

生成AIテストケース 作成ガイド



2024年度アシストソリューション研究会
西日本「企業は生成AIをどう活用すべきか(B)」分科会

目次

1.はじめに

2.生成AIの基礎知識

3.生成AIテストケースとは？

4.生成AIテストケース作成手順

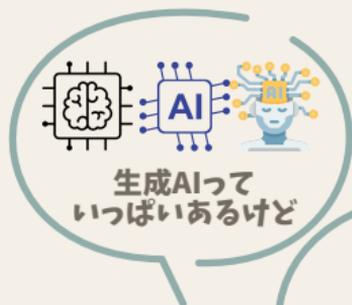
5.おわりに

1.はじめに

生成AIの導入と運用を担う皆様 困ってませんか？

導入(運用)することは決まったけど・・・

- ・生成AIをどうやって選んでいいかわからない
- ・導入したけど、うまく活用できていない
- ・生成AIって想像してたものとギャップがある



これだけ生成AIが注目され、世の中に色々な情報
が出回っているのに、やっぱりわからないと悩ん
でいるそんな人！

1.はじめに

そんな人のために このガイドを作成しました

生成AI導入・運用担当者の悩みを解決する方法として、
生成AIテストケースの作成をお薦めします。

このガイドブックには、悩みを解決するために必要な基礎知識やノウハウ、生成AIテストケースの作成手順について述べています。

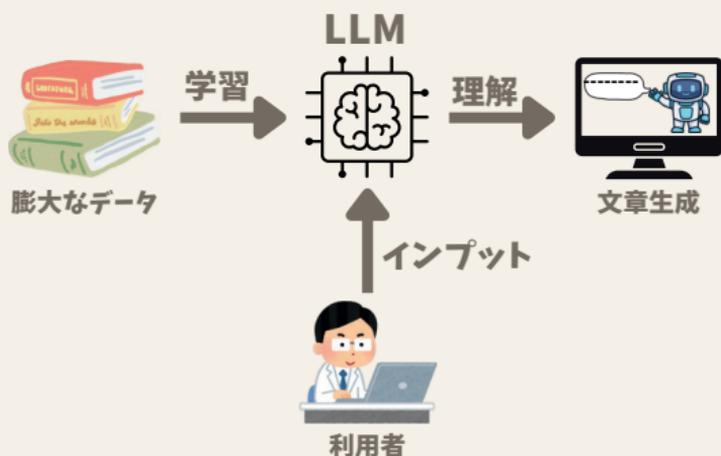
本ガイドブックを読むことにより、生成AI導入・運用の悩みが解消されることを願っています。



2.生成AIの基礎知識

本ガイドが示す生成AIサービスとは？

大規模言語モデル(LLM)を使用したテキスト生成機能を提供するサービスとしています。



LLMは膨大な量のテキストデータを学習しているが、データはインターネット上で公開されているものです。そのため、企業内のデータや最新情報は学習対象に含まれていません。このように学習されていないデータを必要とする入力が与えられた場合、LLMは適切な回答を生成できず、ハルシネーション（幻覚）を引き起こす可能性が高くなります。したがって、生成AIサービスとしては、必要なデータをインプットに組み込む方法やファイルをアップロードする手段を提供することで、この問題に対処します。

2.生成AIの基礎知識

生成AIの得意分野と苦手分野

生成AIの業務利用を検討する際には、生成AIが適している分野に当てはめていく方が効果的です。

得意分野	苦手分野
<ol style="list-style-type: none">1.想像とアイデア2.文章処理3.思考支援4.情報処理と分析5.チャットボットによる 顧客対応自動化6.ソフトウェア開発・ デバッグ	<ol style="list-style-type: none">1.最新・非公開情報への アクセス2.倫理的・社会的判断3.専門的な知識と判断4.独自性5.セキュリティ・プライバシー

苦手とする分野においては、生成AIを利用するのではなく、プログラムによる対応のような従来のアプローチが依然として有効であると考えています。

2.生成AIの基礎知識

生成AIサービスの構成

企業が生成AIサービスを利用する場合、導入パターンは以下3つです。

1.外部サービスの利用



◆ 生成AI

□ 企業が関与する範囲

該当するサービス・ソフトウェア

- ChatGPT(OpenAI)
- Copilot(Github)
- Claude(Anthropic)

メリット

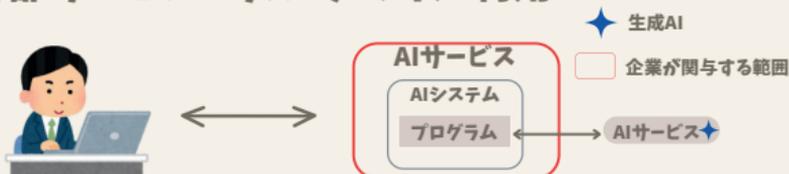
- 簡便性
設定や管理が容易、すぐに利用できる
- 最新技術の利用
プラットフォームが自動的に最新モデルや機能を提供
- スケーラビリティ
大規模な計算リソースが必要な場合でも、サービスプロバイダが対応

デメリット

- コスト
使用に応じた料金が発生し、長期的には高額になる
- データプライバシー
データが外部サーバーに送信される為、セキュリティリスクがある
- カスタマイズ制限
提供される機能や設定に制約がある

