

# セットアップガイド

## 1. はじめに

本書は、Generative AI FW のシステム構築者、運用管理者のための説明書です。

システム概要、およびシステム構築方法について解説しています。導入される前に、必ずお読みください。

本書の内容に関しては、将来予告なしに変更することがあります。

本書の内容の一部または全部を無断で複製・転載・改編することは禁止します。

### 1.1. 用語定義について

用語定義については「スタートアップマニュアル（概要編）」をご確認ください。

## 2. 前提条件

- 本書に記載の手順は全てサーバの管理者ユーザなどの**管理者権限を持つユーザで行う必要**があります。一般ユーザでしかログオンできない環境の場合は以下を実行し、管理者ユーザに昇格させてください。もしくはコマンド実行時に「sudo」を付けてください。

```
1 | sudo -i
```

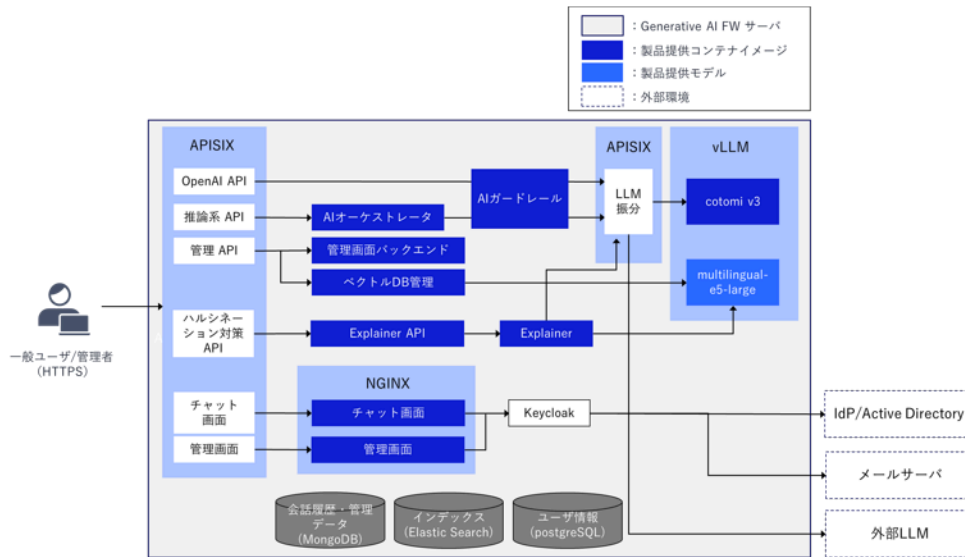
- 本書に記載の手順はサーバにログオンしている状態である必要があります。
- 本ガイドの手順は閉域環境（インターネットに接続できない環境）でのセットアップに対応しております。

## 3. 概要

### 3.1. Generative AI FW とは

Generative AI FW はLLMを使った推論・対話などの機能をUIとAPIで提供しています。機能概要の詳細については別途「スタートアップマニュアル（概要編）」をご確認ください。

### 3.2. システム構成



モジュール名	説明
AIオーケストレータ	LLMを用いて対話・推論を行うアプリケーション
管理画面バックエンド	管理ポータル画面の操作に対する処理を行うアプリケーション
ベクトルDB管理	インデックスに関する処理を行うアプリケーション
チャット画面	LLMを用いて対話を行うためのWebアプリケーション
管理ポータル画面	本製品を利用するために必要な管理操作を行うWebアプリケーション
cotomi v3	本製品で利用可能なLLM。cotomi v3が利用必要。
multilingual-e5-large	インデックスを利用するためのエンベディングモデル
Explainer API	ハルシネーション対策APIを提供するアプリケーション
Explainer	ハルシネーション対策機能を実現するアプリケーション
AIガードレール	LLMを使った推論・対話機能を安全に使用するための機能を提供するアプリケーション
	<div style="background-color: #fff9c4; padding: 5px; border: 1px solid #ccc;"> <p>⚠️ AIガードレール機能を使うにはAIガードレールオプションを別途購入する必要があります</p> </div>

### 3.3. 動作環境

OS	Red Hat Enterprise Linux 9.2 Red Hat Enterprise Linux 9.6
CPU	1.8GHz(8コア) 以上
GPU	使用するcotomiモデルによって異なります。 <ul style="list-style-type: none"> <li>• cotomi v3 (以下のいずれかの条件を満たす必要があります) <ul style="list-style-type: none"> <li>◦ NVIDIA H100 NVL (GPU メモリ-94GB) 以上</li> <li>◦ NVIDIA L40S(GPU メモリ-48GB) 以上</li> <li>◦ NVIDIA RTX PRO 6000 SE(GPU メモリ-96GB) 以上</li> </ul> </li> </ul>
メモリ	64GB以上
ディスク空き容量	使用するcotomiモデルによって異なります。 <ul style="list-style-type: none"> <li>• cotomi v3</li> </ul> var配下 (Podman、コンテナイメージ、ログなど) : 390GB 以上 opt配下 (以下合計) : 254.9GB以上  インストールフォルダ : 100GB以上  DB関連データ(※) : 154.9GB以上  ※DB関連データは稼働に必要な最低限の容量だけを記載しています。後述の「DB関連データのディスク容量について」を一読し、要件に合わせて必要な容量を算出してください。
GPUドライバ	NVIDIA driver(バージョン : 565.57.01以上推奨)
GPU(その他)	NVIDIA Container Toolkit 1.18
コンテナ管理	Red Hat Enterprise Linux 9.2 <ul style="list-style-type: none"> <li>• Podman 4.9以上</li> <li>• podman-compose 1.5.0</li> </ul> Red Hat Enterprise Linux 9.6 <ul style="list-style-type: none"> <li>• Podman 5.4以上</li> <li>• podman-compose 1.5.0</li> </ul>
データベース	MongoDB 7.0

	PostgreSQL 16
検索エンジン	Elasticsearch 8.19
API ゲートウェイ	APISIX 3.14
アプリケーションサーバ	NGINX 1.29
認証	Keycloak 26
LLM	cotomi v3
外部LLM	<ul style="list-style-type: none"> <li>• Azure OpenAI Service (APIバージョン : 2023-05-15)</li> <li>• OpenAIに準拠したLLM</li> </ul>
エンベディングモデル	multilingual-e5-large
その他	<ul style="list-style-type: none"> <li>• etcd 3.6</li> <li>• squid5.5 (proxy環境で使用する場合のみ)</li> <li>• fluentd 6</li> </ul>
連携可能なIdP	<ul style="list-style-type: none"> <li>• Microsoft Entra ID</li> <li>• Active Directory (Windows Server 2022/2025) <ul style="list-style-type: none"> <li>◦ シングルフォレスト・シングルドメイン構成のみサポート</li> <li>◦ 1つのコンテナ、または、1つのOU(組織単位)配下のユーザを連携対象とできます。複数のコンテナや複数のOUと連携する事は出来ません</li> </ul> </li> </ul>
対応可能なWeb検索エンジンサービス	<ul style="list-style-type: none"> <li>• Brave Search</li> <li>• Tavily</li> </ul>
Webブラウザ	<p>Microsoft Edge (バージョン : 137.0.3296.68 以上)</p> <p>Google Chrome (バージョン : 137.0.7151.103 以上)</p>

