HITACHI

JP1 Cloud Service システム管理 生成 AI 連携 利用ガイド

JCSM09-0260-01

前書き

■ 対象サービス

＜ V02-10 以降＞

●システム管理 - スタンダード 小規模

SD-5273182G3 JP1 Cloud Service/System Management - Standard (Lite Edition) 02-40 以降

●システム管理 - スタンダード 15 か月保存

SD-5273181L3 JP1 Cloud Service/System Management - Standard 15 Months Data Retention 02-10 以降

●システム管理 - スタンダード 24 か月保存

SD-5273181M3 JP1 Cloud Service/System Management - Standard 24 Months Data Retention 02-10 以降

●システム管理 - スタンダード 36 か月保存

SD-5273181N3 JP1 Cloud Service/System Management - Standard 36 Months Data Retention 02-10 以降

●システム管理 - エンタープライズ 小規模

SD-5273182J3 JP1 Cloud Service/System Management - Enterprise (Lite Edition) 02-40 以降

●システム管理 - エンタープライズ 15 か月保存

SD-5273181S3 JP1 Cloud Service/System Management - Enterprise 15 Months Data Retention 02-10 以降

●システム管理 - エンタープライズ 24 か月保存

SD-5273181T3 JP1 Cloud Service/System Management - Enterprise 24 Months Data Retention 02-10 以降

●システム管理 - エンタープライズ 36 か月保存

SD-5273181U3 JP1 Cloud Service/System Management - Enterprise 36 Months Data Retention 02-10 以降

■ 輸出時の注意

本製品を輸出される場合には，外国為替および外国貿易法の規制ならびに米国輸出管理規則など外国の輸出関連法規をご確認の上，必要な手続きをお取りください。

なお，不明な場合は，弊社担当営業にお問い合わせください。

■ 商標類

記載の会社名，製品名などは，それぞれの会社の商標もしくは登録商標です。

■ 発行

2026 年 10 月　JCSM09-0260-01

■ 著作権

All Rights Reserved. Copyright (C) 2024, 2026, Hitachi, Ltd.

変更内容

変更内容(JCSM09-0260-01)　JP1 Cloud Service　02-60

追加・変更内容	変更個所
生成 AI 連携呼び出しコマンドに関する記載の追加	1.2，1.8
Azure OpenAI の利用可能モデルに関する注意事項	2.1.2
トラブルシューティングの更新	5.1
次の設定パラメータを追加 • モデルの内部推論／思考モードなどのパラメータ • 検索サービスからの取得件数指定パラメータ • 検索サービスの検索タイプ指定パラメータ • 検索サービスのリランク指定パラメータ • 質問文テンプレートパラメータ	6.1.1

単なる誤字・脱字などはお断りなく訂正しました。

はじめに

このマニュアルは，JP1 Cloud Service 生成 AI 連携のサービス内容について説明したものです。

■ 対象読者

JP1 Cloud Service 生成 AI 連携の導入を検討されている方，JP1 Cloud Service 生成 AI 連携の概要を理解しようとしている方を対象としています。

■ マニュアルの構成

このマニュアルは，次に示す章と付録から構成されています。

第 1 章　JP1 Cloud Service 生成 AI 連携の概要
JP1 Cloud Service 生成 AI 連携で提供するサービスの概要と特長について説明しています。

第 2 章　JP1 Cloud Service 生成 AI 連携の構築
JP1 Cloud Service 生成 AI 連携の構築手順について説明しています。

第 3 章　JP1 Cloud Service 生成 AI 連携の運用
JP1 Cloud Service 生成 AI 連携の運用手順について説明しています。

第 4 章　JP1 Cloud Service システム管理への影響
JP1 Cloud Service 生成 AI 連携を運用する際の JP1 Cloud Service システム管理への影響について説明しています。

第 5 章　トラブルシューティング
JP1 Cloud Service 生成 AI 連携における主なトラブルシューティングについて説明しています。

第 6 章　リファレンス
JP1 Cloud Service 生成 AI 連携が提供する定義ファイルおよびメッセージについて説明しています。

付録 A　連携サービス
連携サービスの利用料金の参考情報について説明しています。

付録 B　プロンプト
JP1 Cloud Service 生成 AI 連携で使用するプロンプトについて説明しています。

付録 C　ハイパーリンク表示の仕様
JP1 Cloud Service 生成 AI 連携が表示するハイパーリンクの仕様について説明しています。

付録 D　ワークフローの定義例

JP1 Cloud Service 生成 AI 連携で使用できるワークフローの定義例について説明しています。

付録 E　各バージョンの変更内容

各バージョンの変更内容について説明しています。

付録 F　用語解説

このマニュアルで使用する用語について説明しています。

■ マニュアルの表記

このマニュアルでは，製品の正式名称と異なる表記があります。次に，このマニュアルでの表記と製品の正式名称が異なる点を次の表に示します。

正式名称	このマニュアルでの表記
JP1/IM3 - Manager [Manager]	JP1/IM3 - Manager

目次

前書き　2
変更内容　4
はじめに　5

1　JP1 Cloud Service 生成 AI 連携の概要　11
1.1　利用可能な Web ブラウザ　12
1.2　機能概要　13
1.3　チャット画面　14
1.3.1　ハイパーリンク表示　15
1.3.2　ワークフロー選択（02-30 以降）　15
1.3.3　レコメンド（02-30 以降）　15
1.4　質問文生成機能　16
1.5　問い合わせ機能　17
1.5.1　問い合わせ送信機能　17
1.5.2　ワークフロー作成機能（02-30 以降）　17
1.5.3　プロンプト作成機能　21
1.5.4　ユーザーコンテンツ取得機能　25
1.5.5　運用管理データ取得機能（02-30 以降）　26
1.5.6　ユーザープログラム呼び出し機能（02-40 以降）　28
1.6　会話履歴　31
1.6.1　会話履歴管理機能　31
1.6.2　会話履歴出力機能　31
1.7　ログ出力機能　33
1.7.1　ログ出力機能の概要　33
1.7.2　ログのダウンロード方法　33
1.8　生成 AI 連携呼び出しコマンド　38
1.8.1　生成 AI 連携呼び出しコマンドの概要　38
1.8.2　生成 AI 連携呼び出しコマンドの仕様　38
1.8.3　使用例：メールに生成 AI の回答を含める　41

2　JP1 Cloud Service 生成 AI 連携の構築　47
2.1　事前に準備するもの　48
2.1.1　システム管理の準備　48
2.1.2　お客様が用意するもの　48
2.1.3　インデックスの作成（任意）　51

2.2 セットアップ 55
2.2.1 生成 AI サービスとの連携設定 55
2.2.2 検索サービスの接続設定 56
2.2.3 ワークフローの設定（02-30 以降） 59
2.2.4 質問文生成機能の設定（任意） 60
2.2.5 質問文の文字数の設定（任意） 61
2.2.6 ログの追加出力の設定（任意） 61
2.2.7 任意のリンク作成の設定（任意） 62
2.2.8 ユーザープログラムの設定（任意）（02-40 以降） 63
2.3 アンセットアップ 65

3 JP1 Cloud Service 生成 AI 連携の運用 66
3.1 設定変更 67
3.1.1 定義ファイルの変更方法 67
3.1.2 シークレットの追加・変更・削除方法 68
3.1.3 連携する生成 AI サービスの変更 68
3.1.4 連携する検索サービスの追加・変更・削除 69
3.1.5 画面設定の変更 70
3.1.6 ログの追加出力の設定変更 71
3.1.7 ワークフローの変更（任意）（02-30 以降） 71
3.1.8 プロンプトの設定（任意）（02-30 以降） 71
3.1.9 ツールプロンプトの設定（任意）（02-40 以降） 72
3.1.10 ユーザープログラムの連携設定の変更（任意）（02-40 以降） 72

4 JP1 Cloud Service システム管理への影響 74
4.1 SID の追加 75

5 トラブルシューティング 76
5.1 トラブルシューティング 77
5.1.1 新規スレッドを作成できない 77
5.1.2 チャット画面のバナーに「保存できる会話履歴の上限に達しました。不要なスレッドを削除してください。スレッドが削除されるまで表示されている会話内容は保存されません。」と表示される。 77
5.1.3 質問エリアへの入力が以下のような不安定な状態となる。 77
5.1.4 検索サービスを利用した質問の回答が行われない。 78
5.1.5 任意のリンクを作成するプロンプトに変更したが，作成されない／ハイパーリンク化しない。 78
5.1.6 KNBD20024-E のエラー詳細で，"Failed to retrieve a response from the assistant. Please try again."が頻発する。 78
5.1.7 回答内容が"Sorry, need more steps to process this request"となる。 79
5.1.8 意図しない XML タグが回答に付与される／Markdown 記法で返却される。 79
5.1.9 回答に含まれる URL がハイパーリンク化せず，リンクが途中で途切れてしまう 79

5.1.10 レコメンドが表示されない／精度が低い 79
5.1.11 KNBD20030-E のエラー詳細で，"The input for the DataModel that generated by the AI was invalid."が頻発し，ツールの実行に失敗する 80
5.1.12 利用できる生成 AI モデルが提供終了となった 80
5.1.13 問い合わせ時に「KNBD20024-E 生成 AI 連携基盤による問い合わせ実行処理でエラーが発生しました。エラー内容: Invalid tool is called by the generative AI. Error: The tool call is empty, so the tool cannot be called.または KNBD20010-E 連携する azure openai への接続に失敗しました。接続情報を確認してください。エラー内容：Error code: 400 - {'error': {'message': 'Could not finish the message because max_tokens or model output limit was reached. Please try again with higher max_tokens.', 'type': 'invalid_request_error', 'param': None, 'code': None}}」というエラーが返却される。 80
5.1.14 問い合わせ時に「KNBD20024-E 生成 AI 連携基盤による問い合わせ実行処理でエラーが発生しました。エラー内容: Failed to retrieve a response from the assistant(＜ operator|agent ＞: ＜ノードの名前＞). Please try again.」 81
5.1.15 問い合わせ時に temperature がデフォルト以外で失敗する。 81

6 リファレンス 82

6.1 定義ファイル 83
6.1.1 設定ファイル（ai_assistant.yml） 83
6.1.2 プロンプト定義ファイル（prompts.chat.ja.yml） 103
6.1.3 ツールプロンプト定義ファイル（tools.ja.yml）（02-40 以降） 107
6.2 メッセージ 114
6.2.1 メッセージの概要 114
6.2.2 メッセージ一覧 116

付録 125

付録 A 連携サービス 126
付録 A.1 Azure Open AI 126
付録 A.2 Azure AI Search 127
付録 A.3 Amazon Bedrock 128
付録 A.4 Knowledge Bases for Amazon Bedrock 129
付録 B プロンプト 130
付録 B.1 02-40 以降のプロンプトのデフォルト 130
付録 B.2 02-30 以降のプロンプトのデフォルト 132
付録 B.3 02-20 以前のプロンプトのデフォルト 136
付録 C ハイパーリンク表示の仕様 142
付録 D ワークフローの定義例 143
付録 E 各バージョンの変更内容 146
付録 E.1 02-60 の変更内容 146
付録 E.2 02-50 の変更内容 146
付録 E.3 02-40 の変更内容 146
付録 E.4 02-30 の変更内容 147

付録 E.5　02-20 の変更内容　147
付録 F　用語解説　148

索引　151

1 JP1 Cloud Service 生成 AI 連携の概要

JP1 Cloud Service 生成 AI 連携で提供するサービスの概要と特長について説明しています。

1.1　利用可能な Web ブラウザ

JP1 Cloud Service 生成 AI 連携をご利用可能な Web ブラウザについては，「JP1 Cloud Service システム管理 利用ガイド」の「1.5.1　前提ソフトウェアの準備」の「インテリジェント統合管理基盤でシステムの監視を実施する場合」をご確認ください。

1.2　機能概要

生成 AI 連携が提供する機能の一覧を以下に示します。

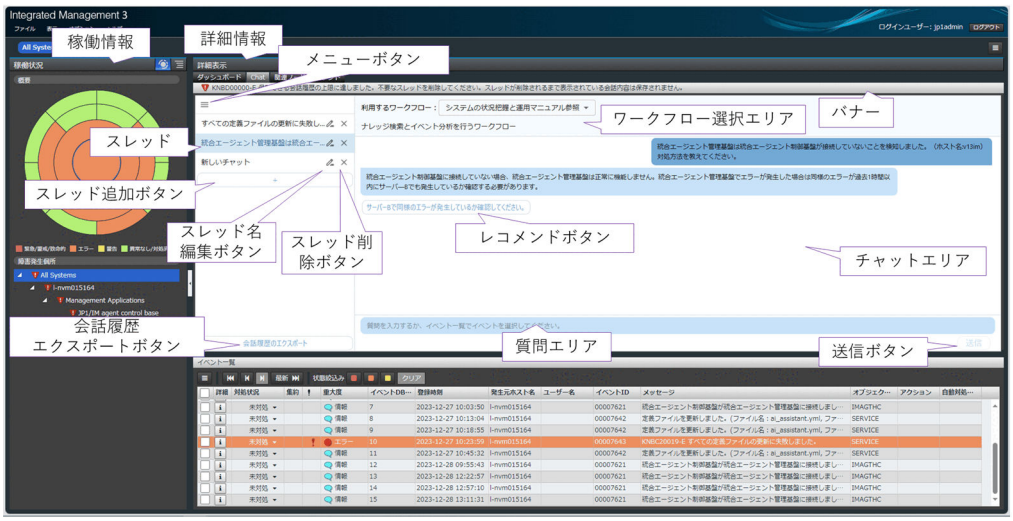
表 1–1　機能一覧

#	機能		説明
1	チャット画面		生成 AI と対話するためのチャット画面です。
2	質問文生成機能		JP1 イベントの情報からテンプレートに従って質問文を生成します。
3	問い合わせ機能		接続先の生成 AI への問い合わせと回答の取得を行います。
4		問い合わせ送信機能	LLM に問い合わせを行い適切な回答を得ます。
5		ワークフロー作成機能	問合せに利用するワークフローの設定を行います。
6		プロンプト作成機能	問合せに適切なプロンプトの設定を行います。
7		ユーザーコンテンツ取得機能	LLM が知らないユーザーコンテンツの情報を取得します。
8		運用管理データ取得機能	LLM が知らない運用管理データを JP1 Cloud Service のシステム管理マネージャーより取得します。
9		ユーザープログラム呼び出し機能	LLM の推論に従い，ユーザープログラムを呼び出します。
10	会話履歴		生成 AI とユーザーの会話内容を記録，管理します。
11		会話履歴管理機能	生成 AI との会話履歴を管理します。
12		会話履歴出力機能	生成 AI との会話履歴を出力する。
13	ログ出力機能		生成 AI 連携のログおよび生成 AI とのやり取りの情報をログに出力します。
14	生成 AI 連携呼び出しコマンド		（02-60 以降） 生成 AI 連携の問い合わせ機能を呼び出すことができるコマンドです。

1.3　チャット画面

生成 AI と対話するためのチャット画面を提供します。チャット画面は統合オペレーション・ビューアーの［稼働状況］領域で IM 管理ノードを選択し，［詳細表示］領域の［タブ］領域に表示される［チャット］タブを選択すると表示されます。選択した IM 管理ノードによらず，すべての IM 管理ノードで同一の内容が表示されます。

図 1–1　チャット画面



画面名称と役割

- チャットエリア：生成 AI への問い合わせ内容と生成 AI からの回答を表示します。
- 質問エリア：生成 AI への問い合わせ内容を入力します。
- ［送信］ボタン：押下すると質問エリアに記載した内容を生成 AI へ問い合わせます。
- メニューボタン：押下すると次のメニューを表示します。
  - シークレットの設定：シークレット一覧画面を開きます。
  - バージョン情報：生成 AI 連携のバージョン情報を表示します（02-50 以降）。
- スレッド：会話履歴が管理される単位。選択したスレッドの内容がチャットエリアに表示されます。
- スレッド追加ボタン：スレッドを追加します。
- スレッド名編集ボタン：スレッド名を編集します。
- スレッド削除ボタン：スレッドを削除します。
- ［会話履歴のエクスポート］ボタン：すべてのスレッドの会話履歴を出力します。
- バナー：エラー発生時の情報を表示します。
- ワークフロー選択エリア：質問の回答に利用するワークフローを選択します。
- レコメンドボタン：回答に加え，次の質問のレコメンドがある場合，ボタンで表示します。ボタン押下でレコメンドの内容を質問エリアに上書きします。

1.3.1　ハイパーリンク表示

チャットエリア内の生成 AI の回答に規定の形で URL の情報が含まれる場合，その箇所はハイパーリンクとして表示されます。ハイパーリンクから生成 AI が参照した情報へ素早くアクセスする事が可能です。デフォルトのハイパーリンクの表示に関する仕様は「付録 C ハイパーリンク表示の仕様」を参照してください。

本機能とプロンプト作成機能を利用する事で，生成 AI の回答内に任意の URL を作成し，ハイパーリンクとして表示することが可能になります。

1.3.2　ワークフロー選択（02-30 以降）

ワークフロー選択エリアで，質問の回答に利用するワークフローを指定します。ワークフローが存在しない場合は質問時にエラーとなります。

1.3.3　レコメンド（02-30 以降）

生成 AI が次にすべき質問があると判断した場合，質問の回答に加えレコメンドを表示します。レコメンドは一つのみ表示されます。

レコメンドの表示契機は「1.5.2 ワークフロー作成機能（02-30 以降）」を参照してください。

1.4　質問文生成機能

JP1 イベントの情報からテンプレートに従って質問文を生成する機能です。選択された JP1 イベントの各種情報を取得し，テンプレートに沿って質問文を生成，生成された質問文を質問エリアに入力します。

本プログラムではデフォルトでは次の質問文が生成されます。

「{イベントメッセージ}。状況を教えてください。」

生成される質問文の変更方法は「3.1.5 画面設定の変更」を参照してください。

1.5　問い合わせ機能

1.5.1　問い合わせ送信機能

［送信］ボタンを押下すると質問エリアの入力内容を生成 AI への問い合わせとして送信します。チャットエリアに問合せ内容と生成 AI からの回答を表示します。生成 AI へ問い合わせを送信する際には「1.6.1 会話履歴管理機能」で保存されている会話内容を含めて生成 AI へ問い合わせを行います。なお，過去の会話内容の中にハイパーリンク化対象の文字列が存在する場合，その文字列は表示されている文字列に置換して問い合わせを行います。

問い合わせの内容は，「1.5.2 ワークフロー作成機能（02-30 以降）」で作成し，「1.3.2 ワークフロー選択（02-30 以降）」で指定したワークフローに従って処理します。

また，生成 AI の回答のレスポンス速度は，連携する生成 AI サービスおよび検索サービスの処理速度に依存することをご留意ください。

1.5.2　ワークフロー作成機能（02-30 以降）

ユーザーが生成 AI を利用して回答を取得するために利用するフローをワークフローと呼びます。ワークフローを適切に定義することで，問い合わせに利用する機能の絞り込みや，利用するナレッジの分離などを行うことができ，適切な回答を得られるようになります。ワークフローは複数定義でき，定義したワークフローはチャット画面より選択し，利用できます。

(1)　ワークフローを構成する要素

ワークフローは次の要素から構成されます。

(a)　ノード

ノードとは特定の役割を遂行する単位です。ノードを組み合わせて一つのワークフローを定義します。ノードは指示を受け取り，生成 AI を利用しながら決められた処理を実行し，ワークフローで定義された次のノードに成果物を渡します。

生成 AI 連携機能でサポートされるノードはオペレーターノードとエージェントノードに大別されます。

■ オペレーターノード

オペレーターノードはユーザーからの質問や会話履歴，エージェントノードの回答をもとに，回答または次の遷移先を決定する役割を持ちます。

■ エージェントノード

エージェントノードはオペレーターノードからのリクエストを受け取り，エージェントノードが利用できる機能を利用して調査・回答を作成する役割を持ちます。

(b) ツール

ツールとは，スキーマ（ツール名：name，説明：description，引数の JSON スキーマ情報：args）とその処理ロジックから構成されるオブジェクトです。単一あるいは複数のツールを組み合わせ，特定の機能を実現します。

ツールはノードが機能を実現するために利用されます。ノードが利用するツールは，ノードのプロンプト定義により指定されます。

(c) プロンプト

プロンプトとは，生成 AI の回答の正確さを高めるために与える指示や命令です。

ユーザーの作成した質問内容だけではなく，生成 AI の役割や使用可能なツールの情報をプロンプトに追加することで，より適切な回答を得られるようにします。プロンプトはノードやツールごとに指定できる情報が異なります。

(2) サポートする構成

(a) ノード

■ ノード種別

生成 AI 連携機能でサポートするノードは以下の通りです。

表 1–2 ノード種別一覧

#	区分	ノード	説明
1	オペレーターノード	Secretary	問合せのエントリーポイントとなるノードです。 ユーザーの質問が簡単なあいさつなど，エージェントノードを利用した調査が不要な場合や，ワークフロー定義に Supervisor が存在しない場合は，ユーザーへ回答を返します。 ユーザーの質問がエージェントノードを利用した調査が必要で Supervisor が定義されている場合は，Supervisor に調査を依頼します。
2		Supervisor	ユーザーの質問を観察して，適切なエージェントノードに処理を振り分けるノードです。問い合わせ実行処理の中で「エージェントノードが呼ばれていない状態」と「エージェントノードが呼ばれた状態」で以下の処理を行います。 ■エージェントノードが呼ばれていない状態 1. Supervisor 自体が最終回答を出力する 2. Supervisor の追加の指示と共にエージェントノードを呼び出す 3. Supervisor の追加の指示無しでエージェントノードを呼び出す ■エージェントノードが呼ばれた状態 1. エージェントノードの回答を最終回答として出力する 2. 次の質問のレコメンドと共にエージェントノードの回答を最終回答として出力する

#	区分	ノード	説明
3	エージェントノード	ナレッジ検索エージェント	ユーザーコンテンツの取得機能を実行し，ユーザーコンテンツに登録された情報を参照して回答するエージェントノードです。 運用手順書等，ユーザーコンテンツに登録している情報を利用した回答が必要な場合は，利用するワークフローにナレッジ検索エージェントを含める必要があります。 ワークフローへの設定方法は「2.2.3 ワークフローの設定（02-30 以降）」を参照してください。 前提となるユーザーコンテンツの取得機能の利用には連携する検索サービスの設定が必要です。検索サービスの接続設定は「2.2.2 検索サービスの接続設定」を参照してください。
4		運用管理データ取得エージェント	運用管理データ取得機能を実行し，システム管理マネージャーに蓄積された運用管理データを参照して回答するエージェントノードです。 JP1 イベントの発生状況など，運用管理データを利用した回答が必要な場合は，利用するワークフローに運用管理データ取得エージェントを含める必要があります。 ワークフローへの設定方法は「2.2.3 ワークフローの設定（02-30 以降）」を参照してください。
5		カスタムエージェント	ユーザープログラム呼び出し機能を実行し，ユーザープログラムの結果をもとに回答するエージェントノードです。02-40 以降で利用可能です。 ユーザープログラムの実行により取得した情報を踏まえた回答が必要な場合は，利用するワークフローにカスタムエージェントを含める必要があります。

■ ノードに定義するプロンプト

ノードに適切なプロンプトを与えることで，ノードの役割の定義や，動作・利用するツールの制御，回答のフォーマット指示を指定することができます。定義できるプロンプトや設定方法は「1.5.3 プロンプト作成機能」を参照してください。

(b)　ツール

■ ツール種別

生成 AI 連携機能でサポートするツールは以下の通りです。

表 1–3　ツール種別一覧

#	名称	説明
1	Azure AI Search 検索ツール	ユーザーコンテンツ取得機能で利用できるツールです。 LLM が生成した検索クエリをパラメータとして，Azure AI Search から検索し，検索結果文字列を返します。検索クエリは半角スペース区切りの文字列を指定します。
2	Knowledge Bases for Amazon Bedrock 検索ツール	ユーザーコンテンツ取得機能で利用できるツールです。 LLM が生成した検索クエリをパラメータとして，Knowledge Bases for Amazon Bedrock から検索し，検索結果文字列を返します。検索クエリは文章で指定します。

#	名称	説明
3	ユーザープログラム呼び出しツール	ユーザープログラム呼び出し機能で利用するツールです。 LLM が生成したユーザープログラムの実行引数をパラメータとして，AWS Lambda にデプロイされた Lambda 関数を同期的に呼び出します。

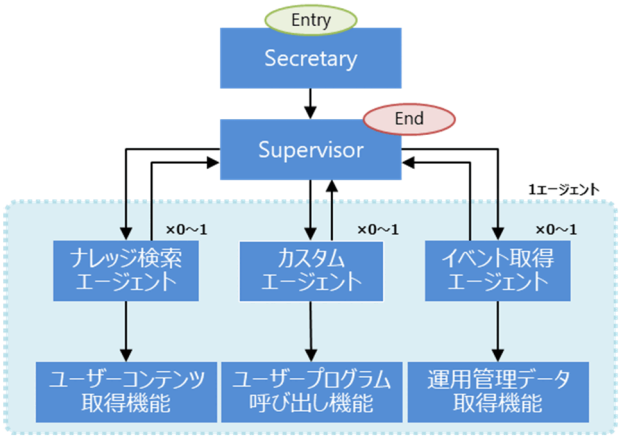
■ ツールに定義するプロンプト

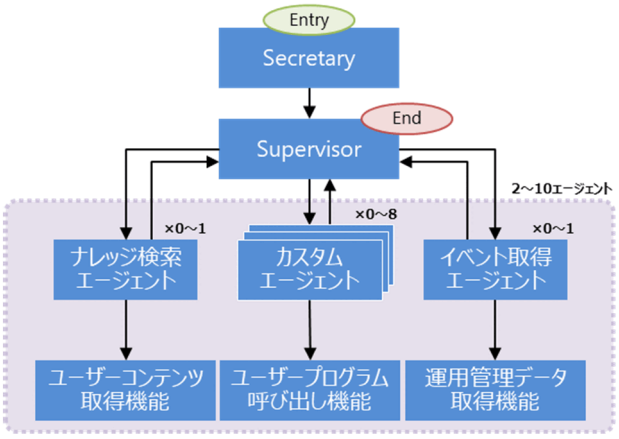
ツールに適切なプロンプトを与えることで，ツールの利用や引数の作成を定義することができます。定義できるプロンプトや設定方法は「1.5.3 プロンプト作成機能」を参照してください。

(c)　ワークフロー

生成 AI 連携機能でサポートするワークフローの構成を以下に示します。

表 1–4　サポートワークフロー構成

#	構成図	説明
1		生成 AI モデルのみを利用し，汎用的な質問への対応や，チケットの作成などを行う構成。 生成 AI モデル単一で回答できる場合に推奨される構成です。
2		1 つのエージェントノードの実行を許可する構成。 生成 AI モデルと，特定の機能を利用することで，汎用的な質問への回答に加え，各エージェントノードに付与された機能を用いた回答も生成できるようにする構成です。 本構成は，指示に対して必要な機能が明確な場合に，エージェントノードを絞ることで，回答精度の向上を図る構成です。 例えば，ユーザーコンテンツの情報のみ必要で，イベントの分析，ユーザープログラムの実行といった操作が不要な場合は，ナレッジ検索エージェントの構成を利用します。 イベント分析のみ行いたい場合や，ユーザープログラムの実行のみ行いたい場合は，それぞれ運用管理データ取得エージェント，カスタムエージェントのみの構成を利用します。

#	構成図	説明
3		2 つ以上のエージェントノードの実行を許可する構成。 生成 AI モデルと，各エージェントノードを複数利用することで，汎用的な質問への回答に加え，各エージェントノードに付与された機能を用いた回答も生成できるようにする構成です。 本構成は，指示に対して必要な機能が不明確な場合に用います。 本構成では，Supervisor が各エージェントノードの役割を踏まえて，適切なエージェントノードを呼び出し，ユーザーの指示に対処します。また，必要に応じて Supervisor はレコメンドを作成します。

(3)　ワークフローの設定方法

ワークフローは要件に合わせて複数作成することができます。「(2)サポートする構成」のワークフロー説明を参照し，必要なワークフローを検討してください。ワークフローの設定方法は「2.2.3 ワークフローの設定（02-30 以降）」を参照してください。

1.5.3　プロンプト作成機能

生成 AI の回答の正確さを高めるため，プロンプト（指示や命令）を定義する機能です。本機能は，問い合わせ送信機能にて入力した質問内容に，生成 AI の役割や使用可能なツールの情報をプロンプトに追加することで，より適切な回答を得られるようにしています。

本機能で設定するプロンプトは「付録 B プロンプト」を参照してください。

(1)　ノードのプロンプト（02-30 以降）

ワークフローのノードごとに指定できる内容が異なります。

パラメータ

以下のパラメータをファイルに定義することで，チューニングや多言語に対応出来るようにします。

- オペレーターノード
  - Secretary

表 1–5　パラメーター一覧（Secretary）

#	パラメータ	項目	設定箇所
1	INSTRUCTION	プロンプトの最初に記載する文字列	リソースファイル
2	TOOLS	AI がアクセスできるツールのリスト	設定ファイル

- Supervisor

表 1–6　パラメーター一覧（Supervisor）

#	パラメータ	項目	設定箇所
1	INSTRUCTION	プロンプトの最初に記載する文字列	リソースファイル
2	TOOLS	AI がアクセスできるツールのリスト	設定ファイル

- エージェントノード
  - ナレッジ検索エージェント，運用管理データ取得エージェント，カスタムエージェント

表 1–7　パラメーター一覧（エージェントノード）

#	パラメータ	項目	設定箇所
1	PREFIX	プロンプトの最初に記載する文字列	リソースファイル
2	TOOLS	AI がアクセスできるツールのリスト	設定ファイル
3	FORMAT_INSTRUCTIONS	AI の応答の型を指示する文字列	リソースファイル
4	DESCRIPTION	エージェントノードの説明	リソースファイル
5	EXAMPLE	エージェントノードの利用例	リソースファイル

設定内容

- オペレーターノード

INSTRUCTION

各ノードに対する動作指示を記載します。ワークフローの構成やノードの活用方法に応じて適切な内容を記載する必要があります。Secretary のみ設定可能です。

- AI の役割

  AI が何を利用でき，どのような事ができるか説明します。

  例：あなたは IT システムのインシデントに対するチケット発行をサポートするアシスタントです。

- リンク作成の定義

  リンク作成を行う条件（例：チケット作成依頼があった場合）や出力形式など，任意のリンクを回答に含める場合のルールを記載します。

TOOLS

AI が利用できるツールの名前を記載します。

Secretary の場合，ワークフローが Secretary 単一の場合は空文字を指定します。それ以外の場合は"__jp1cs_secretary_router"を必ず指定します。