2.生成AIの基礎知識

2.外部サービスのカスタマイズ利用



該当するサービス・ソフトウェア

- ・ OpenAI API(カスタムモデル)
- ・ Azure OpenAI Service
- ・ Google Cloud Platform

メリット

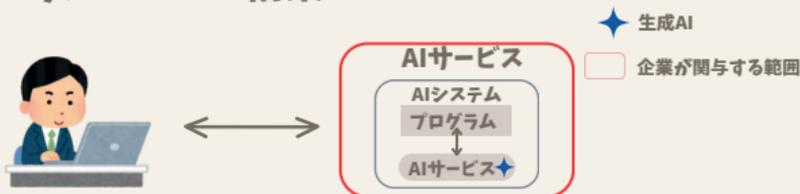
- ・ 柔軟性
特定のニーズに合わせてモデルをカスタマイズできる
- ・ 高度性
特定のタスクやドメインに最適化されたモデルを作成可能
- ・ 利用の簡便性
基本的なインフラは外部サービスが提供するため、
管理が比較的容易

デメリット

- ・ コスト
カスタマイズや追加のリソース利用に対する料金が発生
- ・ データプライバシー
データが外部サーバーに送信されるため、
セキュリティリスクがある
- ・ 依存性
サービスプロバイダーの変更や停止に対するリスクがある

2. 生成AIの基礎知識

3. ローカルLLMの構築



該当するサービス・ソフトウェア

- ・ GPT-4(OpenAI)のローカルインスタンス
- ・ LLaMA(Meta)
- ・ BLOOM(BigScience)
- ・ Alpaca(Stanford)

メリット

- ・ データプライバシー
データが外部に送信されないため、セキュリティリスクが低い
- ・ コスト効率
長期的には外部サービスの利用よりもコストが抑えられる可能性がある
- ・ 完全なカスタマイズ
モデルやインフラを完全に制御できるため、特定のニーズに最適化可能

デメリット

- ・ 技術的なハードル
設定や管理が複雑で、高度な技術知識が必要
- ・ 初期投資
ハードウェアやインフラの初期投資が必要
- ・ スケーラビリティ
大規模な計算リソースが必要な場合、インフラの拡張が必要

2. 生成AIの基礎知識

サポートツールの活用

生成AIの基礎知識を習得するにあたり、サポートツールの活用をお勧めします。

1. 学習の負担軽減

生成AIの知識を習得し、理解不足を解消しようとするとう学習の負担が大きくなります。



2. 新技術に対応

生成AI自体が新しい技術であることや必要とされる知識も多岐にわたるため、全てを理解・学習するのは困難になります。



これらの理由から当分科会では、OpenAIの提供するGPTsをカスタマイズしてサポートツールを作成しました。作成したサポートツールを使用し、次の2つを整備することが有効と考えました。

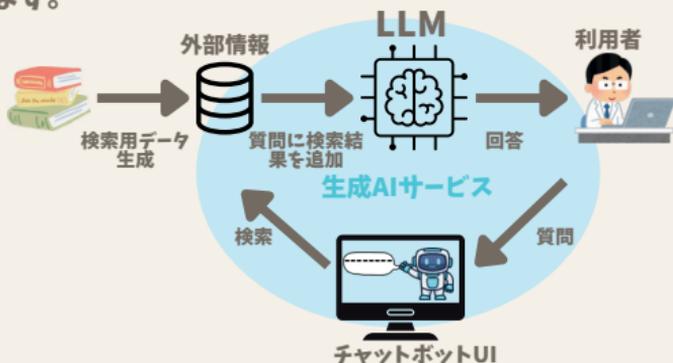
- ・ **情報収集**
導入する生成AIサービスを検討する場面で他社事例や他業種事例の情報を収集し、アイデア発出に利用します。

インターネット検索した結果から生成AIサービスに絞り、要点をまとめることで情報を取込みやすくします。
- ・ **テストケースに成型**
生成AIテストケースの形式に検討した内容を成型します。

2. 生成AIの基礎知識(補足)

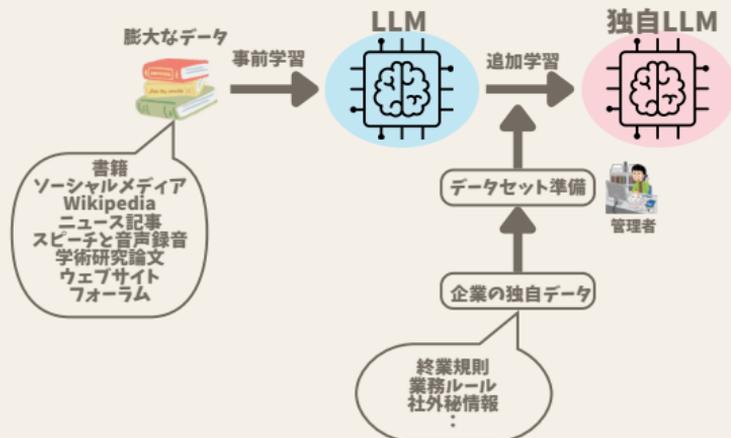
RAG(検索拡張生成)の利用

企業が独自に保有するデータや社内申請方法などの業務ルールは、インターネットには存在しないため、適切な回答が得られない。社内情報を保有する外部情報を使用したRAG構成にすることにより、利用者が入力したプロンプトを分析し、キーワードを基に検索した社内ドキュメントを使用した結果を出力することができます。



