

日本マイクロソフト株式会社 御中

Azure OpenAI Service リファレンスアーキテクチャ賛同パートナープログラム パートナー オリジナルRA申請

株式会社ナレッジコミュニケーション

■シナリオ名

社内文書の要約

■シナリオ概要

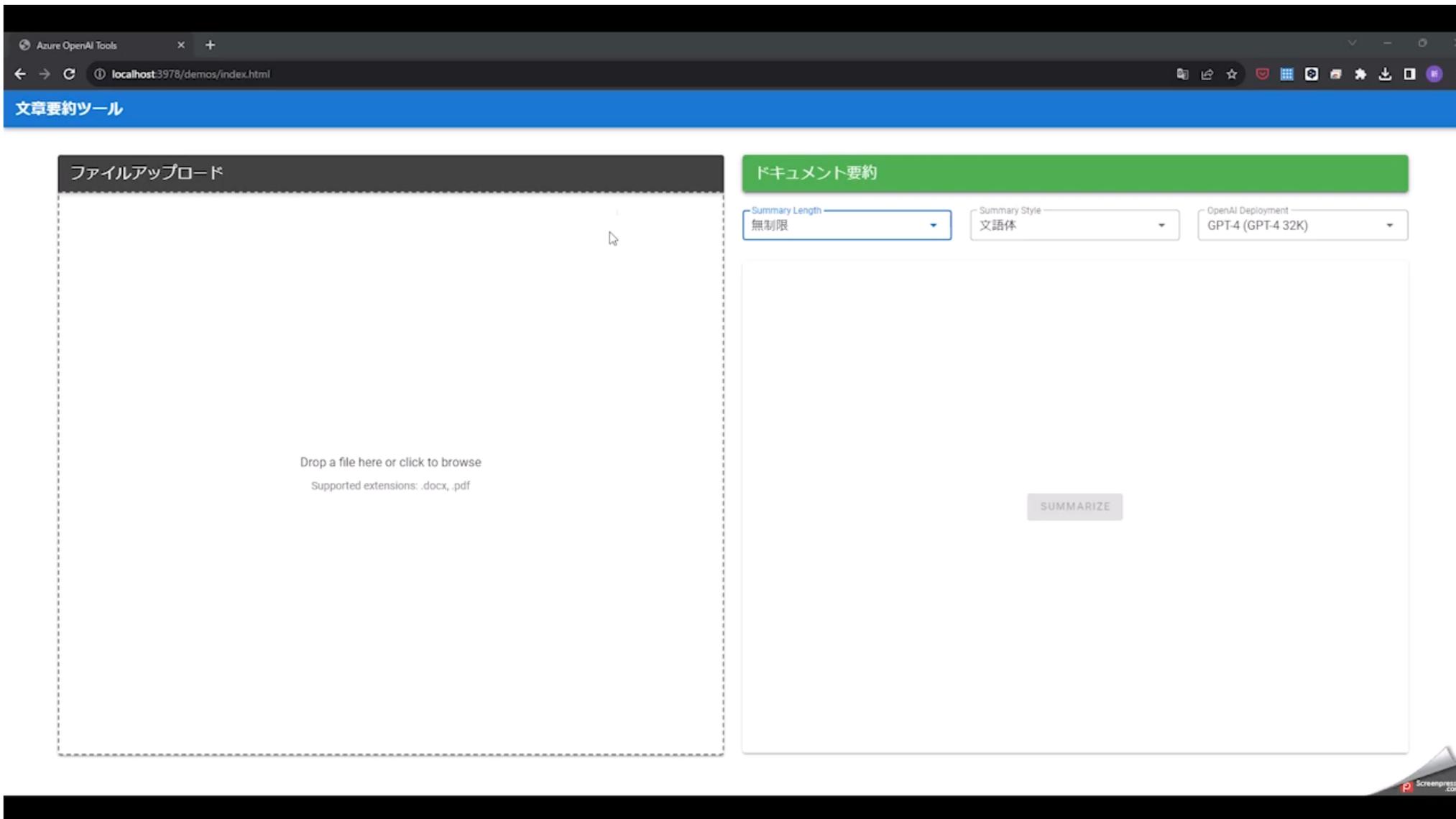
Azure OpenAI Serviceを活用したWordとPDFの社内文書の要約

