

農業経営支援サイト AgriweB における生成 AI × RAG の PoC と 本番導入の実践報告

A Practical Report on PoC and Production Deployment of Generative AI × RAG for AgriweB

齋 藤 広 樹

要 約 本稿では、農業経営支援サイト AgriweB に生成 AI と RAG を適用した対話型 AI「経営アシスト AI」について、PoC から本番導入までの設計と実装を報告する。曖昧な相談を起点に論点を整理し、出典提示付きで専門コンテンツへ誘導する対話設計（追加質問・類似質問を含む）により、従来手法では難しい課題の具体化と深掘りを支援した。PoC で有効性と運用成立性を確認し、本番では認証連携とセキュアな分離構成、UI 実装により継続提供可能なサービスとして実装した。生成 AI を課題整理の伴走役として位置づけることが、業務やサービス適用においては有効である。

Abstract This paper reports the design and implementation of “Management Assist AI,” a dialogue-based AI system that applies generative AI and Retrieval-Augmented Generation (RAG) to the agricultural business support platform AgriweB, from proof of concept (PoC) to production deployment. Starting from vague consultations, the system organizes key discussion points through dialogue and guides users to expert-supervised content with explicit source citations, including follow-up questions and suggestions for similar questions. Thus, it provides support for the clarification and in-depth exploration of issues that are difficult for conventional methods. Through the PoC, we confirmed the effectiveness of this approach and its operational feasibility. In production, we implemented authentication integration, a secure, isolated architecture, and a user interface to enable the continuous provision of service. The case study indicates that positioning generative AI as a companion for structuring issues—rather than an agent that delivers definitive answers—is effective for practical business and service applications.

1. は じ め に

近年、生成 AI は文章理解・要約・対話といった領域で高い性能を示し、業務やサービスへの適用が進んでいる。特に、大規模言語モデル^{*1}（Large Language Model：以下、LLM）は、ユーザーの曖昧な相談に対して文脈を踏まえた応答や追加質問を生成できる点が特徴であり、従来の検索エンジンや FAQ では対応が難しかった課題に対する新たな手段として注目されている。

一方で、LLM は学習時点以降の情報や特定ドメインの専門知識に対する制約を有しており、根拠を示さない回答は業務サービスとしての説明責任を満たしにくい。こうした課題に対し、LLM の外部の信頼できる情報源を検索し、その結果を基に回答を生成する検索拡張生成^{*2}（Retrieval-Augmented Generation：以下、RAG）は、知識の最新性と回答の信頼性を両立する手法として位置づけられる。

農業分野においては、経営環境や制度が複雑かつ変化しやすく、農業経営者が抱える課題も

明確に言語化されていない場合が多い。農業経営支援サイト AgriweB^{*3}には、専門家が監修した多様なコンテンツが蓄積されている一方で、情報量の多さから、ユーザーが自らの状況に即した情報へ到達するまでに一定の負荷を要するという課題があった。

本稿では、AgriweBにおいて、生成 AI と RAG を組み合わせた対話型 AI の概念実証 (Proof of Concept: 以下、PoC) を通じて業務サービスとしての有効性を検証し、「経営アシスト AI」^{*4} を本番導入した取り組みを報告する。特に、課題が曖昧なユーザーに対する対話設計、専門コンテンツとの接続方法、出典提示による信頼性担保、および業務サービスとして成立させるためのシステム構成や運用設計に焦点を当てる。まず 2 章で背景と課題認識を整理して、3 章から 6 章で PoC の目的、設計、実装、検証結果を述べる。7 章では本番導入の構成および PoC の結果を基に実施した対応策について述べる。

なお、本稿では、LLM の性能比較や学習手法の詳細な比較検証、農業制度・補助金制度等の内容解説、UI/UX 設計の一般原則の体系的議論、セキュリティの実装詳細や設定値については記載の対象外とする。

2. 背景と課題認識

本章では、農業経営支援における情報探索の現状と、本プロジェクトが解決すべき課題について整理する。まず、農業経営者が直面している情報収集の困難さと、既存の支援サイトである AgriweB が抱えていた課題を述べる。次に、従来の FAQ や検索手法の限界を指摘する。その上で、生成 AI および RAG を採用した技術的背景と妥当性について、関連研究や他分野の事例を交えて整理し、本稿で扱う対話型 AI^{*5} の位置づけと方向性を示す。

2.1 農業経営における情報探索の課題

農業経営は、労働力不足、市場環境の変化、補助金・制度の多様化などにより、経営判断に必要な情報が増加し、意思決定の難度が高まっている。そのため、経営者が必要情報へ迅速に到達し、状況に応じて判断の質を高める支援の重要性が増している。

しかし現場では、情報が複数の情報源に散在し、ユーザーが「どこを見ればよいかわからない」「情報量が多く読み切れない」「自分の状況に合う判断材料に到達できない」といった状態に陥りやすい。特に課題が曖昧な段階では、検索する文章自体が定まらず、必要情報への入口を作りにくい。