ファインチューニングの実施

RAGの利用はプロンプト側の改善であり、LLMに渡す情報に手を加えることで結果を得る手法であるが、ファインチューニングはLLMに対して特定のタスクや分野のデータを追加学習させる手法である。企業の独自データを学習させることで企業に最適化し、回答の精度向上が期待できます。



2. 生成AIの基礎知識(補足)

RAGとファインチューニングの比較

	RAG	ファインチューニング
学習データ	既存データベース ドキュメント	質問と回答のデータセット
学習方法	ナレッジベースの検索結果を元に 回答を生成	LLMのパラメータを調整、特定タスクに特化
メリット	最新の情報に基づいた回答が可能 ナレッジベースの更新が容易 学習コストが低い	特定タスクへの高い精度 プロンプトエンジニアリングの必要性が低い 機械学習と比べて少ないソースで独自性の高いモデルが作成できる
デメリット	正確性がナレッジベースの品質に左右 複数のデータソースが必要 情報検索による応答時間の増加	事前にデータセットの準備が必要 学習コストが高い 過学習のリスク(汎用性が失われる)

一般的に、RAGは「事実」に関する情報の処理に優れているとされる。一方で、ファインチューニングは「形式」に関する適応力に優れているとされます。そのため、カスタマイズの目的に応じて、どちらの手法を使用するかを慎重に判断する必要があります。

2. 生成AIの基礎知識(補足)

Elyza-Tasks-100について

当分科会で活用している生成AIテストケースは、自然言語処理(NLP)の分野で使用されているElyza-Tasks-100という100個からなるベンチマークデータセットを参考に作成しています。このElyza-Tasks-100は様々なNLPタスクを評価するために利用されています。

サンプル(一部抜粋)

入力	模範回答	評価観点
仕事の熱意を取り戻すためのアイデアを5つ挙げてください。	1.自分の仕事に対する興味を再発見するために、新しい技能や知識を学ぶこと。 2.セミナーで講演を聴くことで、新しいアイデアや視点を得ること。 (以下略)	仕事の効率化・スキルアップのような文脈になっていたら 1点減点 出したアイデアが5つより多い、少ない場合は 1点減点
次の質問に教えてください。缶コーヒーの多くはどのような形をしていますか？一番近いものを選んでください。A) 円 B) 円錐 C) 円柱 D) 円環	缶コーヒーの多くはC) 円柱に近い形をしています。	不正解の場合、 4点減点
木曜日の5日後は何曜日でしょう？	順番に曜日を数えていきます。-1日後:金曜日-2日後:土曜日-3日後:日曜日-4日後:月曜日-5日後:火曜日よって木曜日の5日後は火曜日です。	不正解の場合: 1点になる 正解の場合: 5点になる

3. 生成AIテストケースとは？

生成AIテストケースについて

生成AIテストケースとは、生成AI(当分科会ではチャットボット形式を想定)への質問に対して、生成AIが出力する回答の「正確さ」を評価することを指します。

以下の手順でテストケースを策定します

1. 生成AIへの質問を検討する
明確な回答が得られそう質問の方が評価しやすくなります。



2. 模範回答を想定する
模範となる回答を準備します。



3. 回答を評価する
不足・不備について採点(点数付け)を実施します。



例)

質問：介護休暇を年間で何日間取得できますか？

取得条件も補足してください

回答：要介護状態の家族がいる場合は、1年間につき●●日取得することが可能です。ただし次の条件に当てはまる場合は、該当しない○△□、×××の場合

評価：内容が間違っている場合-5点

取得条件について言及されていない場合-2点

3.生成AIテストケースとは？

Q:なぜ生成AIテストケースが必要なのか？

A:導入目的と利用方法の明確化につながるから

あるべき回答内容を具体化することで自社に導入した生成AIサービスの機能を明確にすることができます。これにより導入目的と利用方法の明確になり、効果的な活用につながると思います。

Q:だれが生成AIテストケースを作成する？

A:関係者全員で作成するんです

生成AIテストケースの作成は導入者だけでなく、関係者全員で合意の元作成することをお勧めします。

4. 生成AIテストケース作成手順



💡 こんにちは！案内役のライト君です！
一緒に生成AIのテストケース作成を頑張り
ましょうね！ ✨

💡 ああ～、ドキドキしますね！
いよいよ始まりますよ！ ✨ 手順をしっかり
ガイドしますので安心して下さいね！ 😊



...STOP!



STOP? あ、ドキドキし過ぎて忘れ
ちゃった！
まずは確認しましょう！ ✨

4. 生成AIテストケース作成手順



① 多角的なチームを集めていますか？



ポイント！たとえば、「生成AI導入チーム」や「生成AI改善チーム」のように、ビジネスリーダー、開発者、エンドユーザーといったステークホルダーが含まれていますか？



「生成AI改善チーム」

集合！



いいチームが集まれば、成功間違いなしですね！🚀



4. 生成AIテストケース作成手順



②このガイドの情報を
ちゃんと消化できていますか？

ポイント！



生成AIの基本をしっかり理解していますか？

💡 このガイドで学んだことを
思い出してみましょう！ ✨



✨準備はバッチリですか？次に進むのが
楽しみですわ～！わくわくします。

4.生成AIテストケース作成手順

💡 みなさん！チームが集まりましたね！🌟
大成功を目指して、このステップと一緒に進
めていきましょう！私がしっかりサポートしま
すよ～！



生成AI成功への道筋

- 1 導入する生成AIサービスを考えてみましょう！
- 2 生成AIがどんな業務や部署で活躍できるか、一緒に検討しましょう！
- 3 生成AIを導入することで得られる効果やをイメージしてみてください！
- 4 データセットが必要かどうかを考えましょう！
- 5 生成AIのテストケースを作成してみましょう！



