

データ抽出ツールのWebアプリ化 ご提案



2025年7月3日

MLシステム開発チーム
シニアAIアーキテクト
小原 大樹

許諾なく撮影や第三者
への開示を禁止します

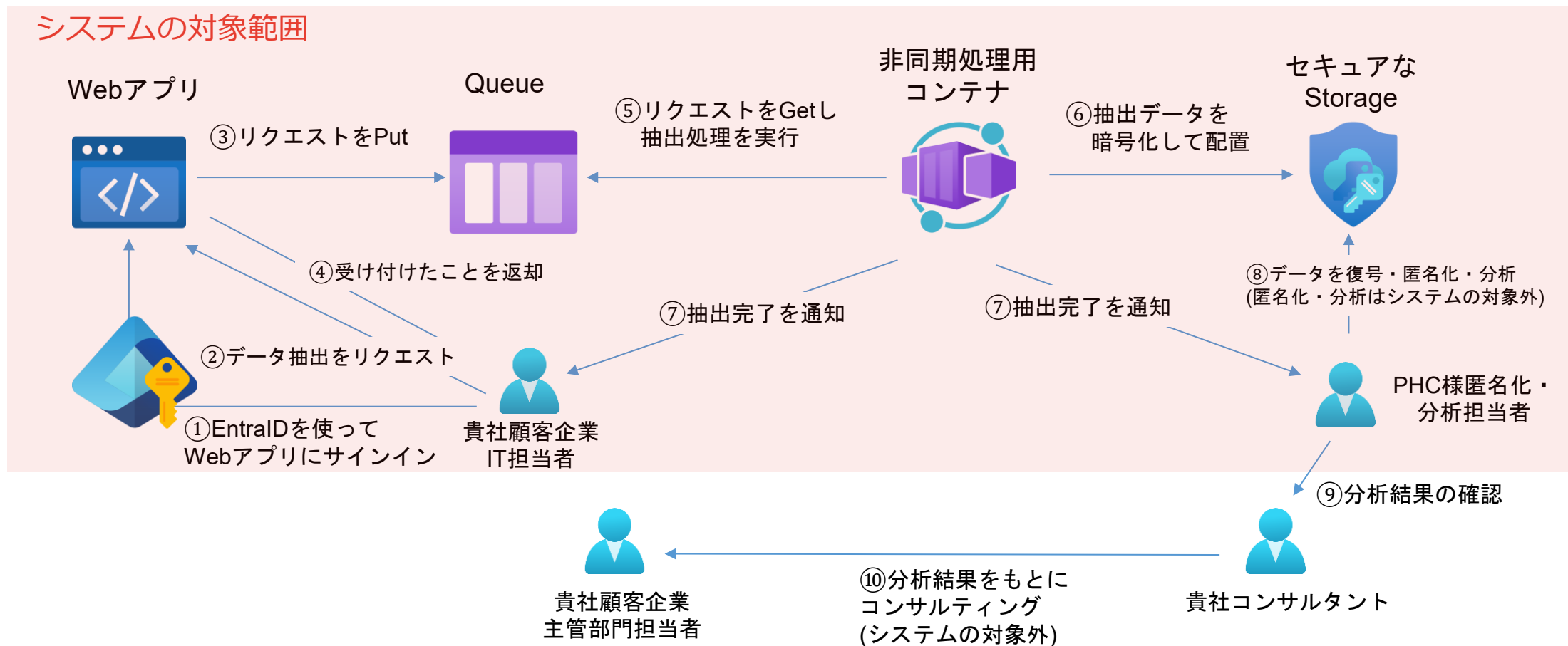
1	プロジェクト概要
2	システム概要
3	画面イメージ
4	セキュリティ要件
5	体制・スケジュール
6	お見積り