2.2 AgriweB の概要・ユーザー特性と課題

AgriweB は、農林中央金庫グループの株式会社 AgriweB (以下、AgriweB 社) が運営する農業経営支援サイトである。AgriweB では、農業経営に関する基礎的な事項を、ユーザーに理解しやすく体系的に整理した情報コンテンツ「基礎知識」を掲載している。さらに、農業経営者の課題意識に寄り添い、経営や技術、制度等に関する知見を分かりやすく解説した記事コンテンツ「コラム」などもある。AgriweB の会員登録者は 3 万名を超えている^[1]。ユーザーは経営者、現場担当者、新規就農者など多様であり、抱える課題も経営計画、販路開拓、資金計画など幅広い。

このようなユーザーに対して AgriweB は専門家が監修した多様なコンテンツを提供しているが、課題は大きく二つに整理できる。

第一に、サイト構成およびコンテンツ閲覧に関する課題である。AgriweB には専門的な情報が蓄積されている一方で、情報量が多く、テキスト中心の構成であるため、ユーザーが必要な情報を探し出し、要点を把握するまでに一定の時間と負荷を要する。特に、ユーザーが複数の記事を比較しながら自身の状況に合う情報を判断する場合、情報探索の負担が大きくなる。

第二に、ユーザー自身の課題認識に関する課題である。農業経営に関する悩みは、経営計画、販路開拓、資金計画など複数の論点が絡み合うことが多く、ユーザーが最初から自身の課題を明確に言語化できるとは限らない。そのため、検索語や質問文を適切に作ることが難しく、結果として必要な情報への入口を見つけにくい場合がある。

以上のように、AgriweB における課題は、専門コンテンツを十分に活用するための情報探索上の課題と、ユーザーが自身の課題を明確化しにくいという課題の二面から捉えられる。本 PoC では、これらの課題に対して、生成 AI と RAG を組み合わせることで、ユーザーの曖昧な相談を起点に論点を整理し、関連する専門コンテンツへ誘導することを目指した。

2.3 先行手法 (FAQ/検索) における課題と、生成 AI 活用に向けた知識付与手法

近年、自治体等において、住民向け FAQ チャットボットの導入や活用が進んでいる^[2]。FAQ 型 AI^{*6} や検索型 AI^{*7} は、定型적인問い合わせ対応には有効である一方、ユーザーが課題を適切に言語化できることを前提とする。そのため、課題が曖昧な相談では、質問が定まらず、対話が単発で終わりやすい。

また、生成 AI に業務ドメイン固有の知識を適切に組み込む手法としては、モデルの再学習^{*8} (Fine-Tuning 以下、ファインチューニング) と LLM が有していない外部知識を活用する RAG が主要な選択肢となる。Microsoft が 2024 年に発表した研究では、複数の LLM と知識集約タスク^{*9} を用いた比較実験により、新規知識や更新頻度の高い情報に対してはファインチューニングよりも RAG が一貫して高い性能を示すことが報告されている^[3]。

さらに、医療分野では富士通の「AI 問診支援システム^[4]」のように、RAG を活用して医療ガイドラインや論文を根拠とした回答を生成する事例がある。金融分野でも三菱 UFJ 信託銀行の「AI 規制 FAQ^[5]」のように、法令・規制文書を検索し根拠付きで FAQ 回答を生成するシステムが実用化されている。

このような背景から、AgriweB においても、コンテンツ閲覧時の課題を解決するために、生成 AI と RAG を組み合わせた対話型 AI を採用することを検討して PoC を実施した。従来の検索型 AI や FAQ 型 AI は、ユーザーが明確な質問文や検索語を入力することを前提とするが、農業経営に関する相談では、ユーザー自身が課題を十分に整理できていない場合がある。そのため、ユーザーの曖昧な相談を起点に、対話を通じて論点を整理し、必要な情報へ段階的に導くために対話型 AI が有効であると考えた。

3. 目的と仮説 (PoC)

本 PoC は 2023 年 12 月から 2024 年 3 月に実施した。PoC の目的は、農業経営者からの相談に対して「課題の具体化」および「課題の深掘り」を支援する対話体験を成立させることである。そのため、AgriweB の情報と LLM を組み合わせ、ユーザーの曖昧な相談から関連情報提示や対話を通じて深掘りできる仕組みを構築した。具体的には、AgriweB のサイト上のコンテンツ「基礎知識」「コラム」と、LLM のモデルである GPT-3.5^{*10} の一般知識を併用した。

PoC では、ユーザーが「課題が分からない/言語化できない」状態から出発することを前提に、対話を通じて課題を深掘りすることで明確化し、必要情報へ導く価値を狙った。従来の検索やFAQが「質問が明確であること」を暗黙の前提にしていたのに対して、本設計は「質問が曖昧であること」を入力条件として取り込んだ。

その上で、以下の三つを組み合わせることで、従来のFAQ型AIでは困難であった「課題の具体化」を実現できるという仮説を立てた。

- 1) ペルソナ設計とシナリオ設計
- 2) RAGを用いたAgriweB固有情報の参照
- 3) 追加質問・類似質問の生成によるフォローアップ

4. 提供価値とサービス設計 (PoC)

本章では、3章で立てた仮説に基づき、PoCにおいて想定したユーザー体験と提供価値、およびそのサービス設計について述べる。本PoCでは、農業経営者が抱える悩みとして、新たな販路開拓や事業拡大、計画書作成、事業承継など多岐にわたる課題が存在することを踏まえ、ペルソナ設計およびそれに基づくシナリオ設計を行った。具体的には、「事業・販路拡大を考えている農業経営者」をペルソナとして設定し、課題が曖昧な状態から相談を開始するユーザーを想定している。