- エージェントノード

PREFIX

プロンプトの最初に記載する文字列を定義します。具体的に以下のような指示を記載します。

- AI の能力の概要
  AI が何を得意とし，どのような情報や知識を持っているのか説明します。
  例：あなたは JP1 を用いた管理を行っているシステム管理のエキスパートです。
- AI の役割
  情報提供，質問応答，アドバイス提供など，AI がユーザーに対して果たす役割を記載します。
  例：あなたはユーザーからの質問を解析し，最も適切な情報を提供することを目的としています。
- ツールの実行結果
  ツールの実行結果に含まれる情報の説明，ツールの実行結果を回答に含める場合のルールを記載します。

TOOLS

AI が利用できるツールの名前を記載します

FORMAT_INSTRUCTIONS

AI の挙動（回答全体の形式や，実行可能なツールの情報を含む）のルールを記載します。

DESCRIPTION

エージェントノードが実行可能な動作の概要を記載します。Supervisor に対するエージェントノードの説明に利用されます。

EXAMPLE

どのような場合にエージェントノードを利用すべきか，利用ケースの例を記載します。Supervisor が次の遷移先を決める際に利用されます。

(2) ツールのプロンプト（02-40 以降）

ツールの定義ファイル(tools.ja.yml)ではツール名とツールの説明，引数の JSON スキーマ情報を定義することができます。

パラメータ

以下のパラメータをファイルに定義することで，チューニングや多言語に対応出来るようにします。

表 1–8 ツールプロンプトのパラメーター一覧

#	パラメータ	項目	設定箇所
1	NAME	ツール名	ツールプロンプト定義ファイル

#	パラメータ	項目	設定箇所
2	DESCRIPTION	ツールの説明	ツールプロンプト定義ファイル
3	ARGS	ツールの引数	ツールプロンプト定義ファイル

設定内容

NAME

ツールの名称を定義します。ここで定義したツール名をノードのプロンプトパラメータ TOOLS に指定することで，LLM が利用できるツールとして認識します。

DESCRIPTION

ツールの説明を定義します。LLM がツールを利用すべきか判断するために利用されます。

- ツールの概要
  ツールにより取得できる内容や，どのような場合に利用すべきか，記載します。
- ツールの引数
  ツールの引数について補足する情報や，引数の例を記載します。

ARGS

ツールの引数の JSON スキーマを定義します。本パラメータはユーザープログラム呼び出しツールのみ定義できます。ツールの引数の名称，タイトル，説明，型情報などを定義します。詳細は「6.1.3 ツールプロンプト定義ファイル（tools.ja.yml）（02-40 以降）」を参照してください。

(3)　ノードのプロンプト（02-20 以前）

パラメータ

以下のパラメータをファイルに定義することで，チューニングや多言語に対応出来るようにします。

表 1–9　パラメーター一覧

#	パラメータ	項目	設定箇所
1	PREFIX	プロンプトの最初に記載する文字列	リソースファイル
2	TOOLS	AI がアクセスできるツールのリスト	設定ファイル
3	FORMAT_INSTRUCTIONS	AI の応答の型を指示する文字列	リソースファイル

設定内容

PREFIX

プロンプトの最初に記載する文字列を定義します。具体的に以下のような指示を記載します。

- AI の能力の概要
  AI が何を得意とし，どのような情報や知識を持っているのか説明します。
  例：あなたは JP1 を用いた管理を行っているシステム管理のエキスパートです。

- AI の役割
  情報提供，質問応答，アドバイス提供など，AI がユーザーに対して果たす役割を記載します。
  例：あなたはユーザーからの質問を解析し，最も適切な情報を提供することを目的としています。
- リンク作成の定義
  リンク作成を行う条件（例：チケット作成依頼があった場合）や出力形式など，任意のリンクを回答に含める場合のルールを記載します。
- ツールの実行結果
  ツールの実行結果に含まれる情報の説明，ツールの実行結果を回答に含める場合のルールを記載します。

TOOLS

AI が利用できるツールの名前を記載します。

FORMAT_INSTRUCTIONS

AI の挙動（回答全体の形式や，実行可能なツールの情報を含む）のルールを記載します。

1.5.4 ユーザーコンテンツ取得機能

(1) 概要

ユーザーコンテンツの検索情報が格納されている検索サービス上のインデックスに対して，検索クエリを送信する機能です。この機能により，生成 AI 連携が連携する LLM が学習済みのデータだけではなく，エンドユーザーが独自に所有しているマニュアルや FAQ，ナレッジベースなどの情報と，それらの情報へのアクセス URL も踏まえた回答を可能にしています。

また，本機能で利用する検索サービスは，生成 AI サービスのベンダーに依存しません（例：Azure OpenAI Service と Knowledge Bases for Amazon Bedrock）。

> **重要**
>
> ユーザーコンテンツにパーソナルデータや機密データは配置しないでください。配置した場合，回答内容およびログに上記データを含むことがあります。

（02-40 以降）

ユーザーコンテンツ取得機能は Azure AI Search 検索ツールまたは Knowledge Bases for Amazon Bedrock 検索ツールを一つまたは複数利用して実現します。複数利用することで，複数のナレッジ（例えば，システムごとに運用手順書を格納するインデックスを分けている場合など）を横断して検索・回答することができます。

本機能の仕様は以下の通りです。

表 1–10　ユーザーコンテンツ取得機能仕様

#	項目	仕様
1	ツール上限	1 つのナレッジ検索エージェントに対して，最大 10 個まで連携可能です。ナレッジ検索エージェントのノードプロンプトのパラメータ TOOLS にて利用するツール名を指定し，複数定義します。
2	ツールの実行回数	1 回の問い合わせ実行につき実行する検索ツールの回数は原則 1 回であり，実行回数をナレッジ検索エージェントのプロンプトにより制御しています。変更することで複数回実行することも可能ですが，応答速度の遅延やタイムアウト発生の可能性が高くなるため，用途に応じてチューニングを行う必要があります。

(2)　ユーザー利用方法

本機能を利用するためには，「2.2.2 検索サービスの接続設定」を参照してください。

1.5.5　運用管理データ取得機能（02-30 以降）

(1)　概要

JP1 Cloud Service のシステム管理マネージャーに蓄積された運用管理データから情報を取得する機能です。この機能により，生成 AI 連携が連携する LLM が学習済みのデータだけではなく，システム管理マネージャーに蓄積された JP1 イベントの情報を踏まえた回答を可能にしています。

本機能の仕様は以下の通りです。

表 1–11　運用管理データ取得機能仕様

#	項目	仕様
1	参照範囲	統合オペレーション・ビューアーにログインしている JP1 ユーザーが参照できる範囲かつ，統合オペレーション・ビューアーの［稼働状況］領域で選択している構成要素で発生した JP1 イベントが参照できます。
2	検索条件	イベントは「時間（B.TIME）」と「ホスト名（E.JP1_SOURCEHOST）」に基づいて取得します。時間とホスト名は生成 AI が質問や会話履歴の情報から生成します。
3	時間範囲	時間範囲が指定されていない場合，過去 1 日分検索します。質問などに含まれる時刻は JST と解釈します。
4	件数	検索条件に合致した JP1 イベント 200 件を上限に取得します。
5	メッセージ	基準となるメッセージ文と類似したイベントメッセージを持つイベントの情報を取得します。質問文や会話履歴に，基準となるメッセージ文に関する情報が含まれていない場合，取得するイベント件数は 0 件となります。基準となるメッセージは生成 AI が質問や会話履歴の情報から生成します。

仕様を満たす運用管理データ取得エージェントへの質問例を示します。

■仕様を満たし，回答が正しく返却されやすいケース

1. 時間範囲やホストが明確に示されている

```
（質問例）
2025/03/31 10:00～2025/04/01 10:00にhoatAで「XXX」というイベントが発生しているか確認してください。
```

2. 運用マニュアルにエスカレーション条件が記載されており，検索する条件が明確である。

```
（質問例）
以下のイベントが発生しています。運用マニュアルを取得して、エスカレーションすべきか判断してください。
<イベント情報>
メッセージ：xxx
発生時刻：2025年03月31日　21時10分
発生元ホスト名：yyy

（運用マニュアル例）
メッセージ：xxx
エスカレーション条件：
　以下の条件のどちらかまたは両方を満たす場合にエスカレーションしてください。
　・条件1．イベントが平日の20時以降に発生しているか確認する。
　・条件2．イベントの発生元ホストがyyyまたはzzzで発生している。
アクション：
　・Aさんに電話してください。
```

■仕様を満たさないため，うまく回答がされないケース

1. 時間範囲が明確に定義されていない

```
（質問例）
過去にxxxイベントは発生していますか。
（検索結果）
過去1日の範囲のみ検索される。
```

2. 参照するイベントが大量で，上限(200 件)を超過する

```
（質問例）
HostAで過去1ヶ月にXXXイベントは1000回以上発生していますか。
（検索結果）
必ず発生していない判定になる。
```

3. 基準となるメッセージに関する情報が存在しない

```
（質問例）
HostAにおける、2025年3月31日1:00から2025年3月31日1:05のイベント発生状況をまとめてください。
（検索結果）
基準となるメッセージが存在しないため、取得イベントは0件となる。
```

(2)　ユーザー利用方法

本機能を利用するために必要な設定はありません。

1.5.6　ユーザープログラム呼び出し機能（02-40 以降）

(1)　概要

ユーザープログラムがデプロイされているプログラム実行サービスに対して，ユーザープログラムの実行を要求する機能です。この機能により，生成 AI 連携機能が連携する LLM の判断により，ユーザープログラムの呼び出しを可能とし，ユーザーによるエージェントノードの機能拡張を実現します。本機能がサポートするプログラム実行サービスは，AWS Lambda です。

> **重要**
>
> ユーザープログラムの実行処理中のエラーメッセージや，返却内容にパーソナルデータや機密データを出力しないでください。
>
> 出力した場合，回答内容およびログに上記データを含むことがあります。

本機能の仕様は以下の通りです。

表 1–12　ユーザープログラム呼び出し機能仕様

#	項目	仕様
1	ツール上限	1 つのカスタムエージェントに対して，最大 10 個まで連携可能です。上限超過時は警告メッセージを出力します。カスタムエージェントのノードプロンプトのパラメータ TOOLS にて利用するツール名を指定し，複数定義します。
2	ユーザープログラムの実行	ユーザープログラムの実行要求とともに，実行引数を送信します。実行引数が送信されるまでの流れは以下の通りです。 1. カスタムエージェントは，ツールの説明と，その実行引数の JSON スキーマを LLM に送信します。 2. LLM は実行引数の JSON スキーマに従って実行引数を生成し，返却します。 3. カスタムエージェントは，返却された実行引数を付与してユーザープログラムの実行要求を送信します※1。 4. ユーザープログラムは指定された実行引数で処理を行い，結果を返却します。 5. カスタムエージェントはツール実行結果としてユーザープログラムの実行結果を取得します。
3	ユーザープログラムの実行引数	実行引数の JSON スキーマはツールプロンプト定義ファイル（tools.ja.yml）に設定情報を記載して定義します。※2 指定できる引数の数や型には制限があります。詳細は「6.1.3 ツールプロンプト定義ファイル（tools.ja.yml）（02-40 以降）」を参照してください。 想定の JSON スキーマが作成されているかは設定ファイル（ai_assistant.yml）の該当ツールの service.langchain.tools[].aws_lambda_function.handle.logging_args_schema に true を指定することで，ログから確認できます。
4	ユーザープログラムへの入力	LLM が作成した実行引数の情報は，シリアライズされた JSON 文字列として送信します。送信される JSON オブジェクトは，ユーザーが定義した JSON スキーマ

#	項目	仕様
4	ユーザープログラムへの入力	に準拠したオブジェクトです。ユーザープログラム内でどの型のオブジェクトとして利用できるかはユーザープログラムを実装するランタイムに依存します。
5	ユーザープログラムの戻り値	ユーザープログラム実行要求の応答（HTTP レスポンスのレスポンスボディ）をシリアライズされた JSON 文字列として以下の処理を行い，解釈します。 • 応答を JSON 文字列にデシリアライズ（文字コードは UTF-8） • JSON 文字列を python の次の dict 型に変換 {"content": [{"type": "text", "text": "任意の文字列"}]} ユーザープログラムの戻り値は，上記の処理を経て{"content": [{"type": "text", "text": "任意の文字列"}]}と表される必要があり，この形式を満たさない場合，ユーザープログラム呼び出し機能は異常終了します。ユーザープログラム呼び出し機能は，処理結果として上記の"任意の文字列"に該当する値を LLM に返却します。 ユーザープログラムの処理中にエラーが発生し，service.langchain.tools[].aws_lambda_function.handle.raise_function_processing_error に false が設定されていた場合は，HTTP レスポンスボディに含まれる"errorMessage"フィールドの値を LLM に返却します。 本機能で受信できるサイズ（HTTP レスポンスヘッダの Content-Length）は 80KB です。
6	ログ出力	ユーザープログラムの呼び出しに合わせて下記のログを出力します。 • ツールを利用可能なワークフローが呼び出されたとき，ツールに定義された JSON スキーマを以下のメッセージ ID のログに出力します。 KNBD20037-I • カスタムエージェントがツールを実行したとき，ツールの識別子を以下のメッセージ ID のログに出力します。 KNBD20036-I • カスタムエージェントの処理が正常に終了したとき，実行したツールの情報を以下のメッセージ ID のログに出力します。 KNBD20003-I これらのログに出力されるリクエスト ID は，AWS によって発行されるリクエストの識別子です。 なお，ユーザープログラムの戻り値は，ユーザープログラム呼び出し機能では明示的にロギングされません。

注※1

AWS Lambda の呼び出しは，同期（RequestResponse）呼び出しで，タイムアウトは 90 秒固定です。Lambda 関数の処理のタイムアウト値は，90 秒以下の値を設定することを推奨します。

注※2

指定できる JSON スキーマのデータ型は以下の通りです。

データ型	Enum での指定	備考
オブジェクト	−	オブジェクトの配下には，オブジェクト以外のデータ型を指定できます。
配列	−	配列の要素に関しても，データ型(オブジェクト，数値，文字列)を定義できます。要素を Enum で指定することも可能です。

データ型	Enum での指定	備考
数値	○	整数または浮動小数点数
文字列	○	
ブール値	–	
null	–	

(2) ユーザー利用方法

本機能を利用するためにはユーザープログラム（＝ AWS Lambda 関数）の作成と，ユーザープログラムを実行するための生成 AI 連携機能への定義が必要です。

作成する Lambda 関数は上記仕様に則って作成してください。

ユーザープログラムを実行するための定義は「2.2.8 ユーザープログラムの設定（任意）（02-40 以降）」を参照してください。

1.6　会話履歴

1.6.1　会話履歴管理機能

生成 AI との会話履歴はスレッド形式で管理されます。スレッドを選択すると過去の会話履歴を確認できます。同時に存在するスレッドは 1 ユーザー 10 スレッドを上限とし，スレッドを削除すると会話履歴も削除されます。チャット画面表示時の挙動は次の通りです。

会話履歴はブラウザのキャッシュに統合オペレーション・ビューアーにログインしている JP1 ユーザーごとに保存されます。そのため，同一 JP1 ユーザーでログインしても異なるブラウザを使用した場合は，会話履歴が異なります。また，ブラウザで共有するキャッシュサイズの上限を超過するとそれ以降会話履歴は保存されず，上限超過時はチャット画面にバナーで上限超過した旨と対処方法を提示します。

> **重要**
>
> キャッシュはブラウザで共有のため，とある JP1 ユーザーが大量の会話履歴を残し上限超過した場合は，同じマシンで同一ブラウザを使用する他の JP1 ユーザーも会話履歴を保存できない状態となります。

1.6.2　会話履歴出力機能

ログインしている JP1 ユーザーの全スレッドの会話履歴を JSON ファイル（UTF-8，BOM 無し，ファイル名が export.json）に出力します。出力項目は次の通りです。

表 1–13　会話履歴の出力項目

#	項目名	内容	備考
1	version	エクスポートのバージョン	02-20 までは 1 が，02-30 は 2 が，02-40 では 3 が指定されます。
2	history	会話履歴のリスト。スレッドの数だけ存在します。	
3	threadUniqueKey	スレッドのキー	
4	threadName	スレッドの名前	
5	messages	メッセージの内容。画面に表示されている内容がすべて存在します。	ブラウザに保存されている内容とは一致しません。
6	role	チャットの送信者。user または ai いずれかの値です。	
7	content	チャットの内容	ハイパーリンク化対象の文字列も置換されずに，そのまま保存されます。
8	workflow	ワークフローの名称	02-30 以降

#	項目名			内容	備考
9			invokeId	生成 AI の回答の識別子	02-40 以降で，生成 AI の回答にのみ付与されます。

以下に 02-40 の出力例を示します。

```
{
  "version": 3,
  "history": [
    {
      "threadUniqueKey": "c7f8ec26-301f-16b1-983e-a0f6f1a3eac7",
      "threadName": "thread-1",
      "messages": [
        {
          "role": "user",
          "content": "ホストAでCPU使用率が90%を超過しました。考えられる原因を教えてください。",
          "workflow": "default_workflow"
        },
        {
          "role": "ai",
          "content": "ホストAでCPU使用率が90%を超える原因はいくつか考えられます。以下にいくつかの可能性を挙げます。¥n1. 高負荷のプロセスまたはアプリケーション…",
          "workflow": "default_workflow",
          "invokeId": "569468b6-962d-461a-b9f9-9c11d39e1ebd ",
        }
      ]
    }
  ]
}
```

1.7　ログ出力機能

1.7.1　ログ出力機能の概要

ログ出力機能は，生成 AI 連携基盤の機能の 1 つとして，生成 AI 連携基盤内の各機能のログを，テキスト形式のファイル（以降，ログファイル）に出力する機能を提供します。

また，ユーザーが入力した質問，生成 AI の回答，ユーザーコンテンツの検索結果に関しては，それらの情報をログとして出力するか否かを設定ファイルにて制御できます。デフォルトの設定値では，上記の情報はログとして出力されません。出力が抑止されるログの対象については，「6.2.1(2)メッセージの出力先」をご確認ください。

監査目的や，生成 AI の思考過程や質問と回答の精度等の分析目的などでログを出力したい場合は，設定ファイルにて出力抑止の設定をデフォルト値（有効）から変更してください。出力抑止を無効にし，ログが追加で出力されるように設定した場合，JP1 Cloud Service に上記の情報が残ることに注意してください。

なお，運用管理データ取得機能およびユーザープログラム呼び出し機能の取得結果に関しては，ログファイルへの出力の対象外となります。

1.7.2　ログのダウンロード方法

ユーザーはログを JP1 Cloud Service サービスポータルの「ログのエクスポート」機能※でダウンロードできます。

注※

「JP1」タブの「ログ一覧」からログを選択して，「エクスポート」を押下して実行する機能を指します。詳細は，「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」の「9.1.3 エクスポート」を参照してください。

(1)　一覧表示

ログ一覧に表示する項目と説明を以下に示します。ログの一覧表示の詳細は，「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」の「9.1.1 一覧表示」を参照してください。

表 1–14　ログ一覧の表示項目

項目	説明
［チェックボックス］	操作対象を示すチェックボックスです。
ログ名	ログの種類を表す名称が表示されます。ログ名のリンクをクリックすると，ログの詳細情報を参照することができます。

項目	説明
ログ名	生成 AI 連携 メッセージログ 生成 AI 連携が出力するログを出力します。ログの出力抑止の設定をデフォルト設定値から変更することで，ユーザーの入力内容およびそれに対する応答のログも出力されます。 生成 AI 連携 LangChain 思考過程ログ プロンプトにより成形された質問，および，intermediate_steps，最終回答といった LangChain の思考過程を出力します。本ログファイルに出力されるログは全て出力抑止対象のログであるため，デフォルト設定値では，何も出力されません。
最終更新日時	ログの最終更新日時が表示されます。

(2) 詳細表示

ログの詳細を表示します。ログの詳細表示の詳細は，「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」の「9.1.2 詳細表示（V01-11 以降）」を参照してください。

(3) エクスポート

ログをエクスポートします。ログのエクスポートの詳細は，「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」の「9.1.3 エクスポート」を参照してください。

ダウンロードしたファイル中に格納されている各ログのファイル名を以下に示します。

表 1–15 ダウンロードできるログのファイル名

ログ名	ファイル名	ファイル名の例
生成 AI 連携 メッセージログ	**<タスク ID>**_ai_assistant_service.csv	0000000001_ai_assistant_service.csv
生成 AI 連携 LangChain 思考過程ログ	**<タスク ID>**_ai_assistant_service_intermediate.csv	0000000001_ai_assistant_service_intermediate.csv

ログ名「生成 AI 連携 メッセージログ」の出力形式を以下に示します。

表 1–16 生成 AI 連携 メッセージログの出力形式

出力形式	”**<通番>**△**<日時>**△**<時刻>**△**<AP 名>**△**<プロセス ID>**△**<スレッド ID>**△**<メッセージ ID>**△**<メッセージテキスト>**”
説明	出力項目の内容は，「表 1-17 生成 AI 連携 メッセージログの出力項目」を参照してください。△は半角スペースを表します。 • 文字コードは UTF-8 です。 • 改行コードは LF です。 • ヘッダ行はありません。 • メッセージテキスト内の日時表示は UTC ＋タイムゾーンです。

出力例 (02-40)	0002 2025-08-20 12:27:20.208+0900 ai_assistant 33057 33455 KNBD20000-I {"question": "ITシステム「サンプルシステム0」では，どのポートを監視すべきですか。検索して教えてください。", "user": "jp1admin", "workflow": "default_workflow"} 0003 2025-08-20 12:28:16.002+0900 ai_assistant 33057 33455 KNBD20001-I {"answer": "検索結果に基づいて，ITシステム「サンプルシステム0」で監視すべきポートについて回答いたします。 ITシステム「サンプルシステム0」では，以下の2つのポートを監視する必要があります： 1. サーバA0: ポート443 (TCPプロトコル)<略> ", "recommends": ['ITシステム「サンプルシステム0」の監視ポート（サーバA0のポート443とサーバB0のポート3306）に関連するイベントの発生状況を調査してください。<略>'], "workflow": "default_workflow"}
出力例 (02-30)	"0011 2023-11-17 10:20:30.000+0900 ai_assistant 12345 00001 KNBD20000-I {"question": " ITシステム「サンプルシステム」では，どのポートを監視すべきですか。検索して教えてください。", "user": "jp1admin", "workflow": "default_workflow"}" "0012 2023-11-17 10:20:34.000+0900 ai_assistant 12345 00001 KNBD20001-I {"answer": "サンプルシステムでは，サーバAの443番ポートと，サーバBの3306番ポートを監視する必要があります。ネットワークモニタリングツールや監視スクリプトを使用することができます。 [sourcepage:Sample_System_boot.pdf-page.1]", "recommends": ["監視スクリプトについて教えてください。"], "workflow": "default_workflow"}"
出力例 (02-20 以前)	0116 2024-01-16 16:34:07.016+0900 ai_assistant 53702 53708 KNBD20000-I {"question": "ITシステム「サンプルシステム」では，どのポートを監視すべきですか。検索して教えてください。", "user": "string"} 0117 2024-01-16 16:34:18.016+0900 ai_assistant 53702 53708 KNBD20001-I {"answer": "サンプルシステムでは，サーバAの443番ポートと，サーバBの3306番ポートを監視する必要があります。ネットワークモニタリングツールや監視スクリプトを使用することができます。 [sourcepage:Sample_System_boot.pdf-page.1]"}

ログ名「生成 AI 連携 メッセージログ」の出力項目を以下に示します。

表 1–17　生成 AI 連携 メッセージログの出力項目

#	項目	説明
1	通番	ログの番号。
2	日時	ログの日時。形式は yyyy-MM-dd
3	時刻	ログの時刻。形式は hh:mm:ss.SSSZZZ
4	AP 名	アプリケーション識別名。ai_assistant を指定します。
5	プロセス ID	実行したプロセスの ID。
6	スレッド ID	実行したスレッドの ID。
7	メッセージ ID	本ログのメッセージの ID。ID の形式および具体的な ID は，「6.2.1 メッセージの概要」を参照してください。
8	メッセージテキスト	詳細なメッセージテキスト。出力される情報は「6.2.2 メッセージ一覧」を参照してください。
9	改行	LF

ログ名「生成 AI 連携 LangChain 思考過程ログ」の出力形式を以下に示します。

表 1–18　生成 AI 連携 LangChain 思考過程ログの出力形式

出力形式	**”<通番>△<日時>△<時刻>△<AP 名>△<プロセス ID>△<スレッド ID>△<メッセージ ID>△<メッセージテキスト>”**
説明	出力項目の内容は，「表 1-19 生成 AI 連携 LangChain 思考過程ログの出力項目」を参照してください。△は半角スペースを表します。 • 文字コードは UTF-8 です。 • 改行コードは LF です。 • ヘッダ行はありません。 • メッセージテキスト内の日時表示は UTC ＋タイムゾーンです。
出力例 (02-40)	0002 2025-08-20 12:27:22.749+0900 ai_assistant 33057 33455 KNBD20002-I {”prompt_question”: ”ITシステム「サンプルシステム0」では，どのポートを監視すべきですか。検索して教えてください。”, ”workflow”: ”default_workflow”, ”nodes”: [{’node’: ’secretary’, ’llm’: ’aws_llm’, ’prompt’: ’__jp1cs_secretary’}, {’node’: ’supervisor’, ’llm’: ’ aws_llm ’, ’prompt’: ’__jp1cs_supervisor’}, {’node’: ’knowledge_search_agent’, ’llm’: ’ aws_llm ’, ’prompt’: ’knowledge_search_agent’}, {’node’: ’event_analysis_agent’, ’llm’: ’ aws_llm ’, ’prompt’: ’__jp1cs_event_analysis_agent’}], ”invoke_id”: ”819ea05a-7b2b-4099-8986-7ea9c24c9343”} 0003 2025-08-20 12:27:24.391+0900 ai_assistant 33057 33617 KNBD20032-I {”node”: ”secretary”, ”message”: ””, ”invoke_id”: ”819ea05a-7b2b-4099-8986-7ea9c24c9343”} 0004 2025-08-20 12:27:26.702+0900 ai_assistant 33057 33617 KNBD20032-I {”node”: ”supervisor”, ”message”: ”ITシステム「サンプルシステム0」の監視すべきポート情報を検索してください。”, ”invoke_id”: ”819ea05a-7b2b-4099-8986-7ea9c24c9343”} 0005 2025-08-20 12:28:11.712+0900 ai_assistant 33057 33617 KNBD20003-I {”node”: ”knowledge_search_agent”, ”intermediate_steps”: [{’AIMessage’: {’tool_calls’: [{’name’: ’user_contents’, ’args’: {’arg’: ’サンプルシステム0 監視 ポート’}, ’id’: ’toolu_bdrk_013JB9TsEUKpvvj5ciGVcvBM’, ’type’: ’tool_call’}]}}, {’ToolMessage’: {’name’: ’user_contents’, ’tool_call_id’: ’toolu_bdrk_013JB9TsEUKpvvj5ciGVcvBM’, ’content’: ’sourcepage:Sample_System_0_boot.pdf-page.1¥ncontent:サンプルシステム0の起動手順<略>’}}], ”invoke_id”: ”819ea05a-7b2b-4099-8986-7ea9c24c9343”} 0006 2025-08-20 12:28:15.998+0900 ai_assistant 33057 33455 KNBD20004-I {”final_answer”: ”検索結果に基づいて，ITシステム「サンプルシステム0」で監視すべきポートについて回答いたします。 ITシステム「サンプルシステム0」では，以下の2つのポートを監視する必要があります： 1. サーバA0: ポート443（TCPプロトコル)<略> ”, ”recommends”: [’ITシステム「サンプルシステム0」の監視ポート（サーバA0のポート443とサーバB0のポート3306）に関連するイベントの発生状況を調査してください。<略>’], ”invoke_id”: ”819ea05a-7b2b-4099-8986-7ea9c24c9343”}
出力例 (02-30)	”0021 2023-11-17 10:20:31.000+0900 ai_assistant 12345 00001 KNBD20002-I {”prompt_question”: ” ITシステム「Aシステム」では，どのポートを監視すべきですか。検索して教えてください。”, ”workflow”: ”default_workflow”, ”nodes”: [{”node”: ”secretary”, ”llm”: ”use_defined_name”, ”prompt”: ”jp1cs_secretary”}, {”node”: ”supervisor”, ”llm”: ”use_defined_name”, ”prompt”: ”jp1cs_supervisor”}, {”node”: ”knowledge_search_agent”, ”llm”: ”use_defined_name”, ”prompt”: ”knowledge_search_agent”}, {”node”: ”event_analysis_agent”, ”llm”: ”use_defined_name”, ”prompt”: ”jp1cs_event_analysis_agent”}]}” ”0022 2023-11-17 10:20:33.000+0900 ai_assistant 12345 00001 KNBD20032-I {”node”: ”secretary”, ”message”: ”ユーザーの質問に答えてください”}” ”0023 2023-11-17 10:20:33.000+0900 ai_assistant 12345 00001 KNBD20032-I {”node”: ”supervisor”, ”message”: ”ITシステム「Aシステム」の監視対象のポートを調査してください”}”

出力例 (02-30)	”0024 2023-11-17 10:20:33.000+0900 ai_assistant 12345 00001 KNBD20003-I {”node”: ”knowledge_search_agent”, ”intermediate_steps”: [{”AIMessage”: {”tool_call”: {”name”: ”user_contents”,”args”: {”__arg1”: ”Aシステム ポート監視”},”id”: ”toolu_bdrk_013ryNHoxrW7F4nPcnXQcsKC”}}},{”ToolMessage”: {”content”: ”sourcepage:Asystem_boot.pdf-page.1¥ncontent: A システムの起動手順<略>”,”name”: ”user_contents”,”tool_call_id”: ”toolu_bdrk_013ryNHoxrW7F4nPcnXQcsKC”}}]}” ”0024 2023-11-17 10:20:34.000+0900 ai_assistant 12345 00001 KNBD20004-I {”final_answer”: ”検索結果から，Aシステムでは以下のポートの監視が必要であることがわかります。<略>”, ”recommends”: [”ポートの監視方法を教えてください”]}
出力例 (02-20 以前)	0021 2023-11-17 10:20:31.000+0900 ai_assistant 12345 00001 KNBD20002-I {”prompt_question”: ” ITシステム「Aシステム」では，どのポートを監視すべきですか。検索して教えてください。”} 0022 2023-11-17 10:20:33.000+0900 ai_assistant 12345 00001 KNBD20003-I {”intermediate_steps”: ”[{”AIMessage”: {”tool_call”: {”name”: ”user_contents”,”args”: {”__arg1”: ”Aシステム ポート監視”},”id”: ”toolu_bdrk_013ryNHoxrW7F4nPcnXQcsKC”}}},{”ToolMessage”: {”content”: ”sourcepage:Asystem_boot.pdf-page.1¥ncontent: A システムの起動手順<略>”,”name”: ”user_contents”,”tool_call_id”: ”toolu_bdrk_013ryNHoxrW7F4nPcnXQcsKC”}}] 0023 2023-11-17 10:20:34.000+0900 ai_assistant 12345 00001 KNBD20004-I {”final_answer”: ”検索結果から，Aシステムでは以下のポートの監視が必要であることがわかります。<略>