コマンドラインツールで実装したデータ抽出機能のWebアプリ化

- ・ 貴社顧客の利用を想定しているためセキュリティに配慮した設計で構築
- ・ ニーズが不透明なためミニマムの機能で素早くリリースしフィードバックを得る

抽出処理に時間が掛かるためコストメリット等もある非同期処理を採用
セキュリティを担保したアーキテクチャを採用

システムの対象範囲



※処理の流れの説明資料のため、アーキテクチャは変更する可能性があります

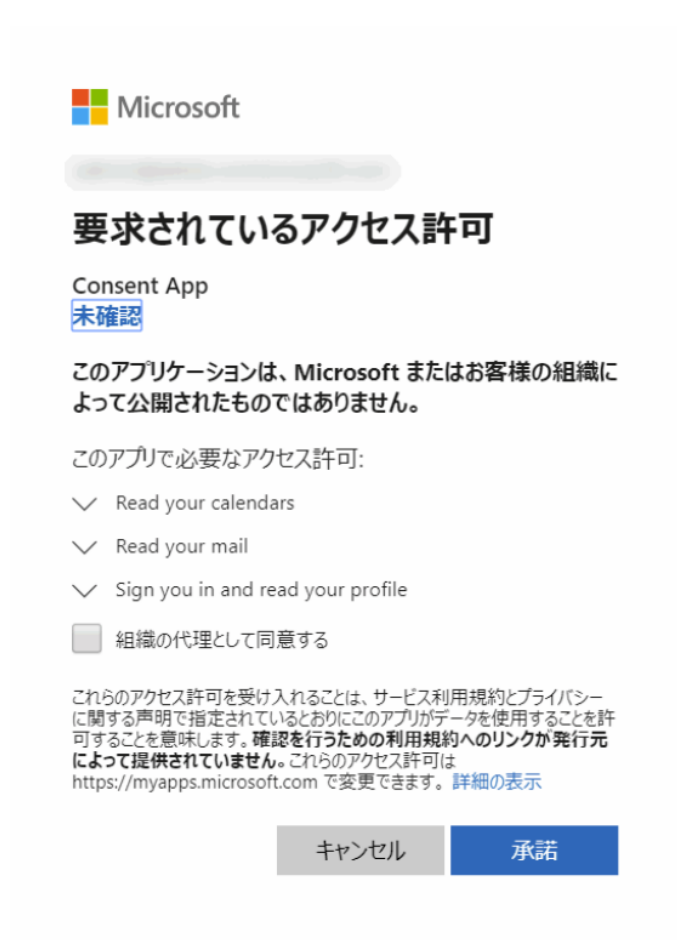
サインイン画面



- ・ ホワイトリストによるIP制限で利用企業のみアクセスを限定
- ・ EntraIDでのサインインのみ許可し利便性とセキュリティを両立

※現時点のイメージのためヒアリング等を受けてUIが変わる可能性があります

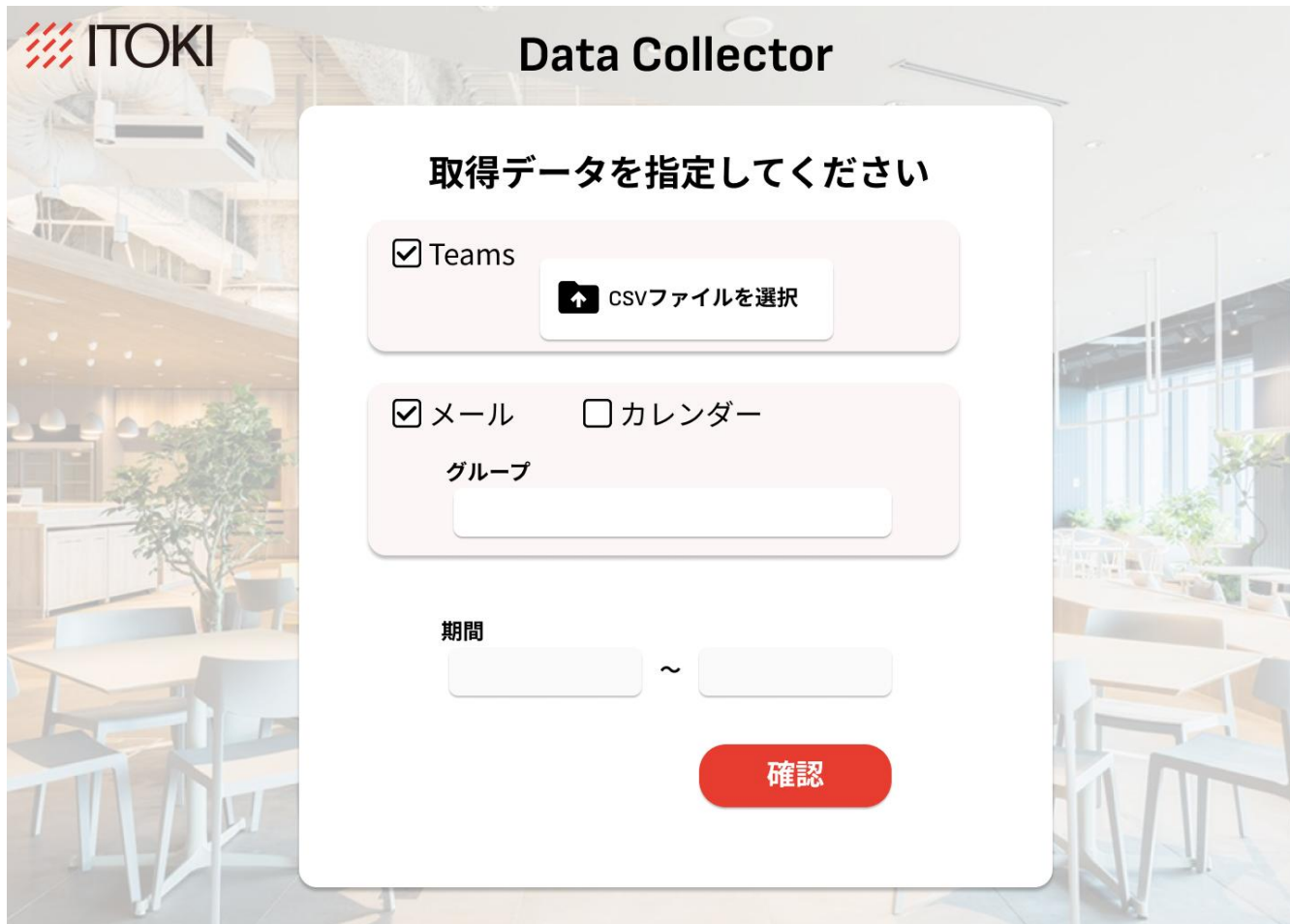
同意画面



- ・ 標準の認証認可のフローに則りサインイン時に同意を取得

※現時点のイメージのためヒアリング等を受けてUIが変わる可能性があります

データ抽出リクエスト画面



The screenshot shows the 'Data Collector' interface for ITOKI. The title 'Data Collector' is at the top. Below it, the instruction '取得データを指定してください' (Please specify the data to be collected) is displayed. There are two main sections for selection:

- Teams:** A checkbox labeled 'Teams' is checked. Below it is a button with an upload icon and the text 'CSVファイルを選択' (Select CSV file).
- メールグループ (Email Group):** A checkbox labeled 'メール' (Email) is checked, and a checkbox labeled 'カレンダー' (Calendar) is unchecked. Below these is a text input field labeled 'グループ' (Group).

At the bottom, there is a '期間' (Period) section with two date input fields separated by a tilde '~'. A red button labeled '確認' (Confirm) is at the bottom right.