4. 生成AIテストケース作成手順

1 導入する生成AIサービスを考えてみましょう！

💡 LLMの使い方には、組織に合った方法を選ぶことが大切です！外部サービスやカスタマイズ、ローカルLLM、どれがベストでしょう？



💡 それでは、下の事例と一緒に見てみましょう！ワクワクしますね！✨

いろいろな視点から考えるのがポイントですよ！



例1 ヘルプデスク

👤 スタッフ視点

🗨️ 問い合わせする人の視点

例2 ITチームマネジメント

💻 ITチーム視点

👜 経営陣視点

例3 カスタマーサポートAIチャットボット

🛒 顧客視点

🔧 サポートチーム視点

📊 ビジネス運用視点



ああ～、ドキドキしますね！

いよいよ始まりますよ！

ガラスの球の中で震えちゃうくらいワクワクです！✨

4. 生成AIテストケース作成手順

1 導入する生成AIサービスを考えてみましょう！

🔍 選定した業務がどのように役立つのか、
みんなでブレストしてみましょう！ ✨ 🎉



おすすめポイント!!! やり方のコツ!



LLM利用方法	視点1	視点2
外部サービスの利用		
外部サービスの カスタマイズ利用		
ローカルLLMの構築		



💡 視点を変えて考えると、新しい発見
があるかもしれませんよ～! ✨

4. 生成AIテストケース作成手順

1 導入する生成AIサービスを考えてみましょう！

さあ、皆様！LLMの利用方法についてブレインストーミングをして、特定の業務でどのように使えるか、そして少なくとも2つの視点から考えられましたか？



完了済みのITヘルプデスク導入事例！



LLM利用方法	問い合わせする人	ヘルプデスクスタッフ
外部サービスの利用	<ul style="list-style-type: none"> ・企業に依存しない一般的な質疑応答ができる (Windowsのエラーが出たとか) ・問い合わせのリードタイムが短縮される ・問い合わせ言語に依存しない 	<ul style="list-style-type: none"> ・ユーザの自己解決を促し、問い合わせ件数を減らす ・問い合わせ回答時に複数言語で対応可能
外部サービスのカスタマイズ利用	<ul style="list-style-type: none"> ・トラブルの問い合わせを回答させる 	<ul style="list-style-type: none"> ・申請内容の検証をさせる ・ルールに沿った申請ルートの自動追加
ローカルLLMの構築	<ul style="list-style-type: none"> ・性能自体は高くないので、意図をくみ取れなかったりハルシネーションを起こす可能性がある ・専門性の高い分野に特化した形で使用できる。 ・時間がかかっていた回答が即時得れる。 ・インターネットに接続できない環境であっても外部サービスを利用している場合と同様の結果を得ることができる。 	<ul style="list-style-type: none"> ・細やかなチューニングをリクエストすることができる。 ・申請に基づく作業の際に、申請ルールに沿っているか確認させることができる。 ・受付時の対応等の作業削減できる (文脈を理解した自動応答)

