

# AIチャットボットに関する 責任あるAI 影響評価

影響評価の各セクションに関するご質問は、影響評価ガイドをご参照ください。

## 第1章: システム情報

### システム概要

1.1 以下にシステム情報を記入してください

システム名	マネジメントオフィスいまむら チャットボット
チーム名	-

変更履歴を以下に記録してください

作成者	代表取締役 今村 敦剛
最終更新日	2025年2月26日

影響評価の完了後にレビューを行う担当者を特定してください。

レビュー担当者	代表取締役 今村 敦剛
---------	-------------

## システムのライフサイクル段階

1.2 システムの各ライフサイクル段階における予定リリース日（完了予定日）を記入してください。

日付	ライフサイクル段階
2025/2/2	計画・分析
2025/2/4	設計
2025/2/9	開発
2025/2/16	テスト
2025/2/26	実装・導入
2025/5/31	保守
未定	廃止

## システムの説明

1.3 開発しているシステムについて、専門用語を避け、わかりやすい言葉で簡潔に説明してください。これにより、レビュー担当者がシステムの概要や動作環境を理解しやすくなります。

システムの説明
<p>このシステムは、WordPressで作成したホームページ上に実装されたAIチャットボットです。「AI Engine」プラグインを活用し、OpenAI APIと連携することで、ユーザーの質問に対してリアルタイムで自然な応答を提供します。</p> <p>本システムは、カスタマーサポート、FAQ対応、ユーザーの問い合わせ対応などを目的としており、Webサイト上でスムーズな対話体験を実現します。自然言語処理（NLP）を活用し、ユーザーの意図を理解した適切な回答を提供することで、業務の効率化や顧客満足度の向上を目指します。</p>

システムに関する補足情報（デモ動画、機能仕様書、スライド資料、システムアーキテクチャ図など）がある場合は、以下にリンクを記載してください。

補足情報の説明	リンク

## システムの目的

1.4 システムの目的や主な機能について、利用者のニーズをどのように満たすのかに焦点を当てて簡潔に説明してください。また、AI技術がどのようにこれらの目的達成に貢献するのも併せて記載してください。

### システムの目的

本システムの目的は、Webサイト上でユーザーの質問に迅速かつ適切に対応し、情報提供を効率化することです。主に、以下のようなニーズに応えます。

- ・ ユーザーの疑問を即座に解決（カスタマーサポートやFAQ対応の自動化）
- ・ 24時間対応の利便性向上（営業時間外でも問い合わせが可能）
- ・ 業務の負担軽減（オペレーターの対応時間を削減し、効率化を図る）

本システムは、WordPressの「AI Engine」プラグインとOpenAI APIを活用し、以下のような主要機能を提供します。

- ・ 自然言語処理（NLP）によるユーザーの意図理解
- ・ 適切な回答をリアルタイムで生成
- ・ 過去の対話履歴を活用したパーソナライズ対応
- ・ 複数の言語に対応可能（必要に応じて拡張）

### AI技術の貢献

本システムでは、OpenAIの高度な自然言語処理技術を活用することで、あらかじめ設定されたルールやシナリオに基づいて動作するチャットボットでは対応が難しかった柔軟な対話や曖昧な質問への応答が可能となります。これにより、より自然でスムーズなユーザー体験を実現し、利便性を向上させます。

## システムの機能

1.5 システム全体に焦点を当て、現在備わっている主な機能や、今後のリリースで追加予定の機能について簡潔に説明してください。

現在のシステム機能	今後のリリースで追加予定の機能
AIチャット機能（メインの機能） <ul style="list-style-type: none"> <li>・ FAQ自動応答（事前に登録した学習データを参照して、適切な回答を返す）</li> <li>・ 意図理解（NLP）（ユーザーの入力を解析し、適切な回答を生成）</li> <li>・ テキスト生成（新しい文章を自動で作成し、質問に答える）</li> <li>・ 多言語対応（異なる言語でのチャットが可能）</li> </ul>	
管理機能 <ul style="list-style-type: none"> <li>・ WordPress連携・プロンプト設定・API管理など、ユーザーが直接利用するものではなく、設定側の機能</li> </ul>	

このシステムが他のシステムや製品とどのように関連しているかを簡潔に説明してください。例えば、他のシステムのモデルを組み込んでいる場合などについて記載してください。

他のシステム／製品との関係
<p>このAIチャットボットシステムは、WordPressの「AI Engine」プラグインを使用し、OpenAI APIと連携することで動作しています。具体的には、OpenAIの事前学習済みモデル（GPTベース）を利用し、さらにファインチューニングを行うことで、特定の用途に適した応答ができるようになっています。</p> <p>このシステムはWordPressサイトと統合されており、Webページ上のコンテンツやFAQと連携することで、ユーザーの問い合わせに対し適切な回答を提供します。</p>