### 3.3.1. DB関連データのディスク容量について

DB関連データは初期容量(1GB)に以下の基準値ごとの必要ディスク容量を足すことで算出しています。以下の項目について、基準値以上での動作が必要な場合は、適宜必要なディスク容量

を加算して必要なディスク容量を算出してください。なお、以下の項目には諸元(上限値)があります。詳細は「運用ガイド」の「諸元」をご確認ください。

- ① 会話履歴については以下の必要なディスク容量を超えると古い履歴から自動で削除されます。2000文字以上で頻繁に会話する、拡張対話でファイルやURL、Web検索を使う対話を多く行う場合は容量の消費が速くなるため、必要に応じて会話履歴の容量上限を拡張してください。

拡張の方法は運用ガイドの「会話履歴の最大保存期間を変更したい」を参照してください。

項目	基準値	必要ディスク容量
ユーザ数	1000人	0.4GB
会話履歴の最大保存日数	30日(2000文字程度の会話を1日12時間稼働させた場合)	66GB
テンプレート登録数	1000	1.5GB
インデックス(登録文書の総数)	100ファイル	25GB
監査ログ	30日(対話履歴のログは2000文字程度の会話を1日12時間稼働させた場合)	45GB
ログ	7日(対話履歴のログは2000文字程度の会話を1日12時間稼働させた場合)	16GB

## 4. セットアップ

本節では本製品のセットアップについて説明します。

まずGenerative AI FWのサーバにログオンします。

- ① OS管理者のアカウント・パスワードは別途「初期アカウント一覧」を確認してください。

## 4.1. 同梱物について

本製品のインストールフォルダ(/opt/nec/genai)の配下は以下の通りです。

以下は動作上の主要なフォルダ・ファイルです。データベースなどで上記以外のフォルダ・ファイルがある可能性があります、問題ありません。

```
1 /opt/nec/genai
2 |— modules : 製品モジュール関連フォルダ
3 |   |— admin-backend
4 |     |   └─ genai-admin-backend.tar.gz
5 |   |— admin-frontend
6 |     |   └─ create_config.sh
7 |     |   └─ genai-admin-frontend.tar.gz
8 |   |— ai-orchestrator
9 |     |   └─ genai-ai-orchestrator.tar.gz
10 |  |— chat-ui
11 |    |   └─ create_config.sh
12 |    |   └─ genai-chat-ui.tar.gz
13 |  |— explainer-api
14 |    |   └─ genai-explainer-api.tar.gz
15 |  |— guardrail-server
16 |    |   └─ genai-guardrail-server.tar.gz
17 |  |— llm-explainer
18 |    |   └─ genai-llm-explainer.tar.gz
19 |  └─ vector-db
20 |    └─ genai-vector-db.tar.gz
21 |
22 |— models : LLMモデル関連フォルダ
23 |   |— cotomi : LLMのコンテナイメージ格納場所
24 |     |   └─ e5 : multilingual-e5-largeのモデル
25 |     |     └─ . . .
26 |
27 |— setup : 工場構築関連スクリプト関連フォルダ
28 |   |— image_lists : 各バージョンで使用しているOSSコンテナイメージ名
29 |     └─ . . .
30 |   |— mongo : mongoDB初期設定関連
31 |     └─ . . .
32 |   |— v2_scripts : v2からのアップデート時に差し替えるスクリプト
33 |     └─ . . .
34 |   |— admin_api_setup.sh : 管理API公開用スクリプト
35 |   |— download_oss.sh : OSSダウンロードファイル
36 |   |— environment_setup.sh : 環境構築構築スクリプト
37 |   |— fluentd_functions.sh : Fluentd管理用共通関数ライブラリ
38 |   |— genai : サービス環境変数ファイル
39 |   |— genai_no_llm.service : ユニットファイル
40 |   |— genai.service : ユニットファイル
41 |   |— get_setting_file.sh_base : 設定情報取得スクリプト(ベース)
42 |   |— idp_setup.sh : カスタム認証設定スクリプト
43 |   |— init_setup.sh : 初期セットアップスクリプト
44 |   └─ restart_setup.sh_base : 再起動セットアップスクリプト(ベース)
45 |
46 |— config : 設定ファイル
47 |   |— keycloak : Keycloakテーマファイル関連
48 |     └─ . . .
49 |   |— apisix-internal.yml : APISIX設定ファイル(内部APISIX)
50 |   |— apisix-public.yml : APISIX設定ファイル(外部APISIX)
51 |   |— docker-compose-cotomi-v3.yml : コンテナ構成情報ファイル(cotomi v3)
52 |   |— docker-compose.yml : コンテナ構成情報ファイル
53 |   |— fluentd.conf_base : fluentd設定ファイル
54 |   |— genai_fluentd.conf : fluentd設定ファイル(モジュール単位)
55 |   |— genai_info-cotomi-v3.json : 画面情報定義ファイル(cotomi v3)
56 |   |— genai.env : 全モジュール共通envファイル
57 |   |— internal-route.yml : 内部ルーティング定義ファイル
58 |   |— nginx.conf : NGINX設定ファイル
59 |   |— quarkus.properties : 監査ログ(アクセスログ)設定ファイル
60 |   └─ squid.conf_base : proxy設定スクリプト用のsquid設定ファイル
61 |
62 |— operation : 作業用スクリプト関連フォルダ
```

```

63 | |   └─ access_update.sh   : アクセス先変更スクリプト
64 | |   └─ genai_backup.sh   : バックアップスクリプト
65 | |   └─ genai_info_update.sh : 画面情報変更スクリプト
66 | |   └─ genai_proxy.sh   : proxy設定スクリプト
67 | |   └─ genai_restore.sh  : リストアスクリプト
68 | |   └─ genai_verup.sh   : バージョンアップスクリプト
69 | |   └─ port_update.sh   : ポート変更スクリプト
70 | |   └─ set_apikey.sh    : APIキー設定スクリプト
71 | |
72 | └─ uninstall : 作業用スクリプト関連フォルダ
73 |     └─ uninstall.sh : アンインストールスクリプト
74 |
75 | └─ documents : マニュアル関連フォルダ
76 |     └─ index.html  : マニュアル概要ファイル
77 |         └─ . . .

```

## 4.2. OSのネットワーク関連セットアップ

出荷時には関連サービスの自動起動が無効になっています。以下の順でサービスを有効化に変更してください。

1. genaiサービスの自動起動を有効にします。

```
1 | systemctl enable genai
```

2. genaiサービスを開始します。

```
1 | systemctl start genai
```

## 4.3. 閉域化対応

本手順はGenerative AI FWを閉域環境で使用する場合の手順を記載します。閉域化対応が不要な場合は本手順の実施は不要です。なお、閉域化を有効にした場合、以下の機能が利用できません。

- 初期認証（メールサーバ連携有）のセットアップ（社内ネットワークのメールサーバの場合は引き続き利用可能）
- 外部LLMを使った機能
- 拡張対話（Webコンテンツ）
- 拡張対話（Web検索）

閉域化を行う場合、以下の手順を実施してください。

1. Generative AI FWのサーバにログインします。
2. /opt/nec/genai/config/genai.envを開きます。

```
1 | vi /opt/nec/genai/config/genai.env
```