農業経営という専門性の高いドメインにおいて、本システムでは、ユーザーの相談に対して対話を通じて課題を深掘りしつつ、RAGにより参照したAgriweBコンテンツを出典として提示する構成とした。さらに、追加質問や類似質問の提示により、ユーザーの思考を継続的に支援する仕組みを備えている。これらの設計背景と、PoCで検証対象とした機能および導線について整理する。

4.1 ユーザー課題と提供価値 (PoCで狙った価値)

農業経営の分野では、事業計画・資金計画・販路開拓といった経営情報に加え、栽培技術、病虫害、気象、市況など生産・販売判断に関わる多様な専門情報を扱う必要があり、さらに制度や市場環境の変化に伴い内容が更新される領域においては、モデル内部への知識固定化よりも、信頼性のある既存コンテンツを柔軟に参照・活用できる構成が適している。そのため、本事例ではAgriweBが有する専門コンテンツを最大限活用するため、RAGを中核とした対話型AIによる生成AIアーキテクチャを採用した。

対話を通じてユーザーの課題やニーズを具体化し、信頼性のある専門コンテンツへ誘導することで、FAQ型AI/検索型AIなどの従来型サービスでは実現するのが難しかった価値を提供する。

4.2 PoCで検証した機能と導線 (PoC範囲)

PoCでは、生成AIとRAGを組み合わせた対話型AIを実装した。ユーザーの相談を受けて課題を深掘りし、参照したAgriweBコンテンツを出典として提示し、フォローアップとして追加質問・類似質問を提示する。主な機能は以下の表1の通りである。

表 1 各アプローチにおける主な機能

No	アプローチ	機能	説明	評価観点
1	ユーザーの曖昧な悩みを AI が段階的に具体化する.	チャットボット形式 (対話 UI)	ユーザーの入力に対して追加の質問を返し、課題の明確化を支援する.	対話が継続するか/ユーザーが課題を明確化できるか
2	回答の根拠を明示し、信頼性を担保する.	RAG (検索 + 生成) / 出典提示 (URL 提示)	AgriweB コンテンツを検索し、参照箇所および記事 URL を回答と併せて提示する.	出典が常に提示されるか/参照元と回答の整合が取れているか
3	課題を多角的に掘り下げ、次の行動につなげる.	追加質問・類似質問	回答後に、深掘りのための追加質問・類似質問を提示し、次の問いを作る.	深掘りが促進されるか/導線 (クリック等) が機能するか

5. システム構成と実装 (PoC)

本章では、4 章で示したサービス設計を実現するための技術的アプローチと実装内容について述べる。PoC 環境として採用したクラウド基盤および LLM の選定、ならびに対話制御を担うアプリケーション基盤の役割を整理する。特に、本システムの核となる RAG については、検索精度と回答の信頼性を両立させるためのデータ前処理やプロンプト設計、さらにユーザーの課題の深掘りを支援する機能の実装について示す。

5.1 技術的アプローチ (PoC)

システムを構築するために PoC 用に Microsoft Azure (以下、Azure) 環境を利用してプロトタイプ環境を構築した。PoC 環境であり一般公開しないためアクセス制限 (IP 制限) をかけて、セキュリティを確保した。また、LLM は Azure のリソースの Azure OpenAI Service の GPT-3.5 を使用した。

5.1.1 チャットボット形式

対話型 AI の UI および会話制御には、BIPROGY 株式会社 (以下、BIPROGY) が提供している対話型 AI アプリケーション基盤 (RinzaTalk^{*11}) を利用し、チャットボット形式のプロトタイプを構築した。RinzaTalk を利用することで、UI 開発や会話制御機能の個別実装に工数・コストをかけることなく、RAG による回答生成や追加質問・類似質問生成といった検証対象機能の実装・テストに注力することができた。

5.1.2 RAG 設計 (データ前処理, ハイブリッド検索, プロンプト設計)

RAG に投入するコンテンツのデータ前処理では、AgriweB のコンテンツを分割・タグ付けし、検索精度を高めるためのハイブリッド検索^{*12}を実装した。プロンプト設計では、AI に「農業経営者に寄り添うアドバイザー」としての役割を与えて、AI が誤った回答をしないように命令することで、出典提示や回答形式を制御した。これにより、回答の信頼性と一貫性を担保し、ユーザーが根拠を確認できるよう設計した。

図 1 に示す回答生成フローは、ユーザー入力における質問から RAG による検索、生成 AI による回答生成、出典提示までの一連の流れを示している。表 2 は図 1 の各処理ステップを対

応付けて説明したものである。AgriweB の専門コンテンツを検索可能な形式にするため、AgriweB 社から受領したコンテンツに対して前処理を実施した。具体的には、HTML タグ等のメタデータを除去し、本文テキストを検索投入用に整形した。メタデータを除去した検索用データを 800 文字単位でチャンク化^{*13}し、本文テキストに対して Embedding^{*14}を生成し、検索インデックスへ格納した。RAG におけるベクトル検索用の Embedding 生成には、Azure OpenAI Service が提供する text-embedding-ada-002 を用いた。これにより、ユーザー入力との意味的類似度に基づく検索ができるようになり、キーワード検索と組み合わせたハイブリッド検索の一要素として利用した。

