

# スタートアップマニュアル（概要編）

本書は、Generative AI Cloud (SaaS)（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。

## 提供機能一覧

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）
- 添付したファイルやWebコンテンツを利用した対話API（拡張対話機能）
- ハルシネーション対策API

#### 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

- ユーザ管理
- グループ
- インデックス/文書管理
- 利用状況確認

#### チャットUI

- 一般対話

- 検索対話 (RAG)
- 拡張対話
  - ファイル添付
  - Webコンテンツ

## チャットボット

- 一般対話
- 検索対話

## Agentic AI for Search用チャットUI

- ElasticsearchTool (ベクトル検索)
- WebSearchTool (Web検索)

## コマンド機能

- 非同期文書登録

## 機能概要

以下に説明されていない機能については、各機能ガイドおよびチュートリアルで確認いただけます。

## Agentic AI for Search

Agentic AI for Search機能はAgentic AIを検索用途にカスタマイズし、本サービス上で利用できるようにした機能です。従来のチャットUIと別にAgentic AI for Search専用のチャットアプリケーションを用いることで利用できます。

Agentic AI for Searchでは、ユーザの要求内容からRAGの実施判断をAgentic AI内のElasticsearchToolが自動で行い、LLMが回答します。また、Agentic AI 管理ポータルからWebSearchToolにWeb検索用APIキーの登録と有効化を実施することで、Web検索の実施判断をAgentic AI内のWebSearchToolが自動で行いLLMが回答します。詳細はAgentic AI for Searchの利用ガイドを参照してください。

## 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（組織管理者）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
度数	従量課金を計算する際の数量で、単金×度数が従量課金の請求額となる。 度数は利用したモデルとトークン数毎に算出式が異なる。
Agentic AI for Search	NECが提供するNEC Agentic AIを検索用途にカスタマイズした機能
回答根拠の確認	LLMの生成文章に対するハルシネーション対策機能