3. GENAI\_CLOSED\_NETWORKの値を1に変更します。

```
1 | GENAI_CLOSED_NETWORK=1
```

4. genaiサービスを再起動します。

```
1 | systemctl restart genai
```

#### 4.4. アクセス先の変更

サーバのアクセス先(ドメイン名)を適切に変更します。手順は「運用ガイド」の「アクセス先・HTTPS証明書の更新」を参照してください。

**i** 「アクセス先・HTTPS証明書の更新」にはKeycloakの管理者ユーザのパスワードが必要です。Keycloakの管理者ユーザのパスワードは別途「初期アカウント一覧」を確認してください。

#### 4.5. 出荷環境での動作確認

構築が正常にできているかどうかを管理ポータル画面、チャット画面、APIの観点で確認していきます。

正常性確認手順は「正常性確認ガイド」をご確認ください。

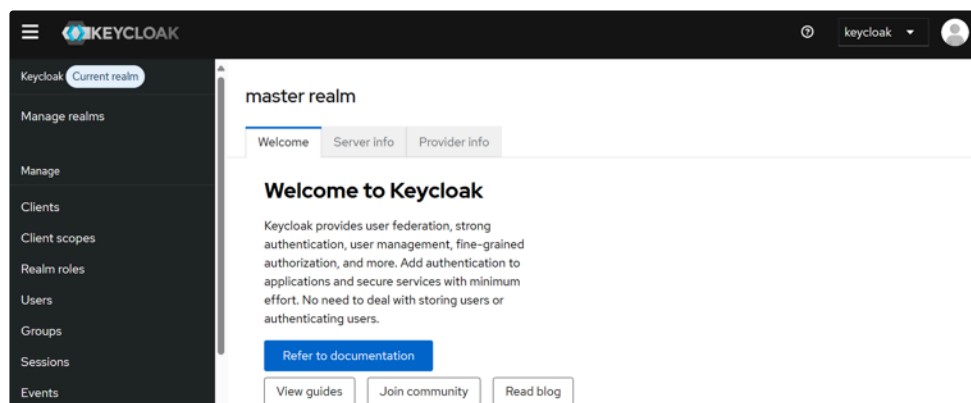
構築時にLLMをセットアップしている場合は対話機能も利用可能です。

#### 4.6. Keycloakの初期パスワードの変更

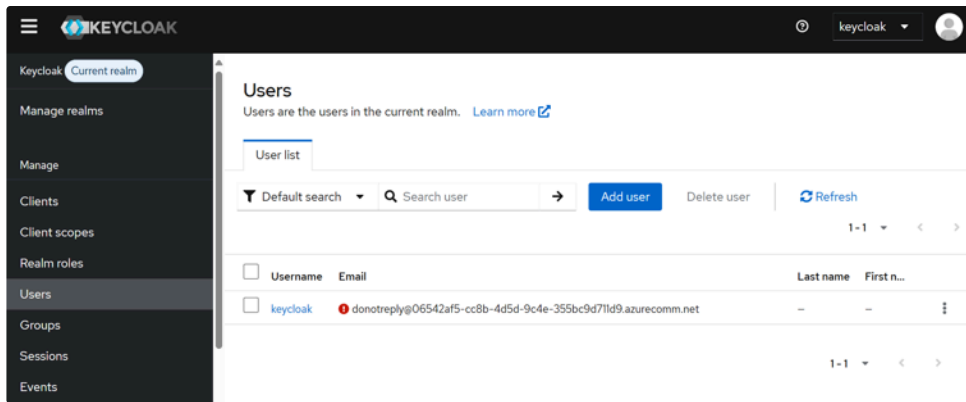
Keycloak管理者ユーザの初期パスワードを変更します。Keycloakの管理画面にログインする方法は「正常性確認ガイド」の「Keycloakの確認」をご確認ください。

**i** Keycloak管理者ユーザのユーザ名、パスワードは忘れないように保管することを推奨します。

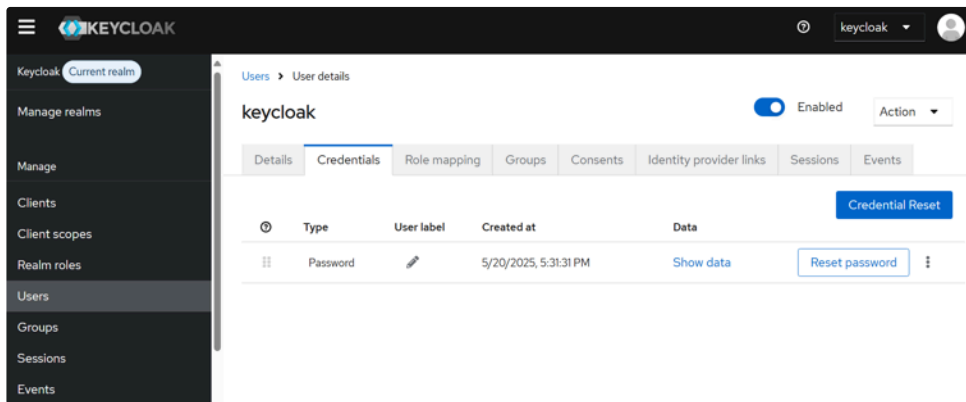
1. Keycloakの管理画面にログインします。master realmが表示されることを確認します。



2. 左メニューから「Users」を選択し、keycloakユーザのリンクをクリックします。



3. Credentialsタブをクリックし、「Reset password」を押します。



4. 新しいパスワードを設定、TemporaryをOffの状態にして、Save を押します。

### Reset password for keycloak

Password \*

New password confirmation \*

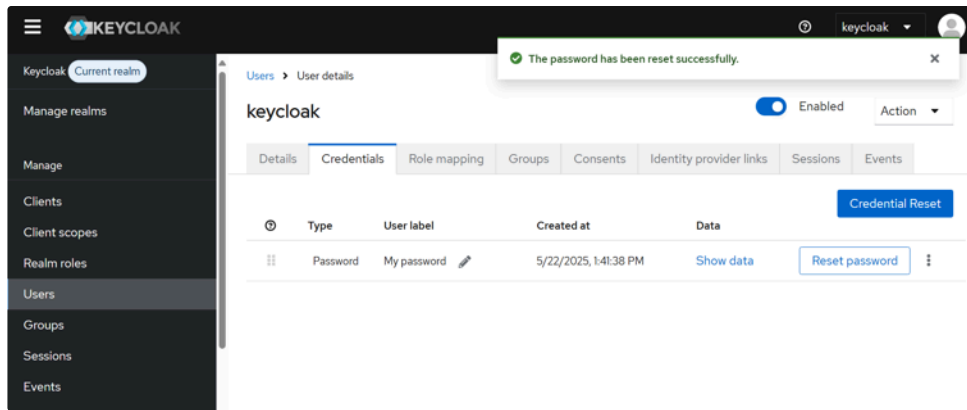
Temporary  Off

5. 確認ダイアログが表示されるので、Reset passwordを押します。

### Reset password?

Are you sure you want to reset the password for the user keycloak?