ログ名「生成 AI 連携 LangChain 思考過程ログ」の出力項目を以下に示します。

表 1–19　生成 AI 連携 LangChain 思考過程ログの出力項目

#	項目	説明
1	通番	ログの番号。
2	日時	ログの日時。形式は yyyy-MM-dd
3	時刻	ログの時刻。形式は hh:mm:ss.SSSZZZ
4	AP 名	アプリケーション識別名。ai_assistant を指定します。
5	プロセス ID	実行したプロセスの ID。
6	スレッド ID	実行したスレッドの ID。
7	メッセージ ID	本ログのメッセージの ID。ID の形式および具体的な ID は，「6.2.1 メッセージの概要」を参照してください。
8	メッセージテキスト	詳細なメッセージテキスト。出力される情報は「6.2.2 メッセージ一覧」を参照してください。
9	改行	LF

1.8　生成 AI 連携呼び出しコマンド

1.8.1　生成 AI 連携呼び出しコマンドの概要

生成 AI 連携の問い合わせ機能を呼び出すことができるコマンドです。

本コマンドは，JP1 Cloud Service システム管理の「自動アクションによるコマンド実行」または「対処アクションの自動実行と手動実行」を介して呼び出せます。

JP1 Cloud Service システム管理が提供するメール通知機能と，本コマンドを組み合わせることで，メールの通知文に生成 AI の回答を含めることができます。詳細は「1.8.3 使用例：メールに生成 AI の回答を含める」をご確認ください。

1.8.2　生成 AI 連携呼び出しコマンドの仕様

(1)　機能

生成 AI 連携呼び出しコマンドのコマンド名は"ai_assistant_cli"です。

コマンドライン引数で指定された質問文を生成 AI 連携基盤へ送信し，問い合わせ機能を用いて生成 AI から回答を取得します。

本コマンドにより生成 AI 連携基盤の問い合わせ機能を呼び出した場合，1.7 ログ出力機能により，チャット画面を介して問い合わせ機能を呼び出した場合と同様のログが出力されます。

注意事項

- 本コマンドを用いた生成 AI 連携基盤への問い合わせの同時実行上限数は 1 です。上限を超えて実行要求を行った場合，エラーとなります。
- 本コマンドで出力できる回答の最大文字数は 4096 文字です。生成 AI 連携基盤から 4096 文字以上の回答が返却された場合，4096 文字になるように打ち切って出力します。
- 本コマンドを利用する場合は環境変数ファイルを指定してください。指定方法は使用例をご確認ください。

(2)　入力形式

生成 AI 連携呼び出しコマンドは，以下のような形式で呼び出します。

```
ai_assistant_cli --query 〈質問文〉 --workflow 〈ワークフロー名〉
 --read_timeout 〈実行タイムアウト値〉 --user 〈ユーザー名〉
```

各引数の詳細な説明は以下の通りです。

```
--query 質問文　（必須）
  生成AI連携基盤に問い合わせを行う際の質問文を設定してください。
  指定可能な値：1～2048文字（UTF-8）
  注意: 本コマンド呼び出し元でエスケープが必要な文字が存在する場合は、適切にエスケープ処理を行い、想定通りの質問文が引数に指定されるようにしてください。例えば、シェルスクリプトで本コマンドの呼び出しを実装する場合は、実装の方式によってはダブルクォーテーションをエスケープして、シェルスクリプトとして正しく実装する必要があります。

--workflow ワークフロー名　（任意）
  生成AI連携基盤に問い合わせを行う際に使用するワークフロー名を設定してください。デフォルト値は knowledge_search_workflow です。
  注意: 運用管理データ取得エージェントを含むワークフローはサポートされません。指定されたワークフローに運用管理データ取得エージェントが含まれている場合、エラー(RC: 11)となります。
  指定可能な値：半角英数字（小文字）とアンダースコア　1～128文字（UTF-8）

--read_timeout タイムアウト値　（任意）
  生成AI連携基盤へのリクエストの読み取りタイムアウト値です。デフォルト値は150 秒です。
  注意: 150秒未満の値を指定した場合、read_timeout値に達した時点で生成AI連携基盤側では処理を継続している可能性があります。その場合、時間をおかずに再度生成AI連携呼び出しコマンドを実行すると、同時実行数の上限に達してエラー(RC: 14)になります。
  指定可能な値：30～600（UTF-8）

--user ユーザー名　（任意）
  生成AI連携基盤のログ上で、呼び出し元の識別情報として使用されるユーザー名です。
  デフォルト値はコマンド名 ai_assistant_cliです。
  指定可能な値：半角英数字（小文字）とアンダースコア　1～31文字（UTF-8）
```

(3)　出力形式

生成 AI 連携呼び出しコマンドは、以下のスキーマに従った JSON 形式の文字列（UTF-8）を標準出力(エラーが発生した場合は標準エラー出力)へ出力します。

```
{
    ”schema_version”: ”出力スキーマのバージョン”
    ”status”: ”実行ステータス”
    ”ai_response”: {
        ”invoke_id”: ”問い合わせ識別子”
        ”answer”: ”生成AIの回答”
    }
    ”error”: {
        ”message”: ”エラーメッセージ”
}
```

各パラメータの詳細な説明は以下の通りです。

#	パラメータ	データ型	必須／任意	範囲	説明
1	schema_version	文字列	必須	-	出力スキーマのバージョン情報です。
2	status	文字列	必須	-	実行ステータス。success,error のいずれかを返します。
3	ai_response	オブジェクト	任意	-	ai_response オブジェクト。 status:success の時のみ存在します。

#	パラメータ	データ型	必須／任意	範囲	説明
4	error	オブジェクト	任意	-	error オブジェクト。status:error の時のみ存在します。
ai_response オブジェクト					
5	invoke_id	文字列	必須	0 文字 <x<= 36 文字	生成 AI 連携基盤が発行する問い合わせ実行の識別子です。
6	answer	文字列	必須	0 文字 <=x<= 4096 文字	問い合わせの回答です。
error オブジェクト					
7	message	文字列	必須	0<x	エラー内容を示すメッセージです。

(4) エラーメッセージ

本コマンドにおいて，エラーが発生した場合のメッセージおよび終了ステータスは，以下の通りです。

#	メッセージ内容	終了ステータス	契機	想定される対処
1	生成 AI 連携呼び出しコマンドの引数が正しくありません。引数を見直してください。（RC: <戻り値>, 引数名: "<引数名>"）	10	コマンド引数が不正なとき	コマンド引数を見直してください。
2	生成 AI 連携基盤でエラーが発生しました。詳細は生成 AI 連携基盤のログを確認してください。（RC: <戻り値>, レスポンスタイムスタンプ: "<timestamp>", 生成 AI 連携基盤メッセージ（概要）: "<メッセージ>", ワークフロー: "<ワークフロー識別子>", 質問文: "<質問文>"）	11	生成 AI 連携基盤の問い合わせ機能で想定されるエラー（設定・認証エラー，外部サービスのエラー等）が発生したとき	生成 AI 連携基盤のログを確認し，生成 AI 連携基盤の設定または外部サービスの設定を見直してください。一時的なエラーである場合は、再度実行してください。
3	生成 AI 連携基盤の処理が規定時間内に完了しませんでした。使用するワークフローまたは質問文を確認してください。（RC: <戻り値>, レスポンスタイムスタンプ: "<timestamp>", ワークフロー: "<ワークフロー識別子>", 質問文: "<質問文>"）	12	生成 AI 連携基盤の処理がタイムアウトしたとき	使用するワークフローもしくは質問文を見直してください。
4	生成 AI 連携呼び出しコマンドに指定された読み取りタイムアウト時間内に生成 AI 連携基盤から応答が得られませんでした。（RC: <戻り値>, ワークフロー: "<ワークフロー識別子>", 読み取りタイムアウト値: <読み取りタイムアウト値>, 質問文: "<質問文>"）	13	生成 AI 連携呼び出しコマンドで指定した読み取りタイムアウト時間以内に生成 AI 連携基盤の処理が完了しなかったとき	質問文，使用するワークフロー，または読み取りタイムアウト値を見直してください。
5	生成 AI 連携呼び出し処理の同時実行数上限を超過しました。時間をおいて再度実行してください。（RC:<戻り値>, レスポンスタイムスタンプ: "<timestamp>", 質問文: "<質問文>"）	14	生成 AI 連携基盤への問い合わせの同時実行数上限を超過したとき	時間をおいて再度実行してください。必要に応じてチャット画面から問い合わせしてください。

#	メッセージ内容	終了ステータス	契機	想定される対処
6	生成 AI 連携基盤へのリクエストのバリデーションに失敗しました。引数に指定した値を見直してください。（RC:<戻り値>, レスポンスタイムスタンプ: "<timestamp>", ワークフロー: "<ワークフロー識別子>", ユーザー名: "<ユーザー名>", 読み取りタイムアウト値: <読み取りタイムアウト値>, 質問文: "<質問文>"）	15	生成 AI 連携基盤へのリクエストが不正な形式であるとき	リクエスト内容を見直してください。
7	生成 AI 連携呼び出しコマンドによる呼び出しに失敗しました。（RC: <戻り値>）	254	生成 AI 連携基盤の接続タイムアウトもしくは，接続先設定が不正なとき	時間をおいて再度実行してください。 頻発する場合は，サポート窓口にご連絡お願いします。
8	生成 AI 連携基盤で予期せぬエラーが発生しました。（RC: <戻り値>, 発生日時: "<レスポンスヘッダの Date>", ワークフロー: "<ワークフロー識別子>", 質問文: "<質問文>"）	255	生成 AI 連携基盤で想定されないエラーが発生したとき	時間をおいて再度実行してください。 頻発する場合は，サポート窓口にご連絡お願いします。
9	生成 AI 連携呼び出しコマンドで予期せぬエラーが発生しました。（RC: <戻り値>）	99	生成 AI 連携呼び出しコマンド内で想定されないエラーが発生したとき	時間をおいて再度実行してください。 頻発する場合は，サポート窓口にご連絡お願いします。

1.8.3　使用例：メールに生成 AI の回答を含める

JP1 Cloud Service システム管理が提供するメール通知機能と，本コマンドを組み合わせることで，メールの通知文に生成 AI の回答を含めるためのスクリプトの実装例を以下に示します。

スクリプトの配置方法は，「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」を参照してください。

スクリプト名：integrate_knowledge_search_and_mail_notification.sh

配置先：システム管理マネージャー上の専用ストレージ

配置先ディレクトリパス：/usrfile/

文字コード／改行コード：UTF-8／LF

```
#!/bin/bash

#-------------------------------------------------------------------------------
# Arguments         : Same arguments as the mail notification feature
# Exit status       : Success: 0
```

```
#                     Failure: 1
#-------------------------------------------------------------------------------

#-------------------------------------------------------------------------------
# Environment setup
#-------------------------------------------------------------------------------
set -u
set -o pipefail

#-------------------------------------------------------------------------------
# Variable initialization
#-------------------------------------------------------------------------------
RC_SUCCESS=0
RC_ERROR=1

MAIL_CMD="mail_notification.sh.x"
AI_CMD="ai_assistant_cli"

#-------------------------------------------------------------------------------
# Argument check
#-------------------------------------------------------------------------------
if [[ $# -ne 8 ]]; then
    usage >&2
    exit 1
fi

#-------------------------------------------------------------------------------
# Store arguments
#-------------------------------------------------------------------------------
EVSEV="$1"
EVIDBASE="$2"
EVSEQNO="$3"
EVHOST="$4"
EVDATE="$5"
EVTIME="$6"
EVMSG="$7"
GROUP_NAME="$8"

#-------------------------------------------------------------------------------
# Initialize processing variables
#-------------------------------------------------------------------------------
FLAG1=false
FLAG2=false
AI_ERR_MSG=""
AI_RC=0
MAIL_RC=0
AI_RESULT_TEXT=""
MERGED_MSG=""
AI_STDOUT=""

#-------------------------------------------------------------------------------
# Build query text
#-------------------------------------------------------------------------------
QUERY=$(printf "以下のメッセージを確認しました。運用手順を教えてください。¥n発生日時: ¥"%s¥"
¥nメッセージ: ¥"%s¥"¥n" ¥
        "$EVDATE" "$EVMSG")

#-------------------------------------------------------------------------------
```

```
# Invoke the ai_assistant command
#-------------------------------------------------------------------------------
AI_OUTPUT=$(
    "$AI_CMD" ¥
        --query "$QUERY" ¥
        2>&1
)
AI_RC=$?

#-------------------------------------------------------------------------------
# Evaluate the ai_assistant command result
#-------------------------------------------------------------------------------
if [[ $AI_RC -eq 0 ]]; then
    AI_RESULT_TEXT=$(printf '%s' "$AI_OUTPUT" | jq -r '.ai_response.answer')
else
    FLAG1=true
    AI_RESULT_TEXT=$(printf '%s' "$AI_OUTPUT" | jq -r '.error.message')
fi

#-------------------------------------------------------------------------------
# Invoke the mail notification script
# Set the 7th argument as a merge of the original message and ai_assistant command result
#-------------------------------------------------------------------------------
if ! $FLAG1; then
    EVMSG="${EVMSG}

     【生成AIの調査結果】
    ${AI_RESULT_TEXT}"
else
    EVMSG="${EVMSG}

     【生成AIの調査結果】
    生成AI連携の呼び出しに失敗しました。"
fi

"$MAIL_CMD" "$EVSEV" "$EVIDBASE" "$EVSEQNO" "$EVHOST" "$EVDATE" "$EVTIME" "$EVMSG" "$GROUP_N
AME" 2>&1
MAIL_RC=$?

#-------------------------------------------------------------------------------
# Evaluate the mail notification result
#-------------------------------------------------------------------------------
if [[ $MAIL_RC -ne 0 ]]; then
    FLAG2=true
fi

#-------------------------------------------------------------------------------
# Output handling
# Control stderr output and exit code based on FLAG1/FLAG2 combinations
#-------------------------------------------------------------------------------
case "${FLAG1}:${FLAG2}" in
    false:false)
        # Case1: no output, exit 0.
        exit $RC_SUCCESS
        ;;
    true:false)
        # Case2: ai_assistant command failed.
        echo "生成AI連携の呼び出しに失敗しました。" >&2
```

```
        echo ”エラー内容: ${AI_RESULT_TEXT} RC:${AI_RC}” >&2
        exit $RC_ERROR
        ;;
    false:true)
        # Case3: mail notification failed.
        echo ”メール通知に失敗しました。” >&2
        exit $RC_ERROR
        ;;
    true:true)
        # Case4: both failed.
        echo ”生成AI連携の呼び出しに失敗しました。” >&2
        echo ”エラー内容: ${AI_RESULT_TEXT} RC:${AI_RC}” >&2
        echo ”メール通知に失敗しました。” >&2
        exit $RC_ERROR
        ;;
esac
```

上記実装例を JP1 Cloud Service システム管理の「自動アクションによるコマンド実行」または「対処アクションの自動実行と手動実行」を介して呼び出すための方法を以下に示します。

1. 自動アクションによるコマンド実行を介して呼び出す方法

   JP1 のマニュアル「JP1 Version 13 統合管理 基本ガイド」における，「6.2 特定のイベントが発行されたら自動でコマンドを実行する方法」を参照し，コマンドが自動で実行されるように設定してください。

   ［アクション詳細設定］画面の「実行内容」に以下の通り指定してください。

#	項目名	指定内容	説明
1	実行ユーザー名	**JP1 ユーザー名**	作成した JP1 ユーザーを指定してください。
2	実行ホスト名	$ACTHOST	-
3	アクション	/usrfile/integrate_knowledge_search_and_mail_notification.sh ”$EVSEV” ”$EVIDBASE” ”$EVSEQNO” ”$EVHOST” ”$EVDATE” ”$EVTIME” ”$EVMSG” ”**グループ名**”	**グループ名**には，「group1」，「group2」，「group3」のどれかを指定してください。 これらのグループは，JP1 Cloud Service のヒアリングシートに記入された通知先メールアドレスのグループに対応付いています。
4	環境変数ファイル	/jp1cs/actenv.conf	-

2. 対処アクションの自動実行を介して呼び出す方法

   JP1 のマニュアル「JP1 Version 13 JP1/Integrated Management 3 - Manager 導入・設計ガイド」における，「7.　対処アクションの自動実行と手動実行（JP1/IM - Agent 連携）」を参照し，対処アクションの自動実行を介して呼び出せるように設定してください。

   統合オペレーション・ビューアーの定義ファイル操作機能を用いて，自動対処アクション定義ファイル（autoactconf.json）をダウンロードし，以下の例を参考に自動対処アクション定義ファイル（autoactconf.json）の編集を行い，編集した定義ファイルをアップロードしてください。

```
{
  ”meta”:{
    ”version”:”1”
  },
```

```
  "actions":[
    {"actionGroup":0,
     "actionId":"1",
     "label":"Errorイベント発生時生成AI調査通知",
     "description":"重大度Errorのイベントに対して自動対処アクション実行。",
     "valid":true,
     "conditions":[
       {"type":"event",
        "key":"E.SEVERITY",
        "ope":"IN",
        "val":["Error"]
       }
     ],
     "action":{
       "type":"cmd",
       "params":{
         "host":"${event:ACTHOST:}",
         "cmd":"/usrfile/integrate_knowledge_search_and_mail_notification.sh ¥"${event:EVSEV
:}¥" ¥"${event:EVIDBASE:}¥" ¥"${event:EVSEQNO:}¥" ¥"${event:EVHOST:}¥" ¥"${event:EVDATE:}¥"
¥"${event:EVTIME:}¥" ¥"${event:EVMSG:}¥" ¥"<グループ名>¥"",
         "envFile":"/jp1cs/actenv.conf"
       }
     }
    }
  ]
}
```

上記実装例と設定によって通知されるメールのタイトルや本文のサンプルは以下の通りです。

メールタイトル

```
[JP1 Cloud Service]システム管理イベント通知【EventID:イベントID No:イベント通し番号】
```

メール本文

```
--------------------------------------------------------------------------------
JP1 Cloud Serviceのシステム管理により本メールを
お送りしています。
--------------------------------------------------------------------------------

ご担当者様

システムがイベントを検知したため、JP1イベント情報をお送りいたします。

【重大度】重大度
【イベントID】イベントID
【発生元ホスト名】発生元ホスト名
【発生日時】発生日時

【メッセージ】
メッセージ内容

【生成AIの調査結果】
生成AIの調査結果
```

```
==============================================
このメールは送信専用となります。
不明な点がございましたら、JP1 Cloud Serviceの
問い合わせ窓口にお問い合わせください。
==============================================

**********************************************************
このメールは、特定の方に送信することを意図したものです。
お心当たりのない方が受信された場合には、誠に申し訳
ございませんが、問い合わせ窓口メールアドレスへ
メールで連絡の上で、受信したメールは削除して頂きますよう
お願い致します。
**********************************************************
```

2 JP1 Cloud Service 生成 AI 連携の構築

JP1 Cloud Service 生成 AI 連携の構築手順について説明しています。

2.1　事前に準備するもの

JP1 Cloud Service 生成 AI 連携を利用するにあたっての事前準備について説明します。

2.1.1　システム管理の準備

生成 AI 連携を使用する場合，システム管理を利用するための環境を構築する必要があります。お客様でシステム管理の利用に必要な環境を構築してください。

システム管理の構築方法については，「JP1 Cloud Service システム管理 利用ガイド」における「2. システム管理を利用するための構築」を参照してください。

システム管理を両現用構成で利用する場合，生成 AI 連携機能の定義ファイルの管理及び同期はお客様側で対応する必要があります。システム管理の両現用構成は「JP1 Cloud Service システム管理 利用ガイド」における「7.　システム管理の両現用構成」を参照してください。

2.1.2　お客様が用意するもの

お客様が用意する必要のあるものを以下に示します。

(1)　連携サービス

本プログラムを利用する際に事前準備が必要な構成要素を以下に示します。生成 AI サービス（Azure OpenAI Service，Amazon Bedrock 等）で提供されるモデルの提供終了（EoL）は，各サービス提供元が決定・通知します。EoL を考慮したモデルの選定・更新・運用は利用者の責任で実施し，必ず提供元の最新ドキュメントをご確認ください。

なお，各連携サービスの利用料金につきましては，お客様負担となります。

表 2–1　本プログラムで必要な構成要素

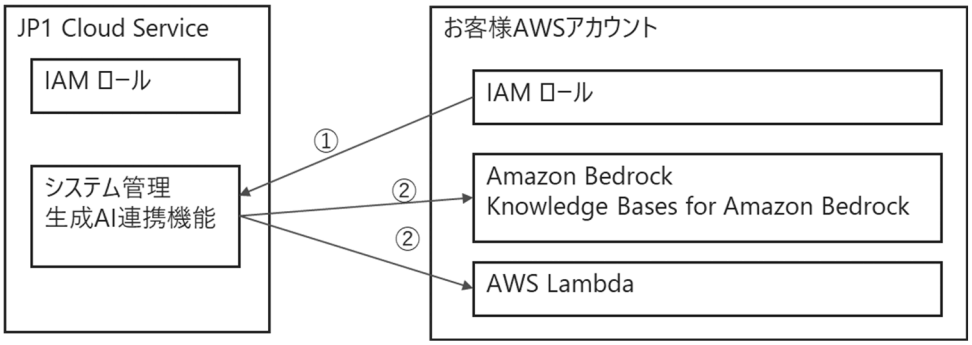
#	構成要素（顧客準備物）	必須／任意	説明
1	Azure OpenAI Service	#3 とどちらか片方が必須	本プログラムが連携する生成 AI サービスとして必要です。GPT 系が利用可能です。 なお，利用料金につきましては，トークン数によって決定いたしますが，お客様が送信された問い合わせ本文に加え，プロンプト作成機能により追加された文章が，トークンの対象となりますので，ご了承ください。利用料金の詳細は，「付録 A 連携サービス」を参照してください。 利用可能な生成 AI サービスの詳細は 6.1.1 設定ファイル（ai_assistant.yml）の generative_ai.azure[].models[].model_name を参照してください。

#	構成要素（顧客準備物）	必須／任意	説明
2	Azure AI Search	任意	本プログラムが連携する検索サービスです。「1.5.4 ユーザーコンテンツ取得機能」を使用して，生成 AI にユーザーコンテンツを参照させて回答させる場合に必要です。本プログラムで指定するスキーマのインデックスを作成する必要があります。利用料金の詳細は，「付録 A 連携サービス」を参照してください。
3	Amazon Bedrock	#1 とどちらか片方が必須	本プログラムが連携する生成 AI サービスとして必要です。Claude 系が利用可能です。 なお，利用料金につきましては，トークン数によって決定いたしますが，お客様が送信された問い合わせ本文に加え，プロンプト作成機能により追加された文章が，トークンの対象となりますので，ご了承ください。利用料金の詳細は，「付録 A 連携サービス」を参照してください。 利用可能な生成 AI サービスの詳細は 6.1.1 設定ファイル（ai_assistant.yml）の generative_ai.aws[].models[].model_id を参照してください。
4	Knowledge Bases for Amazon Bedrock	任意	本プログラムが連携する検索サービスです。「1.5.4 ユーザーコンテンツ取得機能」を使用して，生成 AI にユーザーコンテンツを参照させて回答させる場合に必要です。本プログラムで指定するスキーマのインデックスを作成する必要があります。利用料金の詳細は，「付録 A 連携サービス」を参照してください。 Amazon Bedrock の Rerank モデルを使用して Rerank を行う場合は、Knowledge Bases for Amazon Bedrock サービスロールに「InvokeModel」及び「Rerank」アクションを有効にした許可ポリシーを追加でアタッチする必要があります。
5	Identity and Access Management（IAM）ロール	#3，#4 または #6 を使う場合必須	本プログラムはお客様の AWS リソースとロールを使用したクロスアカウントアクセスにより連携します。連携するために必要な認証・認可リソースとして IAM ロールが必要です。IAM ロールに付与する許可ポリシーは，Amazon Bedrock を使用する場合は「InvokeModel」，Knowledge Bases for Amazon Bedrock を使用する場合は「Retrieve」，AWS Lambda を使用する場合は「InvokeFunction」アクションを有効にしてください。また，前述した IAM ロールの信頼するエンティティに JP1 Cloud Service の AWS アカウントからの AssumeRole アクションを許可してください。
6	AWS Lambda	任意	「1.5.6　ユーザープログラム呼び出し機能（02-40 以降）」を使用して，ユーザープログラムを実行する場合に必要です。本プログラムで指定する入出力形式のプログラムをデプロイする必要があります。

AWS サービスへの接続は次の通り行います。

- 生成 AI 連携機能は JP1 Cloud Service の AWS アカウントに紐づく IAM ロールを使用して AssumeRole アクションを行い，お客様 AWS アカウントの IAM ロールを引き受けます（(1)）。
- 引き受けたお客様 AWS アカウントの IAM ロールを利用して，Amazon Bedrock や AWS Lambda などのサービスを実行します（(2)）。

「表 2-1 本プログラムで必要な構成要素」の#5 IAM ロールは、このお客様 AWS アカウントの IAM ロールを示します。



Azure の各サービスとプライベートエンドポイントを利用したプライベート通信を行なう場合は，システム管理マネージャーとインターネット VPN 接続が可能な状態である必要があります。設定方法の詳細は「JP1 Cloud Service ジョブ管理・システム管理 導入ガイド」の「1.2.3 接続の手順（Azure の場合）」を参照してください。

AWS の各サービスとプライベートエンドポイントを利用したプライベート通信を行なう場合はシステム管理マネージャーと AWS Transit Gateway を経由した接続が可能な状態である必要があります。設定方法の詳細は「JP1 Cloud Service ジョブ管理・システム管理 導入ガイド」の「1.2.2 接続の手順（AWS の場合）」を参照してください。

また，本機能において Azure の各サービスとのプライベート通信を実現するためには，システム管理マネージャーに対して設定を施す必要があります。本機能をご利用中に以下のいずれかを行いたい場合は，サービス窓口へお問い合わせください。

- プライベート通信からパブリック通信へ変更したい。 ※構成要素サービスの URL 自体が変わる場合，お問い合わせは不要です。
- 新規にプライベート通信を設定したい。
- 既存のプライベート通信の IP アドレスを変更したい。

本機能で AWS の各サービスとの認証を実現するためには，システム管理マネージャーに対して設定を施す必要があります。本機能をご利用中に以下の変更を行いたい場合は，サービス窓口へお問い合わせください。

- AWS サービスを利用するアカウント ID を追加／変更したい。

AWS の各サービスと接続するために利用するポート番号の一覧を次に示します。

#	ポート	通信元	通信先	用途
1	443/tcp	システム管理マネージャー※	プライベートエンドポイント	各サービスへの API 実行用。利用するアクションは「表 2-1 本プログラムで必要な構成要素」#5 に示す通りです。

注※

サブ機能用 IP アドレスも許可してください。サブ機能用 IP アドレスについては，「ご利用環境情報（ジョブ管理・システム管理）」をご確認ください。

2.1.3　インデックスの作成（任意）

Azure AI Search または Knowledge Bases for Amazon Bedrock と連携する場合はインデックスの作成が必要です。連携しない場合はこの操作は不要です。検索サービスのインデックスを，以降に記載されたスキーマに沿って作成してください。

インデックスの作成については，作成手順例をまとめたサンプルファイルを「JP1CloudService-生成 AI 連携_サンプルコンテンツ_インデックス作成手順例.zip」にまとめて提供します。サンプルファイルは JP1 Cloud Service の提供ファイルのダウンロード機能を利用してダウンロードできます。ダウンロード方法については「JP1 Cloud Service ジョブ管理・システム管理 サービスポータル 利用ガイド」における「3.2 提供ファイル一覧」を参照してください。

(1)　Azure AI Search のインデックス

生成 AI 連携基盤がユーザーコンテンツの検索先として使用する Azure AI Search のインデックスに関して，必要なインデックス内のフィールドを以下に示します。

表 2–2　Azure AI Search のインデックスの項目

#	フィールド名	説明	設定例	文字数の目安※1
1	id	インデックス内で検索ドキュメントを一意に定めるキーフィールド	"MDAwMQ2"	–
2	content	検索対象の中身をテキスト化して格納するフィールド	"JP1は日立オープンミドルウェアシリーズのひとつです。"	600 文字
3	sourcepage	#2 の元ファイルにおける記載箇所も含めて引用元を表すフィールド。 生成 AI 連携による回答に Azure AI Search から得られた情報を用いる場合，引用元の情報として，本項目の設定内容が使用されます。	"Example.pdf-page.100"	64 文字※2
4	sourcefile	#2 の元ファイルを識別するためのフィールド。 生成 AI 連携による回答に Azure AI Search から得られた情報を用いる場	"Example.pdf"	64 文字※2

#	フィールド名	説明	設定例	文字数の目安※1
4	sourcefile	合，引用元の情報として，本項目の設定内容が使用されます。	"Example.pdf"	64 文字※2
5	document_url	#2 の元ファイルへアクセスする URL（http または https）を格納するフィールド。 生成 AI 連携による回答に Azure AI Search から得られた情報を用いる場合，引用元へのアクセス情報として，本項目の設定内容が使用されます。	'"https://example_file_storage.XXX/Example.pdf?token=YYY" ※3	500 文字※4

注※1

生成 AI 連携として定める文字数の目安を記載しています。文字数が多くなるほど，生成 AI がトークンを消費するため，目安を参考にしてインデックスを作成してください。Azure AI Search としての制限は，Microsoft 社が公開するドキュメントを参照してください。

