

AIエージェント評価・改善プロジェクト

キックオフミーティング

日時: 2025年12月〇日 〇〇:〇〇-〇〇:〇〇

場所: 〇〇会議室 / オンライン

参加者: 業務チーム、開発チーム

1. 本日のゴール（5分）

このプロジェクトは
「**AI**を入れる作業」ではありません

真のゴール

- 散在する社内ナレッジを整理する
- 業務チームの問い合わせ対応工数を削減する
- 「自分たちが楽になる仕組み」を一緒に作る

キックオフで達成したいこと

1. プロジェクトの目的と進め方を共通理解する
2. 業務チームの役割を明確にする
(無理なく続けられる範囲で)
3. 最初に取り組む対象範囲（スコープ）を決める

2. プロジェクトの背景と目的

(10分)

2-1. 現状の課題

📞 問い合わせ対応の負担

- 口頭、チャット、メール...様々な方法で同じ質問が繰り返される
- 「またこの質問か...」と感じることが多い
- 回答する人によって微妙に内容が違うこともある

2-1. 現状の課題（続き）

ナレッジ管理の問題

- 正式な規程があるのに、実際には掲示板のFAQが使われている
- 古いFAQが散在し、どれが最新かわからない
- 「業務部門で管理されるべきナレッジ」が整理されていない状態

2-1. 現状の課題（続き）

不足しているもの

- どの質問が多いのか？（データがない）
- ユーザーは満足しているのか？（評価がない）
- 本当に必要な情報は何か？（優先順位が不明）

2-2. プロジェクトの3つの目的

目的	具体的な成果（KPI）
① 対応工数の削減	問い合わせ件数〇〇%減少 AIによる自己解決率〇〇%以上
② 不足ナレッジの 検出と拡充	「答えられない質問」のリスト化 新規FAQ作成件数〇〇件
③ ユーザー満足度向上	5段階評価で平均〇〇以上 低評価要因の改善サイクル確立

2-3. なぜ「今」やるのか？

✅ すでに基盤が整っている

- 会話ログと5段階評価を記録する仕組み（Dataverse）が構築済み
- 「データ」があれば、客観的に改善できる

⚠️ 放置するとどうなるか？

- 古い情報がさらに散在する
- 新人への引き継ぎが属人化する
- AIを入れても「ゴミを学習させたAI」になる

3. AIエージェントの仕組み

(デモ)

(15分)

3-1. Copilot Studioの基本動作

ユーザーの質問 → ナレッジ検索 → 回答生成 → 5段階評価

2つの回答方式

- **生成AI (Generative Answers)**
規程やマニュアルを読んで自動回答
- **トピック (会話フロー)**
質問を聞き返しながら誘導する

3-2. 実際のデモ（良い例・悪い例）

✅ 良い例：「有給休暇の申請方法は？」

→ 規程を正確に引用して回答

❌ 悪い例：「出張の交通費は？」

→ 掲示板の古い情報を拾って間違った回答をする

❌ 悪い例：「慶弔見舞金について教えて」

→ 「情報が見つかりませんでした」（ナレッジ不足）

3-3. 重要な気づき

「AIは鏡である」

AIは鏡である

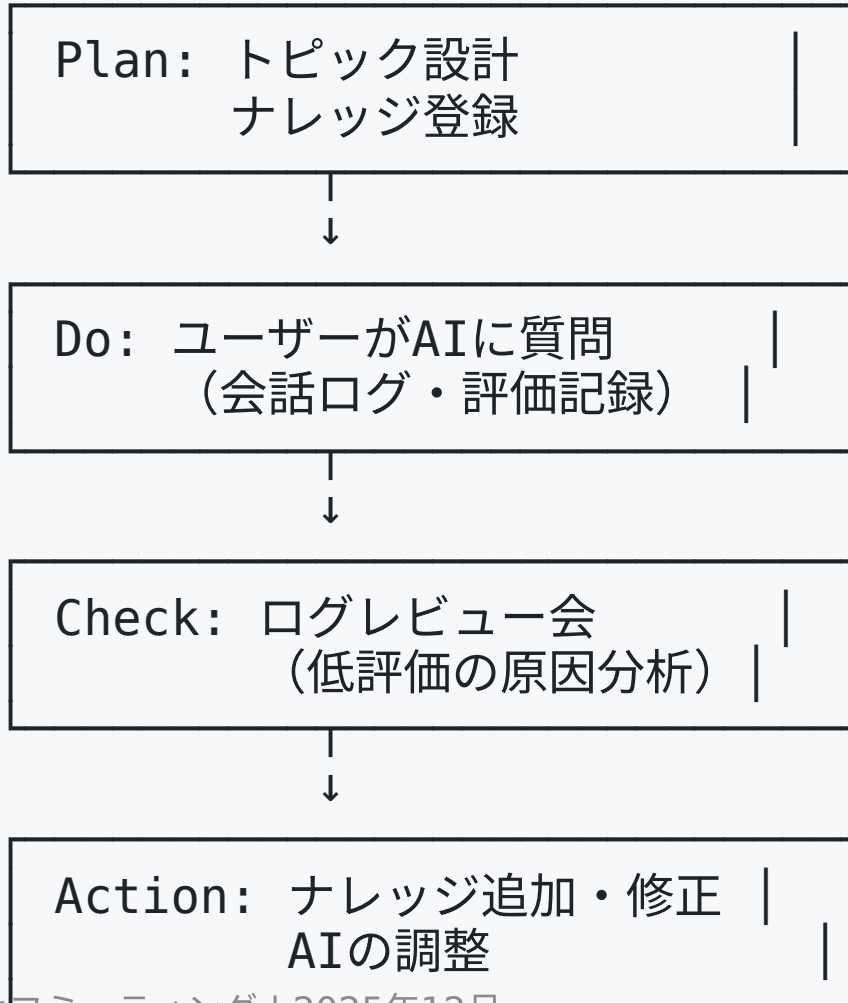
- ナレッジが整理されていれば
→ AIも正確に答える
- ナレッジが散在・矛盾していれば
→ AIも混乱する
- ナレッジが存在しなければ
→ AIは答えられない