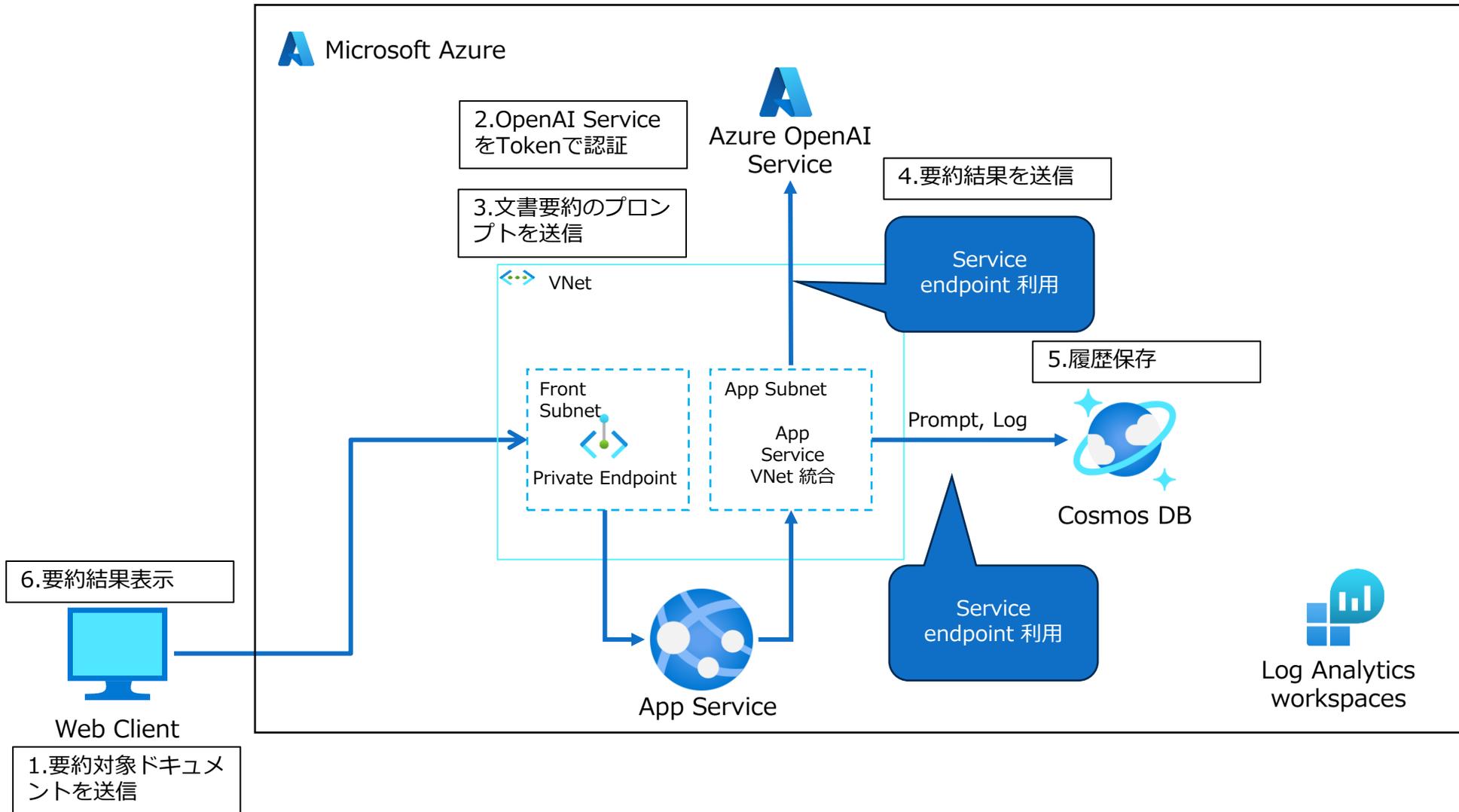
WordとPDF形式の社内ドキュメントを要約し、効率的にインプットすることができます。