注※2

半角英数字と"."（ドット），"-"（ハイフン），"_"（アンダーバー）が使用可能です。

注※3

URL の格納は任意となります。URL を格納しない場合は，"None"という文字列を格納してください。

注※4

URL エンコードされた文字列が使用可能です。

また，各フィールドの属性については以下のように設定してください。

表 2–3　インデックスのフィールド属性

#	フィールド名	キー	型	取得可能	フィルタ可能	ソート可能	ファセット可能	検索可能	アナライザ
1	id	■	String	■	□	□	□	□	–
2	content	□	String	■	□	□	□	■	日本語 - Microsoft※5
3	sourcepage	□	String	■	■	□	■	□	–
4	sourcefile	□	String	■	■	□	■	□	–
5	document_url	□	String	■	□	□	□	□	–

注※5

生成 AI 連携基盤は日本語のみサポート対象としています。

(2) Knowledge Bases for Amazon Bedrock のインデックス

生成 AI 連携基盤がユーザーコンテンツの検索先として使用する Knowledge Bases for Amazon Bedrock のインデックスに関して，必要なインデックス内のフィールドを以下に示します。

表 2–4 Knowledge Bases for Amazon Bedrock のインデックスの項目

#	フィールド名	説明	設定例	文字数の目安※1
1	id	インデックス内で検索ドキュメントを一意に定めるキーフィールド。 Knowledge Bases for Amazon Bedrock がデータソースと同期する際にシステム側で自動で入力されます。	"0d282b7f-aae0-47e4-90e8-d602edbfb0ba"	–
2	content	検索対象の中身をテキスト化して格納するフィールド。 Knowledge Bases for Amazon Bedrock がデータソースと同期する際にシステム側で自動で入力されます。	"JP1は日立オープンミドルウェアシリーズのひとつです。"	600 文字
3	sourcepage	#2 の元ファイルにおける記載箇所も含めて引用元を表すフィールド。 生成 AI 連携による回答に Knowledge Bases for Amazon Bedrock から得られた情報を用いる場合，引用元の情報として，本項目の設定内容が使用されます。	"Example.pdf-page.100"	64 文字※2
4	sourcefile	#2 の元ファイルを識別するためのフィールド。 生成 AI 連携による回答に Knowledge Bases for Amazon Bedrock から得られた情報を用いる場合，引用元の情報として，本項目の設定内容が使用されます。	"Example.pdf”	64 文字※2
5	document_url	#2 の元ファイルへアクセスする URL（http または https）を格納するフィールド。 生成 AI 連携による回答に Knowledge Bases for Amazon Bedrock から得られた情報を用いる場合，引用元へのアクセス情報として，本項目の設定内容が使用されます。	"https://example_file_storage.XXX/Example.pdf?token=YYY"※3	500 文字※4
6	ベクトル値格納用フィールド	#2 をベクトル化した値を格納するためのフィールド。	[0.009516553,0.018477395,0.008…]	–

#	フィールド名	説明	設定例	文字数の目安※1
6	例. embeddings	Knowledge Bases for Amazon Bedrock がデータソースと同期する際にシステム側で自動で入力されます。	[0.009516553,0.018477395,0.008…]	–
7	Bedrock メタデータ格納用フィールド 例. bedrock-metadata	Knowledge Bases for Amazon Bedrock に関連するメタデータを保存するためのフィールド。 Knowledge Bases for Amazon Bedrock がデータソースと同期する際にシステム側で自動で入力されます。	"""{”source”:”s3://knowledgebase/Example.pdf”}"""	–

注※1

生成 AI 連携として定める文字数の目安を記載しています。文字数が多くなるほど，生成 AI がトークンを消費するため，目安を参考にしてインデックスを作成してください。Knowledge Bases for Amazon Bedrock としての制限は，Amazon 社および連携先のベクタストアベンダーが公開するドキュメントを参照してください。

注※2

半角英数字と"."（ドット），"-"（ハイフン），"_"（アンダーバー）が使用可能です。

注※3

URL の格納は任意となります。URL を格納しない場合は，"None"という文字列を格納してください。

注※4

URL エンコードされた文字列が使用可能です。

また，各フィールドについては以下のように設定してください。

- 必要なインデックス内のフィールドは，全て String 型としてください。
- 必要なインデックス内のフィールドは，ベクタストアから値を取得できるものとしてください。

2.2　セットアップ

JP1 Cloud Service 生成 AI 連携のセットアップ方法について説明します。

2.2.1　生成 AI サービスとの連携設定

生成 AI サービスと連携するために接続情報を設定する必要があります。Azure OpenAI，Amazon Bedrock との連携時に必要な情報をそれぞれ以下の表に示します。なお，02-30 以降は連携する生成 AI サービスを複数定義し，利用ケースに合わせて指定する事ができます。

表 2–5　Azure OpenAI との連携に必要な情報

#	項目	内容
1	Azure OpenAI について必要な情報	• Azure OpenAI の API のエンドポイント • Azure OpenAI の API の種類 • Azure OpenAI の API のバージョン • Azure OpenAI のデプロイ名 • Azure OpenAI のモデル名
2	Azure OpenAI との認証方法	• Azure Open AI の API キー

表 2–6　Amazon Bedrock との連携に必要な情報

#	項目	内容
1	Amazon Bedrock について必要な情報	• Amazon Bedrock を実行する IAM ロールの ARN • Amazon Bedrock の API の種類 • Amazon Bedrock のモデル ID またはクロスリージョン推論のプロファイル ID（02-40 以降） • Amazon Bedrock のリージョン

(1)　接続設定

1. 生成 AI サービスについて必要な情報の設定

   設定ファイル（ai_assistant.yml）の以下の項目に必要な情報を設定します。

   - Azure OpenAI を使用する場合：

     generative_ai.azure に Azure OpenAI について必要な情報を設定します。

   - Amazon Bedrock を使用する場合：

     generative_ai.aws に Amazon Bedrock について必要な情報を設定します。

   設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。不要な生成 AI サービスの設定（generative_ai.azure[]/aws[]）はコメントアウトまたは削除してください。

2. Azure OpenAI との認証の設定（Azure OpenAI と連携する場合のみ）

連携する生成 AI サービスとの認証のために API キーの登録が必要です。API キーは生成 AI 連携のチャット画面からシークレットとして登録します。シークレットの登録方法は「3.1.2 シークレットの追加・変更・削除方法」を参照してください。

登録する内容を以下の表に示します。

表 2–7　登録する Azure OpenAI のシークレット

#	シークレットキー	シークレット
1	ai_assistant.generative_ai.azure.<リソース名（小文字）>	Azure portal の Azure OpenAI から該当リソースを選択し，［リソース管理］-［Key and Endpoint］に表示されるキー 1 もしくはキー 2 の値

(2)　利用設定（02-30 以降）

接続設定を行った生成 AI サービスをワークフロー内で利用するため，設定ファイル（ai_assistant.yml）に指定します。ワークフローの設定及びワークフロー内で利用する生成 AI サービスの設定は「2.2.3 ワークフローの設定（02-30 以降）」を参照してください。

(3)　注意事項（02-60 以降）

02-60 以降は，利用可能なモデルは，Azure OpenAI の Responses API に対応しているリージョンにデプロイされたモデルに限ります。詳細は，Microsoft 社が公開しているドキュメントを参照してください。

2.2.2　検索サービスの接続設定

検索サービスを利用する場合と利用しない場合それぞれで必要な作業があります。必要な作業を以降に示します。

(1)　検索サービスを利用する場合

検索サービスと連携するために接続情報を設定する必要があります。Azure AI Search, Knowledge Bases for Amazon Bedrock との連携時に必要な情報をそれぞれ以下の表に示します。

表 2–8　Azure AI Search との連携に必要な情報

#	項目	内容
1	Azure AI Search について必要な情報	• Azure AI Search の API のエンドポイント • Azure AI Search のインデックス名
2	Azure AI Search との認証方法	• Azure AI Search の API キー

表 2–9　Knowledge Bases for Amazon Bedrock との連携に必要な情報

#	項目	内容
1	Knowledge Bases for Amazon Bedrock について必要な情報	• Knowledge Bases for Amazon Bedrock の ID • Knowledge Bases for Amazon Bedrock を実行する IAM ロールの ARN • Amazon Bedrock のリージョン

1. 検索サービスについて必要な情報の設定

   設定ファイル（ai_assistant.yml）の以下の項目に必要な情報を設定します。

   - Azure AI Search を使用する場合：

     service.langchain.tools[].azure_ai_search に Azure AI Search について必要な情報を設定します。

   - Knowledge Bases for Amazon Bedrock を使用する場合：

     service.langchain.tools[].aws_knowledge_bases に Knowledge Bases for Amazon Bedrock について必要な情報を設定します。

   Azure AI Search または Knowledge Bases for Amazon Bedrock のどちらか片方のみを使用する場合は以下の設定も必要です。

   - Azure AI Search のみ使用する場合：

     設定ファイルの service.langchain.tools[]のうち name が user_contents_for_aws_kb の設定をコメントアウトしてください。

   - Knowledge Bases for Amazon Bedrock のみ使用する場合：

     設定ファイルの service.langchain.tools[]のうち name が user_contents の設定をコメントアウトしてください。

   また，設定ファイルの service.langchain.prompts[].tools で指定する値が service.langchain.tools[]の name で指定した値のみになるように変更をしてください。

   設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

2. Azure AI Search との認証の設定（Azure AI Search と連携する場合のみ）

   連携する検索サービスとの認証のために API キーの登録が必要です。API キーは生成 AI 連携のチャット画面からシークレットとして登録します。シークレットの登録方法は「3.1.2 シークレットの追加・変更・削除方法」を参照してください。

   表 2–10　登録する Azure AI Search のシークレット

#	シークレットキー	シークレット
1	ai_assistant.service.langchain.tool.< name>※	Azure portal の Azure AI Search から該当リソースを選択し，［設定］-［Keys］に表示される Primary admin key もしくは Secondary admin key の値

   注※

   <name>には設定ファイル（ai_assistant.yml）で Azure AI Search を指定した tool の name を指定します。

(2)　複数のナレッジを利用する場合（02-40 以降）

検索サービスの複数のナレッジ（Azure AI Search の場合はインデックス，Knowledge Bases for Amazon Bedrock の場合はナレッジ ID）を横断して検索したい場合は複数の検索ツール（Azure AI Search 検索ツールまたは Knowledge Bases for Amazon Bedrock 検索ツール）の定義と利用のための設定が必要になります。

1. 検索ツールのツールプロンプト定義

   ツールプロンプト定義ファイルにツールの名前，説明を定義します。ナレッジ検索エージェントはツールの説明を参照して，実行すべきツールを決定し，検索クエリを作成するため，生成 AI がツールを選ぶための情報と検索クエリの作成方法を記載する必要があります。また，利用するナレッジの数だけ作成する必要があります。

   Azure AI Search 検索ツールを利用する場合はツールプロンプト定義ファイルの azure_ai_search[].name が user_contents を，Knowledge Bases for Amazon Bedrock 検索ツールを利用する場合はツールプロンプト定義ファイルの azure_ai_search[].name が user_contents_for_aws_kb を参考にしてください。

2. 検索ツールの設定

   設定ファイルの service.langchain.tools[]に検索ツールを定義します。ツール名（service.langchain.tools[].name）はツールプロンプト定義ファイルで定義したツール名と一致させる必要があります。Azure AI Search 検索ツールの場合は service.langchain.tools[].azure_ai_search に，Knowledge Bases for Amazon Bedrock 検索ツールの場合は service.langchain.tools[].aws_knowledge_bases にそれぞれ接続情報を定義します。

3. ナレッジ検索エージェントの設定

   ナレッジ検索エージェントをワークフローに定義します。ワークフローの定義は「2.2.3　ワークフローの設定（02-30 以降）」を参照してください。

4. ナレッジ検索エージェントと検索ツールの紐づけ

   ナレッジ検索エージェントが検索ツールを参照するために，ナレッジ検索エージェントのプロンプトで，利用できる検索ツールを指定します。ナレッジ検索エージェントのプロンプト（service.langchain.prompts[]の name が knowledge_search_agent）の tools に 1 で定義した検索ツールの名前を列挙します。

5. 不要なデフォルト設定の削除

   デフォルトで設定ファイルに定義されている検索ツール（service.langchain.tools[].name==user_contents および service.langchain.tools[].name==user_contents_for_aws_kb）が不要な場合はコメントアウトまたは削除してください。

(3)　検索サービスを利用しない場合

- 02-30 以降

設定ファイルのワークフローの定義の Supervisor 配下のエージェントノード（service.langgraph.workflows[].nodes.supervisor.child_agents[]）から knowledge_search_agent を削除してください。また，不要な検索サービスの設定（azure_ai_search または aws_knowledge_bases を指定しているツール定義(service.langchain.tools[])）はコメントアウトまたは削除してください。

- 02-20 以前

検索サービスを利用しない場合は，設定ファイル（ai_assistant.yml）の service.langchain.prompts.tools と service.langchain.tools を変更する必要があります。

service.langchain.prompts.tools には空文字を指定します。以下に例を示します。

```
service:
  langchain:
    prompts:
      - name: sys_chat_prompt
        type: chat
        tools: ””
```

また，service.langchain.tools に記載されている内容をコメントアウトする必要があります。

設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

2.2.3　ワークフローの設定（02-30 以降）

質問への回答に利用するワークフローを設定ファイル（ai_assistant.yml）へ設定します。ワークフローを定義するには下記手順に沿って定義を行います。

(1)　利用するワークフローの定義

チャット画面で選択できるワークフローは，設定ファイルの service.langgraph.workflows[]に定義します。「1.5.2(2)(c)　ワークフロー」に示すワークフローの設定内容は「付録 D　ワークフローの定義例」を参照してください。

(2)　ワークフローで利用するオペレーターノードの生成 AI サービスの定義

ワークフローのノードごとに利用する生成 AI サービスを定義します。接続設定を行った生成 AI サービス（generative_ai.azure[].models[]または generative_ai.aws[].models[]）の name 情報をそれぞれ以下に設定します。

- Secretary に利用する生成 AI サービスを指定する場合
  service.langgraph.workflows[].secretary.llm_name
- Supervisor に利用する生成 AI サービスを指定する場合

service.langgraph.workflows[].supervisor.llm_name

(3)　ワークフローで利用するエージェントノードの生成 AI サービス定義

ワークフロー内で利用するエージェントノードが利用する生成 AI サービスを定義します。エージェントノードに利用する生成 AI サービスはすべてのワークフローで同一の設定です。接続設定を行った生成 AI サービスの name 情報をそれぞれ以下に設定します。

- ユーザーコンテンツ取得を行うナレッジ検索エージェントに指定する場合
  service.langgraph.agents.knowledge_search_agent.llm_name
- 運用管理データ取得機能を行う運用管理データ取得エージェントに指定する場合
  service.langgraph.agents.event_analysis_agent.llm_name
- ユーザープログラム呼び出し機能を行うカスタムエージェントに指定する場合
  service.langgraph.agents.custom_agent.llm_name

(4)　ワークフローのノードが利用するプロンプトの定義（任意）

ワークフローのノードが利用するプロンプトをデフォルト値から変更したい場合は「3.1.8 プロンプトの設定（任意）（02-30 以降）」を参照し，変更してください。

2.2.4　質問文生成機能の設定（任意）

質問文生成機能により生成する質問文を変更する場合，設定ファイル（ai_assistant.yml）の ui.template の設定変更が必要になります。設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。以下に例を示します。

設定ファイル

```
ui:
  template: ”{jp1events}¥n状況を教えてください。”
  jp1event_template: ”イベントID：{B.ID}が{B.TIME}に{E.JP1_SOURCEHOST}で発生。内容は「{B.MES
SAGE}」”
```

選択した JP1 イベント

イベント ID	登録時刻	発生元ホスト名	メッセージ
00007620	2024-03-12T07:04:03Z	HOSTA	CPU 使用率が 90%を超えました
00007621	2024-03-12T08:03:03Z	HOSTB	CPU 使用率が 80%を超えました

質問エリアに表示される内容

```
イベントID：00007620が2024-03-12T07:04:03ZにHOSTAで発生。内容は「CPU使用率が90%を超えました」
イベントID：00007621が2024-03-12T08:03:03ZにHOSTBで発生。内容は「CPU使用率が80%を超えました」
状況を教えてください。
```

なお，テンプレートの中で，"B.TIME"を指定した場合，質問エリアには以下のタイムゾーンの時刻が表示されます。

（02-40 以降）JST

（02-30 以前）UTC

上記の例は 02-30 以前の表示例であり，02-40 以降は，JST に変換された値が表示されます。このタイムゾーンは変更できません。

2.2.5　質問文の文字数の設定（任意）

チャット機能で入力可能な質問文の文字数を変更する場合，設定ファイル（ai_assistant.yml）の ui.chat.max_size の設定変更が必要になります。設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

この設定を変更すると入力可能な文字数を変更できますが，生成 AI にできる質問の文字数は連携する生成 AI サービスのモデルごとに決められた最大トークン数によって異なります。質問文で入力可能なトークン数の最大値は，「モデルの最大トークン数-質問文に追加するプロンプト-2048（AI の回答に許容する文字数）」になり，それより小さい値に設定する必要があります。

2.2.6　ログの追加出力の設定（任意）

ユーザーが入力した質問，生成 AI の回答，ユーザーコンテンツの検索結果に関するログを追加で出力したい場合は，設定ファイル（ai_assistant.yml）の logs.filter の設定変更が必要になります。設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

この設定の対象となるログは，「6.2.1(2) メッセージの出力先」をご確認ください。また，出力を許可する設定にした場合，JP1 Cloud Service にユーザーが入力した質問，生成 AI の回答，ユーザーコンテンツの検索結果に関するログが残ることに注意してください。

2.2.7　任意のリンク作成の設定（任意）

チャット画面のチャットエリア内の生成 AI の回答に任意の URL を作成し，ハイパーリンクとして表示することが可能になります。任意のリンクを作成する場合は，URL 作成が必要な条件と作成すべき情報を生成 AI に指示として与えるためにプロンプト定義ファイル（prompts.chat.ja.yml），ハイパーリンク化する対象の文字列とその置換先文字列の設定と生成 AI への問い合わせで利用するプロンプトの名前を設定するために設定ファイル（ai_assistant.yml）の ui.chat.hyperlink と service.langchain.prompts.[0].name の設定変更がそれぞれ必要になります。プロンプト定義ファイルと設定ファイルの変更方法は「3.1.1 定義ファイルの変更方法」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。プロンプト定義ファイルの詳細は「6.1.2 プロンプト定義ファイル（prompts.chat.ja.yml）」を参照してください。

例として，発生した問題から必要に応じてチケットを作成する URL を，回答に含み，ハイパーリンクで表示するための定義例を示します。

(1)　02-30 以降

例）チケット作成 URL の生成に必要な定義例

プロンプト定義ファイル

- プロンプトでは URL 作成が必要な条件と作成すべき情報を，指示として与えます。

  本例では，チケット作成用 URL を出力に含めるように指示を追加したプロンプトをプロンプト定義ファイル(prompts.chat.ja.yml)のデフォルト値として定義している name: secretary_as_ticket_creator を利用します。プロンプト定義ファイルのデフォルト値の詳細は「付録 B プロンプト」を参照してください。instruction として定義している箇所が該当の指示になります。

設定ファイル

- ハイパーリンク化する対象の文字列とその置換先文字列を定義します。

  プロンプトで指定した出力形式を，設定ファイル(ai_assistant.yml)の ui.chat.hyperlink に追加します。本例のプロンプトの場合は以下のように定義します。

  ```
  regexp: ”【link:¥¥s*(https?:¥¥/¥¥/[^¥¥s]*)】”
  replace: ”【link: LINK】”
  ```

- 生成 AI への問い合わせで利用するワークフローを指定します。

  設定ファイル(ai_assistant.yml)の service.langgraph.workflows[]に Secretary のみのワークフローを定義します。また，このワークフローで利用するプロンプトを service.langchain.prompts[]に定義し，applied_node に定義したワークフローを指定します。

(2)　02-20 以前

例）チケット作成 URL の生成に必要な定義例

プロンプト定義ファイル

- プロンプトでは URL 作成が必要な条件と作成すべき情報を，指示として与えます。

  本例では，チケット作成用 URL を出力に含めるように指示を追加したプロンプトをプロンプト定義ファイル（prompts.chat.ja.yml）のデフォルト値として定義している name: sys_chat_prompt_for_link を利用します。プロンプト定義ファイルのデフォルト値の詳細は「付録 B.3　02-20 以前のプロンプトのデフォルト」を参照してください。name=link として定義している箇所が該当の指示になります。

設定ファイル

- ハイパーリンク化する対象の文字列とその置換先文字列を定義します。

  プロンプトで指定した出力形式を，設定ファイル（ai_assistant.yml）の ui.chat.hyperlink に追加します。本例のプロンプトの場合は以下のように定義します。

```
regexp: ”【link:¥¥s*(https?:¥¥/¥¥/[^¥¥s]*)】”
replace: ”【link: LINK】”
```

- 生成 AI への問い合わせで利用するプロンプトの名前を指定します。

  設定ファイル（ai_assistant.yml）の service.langchain.prompts.[0].name をプロンプト定義で指定した名前（本例では sys_chat_prompt_for_link）に変更します。

2.2.8　ユーザープログラムの設定（任意）（02-40 以降）

プログラム実行サービスと連携するために接続情報を設定する必要があります。AWS Lambda との連携時に必要な情報を以下の表に示します。

表 2–11　AWS Lambda との連携に必要な情報

#	項目	内容
1	AWS Lambda について必要な情報	• AWS Lambda にデプロイされた Lambda 関数の識別子 • Lambda 関数を実行する IAM ロールの ARN • AWS Lambda のリージョン

ユーザープログラム呼び出し機能で利用するユーザープログラムを設定ファイル（ai_assistant.yml）およびプロンプト定義ファイル（prompt.chat.ja.yml），ツールプロンプト定義ファイル（tools.ja.yml）に設定情報を定義します。

定義例をまとめたサンプルファイルを「JP1CloudService-生成 AI 連携_サンプルコンテンツ_ユーザープログラム定義例.zip」にまとめて提供します。

1. ユーザープログラムのツールプロンプト定義

   ツールプロンプト定義ファイルにツールの名前，説明，引数を定義します。カスタムエージェントはツールの説明を参照して，実行すべきツールを決定するため，生成 AI がツールを選ぶための情報を記載する必要があります。

引数はユーザープログラムの実装に合わせた設定にする必要があります。定義例のサンプルは，「JP1CloudService-生成 AI 連携_サンプルコンテンツ_ユーザープログラム定義例.zip」の Lambda 関数と tools.ja.yml ファイルを参照してください。複数のツールを定義する場合は複数定義してください。

2. ユーザープログラムを実行するユーザープログラム呼び出しツールの設定

   設定ファイルの service.langchain.tools[]にユーザープログラム呼び出しツールを定義します。ツール名（service.langchain.tools[].name）はツールプロンプト定義ファイルで定義したツール名と一致させる必要があります。ユーザープログラム呼び出しツールの情報（service.langchain.tools[].aws_lambda_function）には呼び出す対象のユーザープログラムの情報を指定します。(1)に合わせて複数のツールを定義する場合は複数定義してください。

3. カスタムエージェントの設定

   設定ファイルの service.langgraph.workflows[].nodes.supervisor.child_agents.custom_agent にカスタムエージェントを定義します。ワークフローの定義は「2.2.3　ワークフローの設定（02-30 以降）」を参照してください。

4. カスタムエージェントのプロンプト設定

   プロンプト定義ファイルに，(3)で定義したカスタムエージェントのプロンプトを定義します。プロンプトの定義例は「JP1CloudService-生成 AI 連携_サンプルコンテンツ_ユーザープログラム定義例.zip」の prompts.chat.ja.yml を参照してください。

5. カスタムエージェントとプロンプトの紐づけ

   設定ファイルの service.langchain.prompts[]にカスタムエージェントが利用するプロンプトを定義します。この時，以下が一致する必要があります。

   - service.langchain.prompts[].name はプロンプト定義ファイルの custom_agent[].name を指定
   - service.langchain.prompts[].tools は(1)および(2)で定義したツール名を指定します。複数のツールを一つのカスタムエージェントに紐づける場合は複数のツール名をカンマ区切りで指定します。
   - service.langchain.prompts[].applied_node.workflows[]は(3)で定義したワークフロー名を指定
   - service.langchain.prompts[].applied_node.node は(3)で定義したカスタムエージェント名を指定

2.3　アンセットアップ

アンセットアップでは以下の作業を実施します。

- ブラウザのキャッシュの削除

  「1.6.1 会話履歴管理機能」により，生成 AI との会話履歴がブラウザのキャッシュに残ります。削除する必要がある場合は各ブラウザで削除してください。

3 JP1 Cloud Service 生成 AI 連携の運用

JP1 Cloud Service 生成 AI 連携の運用手順について説明しています。

3.1　設定変更

JP1 Cloud Service 生成 AI 連携の設定変更の際に必要な作業について説明します。

3.1.1　定義ファイルの変更方法

設定ファイル（ai_assistant.yml），プロンプト定義ファイル（prompts.chat.ja.yml），ツールプロンプト定義ファイル（tools.ja.yml）の変更には，JP1/Integrated Management 3 - Manager の定義ファイル操作機能を使用します。定義ファイル操作機能の詳細は「JP1 Cloud Service システム管理 利用ガイド」の「4.2.2　監視サービスにおける定義ファイル操作（V02-00 以降）」を参照してください。設定ファイルの詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。プロンプト定義ファイルの詳細は「6.1.2 プロンプト定義ファイル（prompts.chat.ja.yml）」「6.1.3 ツールプロンプト定義ファイル（tools.ja.yml）（02-40 以降）」を参照してください。

設定ファイルを変更する場合に必要な作業を次に示します。なお，設定ファイルは同時に複数更新しないでください。複数同時に更新した場合，更新処理に失敗する可能性があります。更新処理に失敗した場合は生成 AI 連携機能のメッセージログを参照し，エラーが発生していないか確認してください。ログが出力されるまで最大で 15 分かかります。その間，生成 AI 連携機能の設定ファイルのファイル更新は行わないでください。

1. 統合オペレーション・ビューアーから設定ファイルをダウンロードする。

   統合オペレーション・ビューアーの［定義ファイル一覧（ホスト名）］画面から変更するファイルを選択してダウンロードします。［定義ファイル一覧（ホスト名）］画面の詳細は「JP1 Version 13　JP1/Integrated Management 3 - Manager 画面リファレンス」の「2.2.2 ［統合マネージャーの定義ファイル一覧］画面と［定義ファイル一覧（ホスト名）画面］」を参照してください。

2. ダウンロードしたファイルを編集する。

3. 統合オペレーション・ビューアーを使用して，編集したファイルをアップロードする。

   アップロードすると，自動的に生成 AI 連携基盤が再起動して設定が反映されます。

［定義ファイル一覧（ホスト名）］画面の操作により，各ファイルを削除できます。削除した場合，ファイルに設定した内容は破棄され，デフォルト値が記載されたファイルに変更されます。

なお，定義ファイル間の整合性に問題がある場合，メッセージログに生成 AI 連携の起動メッセージ（KNBD20013-I）の直前に，警告メッセージ（KNBD20034-W）が表示されます。複数同時に定義ファイルを更新した場合，複数回の再起動が行われるため，メッセージログに最後に出力された起動メッセージを確認してください。

3.1.2　シークレットの追加・変更・削除方法

シークレットの追加・変更・削除は，生成 AI 連携のチャット画面から実行します。詳細な実行手順については，以下を参照してください。なお，生成 AI 連携が取扱い可能なシークレット数の上限は 100 件です。

- シークレットを追加する場合
  1. チャット画面のメニューボタンから［シークレット設定］を選択します。
  2. シークレット一覧画面の［追加］ボタンを押下します。
  3. シークレット追加ダイアログでシークレットキーとシークレットを入力し［OK］を押下します。
  4. シークレット一覧画面に追加したシークレットが表示されます。
- シークレットを変更する場合
  1. チャット画面のメニューボタンから[シークレット設定]を選択します。
  2. シークレット一覧画面で編集したい行を選択し［編集］ボタンを押下します。
  3. シークレット編集ダイアログでシークレットを入力し［OK］を押下します。
- シークレットを削除する場合
  1. チャット画面のメニューボタンから［シークレット設定］を選択します。
  2. シークレット一覧画面で削除したい行を選択し［削除］ボタンを押下します。
  3. シークレット削除ダイアログで正しい事を確認の上［OK］を押下します。
  4. シークレット一覧画面から削除したシークレットが消えます。

3.1.3　連携する生成 AI サービスの変更

連携する生成 AI サービスを変更する場合に必要な作業を示します。変更内容に合わせて下記に示す作業を実施してください。変更作業では設定ファイル（ai_assistant.yml）を変更します。また，Azure OpenAI Service に変更する場合にはシークレットも変更する必要がある場合があります。設定ファイルの変更方法の詳細は「3.1.1 定義ファイルの変更方法」を参照してください。シークレットの変更方法は「3.1.2 シークレットの追加・変更・削除方法」を参照してください。

- 連携する生成 AI サービスを変更する場合

  モデルの変更等，接続先の生成 AI サービスを変更する場合は，生成 AI サービスのベンダーに合わせて設定ファイル（ai_assistant.yml）の以下の項目を変更します。

  - **Azure OpenAI Service に変更する場合**：generative_ai.azure[]
  - **Amazon Bedrock に変更する場合**：generative_ai.aws[]

  また，Azure OpenAI Service のリソースを変更する場合には，シークレットから古いリソースの API キーを削除し，新しいリソースの API キーを追加します。操作するシークレットは「2.2.1 生成 AI サービスとの連携設定」を参照してください。

- 利用する生成 AI サービスを変更する場合
  - 02-30 以降

    ワークフロー内で利用する生成 AI サービスを変更する場合は，生成 AI サービスの接続設定及び利用設定を行います。接続設定及び利用設定は「2.2.1 生成 AI サービスとの連携設定」を参照してください。
  - 02-20 以前

    生成 AI サービスのベンダーを変更する場合は，generative_ai.azure または generative_ai.aws を削除して，「2.2.1 生成 AI サービスとの連携設定」の内容を元に再度セットアップしてください。Azure OpenAI Service から他ベンダーの生成 AI サービスに変更する場合は，シークレット登録した生成 AI サービスの API キーを削除してください。削除対象のシークレットについては，「2.2.1 生成 AI サービスとの連携設定」を参照してください。
- 生成 AI サービスとの連携を一時的に停止する場合
  - 02-30 以降

