【社員インタビュー】気軽に勉強できる環境を身近に。職種を超えて学び合える文化を作りたい](#)