6. 成功すると、パスワード変更が完了したメッセージが表示されます。



## 4.7. LLMのセットアップ

構築時の環境ではcotomi v3のセットアップが完了しています。追加で外部LLMのセットアップが必要な場合は「運用ガイド」の「LLMの追加・削除」をご確認ください。

### 4.7.1. GPUの設定

GPUの枚数に応じて設定を変更する必要があります。初期値ではGPUの枚数は1に設定されています。LLMに利用するGPUの枚数が1より多い場合は設定を変更します。

1. /opt/nec/genai/config/genai.envを開きます。

```
1 | vi /opt/nec/genai/config/genai.env
```

2. 以下の設定値を変更します。値はGPUの枚数です。2枚の場合は2を指定してください。

```
1 | TENSOR_PARALLEL_SIZE=<GPUの枚数>
```

3. genaiサービスを再起動します。

```
1 | systemctl restart genai
```

## 4.8. proxy対応

本手順は外部と通信する際にproxyが必要な環境で使用する場合の手順を記載します。Azure OpenAI Serviceなどの外部LLMとの対話、URLを用いた拡張対話機能を使用する場合は本設定が必要になります。本対応が必要でない場合は手順の実施は必要ありません。

- 閉域環境の場合、本手順は実施不要です。
- proxyを利用しない設定に戻すことはできません。
- 本設定には「firewalld」サービスが必須です（OSの初期値では利用可能）。  
「firewalld」サービスが動作していない場合、OpenAI Serviceなどの外部LLMとの対話できません。
- firewalld に不要なゾーンや direct ルールが存在している場合、製品が設定するプロキシ向けリダイレクトルールが動作しない可能性があります。本節の手順を行って

も外部通信ができない場合はお問い合わせください。

- ・プロキシサーバ機能を保持するオープンソースソフトウェア「squid」を利用されている場合、設定ファイルが上書きされます。「squid」を利用されている場合、以降の手順で squidの設定更新を行ってください。

#### 4.8.1. RHELリポジトリの有効化

- ・ squidをインストールするためにRHELリポジトリが有効化されている必要があります。
- ・ 以下の処理はRHELのサブスクリプションが登録されていることが前提になります。
- ・ 「4. 証明書の追加」については企業内プロキシや社内リポジトリが独自の認証局（社内CA・自己署名）を使用している場合に実施してください。公的CAで検証可能な環境では証明書の追加は不要です。

##### 1. プロキシ環境変数の設定

- a. 実行環境にプロキシ環境変数が設定されていることを確認します。

```
1 | env | grep HTTP
```

- b. ログインシェルなどで設定済の場合は以下のような設定内容が返却されます。

```
1 | HTTPS_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
2 | HTTP_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
```

- c. 返却されない、または設定内容が有効でない場合は以下のコマンドで環境変数を設定します。

```
1 | export HTTP_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
2 | export HTTPS_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
```

##### 2. リポジトリ確認

- a. リポジトリへ接続してメタデータを取得します。

```
1 | dnf makecache
```

- b. エラーが発生せず、以下のリポジトリ名が表示されれば正常です。表示されない場合は、以降の手順を実施してください。

```
1 | rhel-9-for-x86_64-appstream-rpms   Red Hat Enterprise Linux 9 for x86_64 - AppStream (RPMs)
2 | rhel-9-for-x86_64-baseos-rpms     Red Hat Enterprise Linux 9 for x86_64 - BaseOS (RPMs)
```

##### 3. dnf のプロキシ設定 (/etc/dnf/dnf.conf)

- a. プロキシサーバ情報を設定します。

```
1 | echo "proxy=<proxyサーバ情報>" >> /etc/dnf/dnf.conf
```

- b. 設定後の/etc/dnf/dnf.confファイルは以下のようになります。

```
1 | [main]
2 | gpgcheck=1
3 | installonly_limit=3
4 | clean_requirements_on_remove=True
```

```
5 best=True
6 skip_if_unavailable=False
7 proxy=http://proxy.example.com:8080
```

#### 4. 証明書の追加

- a. 証明書ファイルを配置し信頼ストアを更新します。

```
1 cp <CA証明書ファイル> /etc/pki/ca-trust/source/anchors/
2 update-ca-trust
```

#### 5. RHEL 9 の標準リポジトリ (BaseOS と AppStream) を有効化

- a. リポジトリを有効化します。

```
1 subscription-manager repos --enable rhel-9-for-x86_64-baseos-rpms --enable rhel-9-for-x86_64-appstream-
```

6. 「2. リポジトリ確認」の手順に従い、リポジトリの状態を再確認してください。エラーが発生した場合は、表示されるメッセージの内容を確認し対処してください。

### 4.8.2. squidの利用状況確認

1. squidの利用状況を確認するため以下のコマンドを実行します。

```
1 systemctl is-enabled squid
```

2. 以下のように「enabled」が返る場合、squidがインストール済で利用されています。

```
1 enabled
```

### 4.8.3. 設定変更

1. /etc/sysconfig/genaiを開きます。

```
1 vi /etc/sysconfig/genai
```

2. 以下の設定のコメントアウト(#)を削除し設定を記載してください。不要な設定はコメントアウトのままにしてください。HTTP\_PROXY、HTTPS\_PROXY、NO\_PROXYの設定は必須です。NO\_PROXYの記載済みの値は変更しないでください。

**i** 証明書ファイルは必ず「/opt/nec/genai/certs」に配置してください。

```
1 # proxy settings
2 HTTP_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
3 HTTPS_PROXY=<proxyサーバ情報(例:http://proxy.example.com:8080)>
4 SSL_CERT_FILE=<証明書ファイルのパス>
5 REQUESTS_CA_BUNDLE=<証明書ファイルのパス>
6 CURL_CA_BUNDLE=<証明書ファイルのパス>
7 NO_PROXY=127.0.0.1,localhost,apisix,internal_apisix,etcd, . . . ,genai-guardrail-server
```

3. genai\_proxy.shを実行します。実行は最初の一回だけで構いません。

```
1 cd /opt/nec/genai/operation
2 bash genai_proxy.sh
```

4. 設定が完了すると以下が表示されます。

```
1 proxy configuration succeeded.
```

#### 4.8.4. squidの設定更新

**i** 本手順は「squidの利用状況確認」でsquidが既にインストールされていた場合のみ実施してください。それ以外の場合は実施不要です。

1. squidの設定ファイル (/etc/squid/squid.conf) とバックアップファイル (/opt/nec/genai/config/backup/proxy\_backup\_<スクリプト実行日時>/squid.conf) を比較し、必要な設定をsquidの設定ファイルに反映します。
2. 以下のコマンドを実行しsquidを再起動します。

```
1 | systemctl restart squid
```

3. 以下のコマンドを実行しsquidの状態を確認します。

```
1 | systemctl status squid
```

4. Activeに「active (running)」と表示されていれば、正常に起動できていますので作業完了です。それ以外の場合は、表示内容やsquidのログ (/var/log/squid/) をもとに問題を対策し、1.から再度実行してください。

```
1 | Active: active (running)
```

#### 4.9. 初期ユーザで管理画面にログイン

「正常性確認ガイド」の「管理ポータル画面の確認」の手順に従い、初期ユーザで管理画面にログインしてください。

#### 4.10. 管理ポータル画面から新規ユーザ登録作業

「管理ポータル操作ガイド (ユーザ登録編)」の「ユーザ追加・編集」を参照して**新規の組織管理者ユーザ**を追加してください。

**i** 追加した組織管理者ユーザのユーザ名、パスワードは忘れないように保管することを推奨します。

#### 4.11. ハルシネーション対策機能のセットアップ

本製品では、LLMのハルシネーション対策機能を回答根拠の確認機能として提供しています。回答根拠の確認機能はチャット画面とAPIで利用できます。詳細は「チャット画面利用ガイド」「回答根拠確認チュートリアル」をご確認ください。