    ワークフローの定義(service.langgraph.workflows)に空（[]）を指定します。
  - 02-20 以前

    サービス利用開始後に生成 AI サービスの料金が想定を上回るなど，生成 AI サービスとの連携を一時的に停止したい場合は，設定ファイル（ai_assistant.yml）の項目 generative_ai.azure または generative_ai.aws をコメントアウトし，項目 generative_ai に空（{}）を指定します。また，Azure OpenAI Service を使用しており，API キーを削除したい場合はシークレット登録した生成 AI サービスの API キーを削除してください。削除対象のシークレットについては，「2.2.1 生成 AI サービスとの連携設定」を参照してください。生成 AI サービスとの連携を再開する場合は generative_ai に指定した空（{}）を削除し，コメントアウトを外してください。Azure OpenAI Service を使用しており，API キーを削除した場合はシークレットを再度登録してください。

3.1.4　連携する検索サービスの追加・変更・削除

連携する検索サービスを変更する場合に必要な作業を示します。変更内容に合わせて下記に示す作業を実施してください。変更作業では設定ファイル（ai_assistant.yml）を変更します。また，Azure AI Search に変更する場合にはシークレットも変更する必要がある場合があります。複数の検索サービスを変更する場合は，その数に合わせて設定をしてください。設定ファイルの変更方法の詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。シークレットの変更方法は「3.1.2 シークレットの追加・変更・削除方法」を参照してください。

- 検索サービスを使用することにした場合

  新たに検索サービスを使用する場合は，検索サービスのベンダーに合わせて設定ファイル（ai_assistant.yml）の変更が必要です。設定ファイルの項目 service.langchain.tools の以下の項目のコメントアウトを解除して設定します。
  - **Azure AI Search**： service.langchain.tools.name が user_contents の設定

- **Knowledge Bases for Amazon Bedrock**： service.langchain.tools.name が user_contents_for_aws_kb の設定

項目 service.langchain.prompts.tools には service.langchain.tools.name の値を指定してください。また，Azure AI Search を使用する場合は，シークレットに検索サービスの API キーを登録する必要があります。登録対象のシークレットについては，「2.2.2(1) 検索サービスを利用する場合」を参照してください。

- 検索サービスを変更する場合

  検索サービスを変更する場合は，設定ファイル（ai_assistant.yml）の変更が必要です。検索サービスのベンダーに合わせて設定ファイルの以下の項目を変更してください。

  - **Azure AI Search**：service.langchain.tools.azure_ai_search
  - **Knowledge Bases for Amazon Bedrock**：service.langchain.tools.aws_knowledge_bases

  また，Azure AI Search のリソースを変更する場合には，シークレットから古いリソースの API キーを削除し，新しいリソースの API キーを追加してください。操作対象のシークレットについては，「2.2.2(1) 検索サービスを利用する場合」を参照してください。検索サービスのベンダーを変更する場合は，設定ファイルの項目 service.langchain.tools のコメントアウトを解除して，「3.1.4 連携する検索サービスの追加・変更・削除」の「検索サービスを使用することにした場合」の内容を元に設定をしてください。Azure AI Search から他ベンダーの検索サービスに変更する場合は，シークレット登録した検索サービスの API キーを削除してください。削除対象のシークレットについては，「2.2.2(1) 検索サービスを利用する場合」を参照してください。

- 検索サービスの使用をやめる場合

  検索サービスの使用をやめる場合は，設定ファイル（ai_assistant.yml）の変更が必要です。設定ファイルの項目 service.langchain.tools から使用をやめる検索サービスに関する設定項目 service.langchain.tools.name と項目 service.langchain.tools の使用をやめる検索サービスに関する設定項目をコメントアウトしてください。1 つも検索サービスを使用しない場合は，項目 service.langchain.prompts.tools には空文字を指定してください。また，Azure AI Search の使用をやめる場合は，シークレット登録した検索サービスの API キーを削除してください。削除対象のシークレットについては，「2.2.2(1) 検索サービスを利用する場合」を参照してください。

3.1.5　画面設定の変更

画面の設定を変更する場合に必要な作業を次に示します。

変更可能な項目は以下です。

- 質問文生成機能により生成する質問文
- チャット機能で入力可能な質問文の文字数

変更する場合は設定ファイル（ai_assistant.yml）を変更します。変更方法の詳細については「2.2.4 質問文生成機能の設定（任意）」，「2.2.5 質問文の文字数の設定（任意）」を参照してください。設定ファイルの詳細については，「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

3.1.6　ログの追加出力の設定変更

ログの出力抑止を有効化／無効化し，ログを追加で出力する場合に必要な作業を次に示します。

変更可能な項目は以下です。

- ログの出力抑止の有効化／無効化

  変更する場合は設定ファイル（ai_assistant.yml）を変更します。変更方法の詳細については「2.2.6 ログの追加出力の設定（任意）」を参照してください。設定ファイルの詳細については，「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

3.1.7　ワークフローの変更（任意）（02-30 以降）

ワークフローは複数定義でき，定義できるワークフローの種類は全部で 4 種類です。定義できるワークフローの種類は「1.5.2 ワークフロー作成機能（02-30 以降）」を，ワークフローの種類別の定義例は「付録 D　ワークフローの定義例」をそれぞれ参照してください。

3.1.8　プロンプトの設定（任意）（02-30 以降）

ワークフローに利用するプロンプトを設定ファイル（ai_assistant.yml）およびプロンプト定義ファイル（prompts.chat.ja.yml）に定義します。プロンプトの内容をデフォルト値から変更する場合は以下の設定が必要になります。デフォルト値は「付録 B　プロンプト」を参照してください。

プロンプトを変更するには，以下の 2 つの手順が必要です。

(1)　プロンプトの定義

プロンプトをプロンプト定義ファイル（prompts.chat.ja.yml）に定義します。設定内容（prefix や format_instructions など）の詳細は「1.5.3 プロンプト作成機能」および「6.1.2 プロンプト定義ファイル（prompts.chat.ja.yml）」を参照してください。

(2)　プロンプトとワークフローのノードの紐付け

設定ファイル（ai_assistant.yml）の service.langchain.prompts[]に利用するプロンプトとその適応ノードを定義します。設定内容の詳細は「6.1.1 設定ファイル（ai_assistant.yml）」を参照してください。

例）特定のワークフローで利用するナレッジ検索エージェントのプロンプトを追加し利用する場合

1. プロンプトの定義

   プロンプト定義ファイルの agent.knowledge_search_agent に新規にプロンプト(user_defined_prompt)を定義します

```
agent:
  knowledge_search_agent:
```

```
    - name: user_defined_prompt
      prefix:
        - name: role
（略）
```

2. プロンプトとワークフローのノードの紐付け

   ワークフロー(name: user_defined_workflow)のナレッジ検索エージェントで上記プロンプト(user_defined_prompt)を利用する設定をします。

```
service:
  langchain:
    prompts:
      - name: user_defined_prompt
        type: chat
        tools: ”user_contents”
        applied_node:
          workflows: [”user_defined_workflow”]
          node: knowledge_search_agent
```

3.1.9　ツールプロンプトの設定（任意）（02-40 以降）

エージェントノードが利用するツールのプロンプトを変更する場合は，ツール定義ファイル（tools.ja.yml）を変更します。ツールの内容をデフォルト値から変える場合や，ユーザープログラムの説明や引数を変更する場合は以下の設定が必要になります。

- ユーザーコンテンツ取得機能で利用するツールの説明を変更する場合

  検索サービスのインデックスの中身を変更した場合など，ツールの説明を変更する場合はツールプロンプト定義ファイルの該当ツールの説明（azure_ai_search[].description または aws_knowledge_bases[].description）の変更をします。

- ユーザープログラム呼び出し機能で利用するツールの説明や引数を変更する場合

  ユーザープログラムの引数を追加／削除したい場合など，ツールの説明や引数を変更する場合はツールプロンプト定義ファイルの該当ツールの説明（aws_lambda_function[].description）および引数（aws_lambda_function[].args）を変更します。

3.1.10　ユーザープログラムの連携設定の変更（任意）（02-40 以降）

設定ファイル（ai_assistant.yml）の service.langchain.tools[].aws_lambda_function.handle 以下の設定項目を編集することで，ユーザープログラム呼び出し機能の挙動を変更できます。変更可能な挙動は以下の通りです。

- ユーザープログラムの処理中にエラーが発生した場合の挙動を制御する

  設定ファイルの service.langchain.tools[].aws_lambda_function.handle.raise_function_processing_error を true にすることで，ユーザープログラムの処理中にエラーが発生した場合は，ユーザープログラム呼び出し機能としてエラーが発生したものとして，カスタムエージェ

ントの処理を異常終了します。処理を継続させる場合は，false を設定してください。明示的に設定されていない場合，true が設定されたものとして扱われます。

- ツールに定義された JSON スキーマのログへの出力を制御する

  設定ファイルの service.langchain.tools[].aws_lambda_function.handle. logging_args_schema を true にすることで，JSON スキーマがログへ出力されます。ログへの出力を抑制する場合は，false を設定してください。明示的に設定されていない場合，false が設定されたものとして扱われます。

- ユーザープログラムの呼び出し通知の出力を制御する

  設定ファイルの service.langchain.tools[].aws_lambda_function.handle. logging_start_info を true にすることで，ユーザープログラムの呼び出し通知がログへ出力されます。ログへの出力を抑制する場合は，false を設定してください。明示的に設定されていない場合，true が設定されたものとして扱われます。

4 JP1 Cloud Service システム管理への影響

JP1 Cloud Service 生成 AI 連携を運用する際の JP1 Cloud Service システム管理への影響について説明しています。

4.1 SID の追加

生成 AI 連携が利用する SID を追加します。IM 管理ノードに付与される SID は次の通りです。

- _AIASSISTANT_{ホスト名}/_HOST_{ホスト名}
  {ホスト名}は JP1 Cloud Service のシステム管理マネージャーが稼働するホスト名です。

5 トラブルシューティング

JP1 Cloud Service 生成 AI 連携における主なトラブルシューティングについて説明しています。

5.1 トラブルシューティング

5.1.1 新規スレッドを作成できない

要因

次の要因が考えられます。

1. スレッド数が上限に到達している（スレッド数の上限は 1 ユーザーにつき 10 スレッドです）。
2. 会話履歴の保存上限に達している。

対処

要因に応じて対処してください。

1. 古いスレッドを削除してください。
2. 古いスレッドを削除してください。ブラウザを共有する他のユーザーがスレッド削除をした場合も制限が解除されます。

5.1.2 チャット画面のバナーに「保存できる会話履歴の上限に達しました。不要なスレッドを削除してください。スレッドが削除されるまで表示されている会話内容は保存されません。」と表示される。

要因

会話履歴の保存上限に達している。

対処

古いスレッドを削除してください。ブラウザを共有する他のユーザーがスレッド削除をした場合も制限が解除されます。

5.1.3 質問エリアへの入力が以下のような不安定な状態となる。

- 入力内容が確定した表示となる
- 入力内容が消える
- 入力できない
- コピー内容をペーストできない

要因

IME が ON の状態で質問文の文字数上限を超過して入力またはコピー&ペーストした。

対処

質問エリアからフォーカスを外したのち（チャット画面などを選択し，入力状態を外す），入力しなおしてください。

5.1.4 検索サービスを利用した質問の回答が行われない。

要因

質問の回答精度は使用するモデルに依存します。モデルによっては検索サービスを利用した回答精度が低くなることがあります。

対処

より精度の高いモデルの利用を検討してください。

5.1.5 任意のリンクを作成するプロンプトに変更したが，作成されない／ハイパーリンク化しない。

要因

1. プロンプトの定義が反映されておらず作成／ハイパーリンク化されない
2. 指示が曖昧で生成が必要と判断されず，リンク作成がされない

対処

1. 定義したプロンプトおよび設定ファイルについて以下を確認してください

   ・prompts.chat.ja.yml の prefix に定義されている事

   ・prompts.chat.ja.yml に定義した name が，service.langchain.prompts[].name に指定されている事

   ・ai_assistant.yml の ui.chat.hyperlink にプロンプトで定義した URL にマッチする正規表現の regexp が定義されている事
2. 質問を変えて再度実行してください。

5.1.6 KNBD20024-E のエラー詳細で，"Failed to retrieve a response from the assistant. Please try again."が頻発する。

要因

レガシーモデルの利用により，生成 AI サービスから不正な形式の回答が返却されることがあります。

対処

より精度の高いモデルの利用を検討してください。

5.1.7　回答内容が"Sorry, need more steps to process this request"となる。

要因

連携する生成 AI との問い合わせの繰り返し回数の上限に達した。

対処

もう一度問い合わせてください。頻発する場合は簡単な質問にするなど質問を変えて再度実行してください。

5.1.8　意図しない XML タグが回答に付与される／Markdown 記法で返却される。

要因

生成 AI モデルの特性に依存します。例えば Amazon Bedrock で利用できる Anthropic Claude 系であれば XML タグが，Azure OpenAI Service の GPT 系であれば Markdown が利用される場合があります。

5.1.9　回答に含まれる URL がハイパーリンク化せず，リンクが途中で途切れてしまう

要因

ナレッジ検索エージェントがユーザーコンテンツ取得機能により取得した URL が長く，出力文字数を超過した。

対処

設定ファイルで回答の最大トークン数を変更してください。ただし最大トークン数を変更した場合，可読性や料金に影響があります。

5.1.10　レコメンドが表示されない／精度が低い

要因

レコメンドの精度は使用するモデルに依存します。モデルによってはレコメンドの精度が低くなることがあります。

対処

より精度の高いモデルの利用を検討してください。

5.1.11 KNBD20030-E のエラー詳細で，"The input for the DataModel that generated by the AI was invalid."が頻発し，ツールの実行に失敗する

要因

モデルの回答の最大トークン数が小さい場合，ツールの実行に必要な引数が生成できないことがあります。

対処

設定ファイルで，モデルの回答の最大トークン数を変更してください。

5.1.12 利用できる生成 AI モデルが提供終了となった

対処

利用している生成 AI モデルが提供終了となり，利用したい生成 AI モデルがご使用中の JP1 Cloud Service システム管理 生成 AI 連携のバージョンでサポートされていない場合，生成 AI 連携機能のバージョンアップが必要です。サポート窓口までお問い合わせください。

なお，生成 AI サービス（Azure OpenAI Service，Amazon Bedrock 等）で提供されるモデルの提供終了（EoL）は，提供元のドキュメントをご確認ください。

5.1.13 問い合わせ時に「KNBD20024-E 生成 AI 連携基盤による問い合わせ実行処理でエラーが発生しました。エラー内容: Invalid tool is called by the generative AI. Error: The tool call is empty, so the tool cannot be called.または KNBD20010-E 連携する azure openai への接続に失敗しました。接続情報を確認してください。エラー内容：Error code: 400 - {'error': {'message': 'Could not finish the message because max_tokens or model output limit was reached. Please try again with higher max_tokens.', 'type': 'invalid_request_error', 'param': None, 'code': None}}」というエラーが返却される。

要因

Azure OpenAI Service の推論モデルを Secretary または Supervisor に使用しており，かつ max_tokens が不足している。

（02-60 以降）

Amazon Bedrock の推論モデルでも同様の事象が発生する可能性があります。頻発する場合は，以下の対処を参考にご対応下さい。

対処

max_tokens を増やすか，利用するモデルを変更して下さい。

（02-60 以降）

推論に関連するパラメータの調整もご検討ください。

5.1.14 問い合わせ時に「KNBD20024-E 生成 AI 連携基盤による問い合わせ実行処理でエラーが発生しました。エラー内容: Failed to retrieve a response from the assistant(＜ operator|agent ＞: ＜ノードの名前＞). Please try again.」

要因

Azure OpenAI Service の推論モデルをセクレタリーまたはエージェントノードに使用しており，かつ max_tokens が不足している。

（02-60 以降）

Amazon Bedrock の推論モデルでも同様の事象が発生する可能性があります。頻発する場合は，以下の対処を参考にご対応下さい。

対処

max_tokens を増やすか，利用するモデルを変更して下さい。

（02-60 以降）

推論に関連するパラメータの調整もご検討ください。

5.1.15 問い合わせ時に temperature がデフォルト以外で失敗する。

要因

Azure OpenAI Service の推論モデル（reasoning_effort が none 以外）を使用し，temperature で 1 以外を指定している。

（02-60 以降）

Amazon Bedrock の推論モデルでも同様の事象が発生する可能性があります。頻発する場合は，以下の対処を参考にご対応下さい。

対処

temperature にデフォルト値を指定してください。

6 リファレンス

JP1 Cloud Service 生成 AI 連携が提供する定義ファイルおよびメッセージについて説明しています。

6.1　定義ファイル

JP1 Cloud Service 生成 AI 連携の定義ファイルの形式および文法について説明します。JP1 Cloud Service では定義ファイルとして設定ファイル（ai_assistant.yml）とプロンプト定義ファイル（prompts.chat.ja.yml），02-40 以降はツールプロンプト定義ファイル（tools.ja.yml）を提供します。なお，定義ファイル内に環境依存文字を使用すると，定義に文字化けが発生することがあるため，使用しないでください。

すべての定義ファイルはダウンロードすることができます。ダウンロード方法は「3.1.1 定義ファイルの変更方法」を参照してください。

6.1.1　設定ファイル（ai_assistant.yml）

形式

YAML 形式で記述します。

```
generative_ai:
  azure:
    - endpoint: Azure OpenAIのAPIのエンドポイント
      models:
        - name: Azure OpenAI で利用するモデルの名前
          api:
            type: Azure OpenAIのAPIの種類
            version: Azure OpenAIのAPIのバージョン
            temperature: Azure OpenAIのLLMの回答のランダム性
            max_tokens: Azure OpenAI のLLMの回答の最大トークン数
            reasoning:
              effort: Azure OpenAIのLLMの内部推論の深さ
            verbosity: Azure OpenAIのLLMの出力の簡潔さを指定
          deployment_name: Azure OpenAIのデプロイ名
          model_name: Azure OpenAIのモデル名
  aws:
    - role_arn: Amazon Bedrockを実行するIAMロールのARN
      external_id: Amazon Bedrockを実行するIAMロールの外部ID
      endpoint_url: プライベート接続用のエンドポイントURL
      models:
        - name: Azure Bedrockで利用するモデルの名前
          type: Amazon BedrockのAPIの種類
          model_id: Amazon BedrockのモデルID
          region_name: AWSのリージョン
          role_session_name: AssumeRoleで指定するセッション名
          custom_request_params:
            - key: モデルのリクエストパラメータのキー
              value: モデルのリクエストパラメータの値
          thinking:
            type: モデルの思考モードの指定
            budget_tokens: 推論で使用可能とするトークン数
          output_config:
            effort: 内部推論の深さ
service:
```

```
langchain:
  prompts:
    - name: プロンプト名
      type: プロンプトのtype
      tools: 使用するツールの名前
      applied_node:
        workflows: プロンプトを適用するノードが所属するワークフローの名前
        node: プロンプトを適用するノードの識別子
  tools:
    - name: ツールの名前
      azure_ai_search:
        endpoint: Azure AI SearchのAPIのエンドポイント
        index_name: Azure AI Searchのインデックス名
        top: Azure AI Searchで取得する検索結果の件数
      aws_knowledge_bases:
        knowledge_base_id: Knowledge Bases for Amazon BedrockのID
        role_arn: Knowledge Bases for Amazon Bedrock を実行するIAMロールのARN
        external_id: Knowledge Bases for Amazon Bedrock を実行するIAMロールの外部ID
        endpoint_url: プライベート接続用のエンドポイントURL
        region_name: AWSのリージョン
        role_session_name: AssumeRoleで指定するセッション名
        number_of_results: Knowledge Bases for Amazon Bedrockで取得する検索結果の件数
        override_search_type: Knowledge Bases for Amazon Bedrockで使用する検索タイプ
        additional_configuration:
          rerankingConfiguration:
            type: 使用するRerankモデルのタイプ指定
            bedrockRerankingConfiguration:
              metadataConfiguration:
                selectionMode: Rerankingで使用するメタデータを選択するかの指定
                selectiveModeConfiguration:
                  fieldsToExclude:
                    - fieldName: Rerankingに使用しないメタデータ名
                  fieldsToInclude:
                    - fieldName: Rerankingに使用するメタデータ名
              modelConfiguration:
                additionalModelRequestFields:
                  使用するモデル固有の追加パラメータ名: 追加パラメータの値
                modelArn: 使用するRerankモデルのARN
              numberOfRerankedResults: Rerankモデルが返却するrerank結果の最大数
      aws_lambda_function:
        function_name: AWS Lambda関数のARN
        role_arn: AWS Lambda関数を実行するIAMロールのARN
        external_id: AWS Lambda関数を実行するIAMロールの外部ID
        endpoint_url: プライベート接続用のエンドポイントURL
        region_name: AWSのリージョン
        role_session_name: AssumeRoleで指定するセッション名
        handle:
          raise_function_processing_error: AWS Lambda関数処理中のエラー扱い
          logging_start_info: AWS Lambda関数の呼び出し時のログ出力
          logging_args_schema: AWS Lambda関数の実行引数作成に利用するJSONスキーマの出力

langgraph:
  agents:
    knowledge_search_agent:
      llm_name: ナレッジ検索エージェントに利用するLLMの名前
    event_analysis_agent:
      llm_name: 運用管理データ取得エージェントに利用するLLMの名前
    custom_agent:
```

```
        llm_name: カスタムエージェントに利用するLLMの名前

    workflows:
      - name: ワークフローの名称
        label: ワークフローの表示名
        description: ワークフローの説明
        nodes:
          secretary:
            llm_name: Secretaryノードに利用するLLMの名前
          supervisor:
            llm_name: Supervisorノードに利用するLLMの名前
            child_agents: Supervisorが利用できるエージェントノード識別子
              knowledge_search_agent:
                - name: 利用するナレッジ検索エージェントの名前
              event_analysis_agent:
                - name: 利用する運用管理データ取得エージェントの名前
              custom_agent:
                - name: 利用するカスタムエージェントの名前
                  override_agent_parameter:
                    - key: 上書きするエージェントパラメータのキー
                      value: 上書きするエージェントパラメータの値

ui:
  template: 質問文のテンプレート
  jp1event_template: 質問文のうちJP1イベントに関するテンプレート
  chat:
    max_size: 質問の最大文字数
    max_display_size: チャット画面表示最大文字数
    hyperlink:
      - regexp: ハイパーリンク化する文字列の正規表現
        replace: ハイパーリンク化する文字列の置換後の表示名
logs:
  filter: ログの出力抑止の有効化または無効化
```

ファイル

ai_assistant.yml

定義ファイル編集機能画面のタブ名

aiassistant

説明

生成 AI 連携の設定ファイルです。生成 AI 及びドキュメント検索エンジンの接続情報や，画面設定を定義します。

文字コード

UTF-8（BOM なし）

改行コード

LF

定義の反映時期

JP1/IM3 の定義ファイル操作機能を使ってアップロードした際に，プログラムが再起動して定義が反映されます。

記述内容

項目名	説明	必須/任意	デフォルト値※
generative_ai:	生成 AI 関連の接続先を記載します。	必須	-
azure:	Azure OpenAI Service の接続先を記載します。02-20 以前は generative_ai.aws とどちらか片方のみ設定できます。 **指定可能な要素数** 15	任意	-
- endpoint:	Azure OpenAI API のエンドポイントを記載します。02-20 以前は１つのみ指定可能です。 **指定可能な値** 1〜256 文字の下記形式 `https?://[a-zA-Z0-9_/:%#¥$&¥?()~.=+-]+|endpoint`	必須	-
models:	生成 AI モデルの情報を記載します。02-20 以前は１つのみ指定可能です。 **指定可能な要素数** 10	必須	-
- name:	02-30 以降で定義可能。使用する LLM に生成 AI 連携機能内で利用する名前を指定します。重複した値が存在する場合，いずれか一つが利用されます。 **指定可能な値** "azure_"で始まる，半角英数字とハイフン，アンダースコア 1〜128 文字	必須	-
api:	使用する API について記載します。	必須	
type:	使用する API の種類を記載します。チャット入力候補（Chat Completion）のみ使用することができ，chat を指定します。各モデルで利用可能な API の種類は Azure の公式ドキュメントを参照してください。 **指定可能な値** chat	必須	-
version:	使用する API のバージョンを指定します。設定ファイルの初期値では"2024-10-21"を指定しています。 (02-60 以降)	必須	-

項目名					説明	必須/任意	デフォルト値※
-	-		version:		初期値では"v1"を指定しています。 **指定可能な値** 半角英数字とハイフン 1～32 文字 （02-60 以降） v1 のみ指定可能	必須	-
			temperature		02-40 以降で定義可能。LLM の回答のランダム性を指定します。 **指定可能な値** 0.0～1.0	任意	1
			max_tokens		02-30 以降で定義可能。LLM の回答の最大トークン数を指定します。 **指定可能な値** 1～8192 （02-60 以降） 1～128000	任意	2048
			reasoning		02-60 以降で定義可能。LLM の推論パラメータについて記載します。	任意	
				effort	LLM の内部推論の深さを指定します。 **指定可能な値** none, minimal, low, medium, high, xhigh, max のいずれか	必須	モデルによって指定可能な値が異なります。サービスを提供しているベンダーの公開ドキュメントを参照し，モデルに応じて動作するパラメータを指定してください。
			verbosity		LLM の出力の簡潔さを指定します。 **指定可能な値** low, medium, high のいずれか	任意	
			deployment_name:		Azure OpenAI のデプロイ名を記載します。 **指定可能な値** 半角英数字とハイフン，アンダースコア 1～128 文字	必須	-
			model_name:		Azure OpenAI のモデル名を記載します。（プレビューモデルを含みます。プレビューモデルは Azure OpenAI Service の最新情報をご確認ください） **指定可能な値** 02-60 で指定可能な値 • 02-30 で指定可能な値すべて • 02-40 で指定可能な値すべて	必須	-

項目名				説明	必須/任意	デフォルト値※
	-	-	model_name:	• 02-50 で指定可能な値すべて • gpt-5.6-sol • gpt-5.6-terra • gpt-5.6-luna • gpt-5.5 • gpt-5.4 • gpt-5.4-mini • gpt-5.4-nano • gpt-5.3-chat 02-50 で指定可能な値 • 02-30 で指定可能な値すべて • 02-40 で指定可能な値すべて • gpt-5.2-chat • gpt-5.1-chat • gpt-5-chat • gpt-5-mini • gpt-5-nano 02-40 で指定可能な値 • 02-30 で指定可能な値すべて • gpt-4.1 • gpt-4.1-nano • gpt-4.1-mini 02-30 で指定可能な値 • gpt-4o • gpt-4o-mini • gpt-4 • gpt-4-32k • gpt-35-turbo • gpt-35-turbo-16k	必須	-
	aws			Amazon Bedrock の接続先を記載します。 02-20 以前は generative_ai.azure とどちらか片方のみ設定できます。 **指定可能な要素数** 15	任意	-
	-		role_arn:	Amazon Bedrock を実行する IAM ロールの Amazon リソースネーム（ARN）を指定します。 **指定可能な値** 半角英数字とコロン，スラッシュ，ハイフン，アンダースコア 1～255 文字	必須	-

項目名	説明	必須/任意	デフォルト値※
external_id:	Amazon Bedrock を実行する IAM ロールの AssumeRole の条件に外部 ID を指定している場合は，その外部 ID を指定します。 **指定可能な値** 半角英数字と以下の記号 「+」，「=」，「,」，「.」，「@」，「:」，「/」，「-」 2〜1224 文字	任意	-
endpoint_url:	Amazon Bedrock とプライベート接続する場合に，プライベート接続用のエンドポイント URL を指定します。 **指定可能な値** 1〜256 文字の下記形式 https?://vpce-[a-zA-Z0-9_/:%#¥$&¥?()~.=+-]	任意	
models:	生成 AI モデルの情報を記載します。02-20 以前は 1 つのみ指定可能です。 **指定可能な要素数** 10	必須	-
- name:	02-30 以降で定義可能。使用する LLM に生成 AI 連携機能内で利用する名前を指定します。重複した値が存在する場合，いずれか一つが利用されます。 **指定可能な値** "aws_"で始まる，半角英数字とハイフン，アンダースコア 1〜128 文字	必須	
type:	使用する API の種類を指定します。AnthropicClaude メッセージ API のみ使用することができ，chat を指定します。各モデルで利用可能な API の種類は AWS 公式ドキュメントを参照してください。 **指定可能な値** chat	必須	-
model_id:	Amazon Bedrock のモデル ID またはクロスリージョン推論のプロファイル ID を記載します。 **指定可能な値** 02-60 で指定可能な値 • 02-30 で指定可能な値 • 02-40 で指定可能な値 • 02-50 で指定可能な値 • global.anthropic.claude-sonnet-5 • us.anthropic.claude-sonnet-5	必須	-

項目名	説明	必須/任意	デフォルト値※
model_id:	• eu.anthropic.claude-sonnet-5 • global.anthropic.claude-opus-4-8 • us.anthropic.claude-opus-4-8 • eu.anthropic.claude-opus-4-8 • jp.anthropic.claude-opus-4-8 • au.anthropic.claude-opus-4-8 • global.anthropic.claude-opus-4-7 • us.anthropic.claude-opus-4-7 • eu.anthropic.claude-opus-4-7 • jp.anthropic.claude-opus-4-7 • au.anthropic.claude-opus-4-7 • global.anthropic.claude-sonnet-4-6 • us.anthropic.claude-sonnet-4-6 • eu.anthropic.claude-sonnet-4-6 • jp.anthropic.claude-sonnet-4-6 • global.anthropic.claude-opus-4-6-v1 • us.anthropic.claude-opus-4-6-v1 • eu.anthropic.claude-opus-4-6-v1 02-50 で指定可能な値 • 02-30 で指定可能な値 • 02-40 で指定可能な値 • global.anthropic.claude-sonnet-4-5-20250929-v1:0 • us.anthropic.claude-sonnet-4-5-20250929-v1:0 • eu.anthropic.claude-sonnet-4-5-20250929-v1:0 • jp.anthropic.claude-sonnet-4-5-20250929-v1:0 • global.anthropic.claude-sonnet-4-20250514-v1:0 • global.anthropic.claude-opus-4-5-20251101-v1:0 • us.anthropic.claude-opus-4-5-20251101-v1:0 • eu.anthropic.claude-opus-4-5-20251101-v1:0 • us.anthropic.claude-opus-4-1-20250805-v1:0 • global.anthropic.claude-haiku-4-5-20251001-v1:0	必須	-