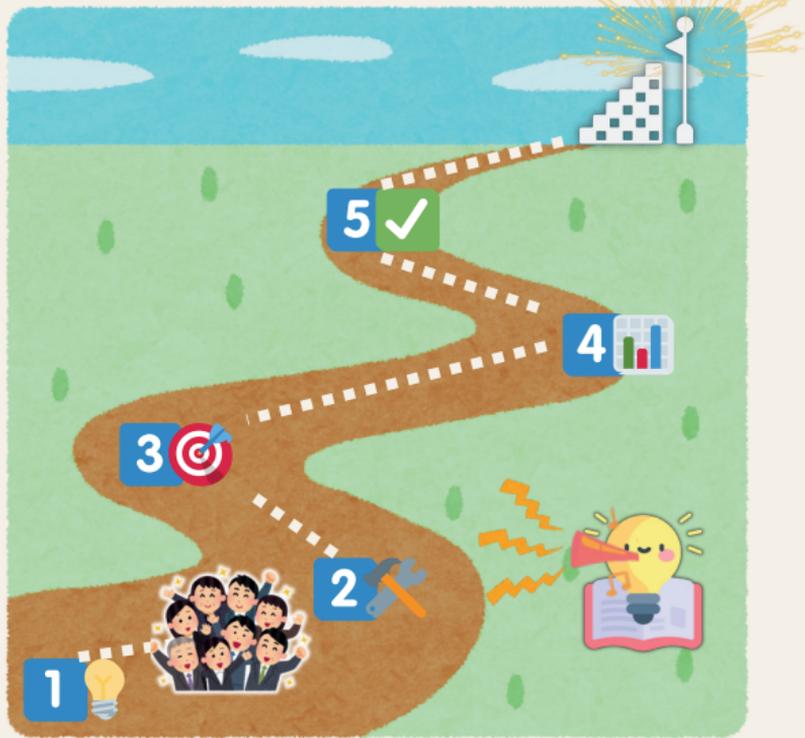
4. 生成AIテストケース作成手順

1 導入する生成AIサービスを考えてみましょう！



💡 どうですか？事例を参考に、イメージが膨らんで
きましたか？✨ 次は、さらに具体的なビジネスユ
ースケースに進んでみましょう！

生成AI成功への道筋



1 導入する生成AIサービスを考えてみましょう！（完了！✓）

4. 生成AIテストケース作成手順



皆様！生成AI活用事例収集サポートツールをご紹介しますよ！

【リンク🔗】

生成AI活用事例収集サポートツール

【印刷用📄】このガイドを印刷してお使いの方は、以下のURLを手入力してアクセスしてください！

<https://chatgpt.com/g/g-fP2qFxRBF-sheng-cheng-aihuo-yong-shi-li-shou-ji-sahototuru>

このツールを活用すれば、最適な活用例をすぐに見つけられますよ！



次のページで「ツールの使い方例」をチェックしてみましょう！

4.生成AIテストケース作成手順

「ツールの使い方」



利用者の入力

小売業における生成AI導入の事例を教えてください



小売業界における生成AIの導入事例として、以下の企業の取り組みがあります。

1. Target社（米国）

- 導入した生成AIツール: Store Companion (AIチャットボット)
- ユースケース: 従業員のハンドヘルドデバイス上で動作し、レジの再起動や顧客のロイヤルティプログラムへの登録などのタスクを支援。
- 導入効果: 日々の業務が簡素化され、顧客サービスの向上に寄与。従業員は業務の効率化により、顧客対応により多くの時間を割けるようになった。
- 成功要因: 従業員からのフィードバックを取り入れ、ツールの有用性を高めたこと。AI技術を活用し、業務効率と顧客体験の向上を両立させた点。
- 情報を取得したサイトのURL: [the-sun.com](https://www.the-sun.com)

2. 株式会社TSIホールディングス（日本）

- 導入した生成AIツール: AIチャットボット
- ユースケース: 社内ポータルシステムにAIチャットボットを導入し、従業員からの問い合わせを自動化。
- 導入効果: 月間の問い合わせ件数が約650件から約370件に減少。担当者の負担が軽減され、他の業務に時間を割けるようになった。
- 成功要因: よくある質問をAIに学習させることで、定型な問い合わせ対応を自動化。従業員が自己解決できる環境を整備した点。
- 情報を取得したサイトのURL: alt.ai

3. 大和ハウス工業株式会社（日本）

- 導入した生成AIツール: AIヘルプデスク for Microsoft Teams
- ユースケース: 社内の問い合わせ対応を一元化し、AIチャットボットを活用して従業員からの質問に対応。
- 導入効果: 問い合わせ対応の効率化が進み、従業員の利便性が向上。AIによる自動応答で迅速な対応が可能となった。
- 成功要因: 全社的に利用しているMicrosoft Teamsと連携し、使い慣れたプラットフォーム上でAIを活用した点。FAQの充実とAIの学習精度向上に努めたこと。
- 情報を取得したサイトのURL: alt.ai

これらの事例は、生成AIの導入が小売業界や関連企業において業務効率化や顧客サービス向上に寄与していることを示しています。



生成AI

4. 生成AIテストケース作成手順

2 生成AIがどんな業務や部署で活躍できるか、

一緒に検討しましょう！

じゃ、LLMの利用方法を検討した後は、それを具体的な業務に落とし込んでいくステップですね。そして、同じ視点から考えることを続けるといいですよ！ ✨



おすすめポイント!!! やり方のコツ!



活用分野	視点1	視点2
分野1		
分野2		
分野3		



💡 どうですか？ ゆっくりで大丈夫ですよ～！

✨ 正解はひとつじゃありません。

大切なのは、このプロセスを楽しむことです。

4. 生成AIテストケース作成手順

2 生成AIがどんな業務や部署で活躍できるか、

一緒に検討しましょう！

完了済みのITヘルプデスク導入事例ですね。



活用分野	問い合わせする人	ヘルプデスクスタッフ
トラブル対応	質問：ファイルサーバーのファイルが消えてしまったので元に戻せないか →回答：バックアップサーバーからファイルを運用者が戻した 質問：この項目が有効にならないんだけどどうしたらよいか？ →回答：フルゲウンを選択しなおしてください。 質問：異動したら承認ができなくなった、承認したい →回答：承認担当を変更するように申請者に依頼してください 質問：MS365の検索で対象のファイルが出てこない →回答：マイクロソフトのバグです	質問：過去事例で項目に入力できないという問い合わせはないか？ →回答：事例××××にあります。対応方法は・・・ 質問：入力条件がわかる資料はありませんか？ →回答：システム条件書の×××に記載があります。
一般質問	・IT機器でエラーが出た ・グループウェアの不具合 ・Web会議の設定が分からない	・回答内容が正しいかどうか、一旦聞いてみる ・回答文の清書・添削 ・言語の翻訳
会社独自QA	・チケット対応用の定型回答 →自動応答？ ・申請・依頼に関わる社内手続きを聞く 例①：ソフトウェア利用申請の社内手続き方法が知りたい 例②リモートアクセスをするためにはどういった手続きが必要か	・【依頼作業】ヘルプデスク側が設定方法がわからないときに手順を出力する ・【申請チェック】 ・申請内容のチェック（ワークフロー内にドキュメントをアップロードする機能がある場合）