回答根拠の確認機能には「推論を用いない回答根拠の確認」と「推論を用いた回答根拠の確認」があり、それぞれ以下のような機能の違いがあります。

- 推論を用いない回答根拠の確認

回答根拠になった文書の内容と関連度のみを確認できる

- 推論を用いた回答根拠の確認

回答根拠になった文書の内容に加え、根拠となった判断理由を提供します。また判断根拠がないなど矛盾がある箇所も指摘します。

上記のうち「推論を用いた回答根拠の確認」は、利用するには以下の条件を満たす必要があります。条件を満たせない場合は利用できません。「推論を用いない回答根拠の確認」は以下の条件に関係なく利用できます。

- Azure OpenAI Service(AOAI)またはOpenAIの環境を持っており、前述のLLMのセットアップ手順にて外部LLMのセットアップが完了していること（閉域環境では使用できません）
- 外部LLM環境にGPT-4oをデプロイしており、LLMのセットアップ時に利用可能なモデルに追加されていること

#### 4.11.1. 推論を用いた回答根拠の確認のセットアップ

「推論を用いた回答根拠の確認」機能は利用するには追加のセットアップが必要です。以下の手順に従ってセットアップを完了させてください。

1. 外部LLMにデプロイしたGPT-4oモデルの名前を確認します。
2. /opt/nec/genai/config/genai.envを開きます。

```
1 | vi /opt/nec/genai/config/genai.env
```

3. ENTAILMENT\_HALLUCINATION\_MODEL\_NAMEの値に手順1で確認した名前を記載します。

```
1 | ENTAILMENT_HALLUCINATION_MODEL_NAME=<GPT-4oのモデルの名前>
```

4. genaiサービスを再起動します。

```
1 | systemctl restart genai
```

5. 管理ポータル画面にログインします。
6. 設定画面を表示し、推論を用いた回答根拠の確認を有効に変更します。



7. 管理ポータル画面をログアウトします。

## 4.12. 拡張対話（Web検索）機能のセットアップ

本製品では、外部の検索エンジンサービスの検索結果を用いてLLMと対話するWeb検索機能を提供しています。Web検索機能はチャット画面とAPIで利用できます。詳細は「チャット画面利用ガイド」「拡張対話チュートリアル」をご確認ください。

- Web検索機能を使うには別途検索エンジンサービスとの契約が必要です。契約はお客様ご自身で行う必要があります。
- 利用可能な外部の検索エンジンサービスは、「Brave Search」「Tavily」です。
- 本機能はインターネット通信ができない閉域環境では使用できません。

Web検索機能をご利用の際は、別途それぞれの検索エンジンサービスの規約に従い、APIキーの払い出しを実施してください。

APIキー払い出し後、以下の手順を踏むことでWeb検索機能が利用可能になります。

1. /opt/nec/genai/config/genai.envを開きます。

```
1 | vi /opt/nec/genai/config/genai.env
```

2. WEBSEARCH\_ENGINE、WEBSEARCH\_API\_KEYを記載します。

「Brave Search」の場合は”0”、「Tavily」の場合は”1”を記載してください。

```
1 | WEBSEARCH_ENGINE="0"  
2 | WEBSEARCH_API_KEY=<払い出した検索エンジンサービスのAPIキー>
```

3. genaiサービスを再起動します。

```
1 | systemctl restart genai
```

## 4.13. 拡張対話（ファイル添付）機能のセットアップ

本製品では、ファイルを添付しファイルの内容についてLLMと対話する機能を提供しています。ただし、ファイルのうち画像ファイルを添付して会話する場合は別途セットアップが必要になります。

画像ファイルでの対話機能が不要な場合は本手順は不要です。

本機能の詳細は「チャット画面利用ガイド」「拡張対話チュートリアル」をご確認ください。画像ファイルで対話するには以下の条件を満たす必要があります。

- Azure OpenAI Service(AOAI)またはOpenAIの環境を持っており、前述のLLMのセットアップ手順にて外部LLMのセットアップが完了していること
- 以下の外部サイトにあるLLMのうち、画像ファイルの入力に対応できるLLMモデルがデプロイしており、LLMのセットアップ時に外部LLMとして利用可能なモデルに追加されていること

- ・本機能はインターネット通信ができない閉域環境では使用できません。

セットアップ手順は以下の通りです。

1. /opt/nec/genai/config/genai.envを開きます。

```
1 | vi /opt/nec/genai/config/genai.env
```

2. GENAI\_ENABLE\_IMAGE\_TYPES を1に変更します。

```
1 | GENAI_ENABLE_IMAGE_TYPES=1
```

3. genaiサービスを再起動します。

```
1 | systemctl restart genai
```

## 4.14. APIキーの更新

Generative AI FW では、Keycloakによるユーザ認証だけでなく、API利用者に向けたAPIキーでの認証をサポートしています。初期のAPIキーから、必ずAPIキーを変更してください。


APIキーの更新は管理ポータル画面から行います。詳細は「管理ポータル操作ガイド（設定画面編）」の「APIセクションの説明」をご確認ください。

## 4.15. 新ユーザ、新規APIキーでの正常性確認

新規ユーザ、新規APIキーの情報が正常に動作するか管理ポータル画面、チャット画面、APIの観点で確認していきます。

正常性確認手順は「正常性確認ガイド」をご確認ください。

なお、チャット画面ではLLMをセットアップ完了しているため、対話ができることも合わせて確認してください。

-  拡張対話（Web検索）機能は正常性確認手順には含まれておりません。必要があればチャット画面利用ガイドを参照し問題なく動作するか確認してください。

## 4.16. 認証方式の検討

Generative AI FWでは複数の認証方式に対応しています。対応している認証方式と機能差異は以下の通りです。カスタム認証の詳細については「カスタム認証セットアップガイド」をご確認ください。

認証方式	ログインID	ユーザ作成時のパスワード	登録時のユーザへ	ログインパスワード	ユーザ・グループの同期機能

		スワード 設定	のメール 連絡	ドの初期 化方法	
初期認証 (メールサ ーバ連携 無)	メールア ドレス	必須	無	管理者によ るログイン パスワード の強制初期 化	無
初期認証 (メールサ ーバ連携 有)	メールア ドレス	必須	有	利用者自身 による初期 化	無
カスタム認 証 (ユーザ ID認証)	ユーザー 名	必須	無	管理者によ るログイン パスワード の強制初期 化	無
カスタム認 証 (Entra ID)	EntraID の設定に 依存	不要	無	EntraID側 で実施	無
カスタム認 証 (Active Directory )	ユーザー 名	不要	無	Active Directory 側で実施	有

「登録時のユーザへのメール連絡」とはユーザを新規に登録した際に、登録ユーザに初回のログイン方法などを記載したメールを自動で通知する機能です。デフォルトでは、初期認証（メールサーバ連携無）に設定されています。「ユーザ・グループの同期機能」とはカスタム認証（Active Directory）のみ利用可能な機能でActive Directoryで作成したユーザ・グループを手動または自動によってGenerative AI FWのユーザ・グループ情報と同期させることができる機能です。初期認証（メールサーバ連携有）に変更する場合は次節を確認してください。カスタム認証への変更については「カスタム認証セットアップガイド」をご確認ください（ユーザ・グループの同期機能についても記載しています）。なお、ログインパスワード忘れた場合の対応方法を「管理者によるログインパスワードの強制初期化」で行う場合の操作方法は別紙の「運用ガイド」をご確認ください。

