

IIJ PaaS活用ソリューション with Microsoft Azure (OpneAI)

会話生成AIチャットのPoC支援



株式会社インターネットイニシアティブ

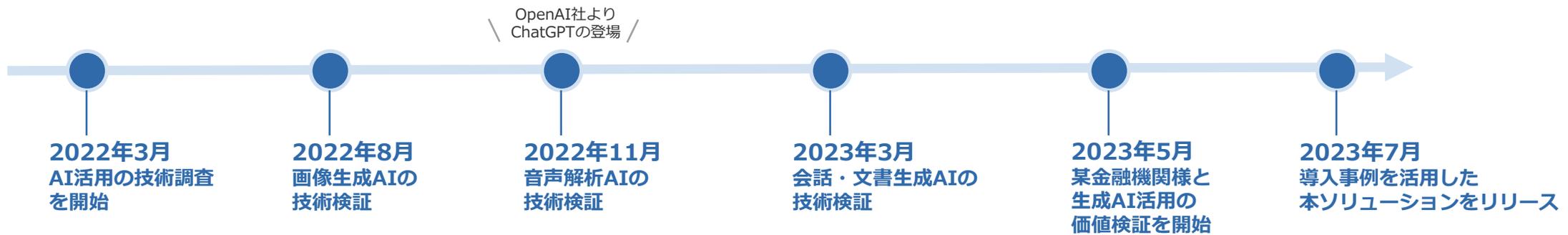
はじめに

■ IIJのAI活用の取り組み

IIJではChatGPTが市場を席捲する以前(2022年)からAIに関する技術情報のトレース、検証など行い積極的に発信を続けてきました。
現在はそのノウハウを活用して会話生成AI(Azure OpenAI Service)導入の支援を行っています。



2022年よりIIJ Engineer BlogにてAIに関するナレッジを発信 (<https://eng-blog.ij.ad.jp/>)



■ IIJのマイクロソフトビジネスのあゆみ

Gold



米「マイクロソフト認定ゴールドパートナー」とは、特定のマイクロソフトソリューション分野における最高レベルの技術力と最高品質のソリューションを提供できるエキスパートとして、マイクロソフトが認め、推薦するパートナー企業です。



米マイクロソフトによる**Microsoft Azure のネットワークサービスに特化した認定プログラム**。Azureのネットワークサービスを利用し、高品質な独自のマネージドサービスを提供できる事業者として認定。

Azure OpenAI Service(AOAI)とは

Azure OpenAI Service(AOAI)とは

AOAIはChatGPTと同等のAIモデルを提供する唯一のパブリッククラウドサービス。
ChatGPTで標準提供されているGPT-3.5に加えて、画像、動画、音楽の生成が可能なGPT-4の利用も可能です。

✓ GPT-3.5とGPT-4の違いについて ※23年7月時点

● 会話の長さが違う

| | |
|---------|-----------------|
| GPT-3.5 | 16,000文字 |
| GPT-4 | 32,000文字 |

● 機能が違う

| | |
|---------|--|
| GPT-3.5 | 自然言語生成による文書作成 |
| GPT-4 | 自然言語生成による文書作成 画像入力による文章生成 コードの生成 |

● 価格が違う

| | |
|---------|--------------------|
| GPT-3.5 | 入力/出力 ¥ 0.29 |
| GPT-4 | 入力¥8.67 出力¥17.3 |

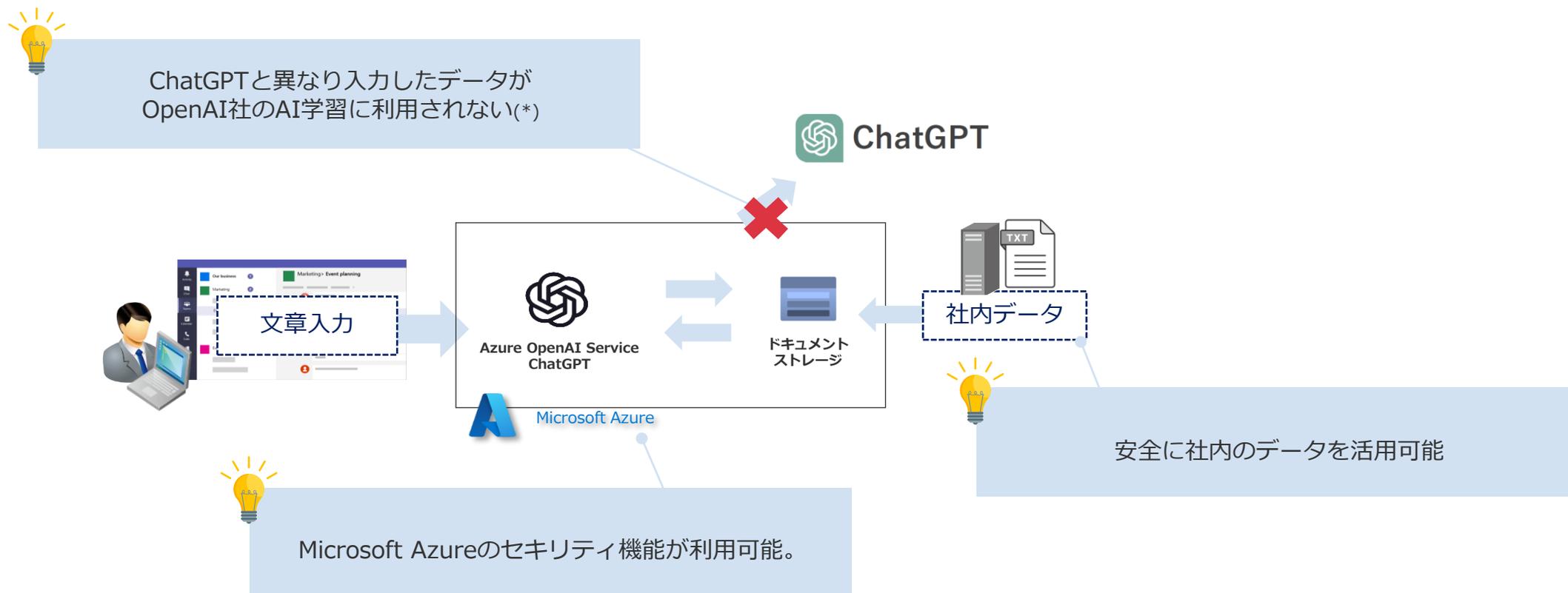


GPT4は3.5 に比べて性能が高く、米司法試験、大学院共通試験、大学レベルの生物学、数学、化学などで好成績 (OpenAI 2023)

GPT-4の方が高機能、ただし価格も高くなる為
GPT-3.5とGPT-4の使い分けが重要!!