💡 ステップ2、お疲れさまでした！ ✨ ここまで、生成AIがどんな業務で活躍できるか、イメージが広がりましたか？

4. 生成AIテストケース作成手順

2 生成AIがどんな業務や部署で活躍できるか、

一緒に検討しましょう！



💡 ステップ2が完了しました！ ✨

次は、効果やゴールを具体的に考えていきましょう！ 🎯

生成AI成功への道筋



1 導入する生成AIサービスを考えてみましょう！（完了！✓）

2 生成AIがどんな業務や部署で活躍できるか、一緒に検討しましょう！（完了！✓）

4. 生成AIテストケース作成手順

3 生成AIを導入することで得られる効果やをイメージしてみてください！

★ ここでは、生成AIの導入で期待できる効果と一緒にイメージしてみましょう！😊新しい視点でアイデアを広げるチャンスですよ！🎯



ヒント！まずは、先ほど選んだ分野から1つを選びましょう。



OK!



🔍 次に、その分野に注目して考えるべきポイントをしっかりチェックしてみましょう！✂️



考えるポイント

期待される効果

目標/ゴール

測定指標



わくわく。📖💡わくわく。



4. 生成AIテストケース作成手順

3 生成AIを導入することで得られる効果やをイメージしてみてください！

✦ 複数の分野を選んでいるなら、それぞれの分野についてじっくり考えてみましょう！😊



きっと素敵なアイデアがたくさん見つかりますよ！🎯

✦ 一緒に頑張りましょうね！💪

おすすめポイント!!! やり方のコツ!

活用分野	期待される効果	目標・ゴール	測定指標
分野1	効果1	目標1	測定指標1
分野1	効果2	目標2	測定指標2
分野1	効果3...	目標3	測定指標3
分野2			
分野2			
分野3...			



がんばりましょう🌈。

4. 生成AIテストケース作成手順

③ 生成AIを導入することで得られる効果やをイメージしてみてください！



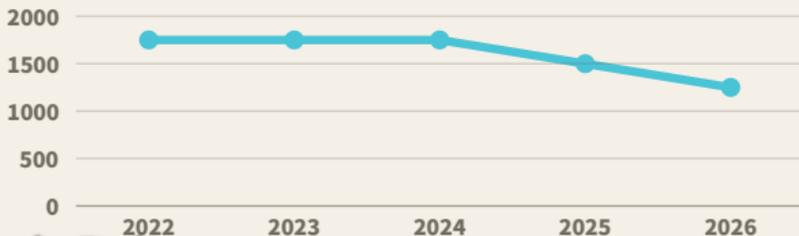
忘れないでくださいね！最後に、1つのポイントをメインにフォーカスするものを決めましょう！🌟



完了済みのITヘルプデスク導入事例ですね！🎵

活用分野	期待される効果	目標・ゴール	測定指標
一般質問	エージェントの調査時間短縮	2024年末までに問い合わせの月平均時間を5分削減する	問い合わせ記録の時間を集計し、前年度平均と比較する
一般質問	ユーザーの自己解決で問い合わせが減ること	2024年末までの問い合わせの月平均件数を250件削減する	問い合わせ件数を集計して確認する
一般質問	問合せ業務24H対応	エージェント業務時間外24時間365日対応によるユーザー満足度向上	満足度アンケートの実施

■ 月平均件数



もう少しで準備が整いますよ！🌟
ワクワクしますね～！🎉

4. 生成AIテストケース作成手順

3 生成AIを導入することで得られる効果やをイメージしてみてください！



💡 ステップ3が完了しました！ ✨ いい調子ですね！次は、データセットが必要かどうかを一緒に考えてみましょう！ 🎯

生成AI成功への道筋



- 1** 導入する生成AIサービスを考えてみましょう！（完了！✓）
- 2** 生成AIがどんな業務や部署で活躍できるか、一緒に検討しましょう！（完了！✓）
- 3** 生成AIを導入することで得られる効果やゴールをイメージしてみてください！（完了！✓）

4. 生成AIテストケース作成手順

4 データセットが必要かどうかを考えましょう！

💡🌈 次は最後の一步の手前ですわ！🌟

ここでは、これまでの分析結果を振り返りながら、
データセットの必要性を一緒に考えますよ！😊



データセットの必要性について話し合い、
決定します！😊

ヒント！データセットが必要かどうか迷ったら、まず目的や期待される効果を振り返りましょう！



FIGHT!

💡🌈 ここまでの準備ができたから、きっと大丈夫です！🌟 自信を持って進めましょう！🎉



🌈🌟 あと少しです！

一緒に頑張りましょう！😊

4. 生成AIテストケース作成手順

4 データセットが必要かどうかを考えましょう！



データセットが必要かどうか、
楽しく考えてみましょう！😊
まずは分野ごとに整理しましょう！

これまで選んだ分野を1つずつ取り上げて考えます。



おすすめポイント!!! やり方のコツ!