#### 4.16.1. 初期認証（メールサーバ連携有）のセットアップ

Generative AI FWではユーザがパスワードを忘れた場合に、KeycloakのForgot Password機能を使ってユーザ自身がパスワードを変更することができます。本機能を有効にするためにはKeycloakとメールサーバを連携させる必要があります。

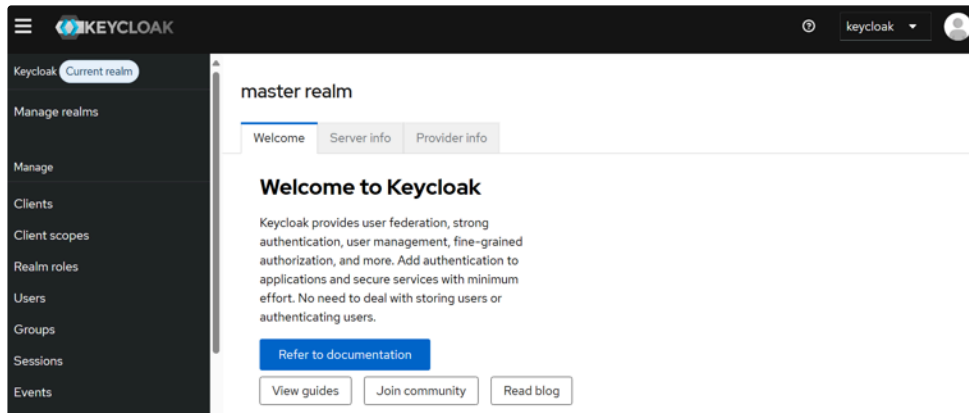
本機能を有効化した場合、「登録時のユーザへのメール連絡」が有効となります。無効化する場合は後述する「登録時のユーザへのメール連絡の無効化」をご確認ください。

なお、本機能を有効にしない場合、初回ログイン以外でユーザのパスワードを変更するためには、keycloakの管理者ユーザでKeycloakの管理画面ログインし、管理者がパスワードのリセットを行う必要があります。詳細は「運用ガイド」の「管理者によるログインパスワードの強制初期化」をご確認ください。

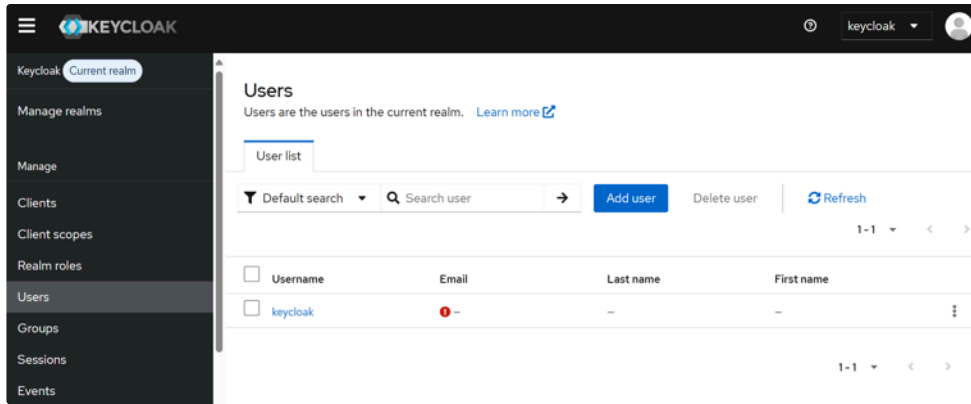
**⚠ カスタム認証利用時は本機能は利用できません。カスタム認証時のパスワード初期化の方法については「カスタム認証利用ガイド」を確認してください。**

Forgot Password機能の設定方法は以下の通りです。Keycloakの管理画面にログインする方法は「正常性確認ガイド」の「Keycloakの確認」をご確認ください。

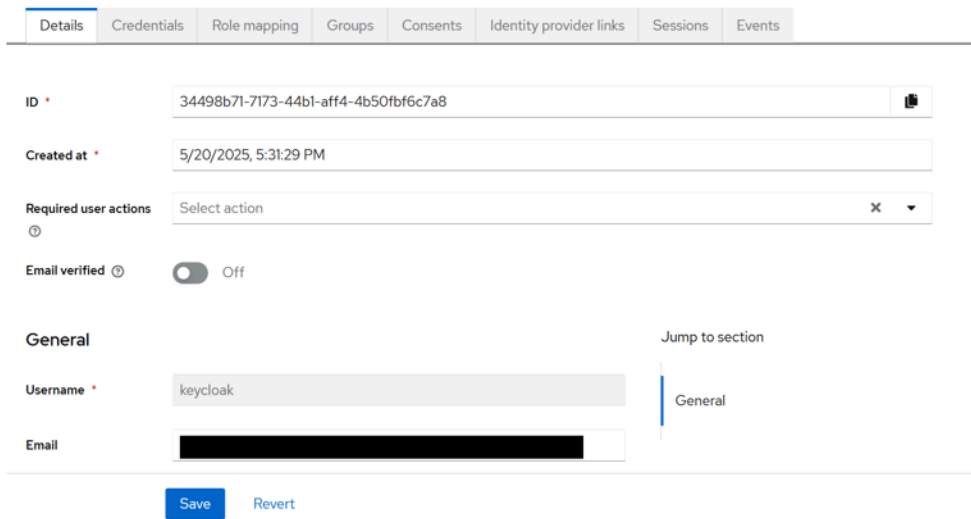
1. Keycloakの管理画面にログインします。



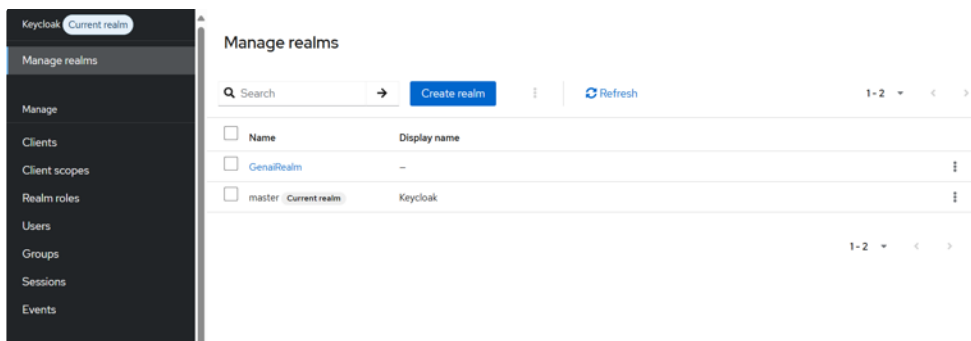
2. 左メニューから「Users」を選択し、keycloakユーザのリンクをクリックします。