つまり、**AIの精度向上 = ナレッジ整理そのもの**

4. 評価・改善のサイクル

(10分)

4-1. PDCAサイクル



4-2. 「低評価」の捉え方

✗ 間違った認識

低評価 = 失敗

✓ 正しい認識

低評価 = 改善の種（宝物）

低評価の会話ログからわかること

- どんな質問に答えられていないか？
- どのナレッジが不足しているか？
- 古い情報が混在していないか？

5. 役割分担とスケジュール

(15分)

5-1. 業務チーム（有識者）の役割

「AIの先生」として、以下をお願いします

タスク	頻度	時間	内容
① ログレビュー	週1回	30分	低評価の会話を確認し判定
② ナレッジ判断	随時	-	「正式採用？捨てる？」決定
③ 原稿作成	随時	-	回答元ネタ作成

ポイント

- 技術的な実装は開発側がやります
- 業務チームは「正解を知っている人」として判断だけをお願いします
- 完璧を目指さず、
小さく始めて少しずつ改善していきましょう

5-2. 開発チーム（実装側）の役割

タスク	内容
① データ分析・可視化	Dataverseのログを集計 Power BIで可視化
② 実装	Copilot Studioに登録 プロンプト調整、トピック作成
③ ファシリテーション	定例会で具体的に問いかけ 判断をサポート

5-3. スケジュール（3ヶ月）

フェーズ	期間	主な活動
Phase 1 準備・トピック設計	1ヶ月目	よく来る質問Top 10リスト作成 既存FAQ棚卸し トピック定義シート記入
Phase 2 実装・試験運用	2ヶ月目	優先トピックの実装 限定ユーザーでテスト 週次ログレビュー開始
Phase 3 本格運用・改善	3ヶ月目	全社展開 KPI測定 継続改善サイクル確立

5-4. 定例会（週次レビュー会）

開催: 毎週〇曜日 〇〇:〇〇-〇〇:〇〇（30分）

アジェンダ:

1. 先週の会話ログ振り返り（低評価トップ3）
2. ナレッジ改善アクション（追加・修正・削除）
3. 次週のタスク確認

6. 最初のターゲット設定 (Scope)

(10分)

6-1. 「小さく始める」重要性

✖ 失敗パターン

最初から全ての規程・FAQを完璧にしようとする
→ 範囲が広すぎて頓挫する

✔ 成功パターン

「まず1つのトピックを完璧にする」
→ 成功体験を作り、横展開する

6-2. 優先トピックの選び方

以下の基準で「Top 3」を選びます

優先順位の基準	具体例
① 問い合わせ頻度が高い	年末調整、交通費申請、有給申請
② 時期が集中する	4月（入社）、12月（年末調整） 3月（年度末）
③ ナレッジが明確	規程がすでにあり 回答内容が一意に定まる

6-3. 【ディスカッション】 最初の対象を決めよう

質問1:

「『またこの質問か...』 とため息が出るものは何ですか？」

質問2:

「新入社員や中途入社者から必ず聞かれることは？」

質問3:

「マニュアルを見れば書いてあるのに電話がかかってくるものは？」

【記入タイム】

付箋やホワイトボードに書き出してください
(3分)



グルーピング・投票で「Top 3」を決定
(5分)

7. 次のアクション

(5分)

7-1. 今日決まったこと

- プロジェクトの目的
ナレッジ整理 + 工数削減
- 役割分担
総務＝判断、開発＝実装
- 最初の対象
【ここに記入】

7-2. 次回までの宿題

業務チーム

1. 「よく来る質問Top 10」リスト作成
(メモ書き・箇条書きでOK)
2. 「トピック定義シート」記入
(まずは3件だけでOK。記入ガイド参照)
3. 既存FAQの質問文リスト
(掲示板にあるQ&Aの「Q」だけ抽出)

7-2. 次回までの宿題（続き）

開発チーム

1. Dataverseのログ集計ダッシュボード作成
2. トピック定義シート回収・レビュー準備

7-3. 次回ミーティング

日時: 2025年12月〇日 〇〇:〇〇-〇〇:〇〇

場所: 〇〇会議室 / オンライン

アジェンダ:

- トピック定義シートレビュー
- 「トピック」か「生成AI」かの仕分け
- 実装スケジュール確定

8. Q&A

(5分)

補足資料

KPI測定項目（案）

KPI	測定方法	目標値
自己解決率	（評価4以上） ÷ 総会話数	70%以上
問い合わせ削減率	導入後 ÷ 導入前	30%削減
ナレッジ拡充数	新規追加 + 統合・削除	50件
低評価改善率	前週比で低評価が減少	毎週改善

補足資料

「気づき」を生むポイント

このプロジェクトを通じて、業務チームに気づいてもらいたいこと:

1. 「**AI**に入れてみて初めて矛盾に気づく」

人間は文脈で理解できるが、AIは正直に古い情報を拾う

2. 「書き散らしの危険性」

掲示板への無秩序なFAQ追記が、実はリスク

3. 「整理すれば楽になる」

「どれが最新か？」と毎回迷う時間が削減される

付録：連絡先・資料リンク

- プロジェクト管理: ○○（開発側）
- 業務窓口: ○○（業務側）
- 資料フォルダ: [SharePoint/Teamsリンク]
- トピック定義シート: [リンク]
- 定例会議事録: [リンク]

ご協力、よろしくお願いいたします！