AOAIの方がセキュリティレベルが高いのでエンタープライズ企業向け

- ✓ ChatGPTでは実現できない自社データの漏洩を防ぐセキュリティ機能が実装可能



* ChatGPTでもオプトアウト申請を行えば、OpenAI社のAI学習からの除外が可能となりましたが、3月の個人情報漏洩事故などの影響で利用を不安視する企業も多いと推測。
(参考: <https://www.itmedia.co.jp/news/articles/2303/25/news051.html>)

(参考) AOAIとChatGPT

- ✓ 企業利用のセキュリティやガバナンスに対応できるのはAOAI

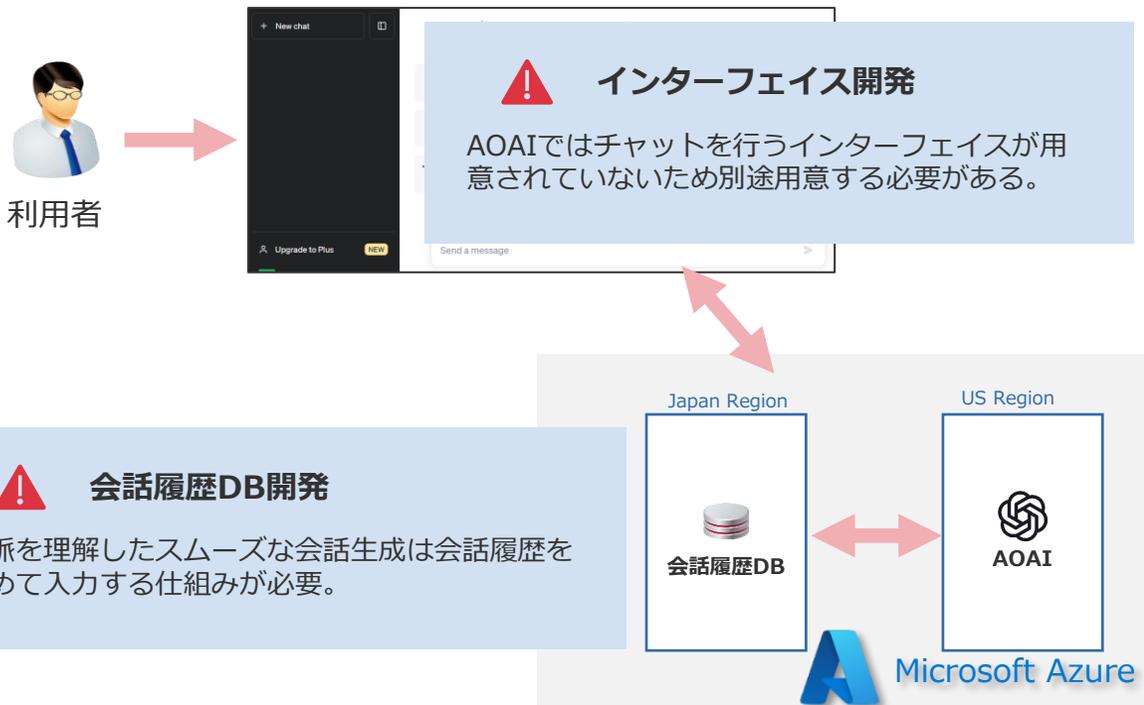
| | Azure OpenAI Service | ChatGPT |
|---------|---|--|
| 提供元 | Microsoft | OpenAI |
| 利用方法 | API連携 | Webサービス API連携 |
| 接続方法 | インターネット接続 閉域接続 | インターネット接続 |
| 入力データ | AIの学習に使用されない | AIの学習に利用する可能性あり (Web) AIの学習に利用しない (API) |
| ガバナンス | Microsoft Entra IDを使ったアクセス制御 禁止ワード等の設定 | OpenAI社に依存 |
| SLA | 99.9% | 無し |
| アカウント管理 | 企業単位 (Azure Entra) | 個人単位 |

AOAI導入の検討 4つのポイント (1/2)

