

Form Recognizer カスタム モデル

Form Recognizer により、高度な機械学習テクノロジーを使用して、フォームやドキュメントのデータを分析し、抽出します。Form Recognizer モデルは、特定のコンテンツを分析するための参照として使用される抽出データの表現です。Form Recognizer モデルには、次の 2 つの種類があります。

- **カスタム モデル**。Form Recognizer カスタム モデルは、ビジネスに固有の "フォーム" から抽出されたデータを表します。カスタム モデルは、個別のフォーム データを分析するようにトレーニングする必要があります。
- **事前構築済みのモデル**。Form Recognizer は、現在、*領収書*、*名刺*、*身分証明書*、および *請求書* の事前構築済みのモデルをサポートしています。事前構築済みのモデルにより、ドキュメント イメージから情報が検出および抽出され、抽出されたデータは構造化された JSON 出力で返されます。

カスタム モデルの機能

Form Recognizer を使用すると、ユース ケースに関連するフォームから情報を抽出するモデルをトレーニングできます。始めるために必要な同じフォームの種類の場合は 5 つのみです。カスタム モデルは、ラベル付きデータセットの有無にかかわらずトレーニングできます。

カスタム モデルの作成、使用、管理

大まかに説明すると、カスタム モデルを構築、トレーニング、および使用する手順は次のとおりです。

[!div class="nextstepaction"] [1. トレーニング データセットを作成する](#)

カスタム モデルの構築は、トレーニング データセットの確立から始まります。サンプル データセットには、同じ種類の完成したフォームが少なくとも 5 つ必要です。異なるファイルの種類にして、テキストと手書きの両方を含めることができます。フォームは同じ種類のドキュメントで、Form Recognizer の [入力要件](#) に従う必要があります。

[!div class="nextstepaction"] [2. トレーニング データセットをアップロードする](#)

トレーニング データを Azure BLOB ストレージ コンテナーにアップロードする必要があります。コンテナーを含む Azure Storage アカウントを作成する方法がわからない場合は、[Azure portal の Azure Storage に関するクイックスタート](#) を 参照してください。Free 価格レベル (F0) を使ってサービスを試用し、後から運用環境用の有料レベルにアップグレードします。

[!div class="nextstepaction"] [3. カスタム モデルをトレーニングする](#)

ラベル付きデータ セットの [ない](#)、または [ある](#) モデルをトレーニングできます。ラベルのないデータセットの場合、Layout API のみに依存し、ユーザーの入力を追加することなく重要な情報を検出および特定できます。ラベル付きデータセットの場合も、Layout API をに依存していますが、特定のラベルやフィールドの場所など、補助的なユーザーの入力が含まれています。ラベル付きデータとラベルなしデータの両方を使用するには、ラベル付きトレーニング データと同じ種類の少なくとも 5 つの完成したフォームから始めて、ラベルなしデータを必要なデータ セットに追加します。

[!div class="nextstepaction"] [4. カスタム モデルを使用してドキュメントを分析する](#)

トレーニング データセットの一部ではないフォームを使用して、新しくトレーニングしたモデルをテストします。カスタム モデルのパフォーマンスを向上させるために、さらにトレーニングを続けることができます。

[!div class="nextstepaction"] [5. カスタム モデルを管理する](#)

いつでも、サブスクリプションに属するすべてのカスタム モデルの一覧を表示すること、特定のカスタム モデルに関する情報を取得すること、アカウントからカスタム モデルを削除することができます。

次のステップ

API リファレンスのドキュメントを参照して、Form Recognizer クライアント ライブラリの詳細について理解します。

[!div class="nextstepaction"] [Form Recognizer API リファレンス](#)