ドキュメントを要約し、重要なキーワードを抽出することで、短時間で情報を把握し、迅速な意思決定が可能になります

以下のような状況で役立つシステムを作成することができます。

- ・ビジネス書、技術書、仕様書、提案書など社内の機密情報が含まれるドキュメントをセキュリティを確保しながら要約することができます
- ・一日の利用料金に制限をかけて、使い過ぎの防止をします





Azure Well-Architected Framework観点での考慮事項 (1)

[Microsoft Azure Well-Architected Framework](#)

信頼性(可用性)

- Azureのサービスでは「可用性ゾーン」や「リージョン」といった単位で可用性を設計しており、これらを適切に組み合わせることで、ビジネスクリティカルなワークロードの信頼性を実現するように設計することが可能です。詳細は「[Azure リージョンと可用性ゾーンとは](#)」をご参照ください。
- このシナリオで用いられている、Functions、CosmosDB、Azure OpenAI、ストレージアカウント等のコンポーネントはそれぞれゾーン冗長やリージョン冗長、geoレプリケーションなど高可用性のオプションや構成を利用可能です。必要となる可用性に応じて導入を検討してください。複数リージョン間/複数ゾーン間でAct-Act構成を取る場合にはAzure Front Door、Azure Application Gatewayなどの利用をご推奨します。

信頼性(回復性)

- アプリケーションの正常性を監視するために、このシナリオでは Log Analytics を使用します。Log Analytics を使用すると、異常時のログを迅速に分析し対応することができます。
- 回復性に関するその他の記事については、「[信頼性の高い Azure アプリケーションを設計する](#)」を参照してください。

セキュリティ

- セキュリティは、重要なデータやシステムの意図的な攻撃や悪用に対する保証を提供します。詳細については、「[セキュリティの重要な要素の概要](#)」を参照してください。
- このシナリオでは、Microsoft Entra IDを使用してユーザーを認証します。
- セキュリティで保護されたソリューションの設計に関する一般的なガイダンスについては、「[Azure のセキュリティのドキュメント](#)」を参照してください。

Azure Well-Architected Framework観点での考慮事項 (2)

コスト最適化

- 一日あたりの利用料の上限をかけることで使いすぎを防止します。
- 不要な費用を削減し、運用効率を向上させる方法をご検討ください。詳しくは、「[コスト最適化の柱の概要](#)」に関する記事をご覧ください。

オペレーショナルエクセレンス

- システムの健全性の担保、トラブルの解決、利用動向の監視を行うためには適切な監視とログ収集が必要となります。詳細は「[ワークロードの監視](#)」をご参照ください。本構成を用いることによりAPIの適切な利用を担保することが可能となります。
- ソフトウェアのアップデートや脆弱性への対応など、ソフトウェア/インフラ設計の改修を円滑に進められるよう、DevOpsプロセスを確立してください。詳細は「[リリース エンジニアリングの継続的インテグレーション](#)」をご参照ください。

パフォーマンス効率

- アプリケーションの負荷が高まることを見越し、スケーラビリティの確保をあらかじめ検討することは重要です。詳細は「[スケーリング用のアプリケーションを設計する](#)」をご参照ください
- Azure App Serviceは負荷に応じて水平にスケールさせることが可能です。詳細については「[自動スケーリングを有効にする方法](#)」をご参照ください。
- 頻出のクエリについてはアプリ側でキャッシュする等のキャッシュ戦略もご検討ください。詳細は「[キャッシュを使用する](#)」をご参照ください。
- また、特定のユーザーにAzure OpenAIの利用が集中することを避けたい場合にはAPI Managementによるスロットリング導入などをご検討ください。詳細は「[Azure API Management を使用した高度な要求スロットル](#)」をご参照ください。