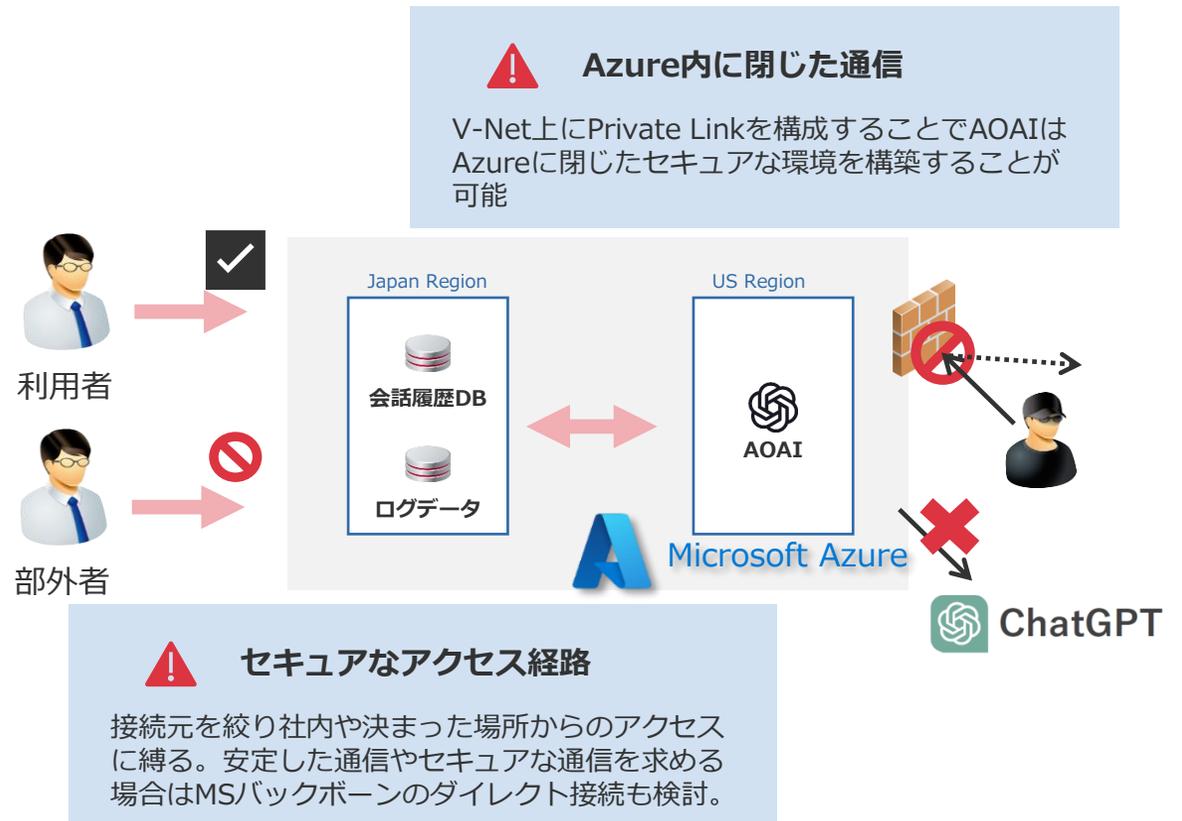


- ✓ AOAIだけではChatGPTのようは会話生成AIチャットは使えない
- ✓ IIJが考えるAOAI導入の検討ポイントは4つ

1. 初期構築が必要



2. データ漏洩対策

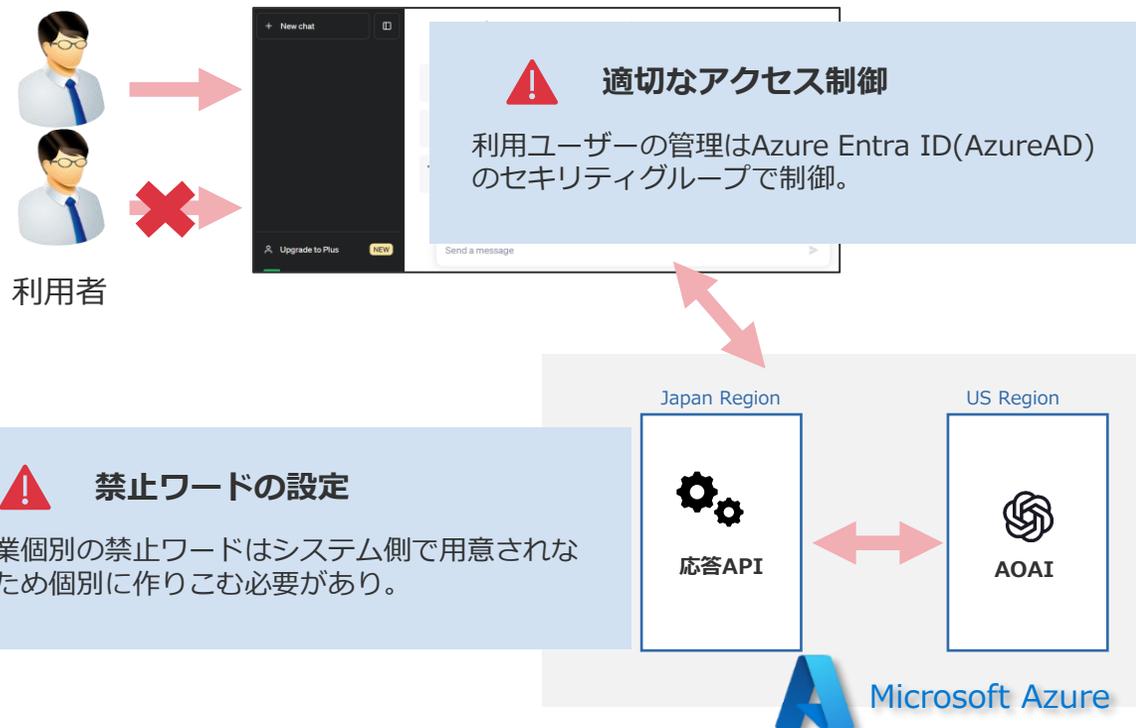


AOAI導入の検討 4つのポイント (2/2)