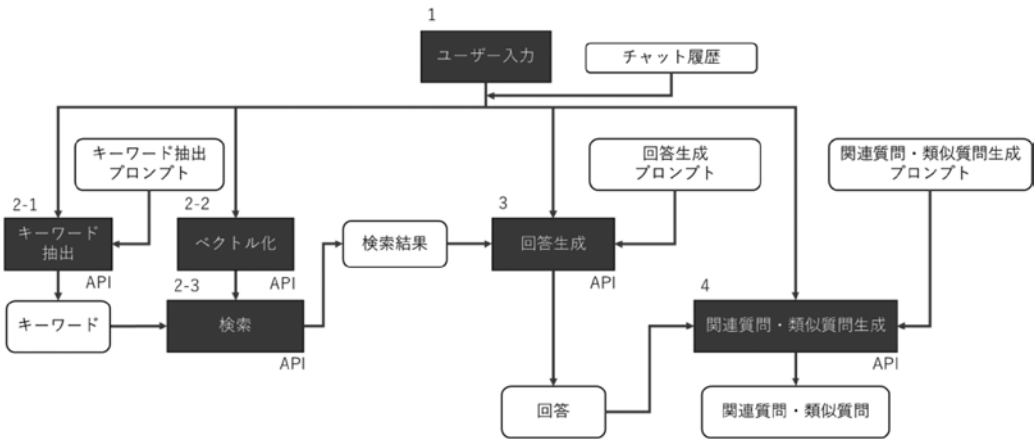


図1 回答生成フロー

表2 回答フローの説明

No	機能	概要
1	ユーザー入力	ユーザーがチャットアプリに質問文を入力し送信する。
2-1	キーワード抽出	No1 のユーザー入力からキーワードを抽出する。
2-2	ベクトル化 ^{*15}	No1 のユーザー入力をベクトル化する。
2-3	検索	No2-1 で抽出したキーワードと No2-2 でベクトル化した値をインプットにハイブリッド検索し、検索結果を出典提示する。
3	回答生成	No2-3 で検索した結果と No1 のユーザー入力から回答を生成する。
4	追加質問・類似質問生成	No3 で生成した回答と No1 のユーザー入力から追加質問と類似質問を生成する。

5.1.3 出典提示

出典提示機能では、RAG の検索結果のヒット時に記事内の文章だけでなく、記事の URL もセットで返すことで参考文献として URL を提示することができる。

一方で、生成 AI の回答には、ハルシネーション^{*16} のリスクがある。本システムではプロンプト設計や出力制御により、ハルシネーションが発生しづらい状況を作った。RAG による出典提示を必須とし、回答ごとに参照元記事を明示することで、ユーザーが一次情報を直接確認できるようにし、信頼性を担保した。また、ユーザーが出典を直接確認できる UI 設計とし、

説明責任を果たす仕組みを導入した。図2の枠内が出典提示機能であり、参考文献というかたちでユーザーが根拠情報を直接確認できる設計となっている。これにより、AIの回答の信頼性が担保される。

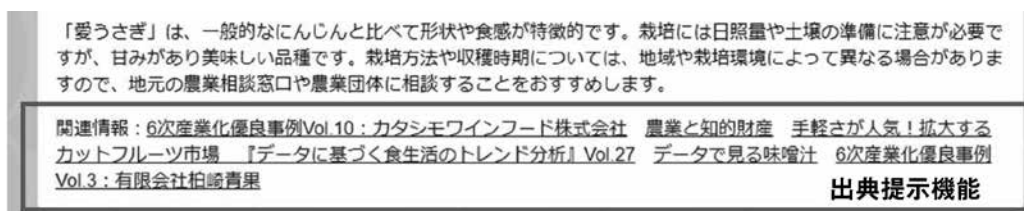


図2 出典提示

5.1.4 追加質問・類似質問

追加質問・類似質問機能はユーザーが課題を深掘りできるよう、AIが追加で質問する追加質問機能や、類似の質問を提示する類似質問機能を実装した。AIは対話履歴を参照し、ユーザーの状況に応じて適切なフォローアップを行う。これにより、ユーザーが自らの課題を解決できる体験を設計した。図3の中央部に表示されている追加質問例が、ユーザーの課題の深掘りを促す設計となっている。

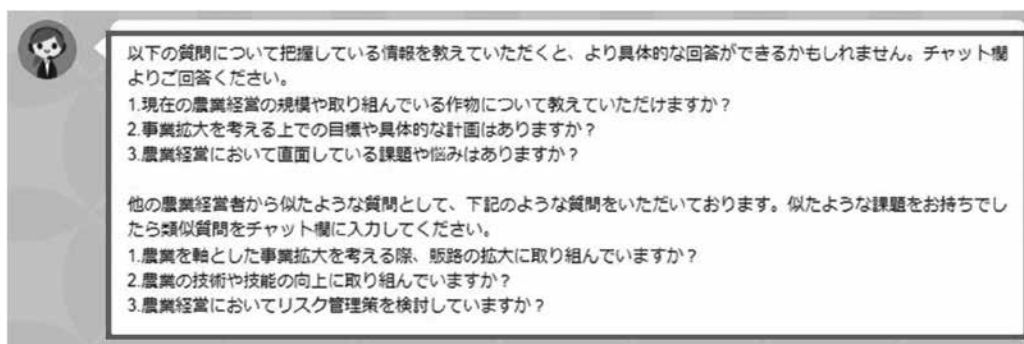


図3 追加質問・類似質問

6. 検証と評価

本章では、AgriweBにおける生成 AI × RAG 活用の PoC を通じて実施した検証と評価について述べる。本 PoC では、定量的な性能評価ではなく、本番導入可否を判断するための定性的な観点での検証を主眼とした。