- ・ 取得対象を指定しデータ抽出をリクエスト

※現時点のイメージのためヒアリング等を受けてUIが変わる可能性があります

確認画面



- ・ 取得対象をその場で確認

※現時点のイメージのためヒアリング等を受けてUIが変わる可能性があります

抽出完了通知

貴社顧客向け

【イトーキ】データ抽出が完了しました 外部 受信トレイ x

システム管理者

To 自分 ▼

〇〇〇社 〇〇様

2025年6月30日 17:00にご依頼されたTeams 50人、メール100人のデータ抽出が完了しました。
データ分析が完了次第弊社コンサルタントよりご連絡差し上げますのでしばらくお待ち下さい。

イトーキ 〇〇システム

- ・ メールで抽出完了を通知

PHC様向け

【イトーキ】【〇〇〇社】データ抽出が完了しました 外部 受信トレイ x

システム管理者

To 自分 ▼

PHC 〇〇様

〇〇〇社より2025年6月30日 17:00にご依頼を受けたTeams 50人、メール100人のデータ抽出が完了しました。
以下よりアクセスして分析を進めてください。

<https://secured.example.com/xxxx>

イトーキ 〇〇システム

- ・ メールで抽出完了を通知
- ・ メール内にセキュアな保存場所へのリンクを添付
- ・ データを復号して分析して頂く

※現時点のイメージのためヒアリング等を受けてUIが変わる可能性があります

包括的なセキュリティ対策が重要なため、開発・運用の各フェーズで以下のような施策を実施予定

開発フェーズ

下記をCI/CDパイプライン等で自動で実施する仕組み作りの構築

- ・ Dependabot/Renovate等によるライブラリの脆弱性検知
- ・ Trivy等を用いたコンテナの脆弱性検知
- ・ 静的コード解析（SAST）の実行
- ・ OWASP ZAP等を用いた脆弱性診断（リリース前に手動実行の想定）

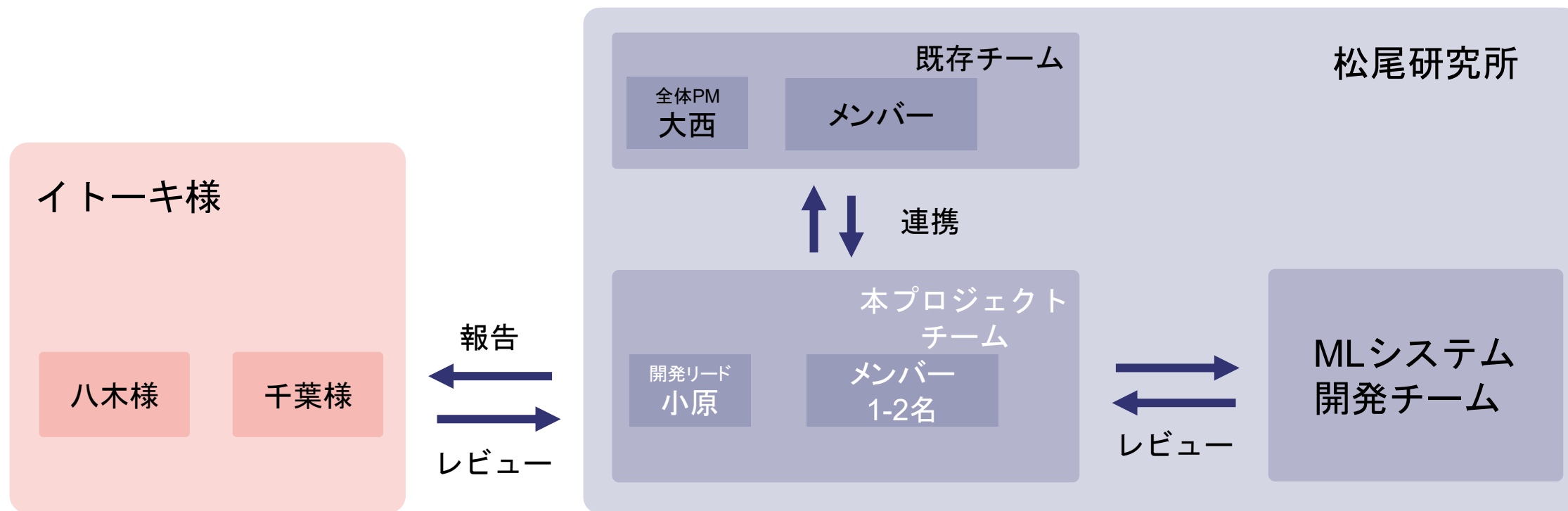
運用フェーズ

- ・ EntraIDでのSSOによる認証認可
- ・ ホワイトリスト式のIP制限により特定顧客企業のみアクセス許可
- ・ Microsoft Defenderシリーズを利用したランタイムの異常検知
- ・ 抽出データは暗号化により保護
- ・ 監査ログとアクセスログの取得により誰がアクセスしたかを追跡可能