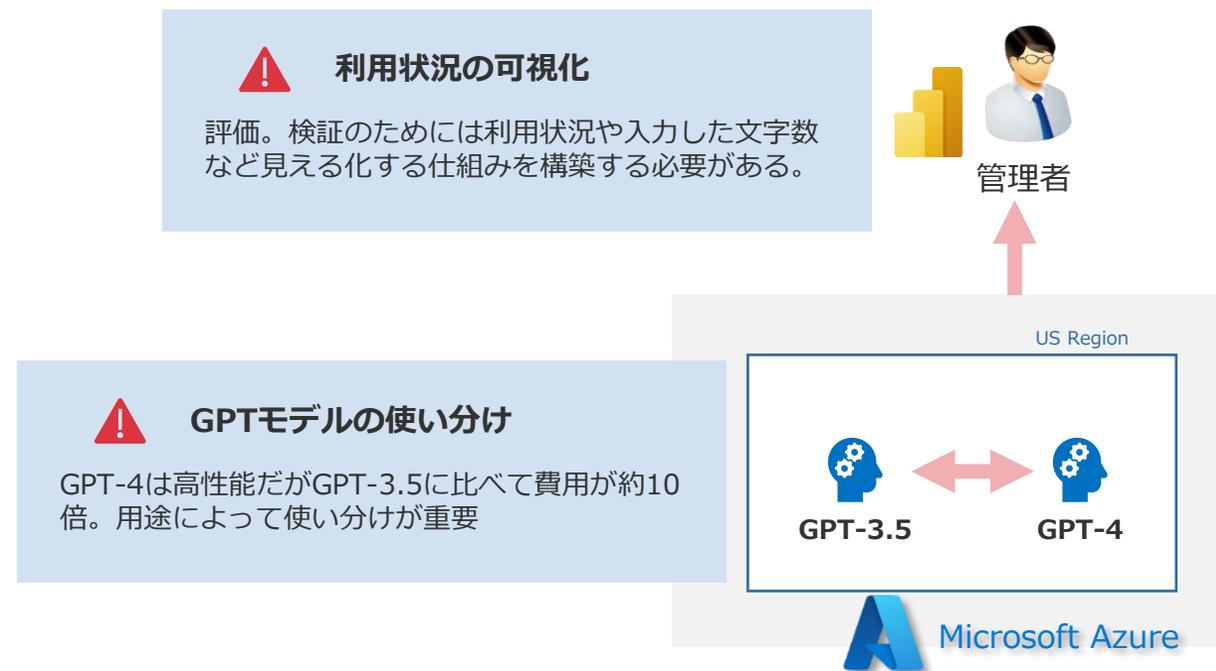


- ✓ AOAIだけではChatGPTのようは会話生成AIチャットは使えない
- ✓ IIJが考えるAOAI導入の検討ポイントは4つ

3. ガバナンス強化



4. コスト・回答精度のバランス

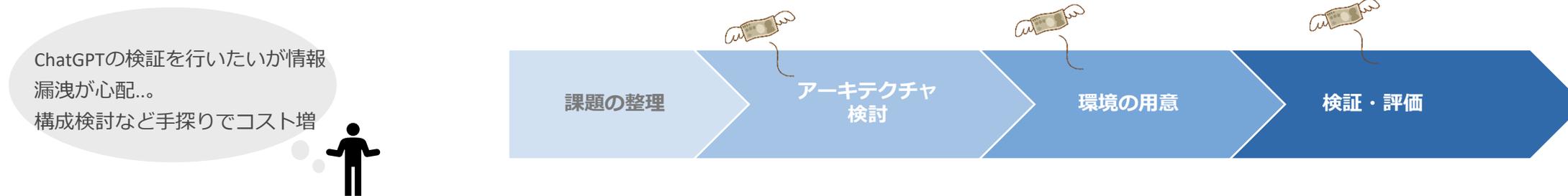


IIJ PaaS活用ソリューション with Microsoft Azure(OpenAI)

本ソリューションの狙い

市場を賑わせているChatGPT。「会話型AIの業務活用の有効性についてPoCを行いたいが、企業利用に耐えうるセキュリティやガバナンスを満たした導入方法が分からない..」というお客様に向けて、AIとMicrosoft Azureの知見を活用し環境構築から今後の活用方法のサポートまで一気通貫で行います。

■ 一般的なPoC



■ IJ PaaS活用ソリューション with Microsoft Azure(for AI)



AOAI導入の流れと本ソリューションのスコープ

AOAI導入は将来の活用を見据えて段階的に拡張する方式がおすすめです。
本ソリューションは、IIJが持つAOAIとMicrosoft Azureの知見を活用してフェーズ1導入をサポートします。
AOAIの評価検証を確実にスピーディーに行いたいお客様に最適なソリューションです。

本ソリューション
ご提供範囲

フェーズ1
検証・評価

最適な検証環境の構築
導入の評価検証を
スピーディーに実現

フェーズ2
本番適用

利用範囲、利用部門の拡大

フェーズ3
応用した利用

自社データを活用
さらなるAI活用を推進

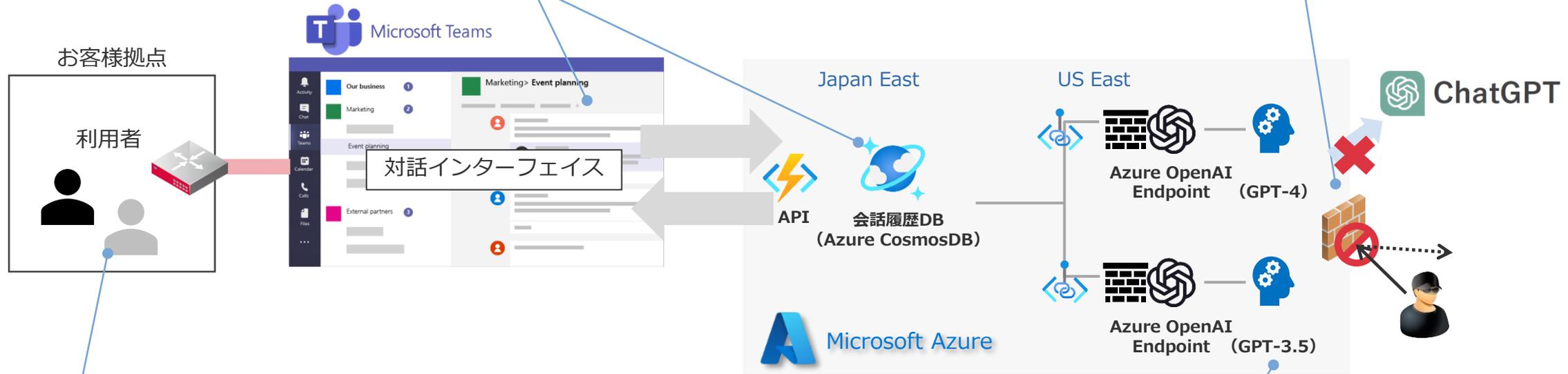
AOAI評価検証に必要な環境を構築してご提供

1. 初期構築・導入サポート

- 対話インターフェイスはMicrosoft Teamsを活用
- チャット文脈を理解させる会話履歴DBを構築
- 利用開始時の利用方法や注意事項のガイダンス

2. データ漏洩対策

- 自社Azureテナントに閉じたデータ活用・アクセス制御
- 対話インターフェイスへはIIJサービスを利用してMSバックボーンに閉じた閉域経路で接続(オプション)