## 対応地域と言語

1.6 システムが導入されている、または導入される可能性のある地域について記載し、言語、法律、文化に関する特別な考慮事項が必要かどうかを明らかにしてください。

現在導入されている地域：	日本
今後のリリースで導入予定の地域：	なし
将来的に導入の可能性のある地域：	なし

自然言語処理システムの場合、対応言語を記述してください：

現在対応している言語：	日本	
今後のリリースで対応予定の言語：	なし	
将来的に対応の可能性のある言語：	なし	

## 導入形態

1.7 システムの導入方法について記録してください。

現在の導入形態	このAIチャットボットシステムは、WordPressの「AI Engine」プラグインを利用してWebサイトに統合されており、OpenAI APIと連携して動作しています。現在、Webサイト上の埋め込み型チャットウィジェットとして提供され、ユーザーがブラウザ経由で利用できる形で展開されています。
今後のリリースで導入形態は変更されますか？ 変更がある場合は、その内容を記載してください。	変更の予定はありません。

## 想定される用途

想定される用途とは、チームが設計・テストを行っているシステムの具体的な使用目的を指します。

これは、誰が・どのような目的で・どのような環境でシステムを利用するのかを説明するものであり、システムの個々の機能とは異なります。ある用途に対して、複数の機能が含まれる場合もあります。

以下の表に、システムの想定される用途を記入してください。

用途の名称	用途の説明
カスタマーサポートの自動化	Webサイトの訪問者が、営業時間外でもFAQやサポート情報を即座に取得できるようにする。繰り返し発生する問い合わせに対応し、サポート業務の負担を軽減する。

## 第2章:意図した用途

想定される用途 #1: [用途の名称] (各用途について繰り返してください)

以下の「想定される用途 #1」のセクションをコピーし、上記で特定した各用途ごとに 質問 2.1~2.8 を繰り返し記入してください。

### システムの目的に合っているかの評価

2.1 システムの使用が、各想定用途における課題をどのように解決するのかを評価してください。なお、課題を解決する方法は一つとは限らず、複数の適切な解決策が存在する可能性があることを考慮してください。

#### システムの目的に合っているかの評価

システムの目的: カスタマーサポートの自動化

解決したい課題:

- ・ ユーザーからの問い合わせが多く、カスタマーサポートの負担が大きい
- ・ 営業時間外の対応ができない

このシステムの解決方法:

- ・ AIチャットボットがFAQや定型的な問い合わせに自動で対応する
- ・ 24時間対応が可能になり、営業時間外でも基本的な問い合わせに応答できる

評価:

- ・ 適合性: 高 → 定型的な質問には即時対応でき、サポート業務の負担軽減に貢献する

### ステークホルダー（利害関係者）と潜在的な利益・リスク

2.2 このシステムの想定用途に関わるステークホルダー（利害関係者）を特定してください。その上で、各ステークホルダーにとっての潜在的な利益と、潜在的なリスク（悪影響）を整理してください。

詳細については、「影響評価ガイド (Impact Assessment Guide)」を参照してください。

ステークホルダー（利害関係者）	システム活用の潜在的なメリット	システム活用の潜在的なリスク
Webサイトの訪問者 / 顧客	<ul style="list-style-type: none"> <li>・ 24時間いつでも問い合わせができる</li> <li>・ 待ち時間なしで迅速に回答を得られる</li> <li>・ 必要な情報を簡単に見つけられる</li> </ul>	<ul style="list-style-type: none"> <li>・ 回答の精度が低い場合、誤った情報を受け取る可能性がある</li> <li>・ 複雑な質問には対応できず、ストレスを感じることもある</li> <li>・ チャットボットの誤回答により、追加の対応が必要になる場合がある</li> </ul>
カスタマーサポート担当者	<ul style="list-style-type: none"> <li>・ 定型的な問い合わせ対応が削減され、本来の業務に集中できる</li> <li>・ 業務負担の軽減による効率向上</li> </ul>	<ul style="list-style-type: none"> <li>・ チャットボットの誤回答により、追加の対応が必要になる場合がある</li> <li>・ カスタマーサポート担当はAIに仕事を奪われ、雇用が危ないと感じる</li> </ul>

Microsoft 責任あるAI 影響評価テンプレート

Webサイト運営者	<ul style="list-style-type: none"> <li>・ 顧客対応のコスト削減</li> <li>・ 顧客満足度の向上</li> <li>・ ユーザーの行動データを収集し、サービス改善に活用できる</li> </ul>	<ul style="list-style-type: none"> <li>・ 初期導入や運用・メンテナンスのコストが発生する</li> <li>・ AIの回答が適切でない場合、企業の信頼性に影響を与える</li> </ul>
システム管理者 / AI運用担当者	<ul style="list-style-type: none"> <li>・ チャットボットの継続的な改善が可能</li> <li>・ データを活用し、AIの精度向上が期待できる</li> </ul>	<ul style="list-style-type: none"> <li>・ AIモデルの管理・改善の負担が増加する</li> <li>・ 不適切な回答が出た場合、迅速な対応が求められる</li> </ul>