まず、6.1 節では検証方法について整理し、シナリオ検証およびフィールド検証の観点から評価の枠組みを示す。次に、6.2 節では各検証において得られた結果について述べ、対話体験の有効性および課題を明らかにする。最後に、6.3 節ではこれらの結果を総合的に評価し、本番導入に向けた課題と位置づけについて整理する。

6.1 検証方法

本 PoC における評価は、効果測定を目的とした定量評価ではなく、本番導入可否を判断す

るための論点整理として実施した。評価は、関係者による定例の打ち合わせの中で、都度デモ環境を用いた試行とフィードバックを繰り返す形で進めた。なお、本 PoC では定性評価における体系的な記録は作成していない。このため、本章で述べる結果は、当時の関係者間の合意に基づく定性的判断である点を本稿の制約として明記する。生成 AI および RAG を農業経営支援サービス「AgriweB」に適用するにあたり、実利用を想定した対話体験の妥当性を確認することを主眼に検証を行った。検証は以下の表 3 のとおり、二つの内容で実施した。

表 3 検証方法

検証内容	確認方法	期待値
シナリオ検証/ 本番導入を見据えた評価	事前に定義したユーザーペルソナおよび相談シナリオに基づき、想定される質問に対する応答内容、ならびに追加質問・類似質問による対話の連続性を確認した。	ユーザーの課題具体化につながることで出典提示と回答内容の整合が取れていること ハルシネーションが発生せず、ユーザーが理解できる日本語で回答が返ってくること
フィールド検証	RAG 構成のもと、実サービスに近い UI および導線で試行を行い、関係者によるレビューを通じて定性的な評価を実施した。	関係者が問題なくサービスを利用できること

6.2 検証結果

本節では、6.1 節で示した検証方法に基づき、シナリオ検証およびフィールド検証の結果について述べる。各検証において、ユーザーの課題具体化支援の有効性、回答内容の信頼性および出典との整合性、ならびに導線や利用体験上の課題について確認した結果を整理する。

6.2.1 シナリオ検証/本番導入を見据えた評価

シナリオ検証/本番導入を見据えた評価を通じて、ユーザーが抱える漠然とした悩みや関心事に対して、生成 AI が対話を通じて論点を整理し、経営課題としての切り口を提示できることを確認した。特に、単発の Q&A に留まらず、追加質問や類似質問を提示することで、ユーザーの課題を深掘りする導線が機能する点が確認された。

一方で、回答内容の具体性は、検索対象データの粒度や前処理状況に大きく影響を受けることが明らかとなり、RAG に投入するデータ設計の重要性が再認識された。検索対象データを一律に増やせば回答品質が向上するわけではないことも分かった。当初はコンテンツの一つである「一問一答」を検索対象データに含め「基礎知識」「コラム」「一問一答」の三つのコンテンツを使用して検証を行っていた。しかし、「一問一答」の情報は具体的かつピンポイントな回答を含むため、本 PoC が目的とした「曖昧な相談を起点に論点を広げる」用途では、検索結果のノイズとなる可能性があった。本件への対応として、「一問一答」の情報を検索対象から除外して検証したところ、回答が改善した。このことから、RAG に投入するデータは、単に量を確保するだけでなく、想定するユーザー体験や対話シナリオに応じて、粒度や対象範囲を設計することが重要であると考えられる。

また、生成 AI の回答には、参照した AgriweB コンテンツの出典を明示する構成とした。この結果、回答内容が AgriweB の「基礎知識」や「コラム」と自然に接続され、出典提示と

回答内容の整合が取れていることと、既存サービスの信頼性を損なうことなく AI を補助的に活用できることを確認した。

これにより、生成 AI が独立した知識提供者として振る舞うのではなく、AgriweB が保有する情報資産への入り口として機能する構成が有効であることが示唆された。

また、ハルシネーションの発生を抑えるプロンプト設計により、検証範囲においてハルシネーションは確認されず、ユーザーが理解できる日本語で回答が返ることが確認された。

そして、PoC 期間中に設計したプロンプト設計、回答生成ロジック、出典提示ならびに追加質問・類似質問生成の仕組みについては、基本構造を維持したまま本番環境へ適用できるとの判断に至った。

6.2.2 フィールド検証

フィールド検証においては、実際の利用を想定した試行を通じて関係者からフィードバックを収集した。その結果、回答内容の充実度や文章量については概ね好評であり、機能として利用したいという意見が得られるなど、一定の有用性が確認された。

一方で、回答生成に時間を要する点や、追加質問・類似質問の位置づけがユーザーにとって分かりにくく、次にどのような行動を取ればよいかが不明確となるケースがあることが指摘された。また、フォローアップ質問の内容自体には納得感があるものの、それを受けてどのように入力すればよいか、またその結果どのような応答が得られるかがユーザーに直感的に理解されにくいという課題が確認された。