3. ガバナンス強化

- アクセス制御や禁止ワード設定で適切な利用環境構築

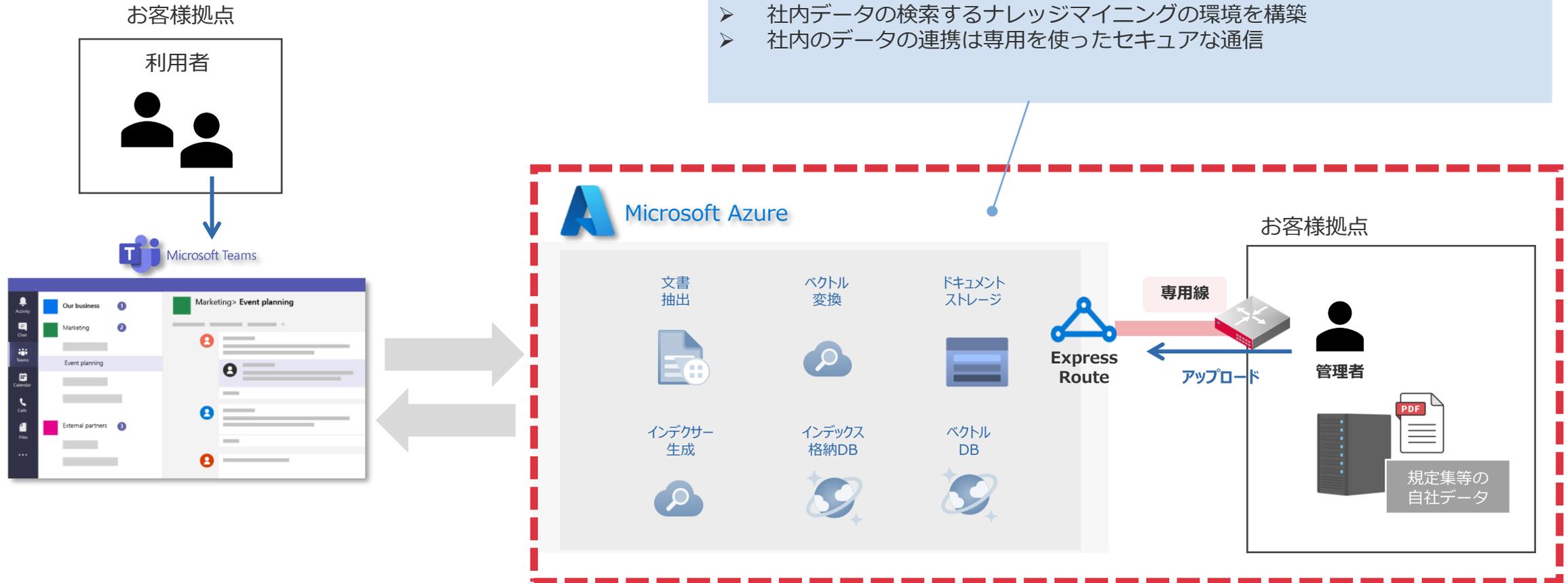
4. コストと回答精度の最適化

- GPT-4はとGPT-3.5の使い分け
- 回答精度の調整

(参考) フェーズ3 : 自社データを活用

自社データを効果的に活用

- 社内データの検索するナレッジマイニングの環境を構築
- 社内のデータの連携は専用を使ったセキュアな通信



ご提供内容

1. 初期構築導入サポート

- ✓ TeamsでUI構築
- ✓ ステートフル会話DB構築
- ✓ Microsoft Azureのサブスクリプション提供・利用申請
- ✓ 利用開始時の利用方法や注意事項のガイダンス

2. データ漏洩対策

- ✓ お客様Azureテナントに閉じたデータ活用
- ✓ 外部からのアクセスを制限
- ✓ 会話履歴やログデータは日本国内に保持

3. ガバナンス強化

- ✓ 禁止ワード制御
- ✓ 利用ログデータ保持
- ✓ アクセス制御

4. コストと回答精度の最適化

- ✓ トークン量の可視化とコストアラート
- ✓ GPTモデルの使い分け

+

基本提供

オプション

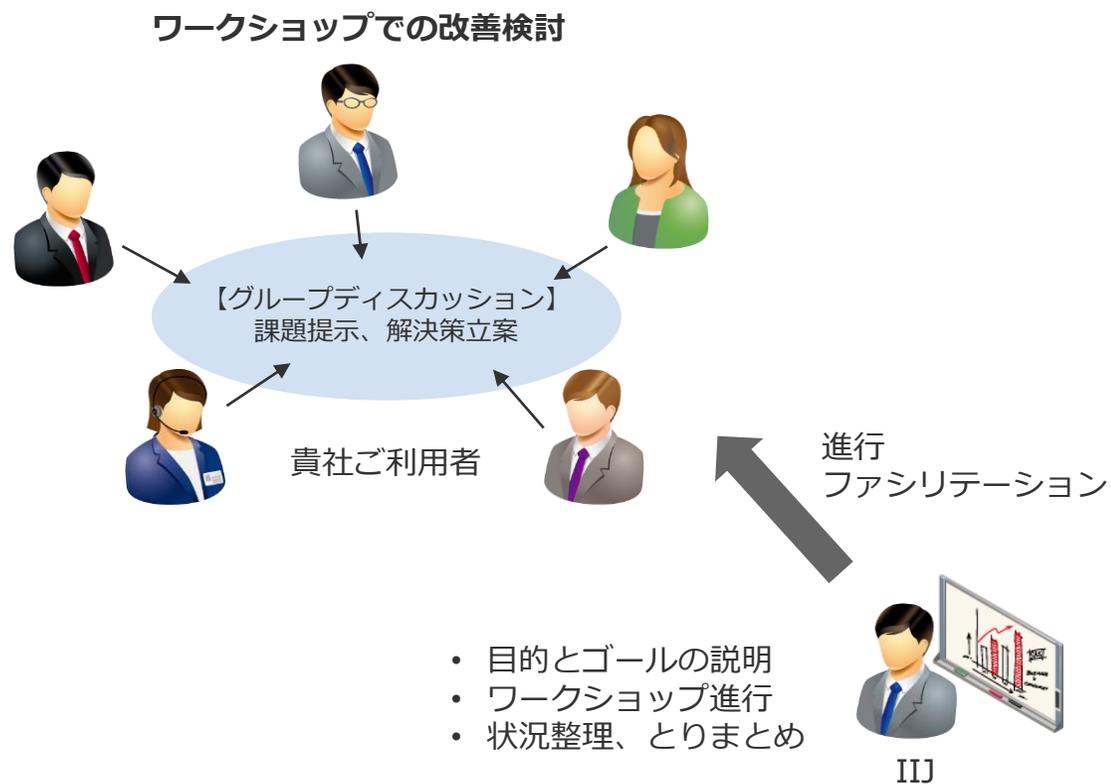
- ✓ 検証環境評価・回答精度改善支援
- ✓ 安心の閉域接続

ご提供機能一覧

| | | | |
|----|-------|---|---|
| 機能 | 利用環境 | <ul style="list-style-type: none"> ■ ユーザインタフェース <ul style="list-style-type: none"> ▶ 会話型AIを「Teams bot」として提供し、従業員によるアプリの追加ができる | <ul style="list-style-type: none"> ■ 提供範囲の制御 <ul style="list-style-type: none"> ▶ アプリ利用のユーザ範囲は、全従業員または特定の従業員を選択できる |
| | 利用者機能 | <ul style="list-style-type: none"> ■ チャット生成 <ul style="list-style-type: none"> ▶ 1対1でチャット応答ができる ▶ 利用者のチャットに対してAIが回答を生成し、チャット応答する ▶ AIと過去3ターン文の会話スレッドを元にした会話ができる ▶ これまでの会話スレッドとは異なるスレッドを開始する際は、直前までの会話スレッドをリセットできる ▶ 会話スレッドは最終投稿から60分で自動リセットされる | <ul style="list-style-type: none"> ■ チャットボット <ul style="list-style-type: none"> ▶ お客様の業務に合わせたキャラクター設定ができる ▶ gpt-3.5-turbo版とgpt-4版の2種類のBotアプリを利用できる ■ チャットログ検索 <ul style="list-style-type: none"> ▶ 利用者は、Teamsの標準機能を用いて過去の会話を参照・検索できる |
| | 管理者機能 | <ul style="list-style-type: none"> ■ 会話履歴保持 <ul style="list-style-type: none"> ▶ 全利用者の会話履歴を2年間保存・検索できる ■ 禁止ワード制御 <ul style="list-style-type: none"> ▶ 登録した禁止ワードを入力すると検知・ブロックされる | <ul style="list-style-type: none"> ■ 利用状況確認 <ul style="list-style-type: none"> ▶ チャットの利用状況（会話数など）を確認できる ▶ 会話毎のトークン量を確認できる ▶ トークン量の集計結果を確認できる ▶ Azureの利用料が設定した価格を超過したらアラートを受け取れる |