## 責任あるAI標準に基づく目標指向の要件に関するステークホルダー

- 2.3 責任あるAI標準の特定の目標では、特定のタイプのステークホルダーを特定することが求められます。これらのステークホルダーは、上記のステークホルダー表に含まれている可能性があります。以下の目標のうち、このシステムに該当するものについて、その目的に関連する具体的なステークホルダーを記載してください。該当しない目標については、表の該当欄に「N/A」と入力してください。

### 目標 A5: 人による監視と制御

この目標はすべてのAIシステムに適用されます。以下の表に記入してください。

運用開始中／後のシステムの管理・運用・監視・制御の責任者	これらのステークホルダーが担う監視・制御の責任
システム管理者 / AI運用担当者	AIチャットボットの動作監視とトラブルシューティング <ul style="list-style-type: none"> <li>モデルの更新やパフォーマンス最適化</li> <li>チャット履歴の分析と改善施策の実施</li> <li>システムの不具合対応やアップデート管理</li> <li>AIの精度向上のための技術的サポート</li> </ul>
カスタマーサポート担当者	AIの回答が適切であるかのモニタリング <ul style="list-style-type: none"> <li>AIの誤回答や未対応の質問への補足対応</li> <li>必要に応じてAIの調整・改善リクエストを行う</li> <li>AIの回答に関するクレーム有無のモニタリング</li> <li>AIの回答にクレームがあった場合の対応</li> </ul>
Webサイト運営者	チャットボットのユーザーエクスペリエンス管理 <ul style="list-style-type: none"> <li>AIの導入効果（問い合わせ削減や顧客満足度向上）の評価</li> <li>利用状況データの収集と活用</li> </ul>

### 目標 T1: 意思決定のためのシステムの説明可能性

この目標は、AIが生成した出力が、人に関する意思決定を支援する目的で使用される場合に適用されます。

この目標が本システムに該当する場合は、以下の表に記入してください。

システムの出力を基に意思決定を行う人	意思決定の対象となる（影響を受ける）人
Webサイトの訪問者 / 顧客 (サービス購入の意思決定)	Webサイトの訪問者 / 顧客自身
Webサイトの訪問者 / 顧客 (問い合わせをする意思決定)	カスタマーサポート担当者
システム管理者 / AI運用担当者 (チャットボットの運用を改善する意思決定)	Webサイトの訪問者 / 顧客自身 カスタマーサポート担当者

## 目標 T2: ステークホルダーへの情報伝達

この目標はすべてのAIシステムに適用されます。以下の表に記入してください。

このシステムを特定の業務に導入するかどうかを決定する人	このシステムと統合する他のシステムを開発・導入する人
代表取締役	<ul style="list-style-type: none"> <li>Webサイト管理者 (WordPressにAI Engineを実装)</li> <li>AIシステム管理者 (データセット作成・前処理、OpenAI APIとの連携・設定を担当)</li> </ul>

## 目標 T3: 人間ではなく、AIとの対話であることの開示

この目標は、AIシステムが人間のようにふるまう場合、または画像・音声・動画を生成・加工し、本物と誤認される可能性がある場合に適用されます。ただし、利用環境や文脈から明らかにAIであるとわかる場合は除きます。

この目標が本システムに該当する場合は、以下の表に記入してください。

このシステムを使用する、または接する可能性のある人は誰か？
<p>Webサイト訪問者 / 顧客</p> <ul style="list-style-type: none"> <li>Webサイト上でAIチャットボットと対話し、問い合わせや情報検索を行う</li> </ul> <p>カスタマーサポート担当者</p> <ul style="list-style-type: none"> <li>AIが対応できない問い合わせにエスカレーションされた場合、AIが提供した情報を確認する</li> </ul> <p>Webサイト運営者 / IT管理者</p> <ul style="list-style-type: none"> <li>チャットボットの運用や改善を行い、ユーザー体験を最適化する</li> </ul> <p>代表取締役</p> <ul style="list-style-type: none"> <li>AIの運用方針を決定し、活用の範囲を管理する</li> </ul>

## 公平性の考慮

- 2.4 このシステムに適用される公平性に関する目標について、以下の点を明確にしてください。1) 関連するステークホルダーを特定する（例：システムの利用者、システムの影響を受ける人）、2) 公平性の考慮が必要な特定の属性グループを特定する（例：社会的に疎外されたグループを含む）、3) 公平性の優先度を設定し、公平性がどのように適用されるかを説明する。公平性の目標がこのシステムに適用されない場合は、最初の列に「N/A」と記入してください。