※ツール名は例です。ご指定のツール等があればご教示ください。

※セキュリティ要件の詳細は機能要件と合わせて確認させてください。

- ・ 既存チームと連携しつつ独立した小さいチームで機動力高く対応予定



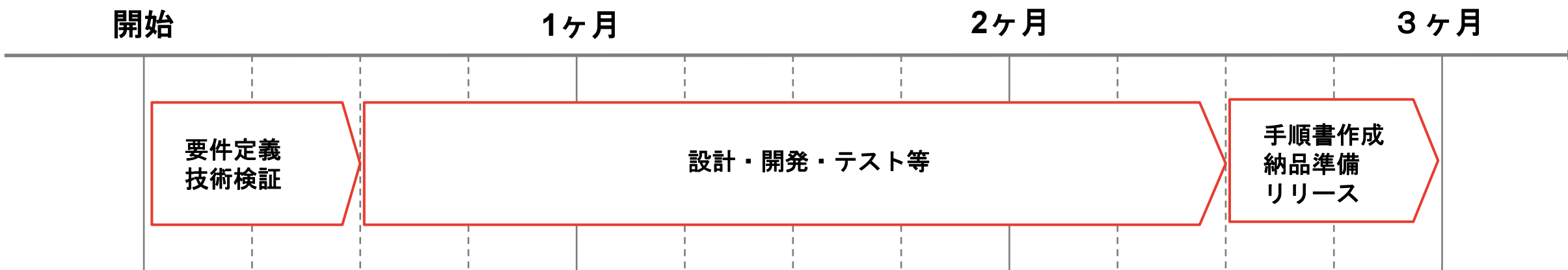
本プロジェクトは小原+メンバー1-2名で対応予定。
(開始タイミング等により担当者が変わる可能性があります)

小原経歴サマリ: 日本及び米シリコンバレーで様々なシステム開発に従事。
医療データやFitbitデータを含む個人情報扱う高いセキュリティ水準のプロダクトも経験。

スケジュールとマイルストーン



- ・最初に一旦要件定義と技術検証を行いますがあジャイル開発を想定しております。
- ・タスクイメージ: 要件定義、技術検証、設計、開発、テスト、納品に向けた手順書等のドキュメント作成・運用引き継ぎ
- ・ヒアリング及び週次または隔週でのレビューにご協力をお願い致します。



今回はセキュリティを担保しつつミニマムでリリースすることを目指します。
以下(一例)はリリース後に必要に応じて追加開発する想定です。

- ・ 貴社のシステム管理者用の画面

- ・ 機能例: 監査ログ・アクセスログのダッシュボードや顧客情報の設定画面等
- ・ スコープ外の理由: 当面は提供顧客も少なく手動で対応可能なため

- ・ データのキャッシュ

- ・ 機能例: 一度取得したデータをDB等に保存し次回以降はそちらを参照する機能
- ・ スコープ外の理由: 当面は提供顧客数が少ないため、また、開発工数が嵩むため
- ・ 補足: 最近Slackでは過去の会話を取得するAPIが50+回/分から1回/分に変更(※)
Teamsも同様の変更を行う可能性もあるため近い将来対応した方が良いと考えています。

※ <https://api.slack.com/changelog/2025-05-terms-rate-limit-update-and-faq>

- ・ お見積り金額: 990万円(330万円/月)
- ・ 期間: 8/1 – 10/31 (調整が必要な場合はご相談ください)