検証環境評価・回答精度改善支援

導入後は、AOAIに対する理解を深め、回答精度を改善しながら本格展開に繋がります。さらに、グループディスカッション形式での課題整理やアイデア出しを開催し、お客様が主体的に業務適用範囲の拡大と将来的な活用を実現できるようご支援します。



| プロセス | 実施内容例 |
|------|--|
| 計画 | <ul style="list-style-type: none"> • 世の中の取り組みや最新情報、事例の収集と共有 • 現状課題の整理 • 評価項目、評価水準、評価方法の案の提示 • 改善案のアイデア出し |
| 実行 | <ul style="list-style-type: none"> • 定例会での状況確認 • 評価プロセスの進捗の確認 • 進行における課題の整理と改善案の検討 • 貴社からの質問への調査や回答 |
| 評価 | <ul style="list-style-type: none"> • 利用者からのヒアリング • 評価基準に基づく評価の支援 • 今後の検討課題の整理と改善案の検討 • 評価用レポート作成 |

安心の閉域接続

IIJのネットワークサービスを使うことでAOAIへの接続において信頼性が高い安定した通信が可能

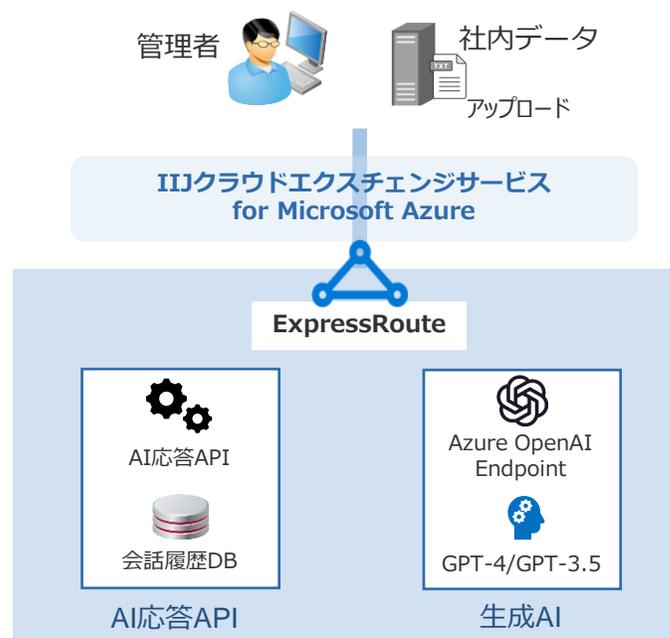
IIJプライベートバックボーンサービス/Smart HUB

お客様環境とMSバックボーンで接続。安定した通信が可能
※活用シーン：フェーズ1(本ソリューション範囲)



IIJクラウドエクステンジサービス for Microsoft Azure

Azureとお客様拠点を閉域でセキュアで安定した通信を実現
※活用シーン：フェーズ3(自社データ活用時)



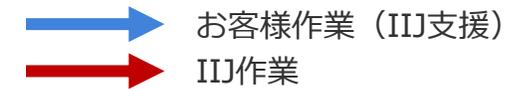
導入の流れとリードタイム

ご発注後、アカウント（Microsoft Azure、Microsoft365）を提供いたします。

お客様にてMicrosoft社に対してAOAIの利用申請をお願いいたします。※IIJでは申請のサポートを行います。

申請完了後、およそ1～2ヶ月で検証環境のご利用を開始いただけます。

■ スケジュール



IIJ PaaS活用ソリューション with Microsoft Azure 費用

初期費用

初期：¥1,500,000

■ 概要

- AOAI検証環境構築
- 利用開始時の利用方法や注意事項のガイダンス

月額費用(システムサポート)

月額：¥200,000

■ 概要

- システムの構成や仕様に関するお問い合わせ
 - 障害復旧作業
 - Microsoftへの問い合わせ、構成管理
- 作業時間：15時間(月)

※税別表記

※初期導入とは別にIIJより提供させていただく以下の品目の費用が別途かかります。

- Microsoft Azure利用料 (月額)
- Azure製品サポート費用 (月額)