### 目標 F1: サービスの品質

この目標は、システムの利用者や影響を受ける人が、異なる属性（年齢、性別、障がいの有無など）によって、サービス品質に差を感じる可能性がある場合に適用されます。その差を解消するために、システムの設計を改善できる場合、本目標を考慮する必要があります。

この目標が本システムに該当する場合は、以下の表を記入してください。

影響を受けるステークホルダー	この目標において優先的に考慮すべき属性グループ	各属性グループがどのような影響を受ける可能性があるか
Webサイト訪問者 / 顧客	非日本語話者	AIが特定の言語に最適化されている場合、対応できる言語に制限があり、正確な情報を取得しにくい可能性がある。
Webサイト訪問者 / 顧客	高齢者 / ITリテラシーが低いユーザー	チャットボットの操作や回答の理解が難しく、適切な情報にたどり着けない可能性がある。
Webサイト訪問者 / 顧客	視覚障がい者	スクリーンリーダー対応が不十分だと、チャットボットの出力を正しく読み上げられず、利用しにくくなる可能性がある。
カスタマーサポート担当者	問い合わせ内容がAIでは解決できないユーザー	AIの回答精度が低い場合、不要な問い合わせが増加し、業務負担が増す可能性がある。

### 目標 F2: 資源と機会の配分

この目標は、金融、教育、雇用、医療、住宅、保険、社会福祉に関連する資源や機会の配分に直接影響を与えるAIシステムに適用されます。

この目標が本システムに該当する場合は、以下の表を記入してください。

影響を受けるステークホルダー	この目標において優先的に考慮すべき属性グループ	各属性グループがどのような影響を受ける可能性があるか
※このAIチャットボットは金融、教育、雇用、医療、住宅、保険、社会福祉に直接影響を与えるシステムではないため、本目標は適用されません		

**目標 F3: ステレオタイプ、軽視、抹消の最小化**

この目標は、AIシステムの出力が、人々や文化、社会に関する記述・描写・表現を含む場合に適用されます。

この目標が本システムに該当する場合は、以下の表を記入してください。

影響を受けるステークホルダー	この目標において優先的に考慮すべき属性グループ	各属性グループがどのような影響を受ける可能性があるか
Webサイト訪問者 / 顧客	女性	当社のホームページは経営者・管理者向けの情報が多いが、AIが学習データのバイアスによって「問い合わせをした人＝男性」と認識する可能性がある。
代表取締役	多様なバックグラウンドを持つユーザー	AIの表現が偏っていると、多様なバックグラウンドを持つユーザーが当社に抱くイメージに悪影響を及ぼす可能性がある。

## 技術成熟度評価

2.5 このシステムの想定用途に関する技術的な成熟度を、最も適切な説明に「X」を付けて示してください。

ひとつ選択	技術成熟度
	このシステムには、基礎研究に支えられたAIが含まれており、同様の用途で大規模な本番環境への導入はまだ行われていない
X	このシステムには、同様の用途で本番環境に導入された実績があり、実現可能性を示す証拠があるAIが含まれている
	この用途で、1つ以上のシステムコンポーネントが初めて適用される。運用可能な条件はまだ完全に定義・評価されていない
	この用途で、システム全体が初めて検証される。運用可能な条件も評価される。または、ほぼ同様のシステムや手法が他の組織によって成功事例として適用されている
	システム全体がすべての想定用途に導入されており、運用条件も本番環境でのテストや使用を通じて確認されている

## タスクの複雑さ

2.6 このシステムの想定用途に関するタスクの複雑さについて、最も適切な説明に「X」を付けて示してください。

ひとつ選択	タスクの複雑さ
	<p><b>単純なタスク</b>：少数の特徴に基づいた分類を行い、明確なカテゴリに振り分けるタスク。人間が簡単に正解を判断でき、システムのミスを容易に識別できる。</p> <p>例：文書のスペルチェックを行う自然言語処理システム。</p>
	<p><b>中程度の複雑さのタスク</b>：少数のカテゴリに分類するが、分類が主観的であるタスク。通常、多くの評価者が同じ答えに到達することで正解が定義される。</p> <p>例：ユーザーが入力中の単語やフレーズを自動補完する自然言語処理システム</p>
X	<p><b>複雑なタスク</b>：多くの特徴を基にしたモデルであり、人間には解釈が難しく、予測が変動しやすいタスク。意思決定基準に明確な境界がなく、人間が最適な答えを決めるのが難しい場合がある。</p> <p>例：ユーザーの入力プロンプトに基づいて文章を生成する自然言語処理システム</p>