加えて、生成 AI の特性上、不適切な回答の生成やそれに伴うリスクに対する懸念も指摘されており、過度な信頼や誤解を防ぐための利用上の説明やガイドの必要性が認識された。

以上より、生成 AI に対して過度な万能性を期待させないための説明や、ユーザーの期待値を適切にコントロールする設計が、本番導入に向けた重要な課題として整理された。

6.3 評価のまとめ

本 PoC で実施したシナリオ検証およびフィールド検証の結果を踏まえ、以下の通り評価する。まず、ユーザーの曖昧な相談に対して対話を通じて課題を具体化し、論点を整理する機能については、回答内容の充実度や文章量に対する評価からも、一定の有効性が確認された。

また、出典提示と回答内容の整合性についても問題なく、既存の AgriweB コンテンツへの導線として機能することが確認された。

一方で、追加質問および類似質問の位置づけがユーザーにとって分かりにくく、次の行動が不明確となる点や、回答生成までの待ち時間が UX 上の阻害要因となり得る点など、インターフェースおよび体験設計に関する課題が明らかとなった。さらに、生成 AI の特性上、不適切回答への懸念が残ることから、ユーザーに過度な期待や誤解を与えないための説明やガイドの整備を含め、利用時の期待値コントロールが不可欠である。

以上の結果より、本 PoC の構成は、農業経営支援において、「正解を提示する AI」ではなく、「課題を整理し、次の行動につなげる伴走役」として、一定の有効性を有することが確認された。よって、PoC で抽出した課題に対して改善を講じること（7.2.5 項で詳述）を前提に、本番導入を進めることにした。

7. 本番導入

本章では、PoC の検証結果および課題を踏まえ、AgriweBにおける生成 AI × RAG の本番導入に向けた取り組みについて述べる。6 章で整理した評価結果をもとに、サービスとして継続的に提供するために必要となるシステム構成、認証連携、UI 設計、および運用面の対応について整理し、PoC から本番環境へ移行する上での設計方針と実装内容を示す。

7.1 本番導入における位置づけ

PoC を通じて、AgriweB に蓄積された情報と生成 AI を組み合わせることで、ユーザーの課題整理や思考支援に一定の有用性が確認できた。一方、本番導入では、サービスとして継続的に提供することを前提に、セキュリティや運用を含めた対応が求められる。PoC で検証した機能や設計を最大限活用しつつ、本番利用に向けていくつかの対応を行った。本番導入する上でのシステム名称は「経営アシスト AI」とした。

7.2 本番導入において実施した主な対応

本節では、PoC で得た検証結果および抽出した課題を踏まえ、本番導入にあたり実施した主な対応について述べる。具体的には、システム構成や認証連携、UI 設計、および運用面の観点から、本番環境として安定的かつ継続的にサービス提供を行うために実施した対応内容を整理する。

7.2.1 本番導入・運用における役割分担

本番導入にあたっては、システム開発および運用を円滑に進めるため、関係者間で役割分担および責任範囲を明確に定義した。本システムにおける主な役割分担は以下の通りである。

1) BIPROGY

BIPROGY は、経営アシスト AI を構成する生成 AI 機能の設計・開発、Azure 上での基盤構築、UI 実装、ならびに RAG やプロンプトを含む回答生成ロジックの実装を担当した。また、本番環境の構築にあたっては、運用実績のある構成を迅速に適用するため、生成 AI 導入のための参照アーキテクチャ/環境構築テンプレート（Azure OpenAI Service スターターセット Plus^{*17} 以下、スターターセット）を利用した。

2) AgriweB の開発ベンダー

AgriweB の開発ベンダーは、経営アシスト AI との認証連携の開発および AgriweB-経営アシスト AI の画面遷移の設計を担当した。

3) AgriweB 社

AgriweB 社は、経営アシスト AI のテストと、RAG へのコンテンツデータ投入作業を含む本番導入後の運用作業を担当した。

PoC での協業関係を踏まえつつ、本番導入・運用においては役割分担を明確化することで、安定的な開発および運用を実現した。図 4 に本番導入の開発スケジュールを記載する。

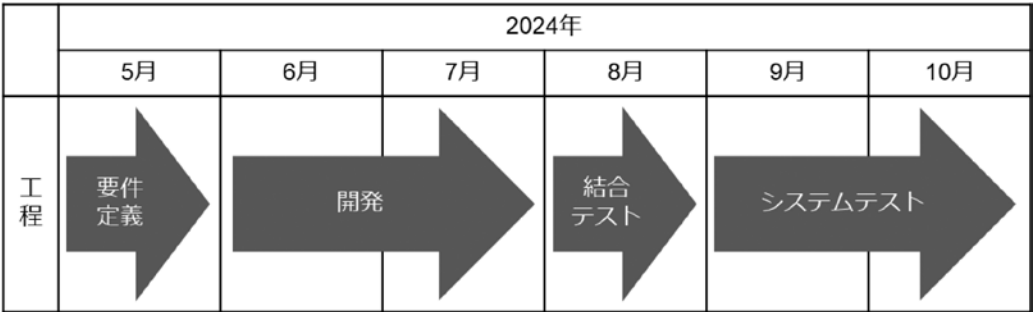


図4 開発スケジュール

7.2.2 全体構成（AgriweB との分離、接続方式、認証連携）

本番導入では、AgriweB と経営アシスト AI を分離した構成とし、セキュリティと運用性を両立させている。AgriweB 側は会員管理・認証・専門コンテンツの提供を担い、経営アシスト AI 側は Azure 上に構築し、RAG による回答生成を含むチャットボット形式のチャットアプリ UI を担当する。両者は API 連携により接続され、ユーザー認証情報を利用する設計とした。ネットワーク分離や認証連携により、外部公開サービスとしての信頼性と責任分界を担保している（図5）。