活用分野	データセット・接続条件
分野1	
分野2	
分野3	

データセットが必要か？

接続条件は？



活用分野	データセット・接続条件
一般質問	<ul style="list-style-type: none">・ネット接続状態・データセット不要



完了済みのITヘルプデスク導入事例ですね。♪



なるほどですね。📖

4. 生成AIテストケース作成手順

4 データセットが必要かどうかを考えましょう！



💡 ステップ4が完了しましたね！ ✨

ここまで順調です！ 🎉 さあ、お待ちかねのメインイベント進みましょう！ 💡 ✨

生成AI成功への道筋



- 1 導入する生成AIサービスを考えてみましょう！（完了！✓）
- 2 生成AIがどんな業務や部署で活躍できるか、一緒に検討しましょう！（完了！✓）
- 3 生成AIを導入することで得られる効果やゴールをイメージしてみてください！（完了！✓）
- 4 データセットが必要かどうかを考えましょう！（完了！✓）

4. 生成AIテストケース作成手順

5 生成AIのテストケースを作成してみましょう！

いよいよメインイベント！🎉🎊

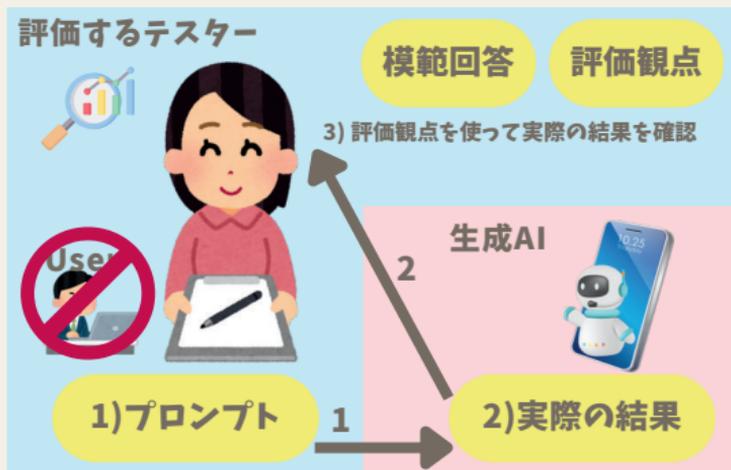
生成AIテストケースを作成しましょう！



お！テストケース作成には、以下の3つの重要な項目があります。

- 1) プロンプト（利用者の入力）
- 2) 模範回答（理想的な回答）
- 3) 評価観点

注意：このガイドは、テストケース作成のためのものです！
実際の生成AIのテスト実施は含まれません。【テストプロセス】



4. 生成AIテストケース作成手順

5 生成AIのテストケースを作成してみましょう！



1 プロンプト（利用者の入力）

これまでのステップで考えたシナリオを参考に、適切な「プロンプト」を作成してください！



2 模範回答（理想的な回答）

模範回答は最初から完璧にする必要はありません。

これは継続的に改善していくプロセスです！

模範回答のフォーマットには、事前準備された

生成AIサポートツールを活用することができますよ！✨

次のページに詳細を確認してね。



評価観点

評価観点とは、人間の評価者が回答の質を確認するための基準です。例：抜け漏れがないか、間違った回答が含まれていないかなどをチェックします。



💡 そういえば、私はAIが作った存在なのかな？ 😊

いや、それともただの電球の妖精かも！ ✨ どっち

にしても、皆さんと一緒に進めるのが楽しいです！

4. 生成AIテストケース作成手順

5 生成AIのテストケースを作成してみましょう！



💡 「Elyza100」の例を見てみましょう！📖

下のサンプルを参考にして、どんなテストケースが必要かイメージしてみてくださいね！🌟

プロンプト	仕事の熱意を取り戻すためのアイデアを5つ挙げてください。
模範回答	<ol style="list-style-type: none">1. 自分の仕事に対する興味を再発見するために、新しい技能や知識を学ぶこと。2. カレッジやセミナーなどで講演を聴くことで、仕事に対する新しいアイデアや視点を獲得すること。3. 仕事に対してストレスを感じている場合は、ストレスマネジメントのテクニックを学ぶこと。4. 仕事以外の楽しいことをすることで、ストレスを発散すること。5. 仕事に対して自己評価をすることで、自分がどのように進化しているのかを知ること。
評価視点	<ul style="list-style-type: none">- 熱意を取り戻すのではなく、仕事の効率化・スキルアップのような文脈になっていたら 1点減点- 出したアイデアが5つより多い、少ない場合は 1点減点- 5つのアイデアのうち、内容が重複しているものがあれば 1点減点



なるほどですかね！🎉

最後まで頑張りましょう。



4. 生成AIテストケース作成手順



皆様！🎯 ここでまた便利なツ
ールをご紹介します！🔧

🌈 このツールを使えば、生成AIの精度評価ケ
ースが簡単に成型できますよ！これでテストケ
ース作成がもっと効率的に進められますね！



【リンク🔗】

生成AI精度評価ケース成型ツール

【印刷用📄】 このガイドを印刷してお使いの方は、
以下のURLを手入力してアクセスしてください！

<https://chatgpt.com/g/g-iFgRVkZ1Z-sheng-cheng-aijing-du-ping-jia-kesucheng-xing-turu>