## 人間の役割

2.7 このシステムの想定用途における人間の役割について、最も適切な説明に「X」を付けてください

ひとつ選択	人間の役割
	<p><b>最低限の監視（問題発生時のみ対応）：</b>システムがエラーを検知した際にのみ、人が対応を行う。通常の運用はシステムが自動で行い、人は監視しない。</p> <p>例：構造化されていないテキストからキーワードを抽出するAIが、ファイル形式の誤りを検知した際にのみオペレーターへ通知を送る</p>
X	<p><b>ほぼ自動運用（必要に応じて人が介入）：</b>システムのほとんどの処理を自動化し、必要に応じて人が引き継げる仕組みがある。</p> <p>例：キーワード抽出AIが信頼度の閾値を下回った場合のみ、管理者が設定した条件に基づいてオペレーターに通知を送る</p>
	<p><b>自動運用だが、特定の条件で人が必ず介入：</b>システムがほぼ自動で運用されるが、特定のケースでは必ず人が対応する必要がある。</p> <p>例：キーワード抽出AIが信頼度の閾値を下回った場合は、管理者の設定に関わらずオペレーターに通知を送る</p>
	<p><b>人間が結果を確認し、必要なら介入できる：</b>システムが出力した結果を人が確認し、問題がなければそのまま処理が進むが、介入することもできる。</p> <p>例：キーワード抽出AIが生成したキーワードをオペレーターが確認し、問題がなければ自動的に確定する</p>
	<p><b>人間の承認がないと処理が進まない：</b>システムが出力した結果を人が確認し、承認しない限り処理が進まない。</p> <p>例：キーワード抽出AIが生成したキーワードは、オペレーターが承認するまで確定されない</p>

## 導入環境の複雑さ

2.8 このシステムの想定用途における導入環境の複雑さについて、最も適切な説明に「X」を付けてください

ひとつ選択	導入環境の複雑さ
	<p><b>単純な環境：</b>システムが導入される環境が固定されており、入力の種類が限られている。また、システムが対応しなければならない予期しない状況がほとんど発生しない。</p> <p>例：研究室内の管理された環境で使用される自然言語処理システム</p>
X	<p><b>中程度の複雑さの環境：</b>システムの導入環境には変動があり、予測できない事態が発生する可能性はあるが、その影響が人に及ぶリスクは低い。問題が起きた際の対処方法も比較的明確である。</p> <p>例：ビジネス向けの自然言語処理システム（社内利用）で、専門的な言葉遣いが主であり、コミュニケーションのルールが大きく変わることが少ない</p>
	<p><b>複雑な環境：</b>システムの導入環境が常に変動し、予測が難しい状況で運用される。また、入力の種類が多岐にわたり、質にもばらつきがある。ユーザーが短時間で判断を下さなければならず、問題が発生した際の対処が難しい場合もある。</p> <p>例：ソーシャルメディアで利用される自然言語処理システム（言葉の使い方やコミュニケーションのスタイルが急速に変化する）</p>