AgriweB と生成 AI サービスは物理的・論理的に分離されており、API 連携によって認証情報のみを安全に受け渡している。分離したことにより、経営アシスト AI のアプリのバージョンアップを行い、デプロイを実施する場合には AgriweB 本体をサービス停止することなく、バージョンアップを実施することができる。

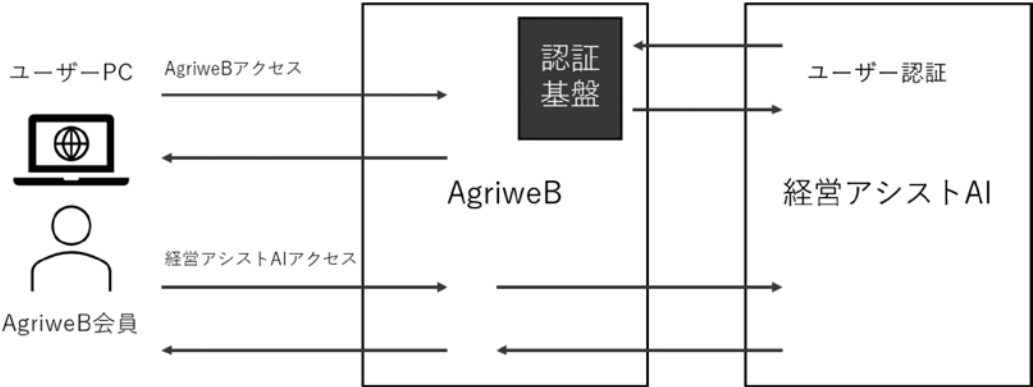


図5 システム概要図

本番開発では、PoC 期間中に検証・調整した以下の要素を基本的にそのまま活用した。

- 1) 回答生成プロセス
- 2) RAG による検索・生成方式
- 3) 追加質問・類似質問生成機能

これにより、PoC で確認された有用性を損なうことなく、本番環境へ移行することができた。また、プロンプト設計の大幅な再設計を行わない方針としたことで、本番導入後の品質変動り

スクを抑制している。

本番導入にあたっては、AgriweB の会員認証基盤と連携したユーザー認証を実装した。これにより、経営アシスト AI では新たなユーザー登録が不要となり、既存の AgriweB 認証方式を用いて、正規の会員のみがサービスを利用できる構成とした。

この設計により、以下の4点を実現した。

- 1) 認証情報の重複管理を回避
- 2) 不正利用リスクの低減
- 3) 既存サービスとの一貫性確保
- 4) ユーザーの別サービスの会員登録作業の省略

7.2.3 UI画面の本番実装

PoC では主に機能検証を目的としたデモ画面を用いていたが、本番導入に際しては、ユーザーおよび管理者の双方に配慮した図6のようなUI画面を構築した。具体的には、スターターセットに含まれる標準UIをベースとし、AgriweB のユーザーの利用を想定したカスタマイズを行った。これにより、UI 開発工数を抑制しつつ、運用実績のある構成を採用することで、安定性を重視した実装とした。



図6 経営アシスト AI のチャット画面

7.2.4 専門家相談の導線の設計

追加質問・類似質問によってユーザーのフォローアップを行った後に具体化した課題を専門家に相談する、という導線を設計した。課題を具体化した後、ユーザーはその課題を解決するためにそのまま対話を繰り返すこともできるし、AgriweB が提供している「経営アシストチャット」機能で専門家に相談することもできる。本番導入の際にユーザーが簡単にこの機能を利用できるよう、UI 上にその機能に遷移できるボタンを作成した。

7.2.5 PoC で抽出したフィードバックと本番導入での対応

PoC では、対話体験の受容性に加え、サービス提供に不可欠な運用上の課題を抽出した。本番導入にあたっては、PoC で得られた示唆を踏まえ、課題ごとに対応方針を整理した。表 4 に、PoC で抽出した課題、原因仮説、本番導入での対応を示す。

表 4 PoC で抽出した課題と本番での対応

PoC で抽出した課題	原因仮説	本番導入での対応
追加質問・類似質問の位置づけがユーザーに伝わりにくく、次の行動が不明確になる点が課題として挙げられた。	UI 上で「なぜ次の質問が提示されるか」の意図が説明されていない。	サービス案内の画面に追加質問・類似質問の概要を記載した。
回答待ち時間が UX 上の阻害要因となり得る。	回答生成中に何も表示されないため、ユーザーからするとシステムが動いているかどうか不安になる。	回答生成中はローディングアニメーションを表示して、ユーザー体験を阻害しない工夫を施した。
不適切回答が懸念される。	LLM の性質上、入力に応じて不適切回答の生成が起り得る。	回答を鵜呑みにしない様に利用規約に記載した。