IIJ PaaS活用ソリューション with Microsoft Azure オプション費用

検証環境評価・改善支援

個別お見積

■ 概要

お問い合わせへの回答や定例会などで検証環境の評価、課題整理をご支援します。お客様と作業内容や対応時間を協議のうえ費用は決定いたします。

- 契約形態：準委任

閉域接続

IIJプライベートバックボーンサービス/Smart HUB

初期：¥230,000

月額：¥331,000

■ 概要

お客様環境とMSバックボーンで接続。Teamsと安定した通信が可能。

- 帯域100M想定
(品目:スマートハブ/東京NAPTセッション (3万) /PBB/PVA)
- お客様環境とPVAをつなぐ回線費用が別途必要です。

※税別表記

参考費用（サービス利用料金）

■前提条件

- ・ GPT-3.5-turboモデルでの会話を想定
- ・ 利用ユーザ100名が1日に3トピックを作成し、8回会話（営業日20日/月）

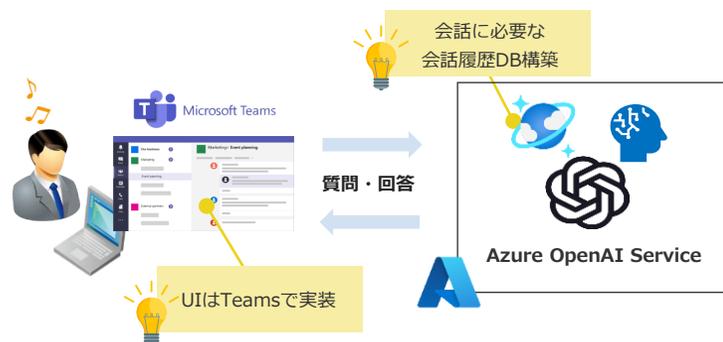
2023年7月時点
1 USD = ¥140.50 で換算

| # | 内容 | サービス名称 | リージョン | スペック | 参考月額 |
|-----------|---------------------------|------------------------------------|-------------------------|--|-----------------|
| 1 | APIサーバ | Azure Functions | Japan East | Premium レベル EP1: 1 コア, 3.5 GB RAM, 250 GB ストレージ 744時間 | ¥25,609 |
| 2 | AI応答 GPT 3.5 | Azure OpenAI Services | East US | Chat GPT (gpt-3.5-turboモデル) 2,000トークン/回 * 3トピック * 更問8回 * 100人 * 20営業日 | ¥26,975 |
| 3 | 会話履歴DB | Azure Cosmos DB | Japan East | 単一マスター、標準プロビジョニングスループット、10GBストレージ | ¥19,211 |
| 4 | USリージョン への通信量 | Bandwidth | Japan East - East US | East US への Japan East からの 100 GB 送信データ転送 | ¥1,067 |
| 5 | プライベート エンドポイント | Azure Private Link | Japan East East US | 100GBのデータ処理、744時間 | ¥3,697 |
| 6 | Microsoft Azure製 品サポート | UOM Azureサポート | — | Microsoft Azureに関するサポート窓口 ※平日9:00~18:00 | ¥30,000 |
| 7 | Teams連携 | スタンドアロン版 Power Virtual Agents | — | 1テナントに2,000セッション/月が付与 足りない場合は追加で¥12,500で1,000セッション/月が付与 ※2023年12月までの特価。それ以降は¥150,000/月 | ¥25,000 |
| 8 | 利用分析用 | Power BI Service | — | 作成したダッシュボードを組織内で閲覧が可能 作成者と閲覧者2名分（¥1,504 ×2） | ¥3,008 |
| 9 | チャット インターフェース | Microsoft 365 Business Standard | — | 利用100ユーザー分×¥1,560 ※お客様でTeamsをお持ちの場合は必要なし | ¥156,000 |
| 合計 | | | | | ¥290,567 |

AI活用の経営効果検証の為のAOAIのPoC環境構築

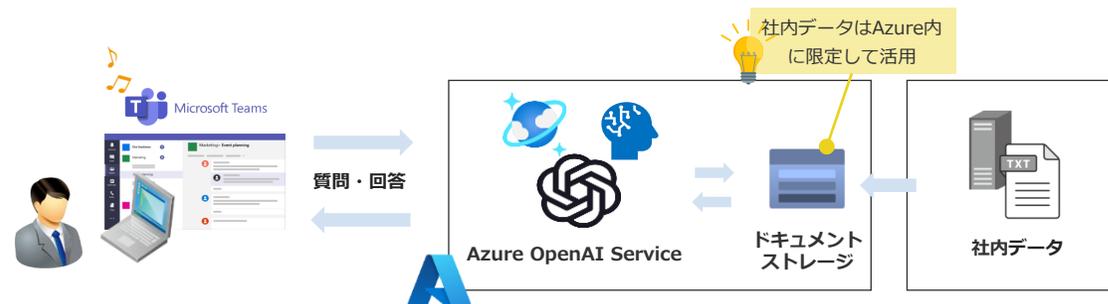
- 段階的な開発フェーズ(フェーズ1・2)を設定。継続的に機能を追加するプロジェクトを提案。
- フェーズ1ではChatGPTの検証環境の構築。AOAIで足りないUI、スムーズな会話のための会話DBを実装。
- フェーズ2では社内データを検索するナレッジマイニングの環境を構築。社内データをAzure内に閉じてセキュアに活用できる環境を実装。

フェーズ1：AOAIの評価検証環境



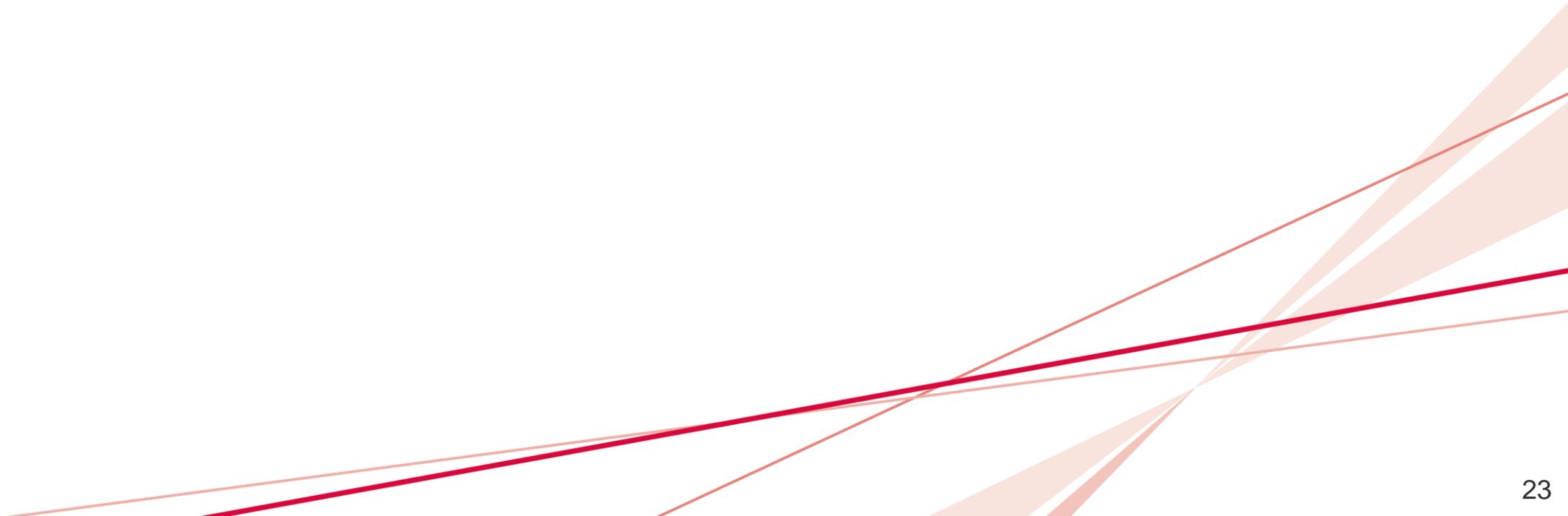
- ✓ AOAIに触れる環境を構築してAIに慣れる
- ✓ AIを活用した経営効果の検証