## 第3章: 想定される負の影響

### 利用制限

3.1 このシステムの使用に関して、法律や社内ポリシーで制限されている用途がある場合は、ここに記載してください。また、それらの制限に従って運用してください

#### 利用制限

個人情報の収集・保存（個人情報保護方針）

- ・ チャットボットが個人情報を保存しないように設定している
- ・ 個人情報を入力しないよう、チャット画面で警告している

### 非対応の用途

3.2 本システムが設計や評価の対象としていない、または使用を避けるべき用途

#### 非対応の用途

- ・ 具体的な経営アドバイスの提供
- ・ 医療・法律・金融アドバイスの提供

### 既知の制限事項

3.3 このシステムに関する既知の制約を説明してください。具体的には、うまく機能しない可能性がある状況、環境要因、運用上の注意点などが該当します

#### 既知の制限事項

複雑な問い合わせへの対応が難しい

- ・ FAQや定型的な質問には対応できるが、イレギュラーなケースや詳細な説明が必要な問い合わせには不十分な場合がある。その場合は、有人対応への引き継ぎが必要。

文脈の理解が不完全な場合がある

- ・ 学習データを考慮して会話を続けることができるが、長時間のやり取りや複雑な文脈では意図を正しく理解できないことがある。

代表ブログ記事の内容を踏まえた回答ができない

- ・ 代表ブログ記事の内容を学習データに含めていないので、その内容を踏まえた回答ができない

専門的な助言は提供できない

- ・ 経営・医療・法律・金融に関する質問は、詳しい状況やデータ等が不明なので、適切な返答ができない

ユーザー入力の誤字・曖昧な表現に対する耐性

- ・ 一般的な誤字には対応できるが、極端に崩れた表現や専門用語には正確に対応できない可能性がある。

多言語対応の精度

- ・ 学習データを日本語で作成しているため、日本語には対応できるが、その他の言語（方言含む）では正確に

<p>応答できない場合がある。</p> <p>インターネット環境に依存する</p> <ul style="list-style-type: none"> <li>AI APIを使用しているため、ネットワークの接続状況によっては応答が遅れる、または利用できなくなることがある。</li> </ul> <p>最新情報の反映に制限がある</p> <ul style="list-style-type: none"> <li>AIの知識は学習データの時点での情報に基づいており、最新のニュースや製品情報には即座に対応できないことがある。</li> </ul> <p>不適切な回答のリスク</p> <ul style="list-style-type: none"> <li>AIの学習データによっては、不適切または誤解を招く回答をする可能性がある。そのため、重要な対応には人間のチェックが必要。</li> </ul>
--

## システムの失敗がステークホルダーに与える影響

3.4 システム全体において想定される失敗（誤検出や見逃しを含む）を定義し、それぞれの想定用途においてステークホルダーにどのような影響を及ぼすかを説明してください

失敗の種類	具体的な例	影響を受けるステークホルダー	影響
誤検出 (False Positive)	チャットボットがユーザーの入力を「有人対応が必要」と誤って判断し、オペレーターにエスカレーションする	カスタマーサポート担当者	不必要な問い合わせが増え、業務負担が増加する
	ユーザーの一般的な問い合わせに対して「対応不可」と誤って判断し、回答を拒否する	ユーザー	ユーザーが求める情報を得られず、満足度が低下する
見逃し (False Negative)	ユーザーが問題を相談しているのに、AIが「対応不要」と判断し、適切な対応につながらない	ユーザー	緊急対応が遅れ、ユーザーが不満を感じる・企業の信用低下につながる
	ユーザーのクレームや不満の表現をAIが認識できず、事務的な対応を継続する	代表取締役	企業の対応が不適切と受け取られ、ブランドイメージが悪化する可能性がある
回答の不正確さ	AIが古い情報や間違っただ情報を提供する	ユーザー	誤情報に基づいた行動を取ってしまう可能性がある
曖昧な回答	ユーザーの質問に対して適切な回答ができず、漠然とした返答をする	ユーザー	サポート品質が低下し、有人対応を求めるケースが増える

## システムの誤使用がステークホルダーに与える影響

3.5 システムの誤使用（意図的・非意図的なものを含む）が、各ステークホルダーにどのような悪影響を及ぼす可能性があるかを説明してください。また、誤使用の影響が社会的に立場の弱いグループ（マイノリティ）に対して異なる影響をもたらすかを検討し、記録してください。

重大な誤使用による影響が特定された場合は、影響の要約に「潜在的なリスク」として記載してください。

失敗の種類	具体的な例	影響を受けるステークホルダー	影響
本来の用途外の利用（非意図的）	ユーザーがカスタマーサポート目的ではなく、経営アドバイスや人事戦略の相談をする	ユーザー	AIが適切な回答をできず、不正確なアドバイスが提供される可能性がある
本来の用途外の利用（意図的）	ユーザーがAIを雑談やエンターテインメント用途で利用する	企業の運営管理者	本来の目的と異なる利用が増え、サーバー負荷やコスト増大の要因となる可能性
意思決定の誤用（非意図的）	AIの回答をそのまま鵜呑みにし、サポートチームの確認なしに判断する	ユーザー / 企業	ユーザーが誤った判断をし、クレームやトラブルにつながる可能性
不適切な質問（意図的）	AIに差別的・暴力的な内容を言わせようとする	企業のブランド管理者	AIが不適切な発言をした場合、企業イメージが損なわれるリスクがある
過剰な負荷をかける利用（意図的）	ボットを使い、大量の問い合わせを行う	システム管理者	システムリソースを圧迫し、他のユーザーが正常に利用できなくなる可能性

## 慎重な対応が求められる用途

3.6 このシステムの利用または誤用が、Microsoftの定める「慎重な対応が必要な利用基準」に該当する可能性があるか検討してください

該当する？	慎重な対応が求められる利用基準
いいえ	<b>法的立場や人生の選択肢への影響</b> このAIシステムの利用または誤用によって、個人の法的な地位や権利、クレジット（信用）、教育、雇用、医療、住宅、保険、社会福祉サービスへのアクセスや、それらの提供条件に影響を与える可能性がある場合
いいえ	<b>身体的・精神的な危害のリスク</b> このAIシステムの利用または誤用によって、個人が重大な身体的または精神的な被害を受ける可能性がある場合
いいえ	<b>人権への脅威</b> このAIシステムの利用または誤用によって、個人の人権が制限・侵害されたり、実質的に損なわれる可能性がある場合。 人権は相互に関連し合っているため、AIはほぼすべての国際的に認められた人権に影響を与える可能性がある