7.3 本章のまとめ

本章では、AgriweB × 生成 AI システムの本番導入にあたり、実施した具体的な開発・運用対応について述べた。PoC で検証した成果を最大限活用しつつ、認証連携、セキュアな Azure 基盤構築、UI 実装といった本番特有の課題に対応することで、サービスとして提供できる生成 AI システムを実現した。

8. お わ り に

本稿では、農業経営支援サイト AgriweB における生成 AI × RAG 活用の実践と設計上の勘所について、PoC から本番導入までの知見を整理した。課題が漠然としやすいユーザーに対し、対話型 AI による課題の具体化、専門コンテンツへの誘導、出典提示による信頼性担保、フォローアップ導線の設計、セキュアな運用構成など、業務適用に不可欠な要素を明らかにした。

最後に、本取り組みの関係者、執筆にあたって指導、協力頂いた方々に深く感謝するとともに御礼して本稿の結びとする。

-
- * 1 大規模言語モデル：大量のテキストデータを用いて学習し、自然言語の理解や生成を行う深層学習モデル。
 - * 2 検索拡張生成：外部データを検索し、その検索結果を入力として大規模言語モデルによるテキスト生成を行う手法
 - * 3 AgriweB：農業経営者向けに、経営・技術・制度等に関する情報を提供する農業経営支援サイト。
 - * 4 経営アシスト AI：農業経営支援サイト AgriweB において提供される、大規模言語モデルを用いて農業経営に関する情報提供や意思決定支援を行う生成 AI 機能。
 - * 5 対話型 AI：ユーザーとの自然言語による対話を通じて、質問応答や情報提供を行う AI システム。
 - * 6 FAQ 型 AI：事前に蓄積された FAQ や文書情報を基に、ユーザーからの質問に対して自動で回答を生成・提示する AI システム。

- * 7 検索型 AI：ユーザーの入力に応じて既存の文書やデータベースを検索し、該当する情報を提示することに特化した AI システム。
- * 8 再学習：事前に学習済みのモデルに対し、特定の目的やドメインのデータを用いて追加学習を行い、性能や挙動を調整する手法。
- * 9 知識集約タスク：外部知識や専門的情報を参照・統合しなければ適切な回答や判断が困難なタスク。
- * 10 GPT-3.5：OpenAI が開発した大規模言語モデルの一種で、文章生成や質問応答などの自然言語処理を行うモデル。PoC 当時は最新のモデルであった。
- * 11 RinzaTalk：BIPROGY が提供する対話型 AI ソリューションで、大規模言語モデルを活用し、業務における質問応答や情報提供を支援する。
- * 12 ハイブリッド検索：キーワード検索（全文検索）と意味ベース検索（ベクトル検索）を組み合わせて、関連性の高い情報を取得する検索手法。
- * 13 チャンク化：長い文章を一定の長さごとに分割し、検索や生成処理に適した単位にすること。
- * 14 Embedding：テキストの意味内容を高次元ベクトルとして表現したもの（埋め込み）。ベクトル間の類似度に基づく意味検索に用いる。
- * 15 ベクトル化：テキストの意味内容を数値ベクトルとして表現し、類似度に基づく検索を可能とする処理。
- * 16 ハルシネーション：生成 AI が、学習データや与えられた情報に基づかず、事実と異なる内容や根拠のない情報を生成してしまう現象。
- * 17 Azure OpenAI Service スターターセット Plus：BIPROGY が提供する、Azure OpenAI Service の利用を前提に、生成 AI 導入の初期検討から環境構築・活用支援までを支援する導入支援サービス。

- 参考文献** [1] 株式会社 Agriweb, 「アグリウェブの使い方」, Agriweb Web サイト,
<https://www.agriweb.jp/merit>
 [2] 渋谷区, 「渋谷区生成 AI チャットボット」, 渋谷区 Web サイト,
https://www.city.shibuya.tokyo.jp/kusei/kocho/ai-chatbot/ai_chatbot.html
 [3] Oded Ovadia, Menachem Brief, Moshik Mishaeli, Oren Elisha, “Fine-Tuning or Retrieval? Comparing Knowledge Injection in LLMs,” Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024), Association for Computational Linguistics, 2024.
 [4] 富士通, 「協業事例インタビュー」, 富士通 Web サイト,
<https://global.fujitsu/ja-jp/about/activities/venture/interview/precision>
 [5] 三菱 UFJ 信託銀行, 「金融市場取引業務における生成 AI を活用した社内問い合わせ対応の自動化実現について」, ニュースリリース (PDF),
https://www.tr.mufg.jp/ippan/release/pdf_mutb/240802_1.pdf

※ 上記参考文献に示した URL のリンク先は、2026 年 6 月 11 日時点での存在を確認。

執筆者紹介 齋藤 広樹 (Hiroki Saito)

2024 年、BIPROGY 株式会社に経験者採用で入社。入社以前は、Web アプリケーションの開発・保守、心臓 MRI 画像を対象とした心臓領域推定 AI モデルの開発、スマートフォンアプリケーション開発などに従事。BIPROGY 入社後は、生成 AI の導入、経理システムの基盤更改、営業店システム開発などを担当している。

