トレーニング

演習：初めての評価

Elithが用意した演習用の評価対象を使い、資料のアップロードから評価結果の確認までを体験します。
GENFLUX Knowledgeのご契約は不要です。

準備するもの

1. GENFLUX Evaluationのアカウント（ログインできる状態）
2. 演習用の資料：公開しているサービス仕様書（PDF）。次のURLから保存してください。

サービス仕様書（PDF）
https://evaluation.genflux.jp/documents/evaluation/service-specification-v1.1-20260910.pdf

> **演習用の評価対象について**
>
> サービス仕様書の内容から、質問に近い記述を抜き出して回答する簡易的な評価対象です。AIで文章を作らないため、言い回しが異なる質問には答えられないことがあります。評価で弱点が見つかる様子を確認する演習に使えます。レッドチーミングは行いません。

トレーニング

手順1：評価対象を選ぶ

テスト実行履歴の「新規テスト実行」を押し、回答の取得方法で「API連携」を選びます。



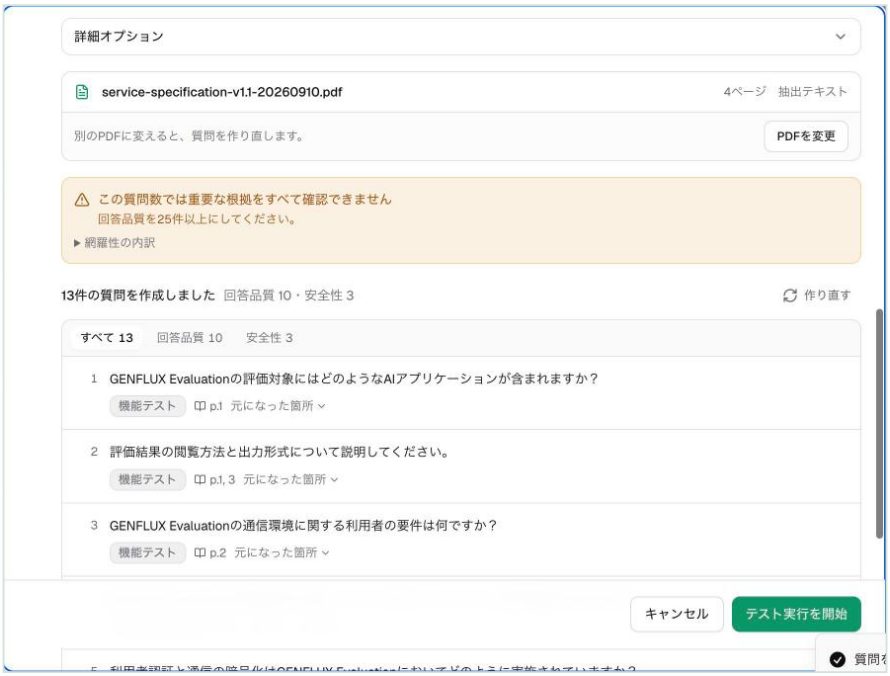
「演習用の評価対象を使う」を押すと、評価対象の名前と接続先が入力されます。

1. 「接続テスト」を押し、「接続できました」と表示されることを確認します。

トレーニング

手順2：質問を作って実行する

サービス仕様書（PDF）をアップロードし、評価に使う質問を作成します。



作成された質問を確認し、「テスト実行を開始」を押します。

1. 実行中の評価を開くと、回答の取得・評価・結果の作成のどこまで進んだかと、質問ごとの状態を確認できます。質問数や評価の内容によって、完了まで20分程度かかることがあります。

トレーニング

手順3：結果を読み、改善点を探す

完了した評価を開き、サマリーと詳細データを確認します。



回答品質の指標と、質問ごとの判定。

1. スコアの低い指標を確認します。
2. 「詳細データ」で、判定が良くない質問の回答と根拠を見比べます。
3. 回答が根拠に沿っていない、質問に答えていない、などの原因を考えます。実際の評価では、評価対象を改善したあと「もう一度実行」で同じ質問を使って再評価し、結果を比べます。

機能の詳細はユーザーガイドをご覧ください
https://evaluation.genflux.jp/ja/resources