## 第4章: データ要件

### データ要件

4.1 このシステムの想定される用途、関係するステークホルダー、および導入予定の地域に応じたデータ要件を定義し、記録してください

#### データ要件

##### データの種類

- ・ Q&Aデータ（自社Webサイトから作成。JSON形式）
- ・ 製品・サービス情報（自社の製品・サービスに関する情報。HTML・テキスト形式）

### 既存（公開されている・他システムで利用されている）データセット

4.2 システムの学習に既存のデータセットを使用する場合、そのデータの量や適切性を評価してください。評価は、前述のデータ要件に照らして行います。

既存のデータセットを使用しない場合は、「該当なし（N/A）」と記入してください

#### 既存のデータセット

##### 該当なし（N/A）

このAIチャットボットシステムでは、ファインチューニング用の学習データとして、自社Webサイトに掲載されている情報に基づいたQ&Aデータのみを使用しています

## 第5章: AIシステム影響のまとめ

### 潜在的なリスクと初期対応策

5.1 これまでの影響評価 (Impact Assessment) で特定した潜在的なリスクを、以下の表にまとめてください。

(ステークホルダーの影響、公平性の考慮点、悪影響の可能性などのセクションを参照し、リスクが記載されている箇所を確認してください。)

また、影響評価ガイド (Impact Assessment Guide) に記載されている対策のヒントを活用し、Microsoftの「責任あるAI基準 (Responsible AI Standard)」が、特定したリスクの一部をどのように軽減できるかを検討してください。

完全に軽減できないリスクについては、チーム内やレビュー担当者と議論し、対応策を検討してください

潜在的なリスク	該当する「責任あるAI基準」の目標 (該当する場合)	初期対応策のアイデア、またはシステム設計で該当の目標をどのように適用するか
不正確な情報の提供 (古いFAQ情報や誤った回答を出す可能性)	Goal F1: サービス品質	FAQデータの定期更新を徹底 (Webサイト更新時にAI学習データもアップデート)
本来の用途外の利用 (カスタマーサポート以外の目的で経営戦略や法律相談を求められる可能性)	Goal T3: AIとの対話であることの開示	プロンプト設計で「このAIはカスタマーサポート専用です」と明示
個人情報の入力リスク (ユーザーが予告なしに個人情報を入力する可能性)	Goal A5: 人による監視と管理	個人情報を入力しないよう、チャット画面に警告を表示 ログを保存しない設定
感情的なユーザー対応の難しさ (クレーム対応時に冷たい印象を与えてしまう可能性)	Goal F1: サービス品質	適切なフレーズを選ぶため、FAQデータの定期更新
ジェンダーバイアスや文化的偏見 (特定の性別や文化に偏った応答をする可能性)	Goal F3: ステレオタイプ・差別的な出力の最小化	学習データの定期チェックを行い、偏った表現がないか確認 性別を問わない表現を推奨し、多様な言葉遣いのデータを取り入れる
意図的な悪用 (ユーザーがAIに攻撃的・差別的な発言をさせようとする)	Goal A5: 人による監視と管理	ChatGPT自体の基本機能として、攻撃的・差別的な発言を防ぐ仕組みがすでに組み込まれている
過剰な負荷によるシステム遅延	Goal T2: ステークホルダーとの適切な情報共有	レートリミットを設定し、短時間での過剰なリクエストを制限
多言語対応の精度のばらつき	Goal F1: サービス品質	主要な問い合わせに関しては、事前に人間が翻訳したFAQを用意し、適切な言語で対応できるようにする

## 目標の適用範囲

5.2 このシステムに適用される目標を評価するために、以下の表を使用してください。

特定の種類のAIシステムにのみ適用される目標については、この影響評価 (Impact Assessment) で対象となるシステムに該当するかを「はい (Yes)」または「いいえ (No)」で示してください。

該当しない場合は、その理由を回答欄に記入してください。

該当する場合は、システム開発時にその目標に関連する要件を満たす必要があります

### 責任に関する目標 (Accountability Goals : A1~A5)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
<b>A1: 影響評価</b> すべてのAIシステムに適用される	はい (システムの適用範囲やリスクを事前に評価する必要があるため)
<b>A2: 重大な悪影響の監視</b> すべてのAIシステムに適用される	はい (誤情報の提供や不適切な対応がブランドイメージに影響を与える可能性があるため)
<b>A3: 目的適合性</b> すべてのAIシステムに適用される	はい (カスタマーサポート用途として適切に機能することが求められるため)
<b>A4: データガバナンスと管理</b> すべてのAIシステムに適用される	はい (FAQや問い合わせ履歴を適切に管理し、最新かつ正確な情報を提供する必要があるため)
<b>A5: 人による監視と管理</b> すべてのAIシステムに適用される	はい (AIの誤回答やエスカレーションが必要なケースでは人間の判断が求められるため)