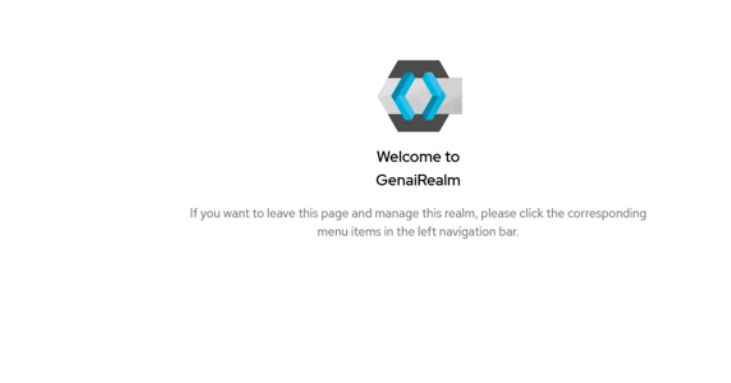
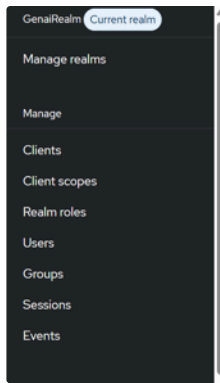


3. Detailsタブを選択、keycloakユーザのEmailアドレスにパスワード初期化で使用する送信元アドレスを設定します。

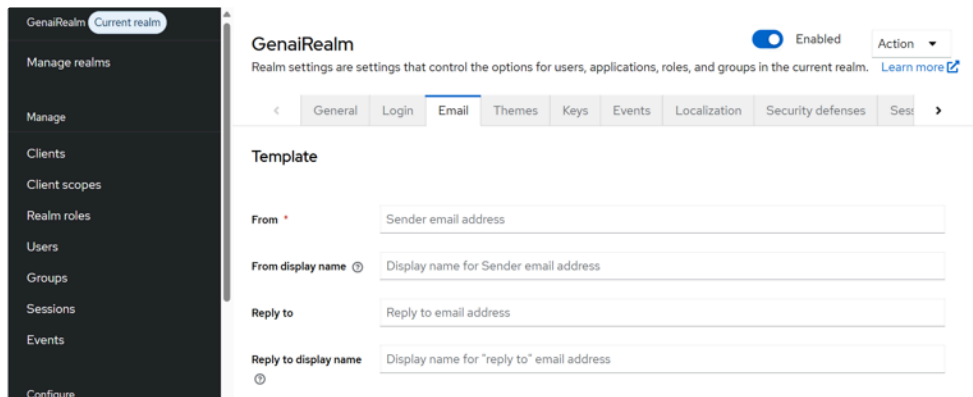


4. 左メニューから「Manage realms」を選択し、GenaiRealmのリンクをクリックし、GenaiRealmに切り替えます。

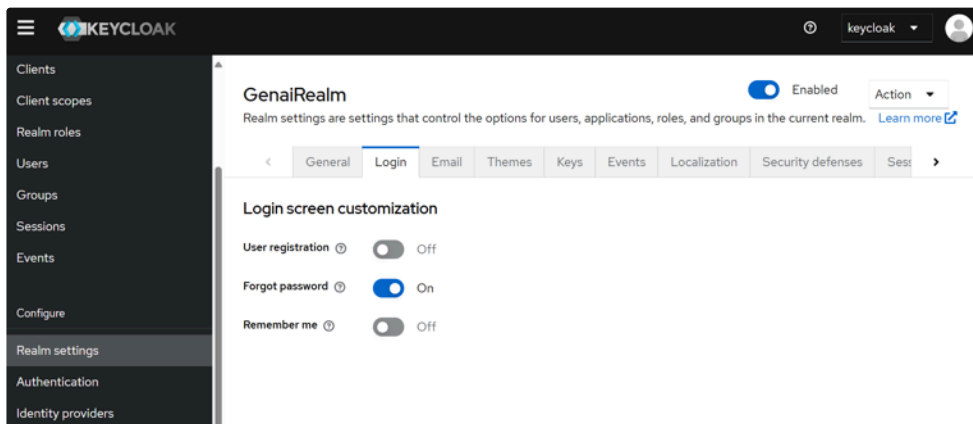




5. 左メニューから「Realm settings」を選択し、Emailタブを開きます。



6. メールサーバに必要な設定を行い、設定を Save してください。設定値の詳細はKeycloakのマニュアル ([Server Administration Guide](#)) をご確認ください。
7. Test connection を押し、連携が成功することを確認します。
8. 次に「Realm settings」のLoginタブを開き、Forgot passwordを有効(On)にします。



9. チャット画面/管理ポータル認証画面に「パスワードをお忘れですか?」のリンクが表示されることを確認してください。



10. パスワードを初期化したいメールアドレスを入力し、送信を押します。メールアドレス宛にパスワード初期化を案内するメールが届いていることを確認してください。



メールは題名が「パスワードのリセット」で以下の内容が届きます。リンクを押してパスワードを再設定してください。

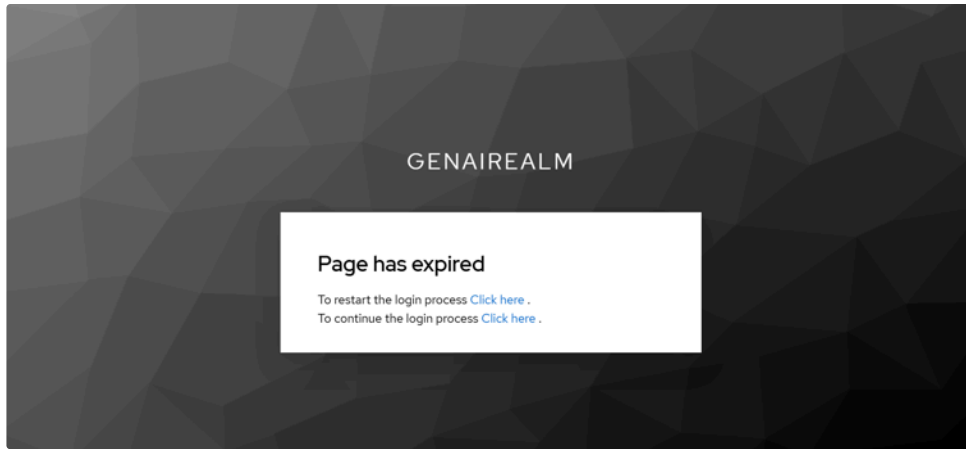
あなたのGenaiRealmアカウントのパスワードの変更が要求されています。以下のリンクをクリックしてパスワードのリセットを行ってください。

[パスワードのリセット](#)