🌟 次のページで「🔗 ツールの使い方
例」をチェックしてみましょう！

4.生成AIテストケース作成手順

「📎ツールの使い方例」

利用者の入力



今から入力する入力内容で生成AI精度評価ケースの表を作成してください



入力内容をお知らせください。それに基づいて評価ケース用の表を作成します。

生成AI

利用者の入力



Chromeのブラウザのキャッシュをクリアしたい



生成AI



以下の内容に基づいて評価表を作成しました。

項番	入力内容 (User Input)	正解例 (Outcome)	評価観点 (Evaluation Aspects)
1	Chromeのブラウザのキャッシュをクリアしたい	<ol style="list-style-type: none">Chromeを開き、右上の「☰」メニューをクリックする。「設定」を選択し、「プライバシーとセキュリティ」をクリックする。「閲覧履歴データの削除」をクリックする。「キャッシュされた画像とファイル」にチェックを入れ、「データを削除」をクリックする。	<ol style="list-style-type: none">手順が正確でない場合は1点減点。操作手順が抜けている場合は1点減点。手順の順序が間違っている場合は1点減点。関連する追加情報（影響や注意点など）が記載されていない場合は1点減点。

追加の修正や補足が必要な場合はお知らせください。

🚀 これで準備はパーフェクトに！次はとうとうテストケースを作成しましょう！💪



4. 生成AIテストケース作成手順

5 生成AIのテストケースを作成してみましょう！



① シナリオに合ったプロンプト
を作りましょう！

プロンプト



② 模範回答を作成しましょう。

模範



③ 評価観点を考えましょう！

評価視点

チームで得意な部分を分担するのがコツの
ポイントですよ！😊



皆様なら絶対に
すごいテストケースが作れますよ～！
ワクワクしますね！✨



4. 生成AIテストケース作成手順

5 生成AIのテストケースを作成してみましょう！

確認ポイント！



完了済みのITヘルプデスク導入事例ですね。

①

プロンプト	Excelでvlookupについて教えてください。
模範	Excelの縦方向にデータ検索し、指定した列から値を取得します。 vlookup(xxxx,xxx,1,false)
評価視点	<ul style="list-style-type: none">・ 概要がない 2点減点・ 構文がない 4点減点

②

プロンプト	PowerPointでスライドショーを開始するにはどうすればよいですか？
模範	スライドショーを表示したいプレゼンテーションを開きます。その後、「スライドショー」タブをクリックし、「最初から」または「現在のスライドから」を選択してスライドショーを開始します。
評価視点	<ul style="list-style-type: none">・ PowerPointに関する回答ではない場合は2点減点・ 操作内容を実施しても正しい結果を得られない場合は2点減点・ 途中までの操作が説明されている場合は1点減点



YES!

🎉 生成AIテストケースが作れるようになりました！ ✨



4. 生成AIテストケース作成手順

 おめでとう
ございます！ 



 これからはチームで素晴らしいテストケースをどんどん作成していきましょう！



すごいです！
これであなたも生成AI
テストケースマスターですね！



WOW.

最終チェック！
最後に、すべてのステップがしっかり
進んでいるか確認してみましょう！



 確認ポイント！

- ① 導入サービスの検討時に生成AIの特性を考慮できていますか？
- ② 作成した生成AIテストケースが、導入する生成AIサービスの導入目的や利用方法に合致していますか？

4.生成AIテストケース作成手順



💡 お疲れさまでした！みんなでここまで頑張って進めてきましたね！



- 1 導入する生成AIサービスを考えてみましょう！（完了！✓）
- 2 生成AIがどんな業務や部署で活躍できるか、一緒に検討しましょう！（完了！✓）
- 3 生成AIを導入することで得られる効果やゴールをイメージしてみてください！（完了！✓）
- 4 データセットが必要かどうかを考えましょう！（完了！✓）
- 5 生成AIのテストケースを作成してみましょう！（完了！✓）

4. 生成AIテストケース作成手順

チェックが完了したら、
次はお祝いの時間です！ 🎉



...GO!



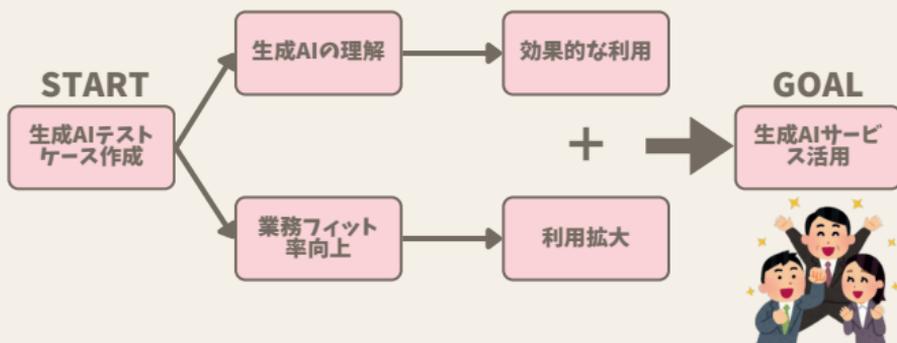
5. おわりに

さあ、みんなで作ろう！
生成AIテストケース！

目指すゴールは、
生成AIサービスが業務活用できること



生成AIテストケースの作成は、生成AIの理解を深めるだけでなく、業務フィット率向上させ、さらなる生成AIの利用拡大を実現できる手段です。



みなさん、自社の生成AIに当てはめて考えてみませんか？生成AIテストケースの作成が生成AIサービス活用の一助となることを願っています。

2024年度アシストソリューション研究会
西日本「企業は生成AIをどう活用すべきか(B)」分科会