## 透明性に関する目標 (Transparency Goals : T1~T3)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
<b>T1: 意思決定のためのシステムの説明可能性</b> 適用対象: AIの出力が、人に関する意思決定や、人に関する意思決定を支援するために使用される場合	いいえ (当社のAIチャットボットは、カスタマーサポートを目的としており、ユーザーや第三者に対して直接的な意思決定を行うものではないため)
<b>T2: ステークホルダーへの情報共有</b> 適用対象: すべてのAIシステム	はい (カスタマーサポート業務において、AIがどのように動作するかを関係者 (顧客・管理者) に適切に説明し、透明性を確保することが求められるため)
<b>T3: AIとの対話であることの開示</b> 適用対象: AIが人間のように振る舞う対話を行う場合 (ただし、利用状況や文脈からAIであることが明らかでない場合)、または、AIが画像・音声・動画を生成・編集し、それが本物であるかのように見える可能性がある場合	はい (ユーザーがチャットボットを人間のオペレーターと誤認しないように、対話がAIによるものであることを明示する必要があるため)

## 公平性に関する目標 (Fairness Goals : F1~F3)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
<b>F1: サービスの品質</b> 適用対象: AIシステムの利用者や、その影響を受ける人々が、年齢・性別・文化的背景などの違いによって、サービスの品質に差が生じる可能性がある場合。システムの設計を工夫することで、この差を解消できる場合に適用される	はい (ユーザーの属性によって応答品質に差が出る可能性があるため)
<b>F2: 資源や機会の分配</b> 適用対象: AIシステムの出力が、金融・教育・雇用・医療・住宅・保険・社会福祉 などに関連する資源や機会の分配に直接影響を与える場合	いいえ (カスタマーサポート用のため、資源や機会の分配には関与しない)
<b>F3: 先入観・差別・偏見の最小化</b> 適用対象: AIシステムの出力に、人々・文化・社会に関する説明や描写が含まれる場合。先入観や差別を助長したり、一部の人々や文化が排除されたりしないようにすることが求められる	いいえ (カスタマーサポート用のため、直接該当しない。ただし文化的偏りを防ぐための配慮は、学習データの更新などの形で、考慮する必要がある)

## 信頼性と安全性に関する目標 (Reliability &amp; Safety Goals : RS1~RS3)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
RS1: 信頼性と安全性に関する指針 すべてのAIシステムに適用される	はい (経産省のAI事業者ガイドライン等を参考に、AIチャットボット運用のための具体的な運用規程を作成する)
RS2: システムの失敗と対応策 すべてのAIシステムに適用される	はい (誤回答やシステムの不具合時に適切な対応を準備する必要がある)
RS3: 継続的な監視・フィードバック・評価 すべてのAIシステムに適用される	はい (回答精度の向上と運用の最適化のために、定期的な評価と改善が必要)。

## プライバシーとセキュリティに関する目標 (Privacy &amp; Security Goals : PS1~PS2)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
PS1: プライバシー基準の遵守 適用対象: Microsoftのプライバシー基準が適用される場合	いいえ (Microsoftの基準には該当しないが、一般的なプライバシー保護対策は必要)
PS2: セキュリティポリシーの遵守 適用対象: Microsoftのプライバシー基準が適用される場合	いいえ (Microsoftの基準には該当しないが、独自のセキュリティ対策を設けるべき)。

## インクルーシブ (包括性) に関する目標 (Inclusiveness Goal)

目標	この目標はこのシステムに適用されますか？ (はい / いいえ)
I1: アクセシビリティ基準の遵守 適用対象: Microsoftのプライバシー基準が適用される場合	いいえ (Microsoftの基準には該当しないが、一般的なアクセシビリティ配慮は必要)

## 影響評価の承認

## 5.3 次のステップに進む前に、必要なレビューを完了し、影響評価の承認を行ってください。

最低限、プロジェクトマネージャー (PM) が影響評価の内容を確認し、完了していることを確認する必要があります。また、開発を開始する前に、組織のルールに従って適切なレビューを実施し、すべての必要な承認を取得してください

レビュー担当者の役職と氏名	チーム内の多様な専門知識 (例: エンジニア、デザイナー、データサイエンティストなど) を活かし、協力して作成されたことを確認します。	レビュー日	コメント
代表取締役 今村敦剛	今村 敦剛	2025年 2月26日	

影響評価 (Impact Assessment) は、最低でも年に1回、新たな用途が追加された際、または次のリリース段階へ進む前に更新・再評価してください。

このドキュメントは、責任あるAI基準 (Responsible AI Standard) の各目標を満たすための重要な参考資料として、今後も活用されます。