項目名				説明	必須/任意	デフォルト値※
			model_id:	• us.anthropic.claude-haiku-4-5-20251001-v1:0 • eu.anthropic.claude-haiku-4-5-20251001-v1:0 • jp.anthropic.claude-haiku-4-5-20251001-v1:0 02-40 で指定可能な値 • 02-30 で指定可能な値 • us.anthropic.claude-sonnet-4-20250514-v1:0 • eu.anthropic.claude-sonnet-4-20250514-v1:0 • apac.anthropic.claude-sonnet-4-20250514-v1:0 • us.anthropic.claude-opus-4-20250514-v1:0 • us.anthropic.claude-3-7-sonnet-20250219-v1:0 • eu.anthropic.claude-3-7-sonnet-20250219-v1:0 • apac.anthropic.claude-3-7-sonnet-20250219-v1:0 • us.anthropic.claude-3-5-sonnet-20241022-v2:0 • apac.anthropic.claude-3-5-sonnet-20241022-v2:0 • us.anthropic.claude-3-5-sonnet-20240620-v1:0 • eu.anthropic.claude-3-5-sonnet-20240620-v1:0 • apac.anthropic.claude-3-5-sonnet-20240620-v1:0 • us.anthropic.claude-3-5-haiku-20241022-v1:0 • us.anthropic.claude-3-sonnet-20240229-v1:0 • eu.anthropic.claude-3-sonnet-20240229-v1:0 • apac.anthropic.claude-3-sonnet-20240229-v1:0 • us.anthropic.claude-3-opus-20240229-v1:0 • us.anthropic.claude-3-haiku-20240307-v1:0 • eu.anthropic.claude-3-haiku-20240307-v1:0	必須	-

項目名					説明	必須/任意	デフォルト値※
			model_id:		• apac.anthropic.claude-3-haiku-20240307-v1:0 02-30 で指定可能な値 • anthropic.claude-3-5-sonnet-20241022-v2:0 • anthropic.claude-3-5-haiku-20241022-v1:0 • anthropic.claude-3-5-sonnet-20240620-v1:0 • anthropic.claude-3-haiku-20240307-v1:0 • anthropic.claude-3-sonnet-20240229-v1:0 • anthropic.claude-3-opus-20240229-v1:0	必須	-
			region_name:		Amazon Bedrock のリージョンを指定します。クロスリージョン推論を利用する場合は，ソースリージョンを指定します。指定可能な値はリージョン一覧を参照してください。	必須	-
			role_session_name:		AssumeRole の際に指定するセッション名を指定します。 **指定可能な値** 以下の正規表現に当てはまる値 [a-zA-Z0-9_+=,.@-]* 2〜64 文字	任意	”jp1cs-genai-bedrock-session-name”
			custom_request_param		本設定値は 02-30 以降で定義可能。	任意	
			-	key	リクエストパラメータのキー。Anthropic Claude 系の以下が指定できます。 **指定可能な値** • モデルの回答の最大トークン数を指定する場合は「max_tokens」 • モデルの回答のランダム性を指定する場合は「temperature」 • モデルの回答の top_p を指定する場合は「top_p」 • モデルの回答の top_k を指定する場合は「top_k」	必須	
				value	リクエストパラメータの値。key に対応する Anthropic Claude 系のパラメータを指定します。 **指定可能な値** • key が max_tokens の場合：1〜4096 • key が temperature の場合：0.0〜1.0	必須	max_tokens: 2048 temperature: 1.0

項目名	説明	必須/任意	デフォルト値※
value	• key が top_p の場合：0.0～1.0 • key が top_k の場合：0～500	必須	max_tokens: 2048 temperature: 1.0
thinking	02-60 以降で定義可能。LLM の思考に関するパラメータについて記載します。	任意	
type	LLM の思考モードを指定します。拡張思考（enabled），アダプティブシンキング（adaptive），思考を無効（disabled）を指定できます。 **指定可能な値** enabled, adaptive, disabled のいずれか	必須	モデルによって指定可能な値が異なります。サービスを提供しているベンダーの公開ドキュメントを参照し，モデルに応じて動作するパラメータを指定してください。
budget_tokens	LLM の拡張思考に使用するトークン数を指定します。 **指定可能な値** 1024～128000	任意	
output_config	02-60 以降で定義可能。LLM の出力に関するパラメータについて記載します。	任意	
effort	LLM の内部推論の深さを指定します。 **指定可能な値** max, xhigh, high, medium, low のいずれか	必須	モデルによって指定可能な値が異なります。サービスを提供しているベンダーの公開ドキュメントを参照し，モデルに応じて動作するパラメータを指定してください。
service:		必須	-
langchain:		任意	-
prompts:	**指定可能な要素数** 80	任意	-
- name:	プロンプト定義ファイルで定義したプロンプト名を英語小文字にて指定します。prompts 内で一意となる値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須	
type:	プロンプトの type を指定します。生成 AI の API の種類に合わせて指定する必要があり，chat のみ指定できます。 **指定可能な値** chat	必須	

項目名	説明	必須/任意	デフォルト値※
- tools:	使用するツールの名前（tools.name の値）をカンマ区切りで記載します。ツールを使用しない場合は空文字を指定します。 **指定可能な値** 半角英数字（小文字）とアンダースコア，カンマ 0～4096 文字	必須	
applied_node:	本設定値は 02-30 以降で定義可能。	必須	
workflows:	プロンプトを適用するノードが所属するワークフローの名前。 **指定可能な値** 半角英数字（小文字）とアンダースコア 1～128 文字のリスト **指定可能な要素数** 15	必須	
node:	プロンプトを適用するノードの識別子。 **指定可能な値** knowledge_search_agent, event_analysis_agent, secretary, supervisor または半角英数字（小文字）とアンダースコア 1～64 文字	必須	
tools:	**指定可能な要素数** 80	任意	-
- name:	ツールの名前です。tools 内で一意となる値を指定します。変更できません。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須	-
azure_ai_search:	Azure AI Search の接続先を記載します。 Azure AI Search 検索ツールを利用する場合は本項目を指定します。	任意	-
endpoint:	Azure AI Search の API のエンドポイントを記載します。 **指定可能な値** 1～256 文字の下記形式 https?://[a-zA-Z0-9_/:%#¥$&¥?()~.=+-]+|endpoint	必須	-
index_name:	Azure AI Search のインデックス名を記載します。	必須	-

項目名	説明	必須/任意	デフォルト値※
- index_name:	**指定可能な値** 半角英数字とハイフン 1～128 文字	必須	-
top	Azure AI Search で取得する検索結果の件数を指定します。 **指定可能な値** 1~100		
aws_knowledge_bases:	Knowledge Bases for Amazon Bedrock の接続先を記載します。配列の同じ要素に azure_ai_search が設定されている場合は無効になります。Knowlede Bases for Amazon Bedrock 検索ツールを利用する場合は本項目を指定します。	任意	-
knowledge_base_id:	Knowledge Bases for Amazon Bedrock の ID を指定します。 **指定可能な値** 半角英数字 1～10 文字	必須	-
role_arn:	Knowledge Bases for Amazon Bedrock を実行する IAM ロールの Amazon リソースネーム（ARN）指定します。 **指定可能な値** 半角英数字とコロン，スラッシュ，ハイフン，アンダースコア 1～255 文字	必須	-
external_id:	Knowledge Bases for Amazon Bedrock を実行する IAM ロールの AssumeRole の条件に外部 ID を指定している場合は，その外部 ID を指定します。 **指定可能な値** 半角英数字と以下の記号 「+」，「=」，「,」，「.」，「@」，「:」，「/」，「-」 2～1224 文字	任意	-
endpoint_url:	Knowledge Bases for Amazon Bedrock とプライベート接続する場合に，プライベート接続用のエンドポイント URL を指定します。 **指定可能な値** 1～256 文字の下記形式 https?://vpce-[a-zA-Z0-9_/:%#¥$&¥?()~.=+-]+	任意	-

項目名	説明	必須/任意	デフォルト値※
region_name:	Knowledge Base for Amazon Bedrock のリージョンを指定します。指定可能な値はリージョン一覧を参照してください。	必須	-
role_session_name:	AssumeRole の際に指定するセッション名を指定します。 **指定可能な値** 以下の正規表現に当てはまる値 [a-zA-Z0-9_+=,.@-]* 2〜64 文字	任意	”jp1cs-genai-aws-kb-session-name”
number_of_results	Knowledge Bases for Amazon Bedrock で取得する検索結果の件数を指定します。 **指定可能な値** 1~100		
override_search_type	Knowledge Bases for Amazon Bedrock で使用する検索タイプを指定します。 **指定可能な値** • ハイブリッド検索を指定する場合は「HYBRID」 • セマンティック検索を指定する場合は「SEMANTIC」		
additional_configuration:	02-60 以降で定義可能。Knowledge Base for Amazon Bedrock の追加設定を指定します。	任意	–
rerankingConfiguration:	Rerank モデルの接続先やその他設定を指定します。	任意	–
type:	使用する Rerank モデルのタイプを指定します。 **指定可能な値** BEDROCK_RERANKING_MODEL	必須	
bedrockRerankingConfiguration:	Knowledge Base for Amazon Bedrock の Rerank モデルの接続先やその他設定を指定します。	任意	–
metadataConfiguration:	Reranking に使用するメタデータについての設定を指定します。	任意	–
selectionMode:	Reranking で使用するメタデータを選択するかを指定します。 **指定可能な値** 全てのメタデータを使用する場合は「ALL」 使用するメタデータについて選択したい場合は「SELECT」	必須	

項目名	説明	必須/任意	デフォルト値※
selectiveModeConfiguration:	Reranking に使用するメタデータもしくは使用しないメタデータを指定します。 selectionMode が「SELECT」の場合のみ指定が有効になります。	任意	-
fieldsToExclude:	**指定可能な要素数** 10	任意	-
- fieldName:	Reranking に使用しないメタデータを指定します。 **指定可能な値** 1～2000 文字	必須	
fieldsToInclude:	**指定可能な要素数** 10	任意	-
- fieldName:	Reranking に使用するメタデータを指定します。 **指定可能な値** 1～2000 文字	必須	
modelConfiguration:	Knowledge Base for Amazon Bedrock の Rerank モデル自体の設定を指定します。	必須	
additionalModelRequestFields:	使用するモデル固有の追加パラメータを指定します。 指定方法は前述の YAML 形式の通りに，Key と Value を指定します。	任意	-
modelArn:	Knowledge Base for Amazon Bedrock で使用する Rerank モデルの Amazon リソースネーム（ARN）を指定します。 **指定可能な値** 1～2048 文字の下記形式 `^(arn:aws(-[^:]+)?:bedrock:[a-z0-9-]{1,20}::foundation-model/(.*))?$`	必須	
numberOfRerankedResults:	使用する Rerank モデルが返却する rerank 結果の最大数をしていします。 **指定可能な値** 1～100	任意	-
aws_lambda_function:	AWS Lambda の接続情報を指定します。ユーザープログラム呼び出しツールを利用する場合は本項目を指定します。	任意	
function_name:	実行対象の Lambda 関数の完全修飾された ARN を指定します。関数のバージョンは任意	必須	

項目名						説明	必須/任意	デフォルト値※
				function_name:		項目です。省略した場合，最新版が指定されます。 **指定可能な値** `arn:<パーティション>:lambda:<リージョン名>:<AWS アカウント ID>:function:<関数名>:<バージョン ※任意>`の形式の文字列。 半角英数字とコロン，スラッシュ，ハイフン，アンダースコア 1〜140 文字	必須	
				role_arn:		AWS Lambda 関数を実行する IAM ロールの Amazon リソースネーム（ARN）指定します。 **指定可能な値** 半角英数字とコロン，スラッシュ，ハイフン，アンダースコア 1〜255 文字	必須	
				external_id:		AWS Lambda 関数を実行する IAM ロールの AssumeRole の条件に外部 ID を指定している場合は，その外部 ID を指定します。 **指定可能な値** 半角英数字と以下の記号 「+」，「=」，「,」，「.」，「@」，「:」，「/」，「-」 2〜1224 文字	任意	
				endpoint_url:		AWS Lambda 関数とプライベート接続する場合に，プライベート接続用のエンドポイント URL を指定します。 **指定可能な値** 1〜256 文字の下記形式 `https?://vpce-[a-zA-Z0-9_/:%#¥$&¥?()~.=+-]+`	任意	
				region_name:		AWS Lambda 関数のリージョンを指定します。指定可能な値はリージョン一覧を参照してください。	必須	
				role_session_name:		AssumeRole の際に指定するセッション名を指定します。 **指定可能な値** 以下の正規表現に当てはまる値 `[a-zA-Z0-9_+=,.@-]*` 2〜64 文字	任意	jp1cs-genai-aws-lambda-session-name
				handle:		AWS Lambda 関数実行時の挙動を指定できます。	任意	
					raise_function_processing_error:	AWS Lambda 関数の処理中に発生したエラーを異常系として扱うかの設定。	任意	true

項目名	説明	必須/任意	デフォルト値※
raise_function_processing_error:	**指定可能な値** true（異常系として扱う），false（正常系として扱う）	任意	true
logging_start_info:	AWS Lambda 関数の実行時の問い合わせ実行識別子とツール識別子をログに出力するかの設定 **指定可能な値** true（出力する），false（出力しない）	任意	true
logging_args_schema:	AWS Lambda 関数の実行変数作成時に JSON スキーマをログに出力するかの設定 **指定可能な値** true（出力する），false（出力しない）	任意	false
langgraph:	本設定値は 02-30 以降で定義可能。	必須	
agents:	エージェントノードに関する設定	任意	
knowledge_search_agent:	ナレッジ検索エージェントの共通設定	任意	
llm_name	ナレッジ検索エージェントに利用する LLM の名前 **指定可能な値** 半角英数字とアンダースコア 1〜128 文字	必須	
event_analysis_agent	運用管理データ取得エージェントの共通設定	任意	
llm_name	運用管理データ取得エージェントに利用する LLM の名前 **指定可能な値** 半角英数字とアンダースコア 1〜128 文字	必須	
custom_agent:	カスタムエージェントの共通設定	任意	
llm_name:	カスタムエージェントに利用する LLM の名前 **指定可能な値** 半角英数字とアンダースコア 1〜128 文字	必須	
workflows:	**指定可能な要素数** 15	必須	
- name:	ワークフローの名称 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1〜128 文字	必須	

項目名	説明	必須/任意	デフォルト値※
label:	ワークフローの表示名 **指定可能な値** 1〜128 文字	必須	
description:	ワークフローの説明 **指定可能な値** 1〜512 文字	必須	
nodes:		必須	
secretary		必須	
llm_name:	Secretary ノードに利用する LLM の名前 **指定可能な値** 半角英数字とアンダースコア 1〜128 文字	必須	
supervisor:		任意	
llm_name:	Supervisor ノードに利用する LLM の名前 **指定可能な値** 半角英数字とアンダースコア 1〜128 文字	必須	
child_agents:	Supervisor が利用できるエージェントノード。02-30 ではリストで指定します。02-40 以降は以降のパラメータで指定します。 **02-30 で指定可能な値** knowledge_search_agent, event_analysis_agent	必須	
knowledge_search_agent:	Supervisor が利用できるナレッジ検索エージェントを指定します。	任意	
- name:	ナレッジ検索エージェントの名前を指定します。 **指定可能な値** knowledge_search_agent	必須	
event_analysis_agent:	Supervisor が利用できる運用管理データ取得エージェントを指定します。	任意	
- name:	運用管理データ取得エージェントの名前を指定します。 **指定可能な値** event_analysis_agent	必須	
custom_agent:	Supervisor が利用できるカスタムエージェントを指定します。 **指定可能な要素数** 8	任意	

項目名	説明	必須/任意	デフォルト値※
- name:	カスタムエージェントの名前を指定します。 **指定可能な値** secretary, supervisor, knowledge_search_agent, event_analysis_agent 以外の，"_"で始まらない半角英数字（小文字）とアンダースコア 1～64 文字	必須	
override_agent_parameter:	上書きするカスタムエージェントのパラメータ	任意	
- key:	カスタムエージェントの上書きするパラメータ名 **指定可能な値** llm_name	必須	
value:	カスタムエージェントの上書きするパラメータの値 **指定可能な値** 半角英数字とアンダースコア 1～128 文字	必須	
ui:	UI に関する設定を定義します。	任意	-
template:	チャット画面の質問エリアに自動的に生成される質問文のテンプレートを定義します。変数（{jp1events}）と制御文字（¥n）を使用可能です。展開に失敗した場合は，変数が文字列（"{jp1events}"など）"として展開されます。 {jp1events}には jp1event_template で指定した内容が入り，イベントが複数存在する場合は改行で結合します。 設定値に空文字を指定した場合は質問文を生成しません。 **指定可能な値** 0～4096 文字	任意	"{jp1events}¥n状況を教えてください。"
jp1event_template	チャット画面の質問エリアに自動的に生成される質問文のテンプレート（ui.template）で{jp1events}を指定した場合に挿入する内容のテンプレートを指定します。 変数と制御文字（¥n）が使用可能です。変数を指定する場合は{}で囲む必要があります。 変数には以下の JP1 イベントの属性が指定可能です。 B.ID, B.MESSAGE, B.SEQNO, B.TIME, E.JP1_SOURCEHOST, E.OBJECT_TYPE, E.SEVERITY, E.USER_NAME	任意	”{B.MESSAGE}”

項目名	説明	必須/任意	デフォルト値※
jp1event_template	(02-60 以降) 02-50 以前に加えて，以下の JP1 イベントの属性が指定可能です。 E.@JP1IM_DISPLAY_MESSAGE	任意	"{B.MESSAGE}"
chat:		任意	-
max_size:	生成 AI へ投げる質問の文字数です。 **指定可能な値** 1～8192	任意	2048
max_display_size:	チャット画面に表示可能な文字数の上限です。 **指定可能な値** 1～50,000	任意	8192
hyperlink:	ハイパーリンク化に関する情報を定義します。最大 10 件まで指定可能です。	任意	-
- regexp:	ハイパーリンク化する文字列を指定します。 **指定可能な値** 1～128 文字	必須	0: "¥¥[document_url:¥¥s*(https?:¥¥/¥¥/[^¥¥[¥¥]]*)¥¥] "
- replace:	ハイパーリンク化する文字列の置換後の表示名を指定します。 **指定可能な値** 1～128 文字	必須	0: "[document_url:URL] "
logs:	ログに関する設定を定義します。	任意	-
filter:	ログの出力抑止の有効化（true）／無効化（false）の設定です。 **指定可能な値** true または false	任意	true

注※　設定ファイルのデフォルト値ではなく，指定がない場合に利用されるデフォルト値です。

リージョン一覧

us-east-1, us-east-2, us-west-1, us-west-2, af-south-1, ap-east-1, ap-south-2, ap-southeast-3, ap-southeast-4, ap-southeast-5, ap-southeast-6, ap-south-1, ap-northeast-3, ap-northeast-2, ap-southeast-1, ap-southeast-2, ap-east-2, ap-southeast-7, ap-northeast-1, ca-central-1, ca-west-1, eu-central-1, eu-west-1, eu-west-2, eu-south-1, eu-west-3, eu-south-2, eu-north-1, eu-central-2, il-central-1, mx-central-1, me-south-1, me-central-1, sa-east-1

6.1.2　プロンプト定義ファイル（prompts.chat.ja.yml）

(1)　02-30 以降

形式と記述内容が 02-20 以前と異なります。それ以外は 02-20 を参照してください。

形式

YAML 形式で記述します。

```
agents:
  knowledge_search_agent:
    - name: プロンプト名
      prefix:
        - name: valueを識別するための任意の文字列
          value: プロンプトに定義する文字列
      format_instructions: AIの応答の型を支持する文字列
      description: エージェントノードの説明
      example: Supervisorが遷移先を決める際に利用するエージェントノードの利用例
  custom_agent:
    - name: プロンプト名
      prefix:
        - name: valueを識別するための任意の文字列
          value: プロンプトに定義する文字列
      format_instructions: AIの応答の型を支持する文字列
      description: エージェントノードの説明
      example: Supervisorが遷移先を決める際に利用するエージェントノードの利用例

operators:
  secretary:
    - name: プロンプト名
      instruction: 動作指示
```

記述内容

項目名	説明	必須／任意
agents:	エージェントノードに利用するプロンプト	必須
knowledge_search_agent:	ナレッジ検索エージェントに利用するプロンプト **指定可能な要素数** 15	必須
- name:	プロンプトの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.prompts.name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
prefix:	プロンプトの最初に記載する文字列です。配列を順番に結合し，文字列を作成します。 **指定可能な要素数**	必須

項目名	説明	必須／任意
prefix:	5	必須
- name:	項目 prefix.value を識別するための任意の文字列を指定します。 **指定可能な値** 半角英数字（小文字）とアンダースコア 1～64 文字	必須
value:	プロンプトに定義する文字列を指定します。 **指定可能な値** 半角英数字（小文字）とアンダースコア 1～8192 文字	必須
format_instructions:	AI の応答の型を指示する文字列を指定します。プレースホルダ"{tools_names}"を含む必要があります。"{tools_names}"には設定ファイル（ai_assistant.yml）で指定した service.langchain.prompts.tools で指定した値が展開されます。 **指定可能な値** プレースホルダ"{tools_names}"を含む 1～8192 文字	必須
description:	Supervisor が参照するエージェントノードの説明を指定します。Supervisor はこの説明を見て，エージェントノードを利用すべきか判断します。 **指定可能な値** 1～1024 文字	必須
example:	Supervisor が遷移先を決める際に利用するエージェントノードの利用例を指定します。Supervisor が遷移先を決める際の補足となる利用例を記載します。 **指定可能な値** 0～1024 文字	必須
custom_agent:	カスタムエージェントに利用するプロンプト **指定可能な要素数** 50	必須
- name:	プロンプトの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.prompts.name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
prefix:	プロンプトの最初に記載する文字列です。配列を順番に結合し，文字列を作成します。 **指定可能な要素数** 5	必須
- name:	項目 prefix.value を識別するための任意の文字列を指定します。 **指定可能な値**	必須

項目名	説明	必須／任意
- name:	半角英数字（小文字）とアンダースコア 1～64 文字	必須
value:	プロンプトに定義する文字列を指定します。 **指定可能な値** 1～8192 文字	必須
format_instructions:	AI の応答の型を指示する文字列を指定します。プレースホルダ"{tools_names}"の指定は任意です。展開される値は"knowledge_search_agent"と同様です。 **指定可能な値** 1～8192 文字	必須
description:	Supervisor が参照するエージェントノードの説明を指定します。Supervisor はこの説明を見て，エージェントノードを利用すべきか判断します。 **指定可能な値** 1～1024 文字	必須
example:	Supervisor が遷移先を決める際に利用するエージェントノードの利用例を指定します。Supervisor が遷移先を決める際の補足となる利用例を記載します。 **指定可能な値** 0～1024 文字	必須
operators:	オペレーターノードに利用するプロンプト	任意
secretary:	Secretary ノードに利用するプロンプト **指定可能な要素数** 15	任意
- name:	プロンプトの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.prompts.name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
instruction	各ノードに対する動作指示を記載します。ワークフローの構成やノードの活用方法に応じて適切な内容を記載する必要があります。 **指定可能な値** 1～8192 文字	必須

デフォルト値は「付録 B.2 02-30 以降のプロンプトのデフォルト」を参照してください。

各項目の説明欄で明示的に許可されていないプレースホルダ("{<文字列>}")形式の文字列は使用できません。プロンプト内にプレースホルダ形式の文字列を記述したい場合は、エスケープを行ってください。

(2)　02-20 以前

形式

YAML 形式で記述します。

```
- name: プロンプト名
  prefix:
    - name: valueを識別するための任意の文字列
      value: プロンプトに定義する文字列
  format_instructions: AIの応答の型を指示する文字列
```

ファイル

prompts.chat.ja.yml

定義ファイル編集機能画面のタブ名

aiassistant

説明

生成 AI 連携で使用するプロンプトを定義するファイルです。

文字コード

UTF-8（BOM なし）

改行コード

LF

定義の反映時期

JP1/IM3 の定義ファイル操作機能を使ってアップロードした際に，プログラムが再起動して定義が反映されます。

記述内容

項目名			説明	必須／任意
-	name:		プロンプトの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.prompts.name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
	prefix:		プロンプトの最初に記載する文字列です。配列を順番に結合し，文字列を作成します。	必須
	-	name:	項目 prefix.value を識別するための任意の文字列を指定します。 **指定可能な値**	必須

項目名			説明	必須／任意
-	-	name:	半角英数字（小文字）とアンダースコア 1～64 文字	必須
		value:	プロンプトに定義する文字列を指定します。 **指定可能な値** 1～8192 文字	必須
	format_instructions:		AI の応答の型を指示する文字列を指定します。プレースホルダ"{tools_names}"を含む必要があります。"{tools_names}"には設定ファイル（ai_assistant.yml）で指定した service.langchain.prompts.tools で指定した値が展開されます。 **指定可能な値** プレースホルダ"{tools_names}"を含む 1～8192 文字	必須

デフォルト値は「付録 B.3 02-20 以前のプロンプトのデフォルト」を参照してください。

6.1.3　ツールプロンプト定義ファイル（tools.ja.yml）（02-40 以降）

形式

YAML 形式で記述します。

```
azure_ai_search: Azure AI Search検索ツールのツールプロンプト定義
  - name: ツールの名前
    description: ツールの説明
aws_knowledge_bases: Knowledge Bases for Amazon Bedrock検索ツールのツールプロンプト定義
  - name: ツールの名前
    description: ツールの説明
aws_lambda_fucntions: ユーザープログラム呼び出しツールのツールプロンプト定義
  -  name: ツールの名前
    description: ツールの説明
    args: ツールの引数
      - field_name: 引数の名前
        schema: 引数のJSONスキーマ
          title: 引数のタイトル
          description: 引数の説明
        annotation: 引数の型情報
          {型に依存した設定}
        nest: オブジェクト型引数の型情報
          - field_name: 引数の名前
            schema: 引数のJSONスキーマ
              title: 引数のタイトル
              description: 引数の説明
            annotation: 引数の型情報
              {型に依存した設定}
```

ファイル

tools.ja.yml

定義ファイル編集機能画面のタブ名

aiassistant

説明

生成 AI 連携で使用するプロンプトを定義するファイルです。

文字コード

UTF-8（BOM なし）

改行コード

LF

定義の反映時期

JP1/IM3 の定義ファイル操作機能を使ってアップロードした際に，プログラムが再起動して定義が反映されます。

記述内容

項目名		説明	必須／任意
azure_ai_search:		Azure AI Search 検索ツールのプロンプト群 **指定可能な要素数** 30	任意
-	name:	ツールの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.tools[].name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
	description:	ツールの説明を指定します。 **指定可能な値** 1～4096 文字	必須
aws_knowledge_bases:		Knowledge Bases for Amazon Bedrock 検索ツールのプロンプト群 **指定可能な要素数** 30	任意
-	name:	ツールの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.tools[].name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
	description:	ツールの説明を指定します。 **指定可能な値** 1～4096 文字	必須

項目名	説明	必須／任意
aws_lambda_function:	ユーザープログラム呼び出しツールのプロンプト群 **指定可能な要素数** 20	任意
- name:	ツールの識別名を指定します。設定ファイル（ai_assistant.yml）の service.langchain.tools[].name で指定した値を指定します。 **指定可能な値** "_"で始まらない半角英数字（小文字）とアンダースコア 1～128 文字	必須
description:	ツールの説明を指定します。 **指定可能な値** 1～4096 文字	必須
args:	ツールの引数設定。引数は 16 個まで定義できます。 **指定可能な要素数** 16	必須
- field_name:	引数の名前 **指定可能な値** 先頭と末尾が"_"になっておらず，かつ，"model_config"という名前ではない半角英字とアンダースコア。 1～32 文字	必須
schema:	引数の JSON スキーマ設定	必須
– title:	JSON スキーマの title 情報。引数を端的に表す文字列を指定する。任意であり，JSON スキーマをログ出力した際にわかりやすくするために利用できる。	任意
description:	JSON スキーマの description 情報。引数の説明を指定する。	必須
annotation:	引数の型アノテーション定義	必須
specify_type:	引数の型を指定します。specify_type に指定できる内容は**指定できる型情報**を参照してください。	必須
specify_opt:	引数に対するオプションを指定します。	
required:	引数の必須か指定します。 **指定可能な値** true（必須），false（任意）	必須
nullable:	引数に null を許可するか指定します。 **指定可能な値** true（null を許可），false（null を許可しない）	必須
nest:	object 型の引数の要素を設定します。specify_type に object 型または object_array 型を指定した場合のみ指定可能です。 **指定可能な要素数** 8	任意

項目名	説明	必須／任意
- field_name:	args の field_name と同じです。	同左
schema:	args の schema と同じです。	同左
title	args の schema.title と同じです。	同左
description	args の schema.description と同じです。	同左
annotation:	args の annotation と同じです。	同左
specify_type:	args の annotation.specify_type と同じです。ただし指定できる型は string 型，integer 型，number 型，boolean 型，enum 型，array 型です。	同左
specify_opt:	args の annotation.specify_opt と同じです。	同左
required:	args の annotation.specify_opt.required と同じです。	同左
nullable:	args の annotation.specify_opt.nullable と同じです。	同左

デフォルト値は「付録 B.3 02-20 以前のプロンプトのデフォルト」を参照してください。

指定できる型情報

annotation.specify_type に指定可能な型と，指定できる情報を以下に示します。

引数の型として string 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** string	必須
max:	最大文字数を指定します。 **指定可能な値** 1～102400	任意
min:	最小文字数を指定します。 **指定可能な値** 1～102400	任意

引数の型として integer 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** integer	必須

項目名	説明	必須／任意
max:	最大値を指定します。 **指定可能な値** -9223372036854775808～9223372036854775807	任意
min:	最小値を指定します。 **指定可能な値** -9223372036854775808～9223372036854775807	任意

引数の型として number 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** number	必須
max:	最大値を指定します。整数部分と小数部分が合計で 15 桁以内になるように指定します。 **指定可能な値** - 999999999999999～999999999999999	任意
min:	最小値を指定します。整数部分と小数部分が合計で 15 桁以内になるように指定します。 **指定可能な値** -999999999999999～999999999999999	任意

引数の型として boolean 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** boolean	必須

引数の型として enum 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** enum	必須
enum_value	enum として指定可能な値のリスト。文字列または整数どちらか一方の要素を指定できます。	必須

項目名	説明	必須／任意
enum_value	**指定可能な値** 文字列の enum の場合は 1〜32 文字 整数の enum の場合は-9223372036854775808〜9223372036854775807 配列の要素数は 1〜32 個	必須

