スタートアップマニュアル (概要編)

本書は、Generative AI API(以降、本サービス)の概要を理解するための情報および利用に関する注意事項について説明します。

提供機能 [🔗](#)

本サービスはAPI機能とUI機能を提供します。

API機能 [🔗](#)

本サービスが提供するAPIは推論用のAPIとモデル管理APIがあります。

推論API [🔗](#)

以下は本サービスが提供するLLMを利用するための推論APIです。APIの詳細はAPIリファレンスを参照してください

- OpenAI API準拠のAPI
- 履歴付き一般対話API (履歴対話機能)
- 登録済のテンプレートプロンプトでの対話API (テンプレート対話機能)
- 文書検索を利用した対話API (検索対話機能)

モデル管理API [🔗](#)

ファインチューニングしたモデルを管理するためのAPIです。ファインチューニング済モデルはオプションメニューを申し込むことで利用できます。

UI機能 [🔗](#)

- 管理ポータル
- チャットUI
- チャットボット

用語定義 [🔗](#)

本サービスのガイドに記載する用語の一覧です

用語	当文書内での意味
cotomi	NECが提供するLLM (大規模言語モデル)
APIキー	APIを利用するための認証キー
管理ポータル	本サービスを利用するために必要な管理操作を行うWebアプリケーションツール
チャットUI	本サービスのLLMを用いて対話を行うためのWebアプリケーション
ユーザ	本サービスの利用者。管理ポータルにアクセスできるシステム管理者と、チャットUIを利用する一般ユーザが存在する
アカウント	本サービスの機能を利用するためのユーザ認証の情報

インデックス	文書検索機能を利用するために作成する、検索対象文書データの集合
ファインチューニング	あらかじめ存在するモデルに利用者が追加学習させて、新しいモデルを生成すること
履歴付き対話APIでの「会話」	同一のhistoryIdを用いて行う一連のやりとり
履歴付き対話APIでの「対話」	会話内の1回のやりとり (LLMへの入力と回答の1回分)
テンプレート	モデルとの会話でよく利用するシステムプロンプト/ユーザープロンプトの定型文を予め作成したもの
システムプロンプト	モデルに会話の背景・目的・ルールなどを指示するために入力するテキスト
ユーザープロンプト	モデルとの対話において利用者が入力するテキスト
システムテンプレート	サービスが提供するテンプレート
ユーザテンプレート	システム管理者が作成したテンプレート
グループ	ユーザ、インデックス、テンプレートが所属できる集合
システムグループ	あらかじめサービスに登録されているグループ
システムユーザ	あらかじめサービスに登録されているユーザ