フェーズ2：AOAIに社内データを学習させる検証



- ✓ 社内情報に基づいた要約、アイデア創出をAIで実現。どのような業務を効率化できるか検証。

appendix



前提条件

■ 提供機能

- ✓ Power Virtual Agents(Power Automate)から外部サイトに接続するためにはプレミアムコネクタが必要となりますが、Power Virtual Agents for Teamsで使用権のあるPower Automateにはライセンスを付与することができません。そのため、Power Virtual Agentsでスタンドアロンのテナントライセンスをご購入いただくことで、Power Virtual Agentsで使用権のあるPower Automateがプレミアムコネクタを利用可能となります。
- ✓ 本システムとChat GPTでは質問に対する会話履歴やコンテキストの利用方法が異なるため、質問で得られる回答の精度に差異が生じる場合があります。その点を加味した評価を実施いただくようお願いいたします。

■ 検証環境導入・構築

- ✓ 構築作業の実施場所は、IIJオフィスまたは作業者のリモートワーク作業場所とさせていただきます。これら以外の作業場所での作業が発生した場合は、出張費・交通費・宿泊費・日当などの経費を弊社規定に基づき、実費にてご請求させて頂くものとします。
- ✓ 構築作業として一時的にお客様オフィスでの作業が必要な場合には、作業人数分の作業場所や作業端末、ネットワーク環境をご提供ください。
- ✓ 検証環境として提供する機能の構築はインテグレーション作業として実施するものです。インテグレーションによる提供範囲について、瑕疵期間は検収後6カ月とさせていただきます。ただし、瑕疵の原因がお客様の責に帰すべき事由の場合は、弊社は当該瑕疵につき本責任を負わないものとさせていただきます。
- ✓ お客様のご都合により、深夜や休日対応が必要な場合は、別途費用が発生いたします。
- ✓ 構築のためにMicrosoft365アカウントやMicrosoft Azureアカウントをご提供いただくことを前提といたします。構築にかかる期間は、作業用アカウントをご提供いただき、構築に必要な申請が完了してからおよそ1カ月を想定しております。なお、各種申請はアカウント準備は弊社にてサポート致します。
- ✓ 本システムは、IPアドレスによるアクセス制限を適用する範囲があります。お客様のIPアドレスに加え、開発や運用保守のため弊社のインターネット出口となるIPアドレスも許可します。
- ✓ Azure OpenAI Serviceは日本リージョンでは提供されておられません（2023年6月時点）。Azure OpenAI Service以外は日本リージョンでの構築が可能ですが、Azure OpenAI Serviceは米国東部リージョンでの構築となります。
- ✓ Azureサービスの品目やスペックについては、現時点での構成を想定して算出しております。適用される可能性がある税金は含まれていません。実際の料金は、Azureのサービスの実際の構成や使用状況によって異なりますことをご了承ください。
- ✓ 構築する検証環境は、Azure OpenAI Serviceを活用した動作を理解・検証するための最小構成でのご提供となります。本番適用に向けて、可用性や性能、セキュリティ等を見直す場合には、構成や費用が検証環境と異なる場合がございます。

前提条件

■ システムサポート

- ✓ 検証環境ご利用開始後のシステム運用保守では所定時間でのサポートを想定しております。お問い合わせはメールにて随時受け付けておりますが、対応時間はIIJ営業時間帯（平日、10:00～17:00）でのベストエフォート対応とさせていただきます。なお、弊社サービスに関するお問い合わせ条件はサービス個別サイトを参照ください。
- ✓ 作業の実施場所は、IIJオフィスまたは作業者のリモートワーク作業場所とさせていただきます。
- ✓ お問い合わせはシステム担当者様からのご依頼を前提とし、エンドユーザー様からのお問い合わせ対応はサポート外です。
- ✓ 検証環境で利用する言語やソフトウェアのバージョンアップの対応は実施対象外といたします。
- ✓ 運用保守作業では、本システムに関する構成管理やお問い合わせ、障害発生時の復旧対応を実施します。月あたりの作業量が弊社想定を継続して超過する場合は、費用について別途協議させていただきます。
- ✓ 本システム構成のバックアップ機能は提供範囲外です。障害発生時にデータが全て消去される可能性があります。また、インフラ構築の自動化の仕組みも構築対象外とするため、Azureの障害発生により復旧が必要になった場合は、手作業にて再構築します。月あたりの作業量が弊社想定を超過する場合は、費用について別途協議させていただきます。
- ✓ 保守費用には、ソースコードや構成の変更を伴う改修作業は含まれておりません。評価時などで改修作業が必要な場合は、オプションメニューのご契約をお願いします。
- ✓ 本システムに障害が発生した際の原因がAzureが提供するサービスの障害に起因する場合には、当該サービスの障害の原因調査や復旧作業は運用保守作業範囲外となります。ただし、Microsoft社への問い合わせや復旧後の動作確認は実施し、お客様にご報告します。
- ✓ システムの設定変更や運用対応などの個別のSEサポートが必要な場合はご相談ください。

■ 検証環境評価・改善支援

- ✓ 定例会を開催する場合、週に1回、1回あたり1時間程度を想定しております。
- ✓ 定例会はMicrosoft Teamsを利用したリモート会議での実施とさせていただきます。ただし、必要に応じて直接対面でのお打ち合わせを検討させていただき、実施場所については都度協議させていただきます。対面で実施する場合にはお客様の参加者全員が対面でご参加いただきますよう、調整をお願い致します。
- ✓ グループディスカッション形式でのアイデア出しや議論を実施する場合、事前にお客様のご担当者と目的や内容、参加者の認識合わせをさせていただきます。
- ✓ 課題管理ツールを利用する場合、弊社にてBacklogをご用意してお客様のアカウントをご提供します（最大3ユーザ）。なお、お客様が利用されている課題管理ツールがございましたらご指定いただくことも可能です。



日本のインターネットは1992年、IIJとともに始まりました。以来、IIJグループはネットワーク社会の基盤をつくり、技術力でその発展を支えてきました。インターネットの未来を想い、新たなイノベーションに挑戦し続けていく。それは、つねに先駆者としてインターネットの可能性を切り拓いてきたIIJの、これからも変わることのない姿勢です。IIJの真ん中のIはイニシアティブ

IIJはいつもはじまりであり、未来です。

本書には、株式会社インターネットイニシアティブに権利の帰属する秘密情報が含まれています。本書の著作権は、当社に帰属し、日本の著作権法及び国際条約により保護されており、著作権者の事前の書面による許諾がなければ、複製・翻案・公衆送信等できません。本書に掲載されている商品名、会社名等は各会社の商号、商標または登録商標です。文中では™、®マークは表示していません。本サービスの仕様、及び本書に記載されている事柄は、将来予告なしに変更することがあります。