引数の型として array 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。 **指定可能な値** array	必須
content_annotation:	配列の要素の型を指定します。指定しない場合，任意の型を許容します。	任意
field_type:	要素の型を指定します。 **指定可能な値** string, integer, number, enum	必須
max:	string, integer, number を指定した場合は指定します。指定可能な値は**指定できる型情報**と同じです。	必須
min:	string, integer, number を指定した場合は指定します。指定可能な値は**指定できる型情報**と同じです。	必須
enum_value:	enum を指定した場合は指定します。指定可能な値は**指定できる型情報**と同じです。	必須
max:	配列の要素数の最大を指定します。 **指定可能な値** 1〜1024	任意
min:	配列の要素数の最小を指定します。 **指定可能な値** 1〜1024	任意

引数の型として object 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。object 型変数の中身は nest に指定します。 **指定可能な値** object	必須

引数の型として object_array 型を指定

項目名	説明	必須／任意
specify_type:		必須
field_type:	型を指定します。object 型変数の中身は nest に指定します。 **指定可能な値** object_array	必須
max:	配列の要素数の最大を指定します。 **指定可能な値** 1～1024	任意
min:	配列の要素数の最小を指定します。 **指定可能な値** 1～1024	任意

6.2　メッセージ

JP1 Cloud Service 生成 AI 連携がログファイルとチャット画面に出力するメッセージについて説明します。

6.2.1　メッセージの概要

(1)　メッセージの出力形式

JP1 Cloud Service 生成 AI 連携が出力するメッセージの形式を次に示します。

```
KNBDnnnnn-Y メッセージテキスト
```

KNBDnnnnn-Y はメッセージ ID です。メッセージ ID を構成する項目の内容を次に示します。

KNBD

KNBD：JP1 Cloud Service 生成 AI 連携から出力されたメッセージであることを示します。

nnnnn

メッセージ番号を示します。

Y

メッセージの種類を示します。

E：エラーメッセージを示します。処理は中断されます。

W：警告メッセージを示します。メッセージ出力後，処理は続行されます。

I：通知メッセージを示します。ユーザーに情報を知らせます。

(2)　メッセージの出力先

メッセージの出力先を表に示します。

表中では，それぞれの出力先を次のように表記しています。

（凡例）

メッセージログ：「生成 AI 連携 メッセージログ」に出力されます。

LangChain 思考過程ログ：「生成 AI 連携 LangChain 思考過程ログ」に出力されます。

画面：チャット機能で提供する画面上に表示されます。

○：出力する

－：出力しない

表 6–1　メッセージの出力先一覧

メッセージ ID	出力先			出力抑止
	メッセージログ	LangChain 思考過程ログ	画面	
KNBD00000	○	–	○	–
KNBD20000～ KNBD20001	○	–	–	対象
KNBD20002～ KNBD20004	–	○	–	対象
KNBD20005～ KNBD20012	○	–	○	–
KNBD20013～ KNBD20015	○	–	–	–
KNBD20016～ KNBD20017	○	–	○	–
KNBD20018	○	–	–	–
KNBD20019～ KNBD20025	○	–	○	–
KNBD20026	–	○	–	–
KNBD20027	○	–	–	–
KNBD20028	○	–	○	–
KNBD20029	○	–	–	–
KNBD20030	○	–	○	–
KNBD20031	○	–	–	–
KNBD20032	–	○	–	対象
KNBD20033	○	–	○	–
KNBD20034～ KNBD20037	○	–	–	–
KNBD20038～ KNBD20040	○	–	○	–
KNBD20041～ KNBD20043	○	–	–	–

6.2.2　メッセージ一覧

(1)　メッセージの記載形式

本マニュアルでのメッセージの記載形式を次に示します。

メッセージ ID

> メッセージ本文（太字は埋め文字を示します）

契機：メッセージが出力される契機を示します。

対処：メッセージが出力されたときの対処方法を示します（メッセージの種類が E または W の場合）。

(2)　メッセージ一覧

KNBD00000-E

> 保存できる会話履歴の上限に達しました。不要なスレッドを削除してください。スレッドが削除されるまで表示されている会話内容は保存されません。

契機：会話履歴の保存上限に達したとき。

対処：不要なスレッドを削除してください。

KNBD20000-I

> {"question": "**質問文**", "user": "**JP1CS ユーザー名**", "workflow": "**ワークフロー名**"}

契機：問合せを送信したとき。

KNBD20001-I

> {"answer": "**回答文**", "recommends": "**レコメンド**", "workflow": "**ワークフロー名**"}

契機：問合せの回答を受信したとき。

KNBD20002-I

> {"prompt_question": "**プロンプトが生成した，はじめの質問**", "workflow": "**ワークフロー名**", "nodes": "**各ノードが使用するプロンプトと生成 AI サービスのリスト**", "invoke_id": "**問い合わせ識別子**"}

契機：問合せを送信したとき。

KNBD20003-I

{"node": "**ノード識別子**", "intermediate_steps": "**intermediate_steps**", "invoke_id": "**問い合わせ識別子**"}

契機：エージェントノードの応答を受信したとき。

KNBD20004-I

{"final_answer": "**生成 AI の最終回答**", "recommends": "**レコメンド**", "invoke_id": "**問い合わせ識別子**"}

契機：問合せの回答を受信したとき。

KNBD20005-E

連携する**接続先**への認証に失敗しました。接続情報およびシークレット情報を確認してください。エラー内容：**エラー詳細**

契機：連携する生成 AI または検索サービスへの認証に失敗したとき。

対処：エラー詳細に「No API key provided.」と表示される場合，生成 AI サービスの API キーをシークレット登録してください。エラー詳細に「Access denied due to invalid subscription key or wrong API endpoint.」と表示される場合，接続先に表示されるサービスの API キーまたはエンドポイントが正しく登録されていることを確認してください。エラー詳細に「Operation returned an invalid status 'Forbidden'」と表示される場合，接続先の API キーが正しく登録されていることを確認してください。

KNBD20006-E

リクエストがタイムアウトしました。エラー内容：**エラー詳細**

契機：リクエストがタイムアウトしたとき。

対処：時間をおいて再度実行してください。繰り返し発生する場合はシステム管理者に連絡してください。

KNBD20007-E

連携する生成 AI サービスによる問い合わせの繰り返し回数上限またはタイムアウトが発生しました。エラー内容：**エラー詳細**

契機：以下のいずれかに該当したとき。

- 連携する生成 AI または検索サービスからの返答が遅い
- 連携する生成 AI との問い合わせの繰り返し回数の上限に達した

対処：質問文を簡単な文章に変更して再度質問してください。繰り返し発生する場合は利用する接続先のモデルやリージョンなどの設定を見直してください。

KNBD20008-E

内部エラーが発生しました。（**エラー詳細**）

契機：内部エラーが発生したとき。

対処：JP1 Cloud Service サポートにお問い合わせください。

KNBD20009-E

連携する生成 AI サービスのトークン上限に達しました。時間をおいて再度実行してください。エラー内容：**エラー詳細**

契機：生成 AI のトークン上限に達したとき。

対処：時間をおいて再度実行してください。

KNBD20010-E

連携する**接続先**への接続に失敗しました。接続情報を確認してください。エラー内容：**エラー詳細**

契機：連携する生成 AI または検索サービスへの接続で問題が発生したとき。

対処：エラー詳細をもとに接続情報や接続先を見直してください。

KNB20011-E

連携する**接続先**の API 実行で次のエラーが発生しました。エラー内容：**エラー詳細**

契機：外部サービスの API 実行に失敗したとき。

対処：エラー詳細をもとに連携するサービスで問題が起こっていないか確認してください。

KNBD20012-E

連携する**接続先**の上限に達しました。エラー内容：**エラー詳細**

契機：外部サービスの何らかの上限に到達したとき。

対処：質問文を簡単な文章に変更して再度質問してください。

KNBD20013-I

生成 AI 連携サービスの起動に成功しました

契機：生成 AI 連携サービスの起動に成功したとき。

KNBD20014-I

生成 AI 連携サービスを終了します

契機：生成 AI 連携サービスを終了するとき。

KNBD20015-E

設定ファイルの読み込みに失敗しました。設定ファイルを見直してください。ファイル名:**ファイル名**, エラー: **エラー詳細**

契機：設定ファイルの読み込み・バリデーションで失敗したとき。

対処：エラー詳細をもとに設定ファイルを修正してください。

KNBD20016-E

設定ファイルにおいて、項目が定義されていません。（ファイル名:**設定ファイル名**，項目名:**項目名**）

契機：設定ファイルに項目が存在せずに失敗したとき。

対処：表示された項目名に関する設定を修正してください。

KNBD20017-E

連携する生成 AI サービスが設定されていません。（生成 AI サービス名:**生成 AI サービス名**）

契機：生成 AI サービスが設定されていないとき。

対処：設定ファイルの連携する生成 AI サービスに関する項目を修正してください。

KNBD20018-W

指定されたツール（ツール名：**ツール名**）が存在しません。スキップします。

契機：対応するツールが存在しない場合。

対処：設定ファイルのツールに関する項目を見直してください。

KNBD20019-E

連携する接続先への接続に失敗しました。ネットワーク設定を確認してください。エラー内容：**エラー詳細**

契機：ネットワーク設定などインフラにより連携する生成 AI または検索サービスへの接続に失敗したとき。

対処：連携するサービスのネットワーク設定を見直してください。

KNBD20020-E

登録対象がすでに存在するため、シークレットが登録できませんでした。(シークレットキー:**シークレットキー**)

契機：登録対象のシークレットがすでに存在するとき。

対処：登録しようとしているシークレットキーを見直してください。

KNBD20021-E

編集対象が存在しないため、シークレットが編集できませんでした。(シークレットキー: **シークレットキー**)

契機：編集対象または削除対象のシークレットが存在しないとき。

対処：編集または削除しようとしているシークレットキーを見直してください。

KNBD20022-E

シークレットの変更に失敗しました。他のユーザーと処理が重複した可能性があります。

契機：シークレットの追加/変更/削除時に排他を取得できないとき。

対処：時間をおいて再実行してください。

KNBD20023-E

バリデーションに失敗しました。（パラメータ：**パラメータ名**）

契機：バリデーションに失敗したとき。

対処：リクエスト内容を見直してください。

KNBD20024-E

生成 AI 連携基盤による問い合わせ実行処理でエラーが発生しました。エラー内容: **エラー詳細**

契機：生成 AI 連携基盤の問い合わせ処理の実行過程で，エラーが発生したとき。

対処：エラー詳細をもとに設定ファイルの見直しや，連携サービスの設定の見直し，リクエストの再実行をしてください。

KNBD20025-E

システムプロンプトの生成に失敗しました。エラー内容：**エラー詳細**

契機：プロンプトの追加または編集時に，プレースホルダを削除したり，想定していないプレースホルダを追加したとき。

対処：プロンプト定義ファイルを見直してください。

KNBD20026-I

ログの出力抑止は**有効**|**無効**です。

契機：サービス起動時。

KNBD20027-W

設定ファイルが存在しませんでした。バックアップからデフォルトの設定ファイルを復元しました。（ファイル名:**設定ファイル名**）

契機：設定ファイルを削除したとき。

KNBD20028-E

指定されたワークフローまたはワークフローに指定した内容が存在しません。設定情報を見直してください。（ワークフロー名：**ワークフロー名**）

契機：質問した時

対処：設定ファイルのワークフロー設定を見直してください。

KNBD20029-E

定義されたワークフローのバリデーションに失敗しました。（ワークフロー名：**ワークフロー名**, エラー内容：**エラー内容**）

契機：設定を変更したとき

対処：設定ファイルのワークフロー設定を見直し，エラー原因を取り除いてください。

KNBD20030-E

エージェントの実行中にエラーが発生しました。（エージェント名: **エージェント名**, エラー内容: **エラー内容**）

契機：質問時にエージェントノードでエラーが発生した場合

対処：エラー詳細をもとに設定ファイルの見直しや，リクエストの再実行をしてください。

KNBD20031-W

指定されたプロンプト（プロンプト名：**プロンプト名**）が存在しません。デフォルトのプロンプトを利用します。

契機：設定ファイルのプロンプト(service.langchain.prompts[])に設定されたプロンプトが，プロンプト定義ファイルに存在しない場合

KNBD20032-I

{"node": "**ノード識別子**", "message": "**次ノードへのメッセージ**", "invoke_id": "**問い合わせ識別子**"}

契機：オペレーターノードの応答を受信した時

KNBD20033-E

指定されたツール（ツール名：**ツール名**）が一つも存在しません。

契機：プロンプトで指定したツール群に関して、設定値が存在するツールが一つも存在しない場合

対処：設定ファイルのプロンプト(service.langchain.prompts[])のツール指定と、設定ファイルのツール(service.langchain.tools[])を確認してください。

KNBD20034-W

設定情報の整合性に問題があります。エラー内容はログを確認してください。警告状態で起動します。

契機：起動時に定義ファイル間の設定の整合性に問題がある場合。

対処：設定ファイルのプロンプト名やツール名がプロンプト定義ファイルやツールプロンプト定義ファイル名に存在するか確認してください。

KNBD20035-E

エージェントに対して割り当て不可能なツールが割り当てられました。（エージェント名：**エージェント名**, ツール名: **ツール名**）

契機：エージェントノードに対して，紐づけ不可能な種別のツールが割り当てられたとき。

対処：設定ファイルのプロンプト(service.langchain.prompts[])のツール指定(tools)と適用するノード指定(applied_node)が正しいか確認してください。

KNBD20036-I

ユーザープログラムを呼び出します。（invoke_id：**問い合わせ識別子**, ツール名: **ツール名**）

契機：ユーザープログラムを呼び出すとき。

KNBD20037-I

ツール名: **ツール名**に設定されたスキーマを出力します。スキーマ: **JSON スキーマ**

契機：ユーザープログラム呼び出し機能を利用できるカスタムエージェントが含まれるワークフローが呼び出されたとき。

KNBD20038-E

ユーザープログラムから不正な値が返却されました。（ツール名：**ツール名**, リクエスト ID: **リクエスト ID**, エラー内容: **エラー内容**）

契機：ユーザープログラムの戻り値が不正なとき。

対処：ユーザープログラムの戻り値を見直してください。

KNBD20039-E

ユーザープログラムの実行処理中にエラーが発生しました。（ツール名：**ツール名**, リクエスト ID: **リクエスト ID**, エラー内容: **エラー内容**）

契機：ユーザープログラムの実行処理中にエラーが発生したとき。

対処：ユーザープログラムの処理や設定を見直してください。

KNBD20040-E

ユーザープログラムの実行サービスでエラーが発生しました。（ツール名：**ツール名**, リクエスト ID: **リクエスト ID**, エラー内容: **エラー内容**）

契機：ユーザープログラムの実行サービスでエラーが発生したとき。

対処：ユーザープログラムの処理や設定，ユーザープログラムの実行サービスの設定を見直してください。

KNBD20041-W

ワークフローにおいて仕様で定められた値を逸脱した設定を検知しました。（ワークフロー名：**ワークフロー名**, 警告内容: **警告内容**）

契機：ワークフローで利用するエージェントノードやツールの設定が仕様から逸脱している場合。

対処：警告内容を確認と仕様を確認してください。

KNBD20042-E

生成 AI 連携基盤が同時に処理できるリクエスト数の上限に達しました。時間をおいて再度実行してください。

契機：生成 AI 連携基盤へのリクエストにおいて，同時実行数の上限を超過したとき。

対処：時間をおいて再度実行してください。

KNBD20043-E

指定されたワークフローは実行できません。(ワークフロー名: **ワークフロー名**, エラー内容: **エラー内容**)

契機：サポートされていないワークフローを呼び出したとき。

対処：エラー内容を確認し，サポートされているワークフローが呼び出されるようにしてください。

付録

付録 A　連携サービス

ここでは，「2.1.2(1) 連携サービス」で提示している連携サービスについて，02-20 以前の利用料金の参考情報を示します。

重要

- 本付録の情報は，2024 年 3 月現在のものです。各サービスの詳細および利用料金の最新情報は，各サービスの公式 HP を参照してください。
- 利用料金の例は，Microsoft 社が提供している料金計算ツールを使用して計算した値または AWS が公開しているモデルの料金表をもとに計算した値です。サポートやオプションの利用料金は含まれません。
- JP1 Cloud Service 生成 AI 連携において，Azure Open AI のモデルは「言語モデル」を想定しています。Amazon Bedrock のモデルは「anthropic.claude-3-haiku」を想定しています。
- JP1 Cloud Service 生成 AI 連携において，Azure AI Search のレベルは「Standard S1」を想定しています。Knowledge Bases for Amazon Bedrock はベクトルストアに OpenSearch Serveless を想定しています。
- 本付録で提示しているトークン数は一例となります。問い合わせ実行ごとに値が変化します。

付録 A.1　Azure Open AI

利用料金に関連する主なパラメータは以下の通りです。

#	パラメータ名	説明
1	リージョン	本サービスを利用するリージョン。
2	モデル（バージョン）	生成 AI のモデル。
3	プロンプトのトークン数	LangChain が生成 AI へ入力するトークン数。1 つの問い合わせ文に対し，LangChain から生成 AI へ入力したトークンの合計数。
4	完了のトークン数	生成 AI が LangChain へ出力するトークン数。1 つの問い合わせ文に対し，生成 AI が LangChain へ出力したトークンの合計数。

例 1

たとえば，下記の場合，利用料金は，1 問い合わせあたり，およそ$0.011 です。

生成 AI 連携において，ドキュメント検索サービスの連携は無い状態とします。

問い合わせ文：

CPU 使用率が高くなっています。考えられる原因を教えてください。

#	パラメータ名	値
1	リージョン	East US
2	モデル名（バージョン）	gpt-4o(2024-05-13)
3	プロンプトのトークン数	982※
4	完了のトークン数	421※

注※

同じ問い合わせ文を 5 回実施した結果の平均値を示します。小数点以下は切り上げています。

例 2

たとえば，下記の場合，利用料金は，1 問い合わせあたり，およそ$0.017 です。

生成 AI 連携において，ドキュメント検索サービスの連携がある状態とします。

ここでのドキュメント検索サービスの連携は，サンプルコンテンツ「JP1CloudService-生成 AI 連携_サンプルコンテンツ_インデックス作成手順例.zip」に沿うものとします。

問い合わせ文：

IT システム「サンプルシステム」では，どのポートを監視すべきですか。検索して教えてください。

#	パラメータ名	値
1	リージョン	East US
2	モデル名 (バージョン)	gpt-4o(2024-05-13)
3	プロンプトのトークン数	2900※
4	完了のトークン数	192※

注※

同じ問い合わせ文を 5 回実施した結果の平均値を示します。

付録 A.2　Azure AI Search

利用料金に関連する主なパラメータは以下の通りです。

#	パラメータ名	説明
1	リージョン	本サービスを利用するリージョン。
2	価格レベル	Azure AI Search の価格レベル。「Basic」や「Standard S1」などのレベルが選択できます。
3	ユニット	使用可能な合計容量の単一増分。
4	時間	デプロイメントした時間。

例 1

たとえば，下記の場合，利用料金は，1 か月あたり，およそ$324.12 となります。

#	パラメータ名	値
1	リージョン	Japan East
2	レベル	Standard S1
3	ユニット	1
4	時間	1 か月

付録 A.3　Amazon Bedrock

利用料金に関連する主なパラメータは以下の通りです。

#	パラメータ名	説明
1	リージョン	本サービスを利用するリージョン。
2	モデル（バージョン）	生成 AI のモデル。
3	プロンプトのトークン数	LangChain が生成 AI へ入力するトークン数。1 つの問い合わせ文に対し，LangChain から生成 AI へ入力したトークンの合計数。
4	完了のトークン数	生成 AI が LangChain へ出力するトークン数。1 つの問い合わせ文に対し，生成 AI が LangChain へ出力したトークンの合計数。

例 1

たとえば，下記の場合，利用料金は，1 問い合わせあたり，およそ$0.0010 です。

生成 AI 連携において，ドキュメント検索サービスの連携は無い状態とします。

問い合わせ文：

CPU 使用率が高くなっています。考えられる原因を教えてください。

#	パラメータ名	値
1	リージョン	us-east-1
2	モデル名	anthropic.claude-3-haiku-20240307-v1:0
3	プロンプトのトークン数	1947※
4	完了のトークン数	384※

注※

同じ問い合わせ文を 5 回実施した結果の平均値を示します。小数点以下は切り上げています。

例 2

たとえば，下記の場合，利用料金は，1 問い合わせあたり，およそ$0.0012 です。

生成 AI 連携において，ドキュメント検索サービスの連携がある状態とします。

ここでのドキュメント検索サービスの連携は，サンプルコンテンツ「JP1CloudService-生成 AI 連携_サンプルコンテンツ_インデックス作成手順例.zip」に沿うものとします。

問い合わせ文：

IT システム「サンプルシステム」では，どのポートを監視すべきですか。検索して教えてください。

#	パラメータ名	値
1	リージョン	us-east-1
2	モデル名（バージョン)	anthropic.claude-3-haiku-20240307-v1:0
3	プロンプトのトークン数	3385※
4	完了のトークン数	316※

注※

同じ問い合わせ文を 5 回実施した結果の平均値を示します。

付録 A.4　Knowledge Bases for Amazon Bedrock

Knowledge Bases for Amazon Bedrock 利用料金に関連する主なパラメータは以下の通りです。

#	パラメータ名	説明
1	リージョン	本サービスを利用するリージョン。
2	ベクトルストア	ベクトルストア
3	データストア	ベクトルストアのデータストア
4	時間	デプロイメントした時間。

例えば，下記の場合，利用料金は 1 か月あたり，およそ$250 となります。

#	パラメータ名	値
1	リージョン	ap-northeast-1
2	ベクトルストア	OpenSearchServerless
3	データストア	S3
4	時間	1 か月

付録 B　プロンプト

JP1 Cloud Service 生成 AI 連携で使用するプロンプトを以下に示します。{} は埋め文字です。

```
System: {システムプロンプト}
{過去会話}
Human: {質問文}
```

埋め文字について内容を以下に示します。

#	項目	内容
1	{システムプロンプト}	設定ファイル（ai_assistant.yml）で指定したプロンプト名（service.langchain.prompts.name）を元にプロンプト定義ファイル（prompts.chat.ja.yml）で設定した値が指定されます。 プロンプト定義ファイルに n 個の prefix を設定した場合，以下の形式でシステムプロンプトが指定されます。 prefix.[0].value prefix.[1].value … format_instructions
2	{過去会話}	質問を送信した際に，同じスレッドですでに会話をしている場合，質問より前の会話履歴が古い順に会話の数だけ以下の形式で指定されます。 HUMAN：ユーザーの入力文 Assistant：生成 AI の回答 スレッド内で初めての質問の場合は何も指定されません。
3	{質問文}	チャット画面にて入力した質問文が指定されます。

付録 B.1　02-40 以降のプロンプトのデフォルト

ナレッジ検索エージェントが利用するデフォルトのプロンプトは以下の通り設定されています。

```
あなたの目的:
あなたはITシステムに関する質問をサポートするアシスタントです。
ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答します。
ツールの実行結果:
1. 出典リスト:
    ツールの実行結果として、出典リストが与えられます。
    ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
    出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
2. 出典リストのフォーマット:
    各出典は、以下の形式で表されます。
    ------------------------
    sourcepage: 出典元ページ
    content: 実際の情報
```

```
    document_url: 出典元URL
    ------------------------
3. 出典を回答に含める:
    回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
        ・出典を参照するには、四角いブラケットを使用して、［sourcepage: 出典元ページ][docume
nt_url: 出典元URL］という形式にします。接頭辞として"sourcepage"と"document_url"をつけること
は重要です。
        ・出典は組み合わせず、各出典を別々に記載してください。例えば、［sourcepage: 出典元
ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2］のよ
うに記載します。
        ・出典の記載以外で四角いブラケット"[]"を使用しないでください。出典内の記述に四角いブ
ラケット"[]"が使われている場合、"<>"などの他の記号に置き換えるか、消して引用して下さい。"
あなたの振る舞い:
    1.質問の理解:
        ユーザーの質問を正確に理解し、適切な応答を準備します。
        質問されたことにのみ回答し、不要な情報を提供しないようにします。
    2.次にすべきことの考察:
        質問に回答するために、次に何をすべきかを常に考えます。
        あなたが行える行動は次の2つです:
            ・適切なツールを呼び出す
            ・ユーザーに回答する
        適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
        ユーザーがチケットの作成を依頼した場合は、チケットの作成を提案してください。
    3.ツールの利用:
        あなたが使えるツールは以下です。
        [{tools_names}]
        ツールを実行し、結果を観察して次に何をすべきか考えます。
        ツールの実行は原則として1回までとします。
        1回のツール実行で得られた情報が回答に十分でない場合のみ、追加でツールを実行しても構
いません。
        ツールの実行は最大10回まで行えます。

インストラクション:
    回答は以下の形式で作成してください
        1.回答には決してマークダウン記法を利用しないでください
            例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
        2.プレーンテキスト形式で出力してください
```

ユーザーコンテンツ取得機能で利用するツールの説明のデフォルトのプロンプトは以下の通り設定されています。

Azure AI Search 検索ツールの場合

```
ITシステムに関する情報(マニュアル、手順書、仕様書 等)を検索できる。複数の単語を入力して検
索。（入力例: JP1 障害）
```

Knowledge Bases for Amazon Bedrock 検索ツールの場合

```
ITシステムに関する情報(マニュアル、手順書、仕様書 等)を検索できる。複数の単語や、文章を入力
して検索。（入力例: JP1の障害について教えて）
```

付録 B.2　02-30 以降のプロンプトのデフォルト

Secretary が利用するデフォルトのプロンプトは以下の通り設定されています。

```
あなたは有能な秘書です。
```

Secretary が次の遷移先を決める際に利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

```
以下の会話における直前の入力に対して
下記の指示に従って返事（'message'）、次の対応者（'next'）、を決定してください。
なお、次の対応者（'next'）は必ず設定してください。
1.直前の入力が感謝の言葉や終わりの挨拶、日常会話の場合:
 'message'に返しの挨拶を設定し、'next'の設定は'FINISH'としてください。
2.直前の入力が日常会話でない場合:
 'message'は設定せず、'next'の設定は'supervisor'としてください。
```

Supervisor が利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

```
あなたは、ワーカーの会話を制御する任務を負ったスーパーバイザーです。
あなたの配下にいるワーカーは以下の通りです:
```
{agent_descs}
```
各ワーカーはタスクを実行し、その結果とステータスを返信します。
```

Supervisor が次の遷移先を決める際に利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

- Secretary から指示を受けて次の遷移先を決める場合

```
以下は会話の履歴です。
会話の最後のメッセージがユーザーからのリクエストになります。これに対して次のいずれかの対応をしてください。
リクエスト対して，そこまでの会話履歴の情報のみであなた自身が回答できると判断した場合は、
`message`にその回答と'next'に'FINISH'を設定してください。
そうでない場合には、ワーカーの中からリクエストに対する回答者を決定し`next`に設定してください。
なおこの場合は`message`にはワーカーへの指示を入力してください。
```

- エージェントノードの回答を受けて次の遷移先を決める場合

```
以下は会話の履歴です。
会話の最後のメッセージはそれ以前のユーザーからのリクエストに対するワーカーの回答になります。
以下に従って対応してください。
Step-1: 直前のワーカーの回答がさらに他のワーカーの調査により回答を最適にできるか判断してください。
例えば、次のような場合他のワーカーの調査が有効になります。
{agent_examples}
Step-2: Step-1で他のワーカーの追加調査が有効と判断した場合、追加調査の内容を具体化し、
具体化した調査内容の指示を`message`に設定してください。
```

実際に次のワークーに回すかはユーザーに確認しますので、`message`も確認の疑問文でなく「～してください」の依頼形でお願いします。

ナレッジ検索エージェントが利用するデフォルトのプロンプトは以下の通り設定されています。

```
あなたの目的:
あなたはITシステムに関する質問をサポートするアシスタントです。
ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答します。
ツールの実行結果:
1. 出典リスト:
    ツールの実行結果として、出典リストが与えられます。
    ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
    出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
2. 出典リストのフォーマット:
    各出典は、以下の形式で表されます。
    -------------------------
    sourcepage: 出典元ページ
    content: 実際の情報
    document_url: 出典元URL
    -------------------------
3. 出典を回答に含める:
    回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
        ・出典を参照するには、四角いブラケットを使用して、 [sourcepage: 出典元ページ][document_url: 出典元URL] という形式にします。接頭辞として"sourcepage"と"document_url"をつけることは重要です。
        ・出典は組み合わせず、各出典を別々に記載してください。例えば、 [sourcepage: 出典元ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2] のように記載します。
        ・出典の記載以外で四角いブラケット"[]"を使用しないでください。出典内の記述に四角いブラケット"[]"が使われている場合、"<>"などの他の記号に置き換えるか、消して引用して下さい。"
あなたの振る舞い:
    1.質問の理解:
        ユーザーの質問を正確に理解し、適切な応答を準備します。
        質問されたことにのみ回答し、不要な情報を提供しないようにします。
    2.次にすべきことの考察:
        質問に回答するために、次に何をすべきかを常に考えます。
        あなたが行える行動は次の2つです:
            ・適切なツールを呼び出す
            ・ユーザーに回答する
        適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
        ユーザーがチケットの作成を依頼した場合は、チケットの作成を提案してください。
    3.ツールの利用:
        あなたが使えるツールは以下です。
        [{tools_names}]
        ツールを実行し、結果を観察して次に何をすべきか考えます。
        ツールを用いて得られた情報が回答に十分な場合でも、念のためもう1度、他のツールを用いて情報を参照することを推奨します。
        ツールの実行は最大10回まで行えます。

インストラクション:
    回答は以下の形式で作成してください
        1.回答には決してマークダウン記法を利用しないでください
            例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
        2.プレーンテキスト形式で出力してください
```

運用管理データ取得エージェントが利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

```
あなたの目的:
あなたは、イベント情報を扱うエージェントで、イベントのエスカレーション要否を判定したり、
イベントが指定の条件に合致するか調べたり、イベントを検索したりできます。
もし選択イベントが指定の条件を満たすかどうかを判定する場合、
必ず次の順番で tool を動かしてください。
  1 retrieve_day_details（必ず実行すること）
  2 check_event_conditions
  3 search_related_events
  4 calc_event_frequency
インストラクション:
  回答は以下の形式で作成してください
      1.回答には決してマークダウン記法を利用しないでください
          例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
      2.プレーンテキスト形式で出力してください
```

- 運用管理データ取得エージェント内のツール(retrieve_day_details)が利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

```
時刻を表す引数から、曜日つきの時刻の文字列を作成して返す。
引数の日時だけを渡してください。
引数の例：
  year: 2024
  month: 01
  day: 05
  hour: 07
  minute: 15
  second: 25
```

- 運用管理データ取得エージェント内のツール(check_event_conditions)が利用するデフォルトのプロンプトは以下の通り設定されています。プロンプトは下記表示順に必要に応じて利用します。これらは変更できません。

```
・イベントが選択されているときに使用すること

イベントが指定の条件を満たすか否かを判定します。
引数として条件(condition)と選択イベント(selected_event)をとります。
条件(condition)は、満たすべき条件の説明文です。条件は、このイベントの属性(時間や発生サーバなど)だけから判定可能な条件だけを書いてください。
選択イベント(selected_event)は、判定対象のイベントです。
```

```
与えたイベントが、条件を満たすか否か判定してください。
サーバ名をチェックするときは、サーバ名条件の正規表現を考えてからチェックしてください。
コードではなく判定結果を返してください。
情報が足りず、このtoolの実行のみで条件が判断できない場合は「関連イベント検索結果が必要です。」としてください。

判定結果を引数のEvaluationResultに入れてください。

 イベント：
```

```
{selected_event}

条件：
{condition}
 補足説明：
Severityには、緊急（Emergency），警戒（Alert），致命的（Critical），エラー（Error），警告（Warning），情報（Information），通知（Notice），デバッグ（Debug）があります。値として、Emergency、 Alert などが格納されます。
重大な順に Emergency >= Alert >= Critical >= Error > Warning > Information >= Notice >= Debug です。
```

- 運用管理データ取得エージェント内のツール(search_related_events)が利用するデフォルトのプロンプトは以下の通り設定されています。プロンプトは下記表示順に必要に応じて利用します。これらは変更できません。

```
・イベント検索が必要なときに使用すること

イベントを検索するための検索条件を作成し、その検索条件を満たすイベントを返します。
検索条件は、検索したいイベントのメッセージ、検索期間、発生元サーバ名についての情報を含めて作成してください。
引数としてタスク(task)と検索条件(message、time_from,time_to,host)をとります。
タスクは、作成したい検索条件の説明文です。タスクには、判定条件から作成した検索条件だけを書いてください。
task: イベントの検索条件。
message: 検索したいイベントのメッセージ。

以下６つの引数は検索期間の始まりの日時です。
引数の入力は任意ですが入力する場合は必ず年から秒の6個はすべて入力してください。
time_from_year: 検索期間の始まりの時間の年を表す。
time_from_month: 検索期間の始まりの時間の月を表す。
time_from_day: 検索期間の始まりの時間の日を表す。
time_from_hour: 検索期間の始まりの時間の時を表す。
time_from_minute: 検索期間の始まりの時間の分を表す。
time_from_second: 検索期間の始まりの時間の秒を表す。

以下６つの引数は検索期間の終わりの日時を示す。
引数の入力は任意ですが入力する場合は必ず年から秒の6個はすべて入力してください。
time_to_year: 検索期間の終わりの時間の年を表す。
time_to_month: 検索期間の終わりの時間の月を表す。
time_to_day: 検索期間の終わりの時間の日を表す。
time_to_hour: 検索期間の終わりの時間の時を表す。
time_to_minute: 検索期間の終わりの時間の分を表す。
time_to_second: 検索期間の終わりの時間の秒を表す。

host: 検索対象のサーバ名。明示されていないなら空文字("")にすること。
```

```
以下のイベント情報から、メッセージが以下の抽出メッセージと同じ内容のものだけを抽出して下さい。
抽出メッセージ：{message}
ただし、メッセージの中に、時刻、何かの名前、などの変数が含まれる場合もあります。変数が異なっていても、それ以外の部分が同じ場合は、同じ内容のメッセージとして抽出してください。

では、以下の手順で考えてみましょう。
1. まず、抽出メッセージに時刻、何かの名前などと思われる文字列が含まれるかを確認します。
2. 含まれる場合は、その文字列以外の部分を参照し、抽出するイベントのメッセージの条件を考えま
```

```
す。
3. 最後にイベント情報の中から、2で考えた条件の通りにイベントを抽出します。

抽出したイベントのDB内通し番号（SEQNO）のリストを引数に入力してください。
{data}
```

```
以下のメッセージの中から、変数(埋め文字)にあたると推定される部分を抽出してください。エラーコードにあたる文字列は抽出しないでください。変数がない場合もあります。
変数にあたる文字列を改行区切りで回答してください。
{message}
```

- 運用管理データ取得エージェント内のツール(calc_event_frequency)が利用するデフォルトのプロンプトは以下の通り設定されています。これは変更できません。

```
イベント件数や発生回数を数えたいとき、イベント発生頻度を計算するときに使うツールです。例えば、イベントが1時間に何回発生しているか計算します。
複数の期間を与えて、期間ごとのイベント件数をまとめて計算するツールです。期間が複数あるときに使うツールです。

以下６つの引数は頻度計算の基準の日時を示す。
time_from_year: 基準時間の年を表す。
time_from_month: 基準時間の月を表す。
time_from_day: 基準時間の日を表す。
time_from_hour: 基準時間の時を表す。
time_from_minute: 基準時間の分を表す。
time_from_second: 基準時間の秒を表す。

引数の event_time_list は、イベントの発生時刻のリストです。
引数の durations は、期間(秒数)を表す整数(int)のリストです。

基準時刻を「T」とします。Tを基準として、TからT-1分前までの期間、TからT-5分前までの期間のイベント件数を計算したいとします。
この場合、durationsは、[60、 300] となります。それぞれ、1分の秒数、5分の秒数です。

返り値の例：
[{{
    'duration': 30分間,
    'count': 7件
}}]
```

付録 B.3　02-20 以前のプロンプトのデフォルト

生成 AI 連携のプロンプト定義ファイル（prompts.chat.ja.yml）ではデフォルトのプロンプトとして「sys_chat_prompt」と「sys_chat_prompt_for_link」の 2 つのプロンプトが以下のように設定されています。

```
- name: sys_chat_prompt
  prefix:
    - name: role
      value: |
        あなたの目的:
```

```
        あなたはITシステムに関する質問をサポートするアシスタントです。
        ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答します。
    - name: tool
      value: |
        ツールの実行結果:
        1. 出典リスト:
            ツールの実行結果として、出典リストが与えられます。
            ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
            出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
        2. 出典リストのフォーマット:
            各出典は，以下の形式で表されます。
            ------------------------
            sourcepage: 出典元ページ
            content: 実際の情報
            document_url: 出典元URL
            ------------------------
        3. 出典を回答に含める:
            回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
                ・出典を参照するには、四角いブラケットを使用して、 [sourcepage: 出典元ページ][document_url: 出典元URL] という形式にします。接頭辞として"sourcepage"と"document_url"をつけることは重要です。
                ・出典は組み合わせず、各出典を別々に記載してください。例えば、 [sourcepage: 出典元ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2] のように記載します。
                ・出典の記載以外で四角いブラケット"[]"を使用しないでください。出典内の記述に四角いブラケット"[]"が使われている場合、"<>"などの他の記号に置き換えるか、消して引用して下さい。
  format_instructions: |
    あなたの振る舞い:
        1.質問の理解:
            ユーザーの質問を正確に理解し、適切な応答を準備します。
            質問されたことにのみ回答し、不要な情報を提供しないようにします。
        2.次にすべきことの考察:
            質問に回答するために、次に何をすべきかを常に考えます。
            あなたが行える行動は次の2つです:
                ・適切なツールを呼び出す
                ・ユーザーに回答する
            適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
        3.ツールの利用:
            あなたが使えるツールは以下です。
            [{tools_names}]
            ツールを実行し、結果を観察して次に何をすべきか考えます。
            ツールを用いて得られた情報が回答に十分な場合でも、念のためもう1度、他のツールを用いて情報を参照することを推奨します。
            ツールの実行は最大10回まで行えます。

    インストラクション:
        回答は以下の形式で作成してください
            1.回答には決してマークダウン記法を利用しないでください
                例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
            2.プレーンテキスト形式で出力してください
- name: sys_chat_prompt_for_link
  prefix:
    - name: role
      value: |
        あなたの目的:
```

```
        あなたはITシステムに関する質問をサポートするアシスタントです。
        ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答します。
    - name: link
      value: |
        チケットの作成の提案:
        ユーザーがチケットの作成を依頼した場合は以下のルールでリンクを作成してください。チケットの作成の意思が明確ではない場合は提案しないでください。不要な作業を行わないことは、ユーザーに役に立つ回答を提供するために重要です。
            ・マルチバイトの括弧を用いて【link: https://redmine.example.com/redmine/projects/system/issues/new?issue[description]=<問題の内容>&issue[severity]=<問題の重大度>】の形式で出力してください。
            ・<問題の内容>の生成は以下のステップバイステップの手順で行います。
                1.会話履歴から問題の内容を簡潔な題名に要約します。
                2.要約した内容をURLエンコードします。決して、URLエンコードを省略しないでください。決められた手順に従うことは、ユーザーに役に立つ回答を提供するために重要です。
                例：<問題の内容>が”問題サンプル”の場合、URLエンコード後の文字列は”%E5%95%8F%E9%A1%8C%E3%82%B5%E3%83%B3%E3%83%97%E3%83%AB”となります。
                3.URLエンコードした内容をURLの<問題の内容>に代入してください。
            ・<問題の重大度>には会話履歴から問題の重大度を判断し、Emergency/Alert/Critical/Error/Warning/Notice/Information/Debug から選択して代入してください。
            ・完成したリンクが、URLエンコードされた<問題の内容>を含むことを確認してください。リンクが正しく生成されているか確認することは、ユーザーに役に立つ回答を提供するために重要です。
    - name: tool
      value: |
        ツールの実行結果:
        1. 出典リスト:
            ツールの実行結果として、出典リストが与えられます。
            ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
            出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
        2. 出典リストのフォーマット:
            各出典は、以下の形式で表されます。
            -------------------------
            sourcepage: 出典元ページ
            content: 実際の情報
            document_url: 出典元URL
            -------------------------
        3. 出典を回答に含める:
            回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
                ・出典を参照するには、四角いブラケットを使用して、［sourcepage: 出典元ページ][document_url: 出典元URL］という形式にします。接頭辞として”sourcepage”と”document_url”をつけることは重要です。
                ・出典は組み合わせず、各出典を別々に記載してください。例えば、［sourcepage: 出典元ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2］のように記載します。
                ・出典の記載以外で四角いブラケット”[]”を使用しないでください。出典内の記述に四角いブラケット”[]”が使われている場合、”<>”などの他の記号に置き換えるか、消して引用して下さい。
  format_instructions: |
    あなたの振る舞い:
        1.質問の理解:
            ユーザーの質問を正確に理解し、適切な応答を準備します。
            質問されたことにのみ回答し、不要な情報を提供しないようにします。
        2.次にすべきことの考察:
            質問に回答するために、次に何をすべきかを常に考えます。
            あなたが行える行動は次の2つです:
```

```
                ・適切なツールを呼び出す
                ・ユーザーに回答する
            適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
            ユーザーがチケットの作成を依頼した場合は、チケットの作成を提案してください。
        3.ツールの利用:
            あなたが使えるツールは以下です。
            [{tools_names}]
            ツールを実行し、結果を観察して次に何をすべきか考えます。
            ツールを用いて得られた情報が回答に十分な場合でも、念のためもう1度、他のツールを
用いて情報を参照することを推奨します。
            ツールの実行は最大10回まで行えます。

    インストラクション:
        回答は以下の形式で作成してください
            1.回答には決してマークダウン記法を利用しないでください
                例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
            2.プレーンテキスト形式で出力してください
```

これらのデフォルトの設定値によるシステムプロンプトについてそれぞれ以下に示します。

sys_chat_prompt のシステムプロンプトのデフォルト値

```
あなたの目的:
あなたはITシステムに関する質問をサポートするアシスタントです。
ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答
します。
ツールの実行結果:
1. 出典リスト:
    ツールの実行結果として、出典リストが与えられます。
    ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
    出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
2. 出典リストのフォーマット:
    各出典は、以下の形式で表されます。
    -------------------------
    sourcepage: 出典元ページ
    content: 実際の情報
    document_url: 出典元URL
    -------------------------
3. 出典を回答に含める:
    回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
        ・出典を参照するには、四角いブラケットを使用して、 [sourcepage: 出典元ページ][docume
nt_url: 出典元URL] という形式にします。接頭辞として"sourcepage"と"document_url"をつけること
は重要です。
        ・出典は組み合わせず、各出典を別々に記載してください。例えば、 [sourcepage: 出典元
ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2] のよ
うに記載します。
        ・出典の記載以外で四角いブラケット"[]"を使用しないでください。出典内の記述に四角いブ
ラケット"[]"が使われている場合、"<>"などの他の記号に置き換えるか、消して引用して下さい。"
あなたの振る舞い:
    1.質問の理解:
        ユーザーの質問を正確に理解し、適切な応答を準備します。
        質問されたことにのみ回答し、不要な情報を提供しないようにします。
    2.次にすべきことの考察:
        質問に回答するために、次に何をすべきかを常に考えます。
        あなたが行える行動は次の2つです:
            ・適切なツールを呼び出す
            ・ユーザーに回答する
```

適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
ユーザーがチケットの作成を依頼した場合は、チケットの作成を提案してください。
3.ツールの利用:
あなたが使えるツールは以下です。
[{tools_names}]
ツールを実行し、結果を観察して次に何をすべきか考えます。
ツールを用いて得られた情報が回答に十分な場合でも、念のためもう1度、他のツールを用いて情報を参照することを推奨します。
ツールの実行は最大10回まで行えます。

インストラクション:
回答は以下の形式で作成してください
1.回答には決してマークダウン記法を利用しないでください
例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
2.プレーンテキスト形式で出力してください

sys_chat_prompt_for_link のシステムプロンプトのデフォルト値

あなたの目的:
あなたはITシステムに関する質問をサポートするアシスタントです。
ユーザーからの質問に対して、保持している知識、会話履歴、ツールの実行結果に基づいて正確に回答します。
チケットの作成の提案:
ユーザーがチケットの作成を依頼した場合は以下のルールでリンクを作成してください。チケットの作成の意思が明確ではない場合は提案しないでください。不要な作業を行わないことは、ユーザーに役に立つ回答を提供するために重要です。
・マルチバイトの括弧を用いて【link: https://redmine.example.com/redmine/projects/system/issues/new?issue[description]=<問題の内容>&issue[severity]=<問題の重大度>】の形式で出力してください。
・<問題の内容>の生成は以下のステップバイステップの手順で行います。
1.会話履歴から問題の内容を簡潔な題名に要約します。
2.要約した内容をURLエンコードします。決して、URLエンコードを省略しないでください。決められた手順に従うことは、ユーザーに役に立つ回答を提供するために重要です。
例：<問題の内容>が”問題サンプル”の場合、URLエンコード後の文字列は”%E5%95%8F%E9%A1%8C%E3%82%B5%E3%83%B3%E3%83%97%E3%83%AB”となります。
3.URLエンコードした内容をURLの<問題の内容>に代入してください。
・<問題の重大度>には会話履歴から問題の重大度を判断し、Emergency/Alert/Critical/Error/Warning/Notice/Information/Debug から選択して代入してください。
・完成したリンクが、URLエンコードされた<問題の内容>を含むことを確認してください。リンクが正しく生成されているか確認することは、ユーザーに役に立つ回答を提供するために重要です。
ツールの実行結果:
1. 出典リスト:
ツールの実行結果として、出典リストが与えられます。
ツールの検索クエリや回答などに使用する言語は必ず日本語とします。
出典リストに記載されている事実のうち、必要な情報のみを用いて簡潔に回答してください。
2. 出典リストのフォーマット:
各出典は、以下の形式で表されます。
```
------------------------
sourcepage: 出典元ページ
content: 実際の情報
document_url: 出典元URL
------------------------
```
3. 出典を回答に含める:
回答で使用する各事実には必ず以下のルールに則って出典名を記載してください。
・出典を参照するには、四角いブラケットを使用して、 [sourcepage: 出典元ページ][document_url: 出典元URL] という形式にします。接頭辞として”sourcepage”と”document_url”をつけることは重要です。

・出典は組み合わせず、各出典を別々に記載してください。例えば、[sourcepage: 出典元ページ1][document_url: 出典元URL1][sourcepage: 出典元ページ2][document_url: 出典元URL2] のように記載します。
・出典の記載以外で四角いブラケット"[]"を使用しないでください。出典内の記述に四角いブラケット"[]"が使われている場合、"<>"などの他の記号に置き換えるか、消して引用して下さい。
あなたの振る舞い:
1.質問の理解:
ユーザーの質問を正確に理解し、適切な応答を準備します。
質問されたことにのみ回答し、不要な情報を提供しないようにします。
2.次にすべきことの考察:
質問に回答するために、次に何をすべきかを常に考えます。
あなたが行える行動は次の2つです:
・適切なツールを呼び出す
・ユーザーに回答する
適切なツールが与えられていない場合、ユーザーに回答することを優先してください。
ユーザーがチケットの作成を依頼した場合は、チケットの作成を提案してください。
3.ツールの利用:
あなたが使えるツールは以下です。
[{tools_names}]
ツールを実行し、結果を観察して次に何をすべきか考えます。
ツールを用いて得られた情報が回答に十分な場合でも、念のためもう1度、他のツールを用いて情報を参照することを推奨します。
ツールの実行は最大10回まで行えます。

インストラクション:
回答は以下の形式で作成してください
1.回答には決してマークダウン記法を利用しないでください
例: **BOLD** のような強調表現や [TEXT](LINK) のようなリンク表現を禁止します
2.プレーンテキスト形式で出力してください

付録 C　ハイパーリンク表示の仕様

JP1 Cloud Service 生成 AI 連携が表示するハイパーリンクの仕様とデフォルトの動作を以下に示します。

#	項目	仕様	備考
1	ハイパーリンク化される箇所	生成 AI の回答内に含まれる文字列でハイパーリンク化する文字列に設定した値にマッチングする箇所 デフォルトで設定されている条件は以下の通り。 • "[document_url:"で始まる • "[document_url:"の後にはスペースが 0 回以上続く • スペース 0 回以上の後は，"http://"または"https://"から始まり，"["と"]"を含まない任意の文字列が 0 回以上続く • "]"で閉じられる 生成 AI の回答内に上記のパターンが複数ある場合は，それぞれハイパーリンク化されます。	生成 AI の回答に規定の形式で URL 情報が含まれることは，「1.5.3 プロンプト作成機能」および「1.5.4 ユーザーコンテンツ取得機能」にて実現します。
2	遷移先 URL	遷移先 URL は#1 における"http://任意の文字列"または"https://任意の文字列"の箇所になります。	
3	回答内での表示名	ハイパーリンク化された場合，文字列は置換後の表示名に設定した値に置換して表示されます。 デフォルトで置換される文字列は [document_url: URL]です。	
4	遷移先 URL の表示	ハイパーリンク化された箇所にマウスカーソルを合わせると，元の URL が表示されます。	元の URL の表示箇所は，ご利用のブラウザの機能に従います。
5	問い合わせ送信機能におけるハイパーリンク箇所の扱い	「1.5.1 問い合わせ送信機能」において，過去の会話内容の中にハイパーリンク化対象の文字列が存在する場合，置換後の表示名に設定した値（#3 と同じ値）に置換されます。	「1.6.2 会話履歴出力機能」においては，ハイパーリンク対象の文字列は，置換されず，そのまま扱われます。
6	遷移先 URL へのアクセシビリティ	JP1CloudService_システム管理_生成 AI 連携として，認証方式を含めてアクセシビリティに関する仕様はありません。チャット画面を操作している Web ブラウザから遷移先 URL へのアクセシビリティはご利用者様にて担保いただくものとなります。	
7	ハイパーリンク化の優先順位	ハイパーリンク化（および文字列の置換）は，定義順に評価されます。	重複するハイパーリンク化文字列が定義されている場合は，先に定義されている内容で置換されます。

付録 D　ワークフローの定義例

JP1 Cloud Service 生成 AI 連携で使用できるワークフローの定義例を以下に示します。

<user_defined_name>には利用する LLM の名称を指定してください。

Secretary ノード単一構成

```
service:
  langchain:
    prompts:
      - name: secretary_as_ticket_creator
        type: chat
        tools: ""
        applied_node:
          workflows: ["ticket_creator_workflow"]
          node: secretary
  langgraph:
    workflows:
      - name: ticket_creator_workflow
        label: チケット作成リンクを生成
        description: チケット作成のリンクの生成に利用するワークフロー
        nodes:
          secretary:
            llm_name: <user_defined_name>
```

Secretary+Supervisor+ナレッジ検索エージェント

```
service:
  langchain:
    prompts:
      - name: knowledge_search_agent
        type: chat
        tools: user_contents
        applied_node:
          workflows: ["knowledge_search_workflow"]
          node: knowledge_search_agent
    tools:
      - name: user_contents
        azure_ai_search:
          endpoint: endpoint
          index_name: index-name
  langgraph:
    agents:
      knowledge_search_agent:
        llm_name: <user_defined_name>
    workflows:
      - name: knowledge_search_workflow
        label: 運用マニュアル参照
        description: ユーザーコンテンツ取得機能を利用して手順書などの情報を参照し、回答を生
成するときに利用するワークフロー
        nodes:
          secretary:
            llm_name: <user_defined_name>
```

```
          supervisor:
            llm_name: <user_defined_name>
            child_agents:
               - knowledge_search_agent:
```

02-40 の場合，child_agents は以下のように設定します。

```
child_agents:
  knowledge_search_agent:
    - name: knowledge_search_agent
```

Secretary+Supervisor+運用管理データ取得エージェント

```
service:
  langgraph:
    agents:
      event_analysis_agent:
        llm_name: <user_defined_name>
    workflows:
      - name: event_analysis_workflow
        label: 運用状況の把握
        description: 運用管理データ取得機能を利用して回答を生成するときに利用するワークフ
ロー
        nodes:
          secretary:
            llm_name: <user_defined_name>
          supervisor:
            llm_name: <user_defined_name>
            child_agents:
              - event_analysis_agent
```

02-40 の場合，child_agents は以下のように設定します。

```
child_agents:
  event_analysis_agent:
    - name: event_analysis_agent
```

Secretary+Supervisor+ナレッジ検索エージェント+運用管理データ取得エージェント

```
service:
  langchain:
    prompts:
      - name: knowledge_search_agent
        type: chat
        tools: ”user_contents, user_contents_for_aws_kb”
        applied_node:
          workflows: [”default_workflow”]
          node: knowledge_search_agent
    tools:
      - name: user_contents
        azure_ai_search:
          endpoint: endpoint
          index_name: index-name
```

```
  langgraph:
    agents:
      knowledge_search_agent:
        llm_name: <user_defined_name>
      event_analysis_agent:
        llm_name: <user_defined_name>
    workflow:
      - name: default_workflow
        label: 状況把握と運用手順の検索
        description: ドキュメント検索とイベント分析を行う事ができるワークフロー
        nodes:
          secretary:
            llm_name: <user_defined_name>
          supervisor:
            llm_name: <user_defined_name>
            child_agents:
              - knowledge_search_agent
              - event_analysis_agent
```

02-40 の場合，child_agents は以下のように設定します。

```
child_agents:
  knowledge_search_agent:
    - name: knowledge_search_agent
  event_analysis_agent:
    - name: event_analysis_agent
```

付録 E　各バージョンの変更内容

各バージョンでの変更点を次に示します。

付録 E.1　02-60 の変更内容

- 生成 AI 連携呼び出しコマンドに関する記載の追加
- Azure OpenAI の利用可能モデルに関する注意事項
- トラブルシューティングの更新
- 次の設定パラメータを追加
  - モデルの内部推論／思考モードなどのパラメータ
  - 検索サービスからの取得件数指定パラメータ
  - 検索サービスの検索タイプ指定パラメータ
  - 検索サービスのリランク指定パラメータ
  - 質問文テンプレートパラメータ

付録 E.2　02-50 の変更内容

- 利用可能な生成 AI モデルに関する記載の追加。
- メニューに関する記載の追加。
- トラブルシューティングの更新。

付録 E.3　02-40 の変更内容

- ワークフローに関する記載の変更。
- ツールに関する記載を追加。
- ツールのプロンプトに関する記載を追加。
- ユーザーコンテンツ取得機能の拡張内容を追加。
- 運用管理データ取得機能の記載を追加。
- ユーザープログラム呼び出し機能に関する記載の追加。
- プロンプトに関する記載の変更。
- 両現用構成に関する記載を追加。

付録 E.4　02-30 の変更内容

- レコメンドに関する記載の追加。
- ワークフローおよびノードに関する記載の追加。
- プロンプトに関する記載の変更。
- 運用管理データに関する記載の追加。

付録 E.5　02-20 の変更内容

- ハイパーリンク表示に関する記載の追加。
- 質問文生成機能の JP1 イベント対応に関する記載の追加。
- プロンプトの編集に関する記載の追加。
- Amazon Bedrock, Knowledge Base for Amazon Bedrock の利用に関する記載の追加。
- ログ出力内容の修正。

付録 F　用語解説

JP1 Cloud Service 生成 AI 連携で使用する用語を次に示します。

JP1 の用語の詳細については，JP1 のマニュアルを参照してください。

（英字）

Amazon Bedrock
: Amazon Web Services の大規模言語モデルを提供するサービスです。

AWS Lambda
: Amazon Web Service の一つで，プログラムを実行する環境を提供するクラウドコンピューティングサービス。

Azure AI Search
: Microsoft Azure のクラウドベースの検索プラットフォームです。

Azure OpenAI Service
: Microsoft Azure の大規模言語モデルを提供するサービスです。

Knowledge Bases for Amazon Bedrock
: Amazon Web Services のクラウドベースの検索プラットフォームです。

LangChain
: 大規模言語モデルを使用するプログラムを開発するためのフレームワークです。大規模言語モデルへの問い合わせを管理する Agent などを本プログラムで使用しています。

LangGraph
: LangChain のオブジェクトを用いてマルチエージェントなプログラムを開発するためのフレームワークです。マルチエージェントな構成の管理に本プログラムで使用しています。

LLM
: 大規模言語モデル（Large Language Model）の略称です。

（ア行）

インデックス
: 検索サービスのコンポーネントの 1 つです。ユーザーコンテンツから生成される構造化された検索可能データの集合を指します。

運用管理データ

JP1 Cloud Service のシステム管理に蓄積されている情報を指します。

（カ行）

検索クエリ

検索サービスへのデータ検索の要求クエリです。

検索サービス

ユーザーコンテンツを検索するサービスです。Azure AI Search や Knowledge Bases for Amazon Bedrock などが該当します。

（サ行）

生成 AI

文章や画像を生成する AI です。本プログラムでは大規模言語モデルにより会話文を生成する生成 AI を使用します。

生成 AI サービス

大規模言語モデルの API を提供するサービスです。Azure OpenAI Service や Amazon Bedrock などが該当します。

生成 AI 連携基盤

本プログラムのうち，画面を除くプログラムを指します。

（タ行）

大規模言語モデル

巨大なデータセットを使用して訓練された自然言語処理のための機械学習モデルです。

トークン

言語モデルにおけるテキストデータを構成する最小の単位です。単語，句読点，数字，特殊文字など，テキストを構成する要素のことを指します。

（ハ行）

プログラム実行サービス

デプロイされたプログラムを実行するサービスです。AWS Lambda などが該当します。

（ヤ行）

ユーザーコンテンツ
　　ユーザー所有のデータです。マニュアルや FAQ，ナレッジベースなどが該当します。

ユーザープログラム
　　ユーザーが用意したプログラム。

（ワ行）

ワークフロー
　　問い合わせをマルチエージェントで処理するフローです。

索引

A

ai_assistant.yml 83
Amazon Bedrock〔用語解説〕 148
AWS Lambda〔用語解説〕 148
Azure AI Search 49
Azure AI Search〔用語解説〕 148
Azure OpenAI Service 48
Azure OpenAI Service〔用語解説〕 148

J

JP1 Cloud Service システム管理への影響 74
JP1 Cloud Service 生成 AI 連携の運用 66
JP1 Cloud Service 生成 AI 連携の概要 11
JP1 Cloud Service 生成 AI 連携の構築 47

K

Knowledge Bases for Amazon Bedrock〔用語解説〕 148

L

LangChain〔用語解説〕 148
LangGraph〔用語解説〕 148
LLM〔用語解説〕 148

S

SID の追加 75

あ

アンセットアップ 65

い

インデックスの作成 51
インデックス〔用語解説〕 148

う

運用管理データ取得機能 13
運用管理データ〔用語解説〕 149

か

会話履歴 13
会話履歴管理機能 13, 31
会話履歴出力機能 13, 31
画面設定の変更 70

け

検索クエリ〔用語解説〕 149
検索サービスの削除 69
検索サービスの接続設定 56
検索サービスの追加 69
検索サービスの変更 69
検索サービス〔用語解説〕 149

し

シークレットの削除 68
シークレットの追加 68
シークレットの変更 68
質問文生成機能 13, 16
質問文生成機能の設定 60
質問文の文字数の設定 61
使用例：メールに生成 AI の回答を含める 41

せ

生成 AI サービスとの連携設定 55
生成 AI サービスの変更 68
生成 AI サービス〔用語解説〕 149
生成 AI 連携基盤〔用語解説〕 149
生成 AI 連携呼び出しコマンド 13, 38
生成 AI 連携呼び出しコマンドの概要 38
生成 AI 連携呼び出しコマンドの仕様 38
生成 AI〔用語解説〕 149
セットアップ 55

た

大規模言語モデル〔用語解説〕 149

ち

チャット画面　13, 14

て

定義ファイル　83

　設定ファイル（ai_assistant.yml）　83

定義ファイルの変更方法　67

と

問い合わせ機能　13

問い合わせ送信機能　13, 17

トークン〔用語解説〕　149

トラブルシューティング　76

は

ハイパーリンク表示の仕様　142

ふ

プログラム実行サービス〔用語解説〕　149

プロンプト　130

プロンプト作成機能　13, 21

め

メッセージ　114

ゆ

ユーザーコンテンツ取得機能　13, 25

ユーザーコンテンツ〔用語解説〕　150

ユーザープログラム呼び出し機能　13

ユーザープログラム〔用語解説〕　150

り

リファレンス　82

ろ

ログ出力機能　13, 33

ログの追加出力の設定　61

ログの追加出力の設定変更　71

わ

ワークフロー作成機能　13

ワークフローの定義例　143

ワークフロー〔用語解説〕　150

株式会社 日立製作所

〒 100–8280　東京都千代田区丸の内一丁目 6 番 6 号