このリンクは5分だけ有効です。

もしパスワードのリセットを行わない場合は、このメッセージを無視してください。何も変更されません。

なお、メール送信後以下の画面が表示されることがありますが初期化メールが届いていれば動作に問題ありません。



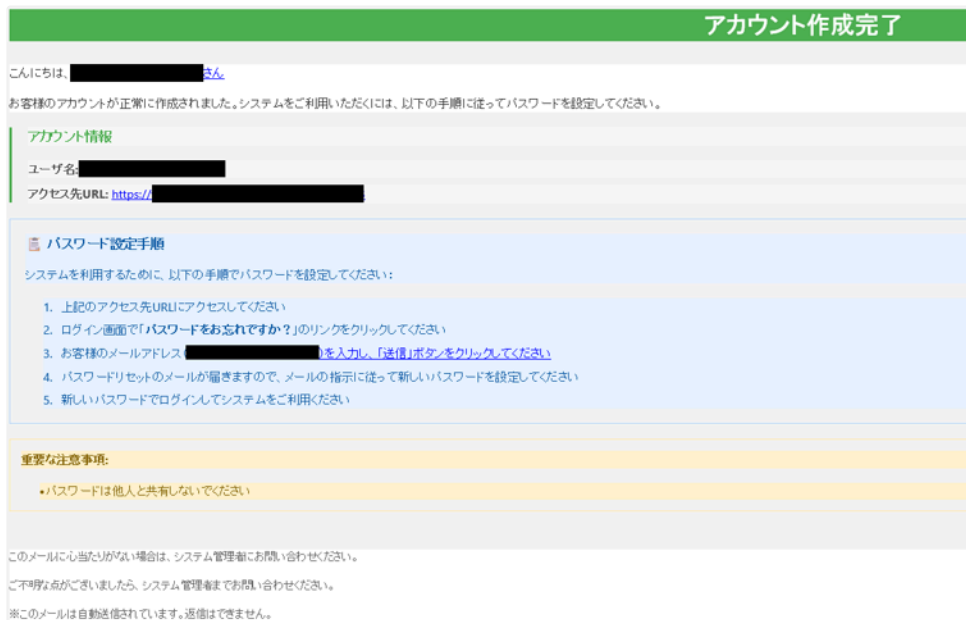
#### 4.16.1.1. 登録時のユーザへのメール文面

初期認証（メールサーバ連携有）のセットアップを実施した際のユーザ登録時のメール内容について説明します。

- ❶ 初期認証（メールサーバ連携有）のセットアップ」を実施した場合、ユーザ登録時のメール送信機能は自動的に有効化されています

デフォルトでは以下のようなメールが登録ユーザへ送信されます。

メールの文面は任意にカスタマイズ可能です。



メールの文面を任意にカスタマイズする場合は以下のファイルをそれぞれ編集してください。編集を行うと自動的に反映されます。

対象ファイルパス	説明
/opt/nec/genai/config/keycloak/themes/welcome-email/email/html/welcome-email.ftl	htmlメール

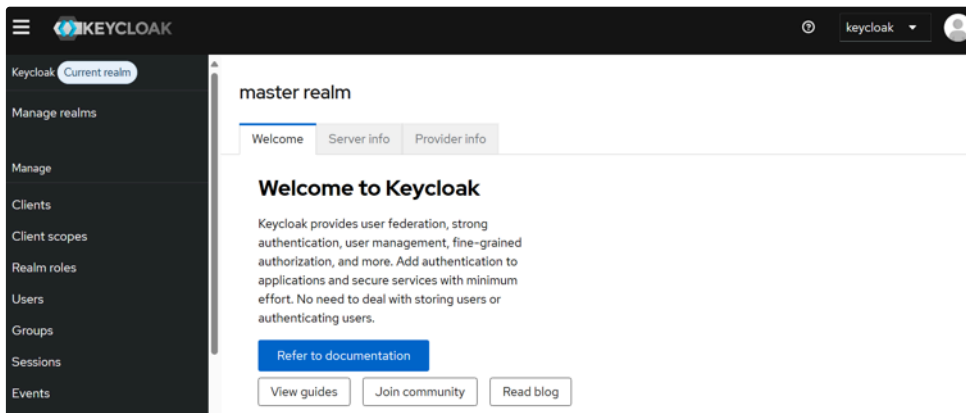
	テンプレート
<code>/opt/nec/genai/config/keycloak/themes/welcome-email/email/text/welcome-email.ftl</code>	textメールテンプレート

#### 4.16.2. 登録時のユーザへのメール連絡の無効化

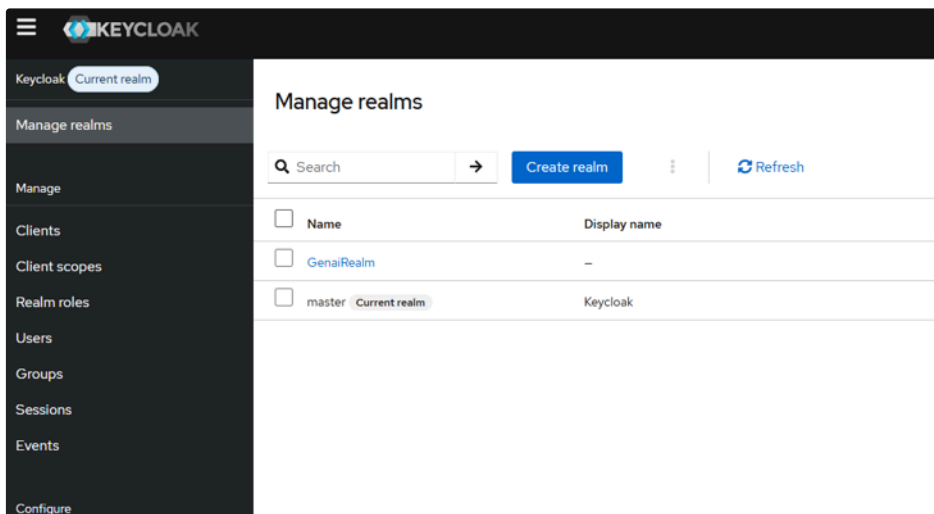
初期認証（メールサーバ連携有）のセットアップを実施した際のユーザ登録時のメール送信を無効化したい場合の手順について説明します。登録時のメール送信を無効化したい場合のみ実施してください。

**i** 初期認証（メールサーバ連携有）のセットアップ」を実施した場合、ユーザ登録時のメール送信機能は自動的に有効化されています

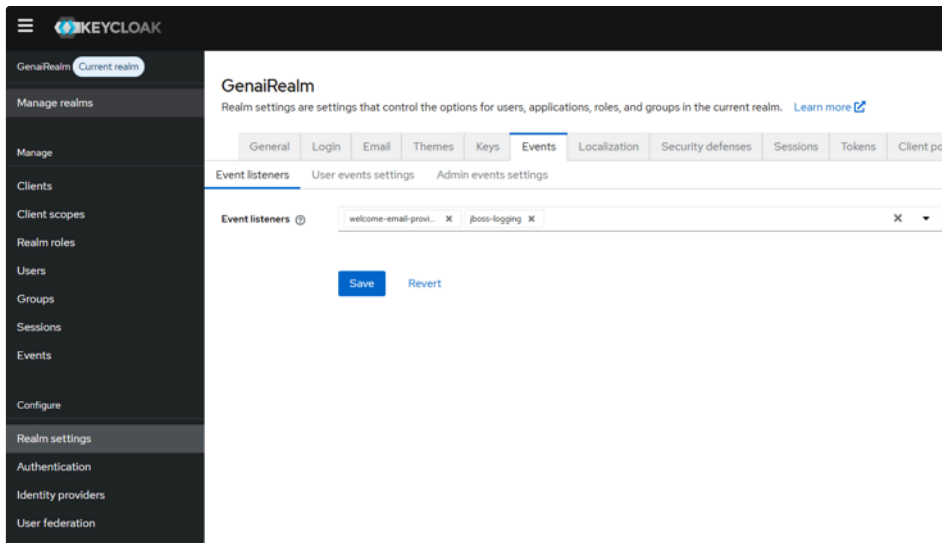
1. Keycloak管理画面にログインします。



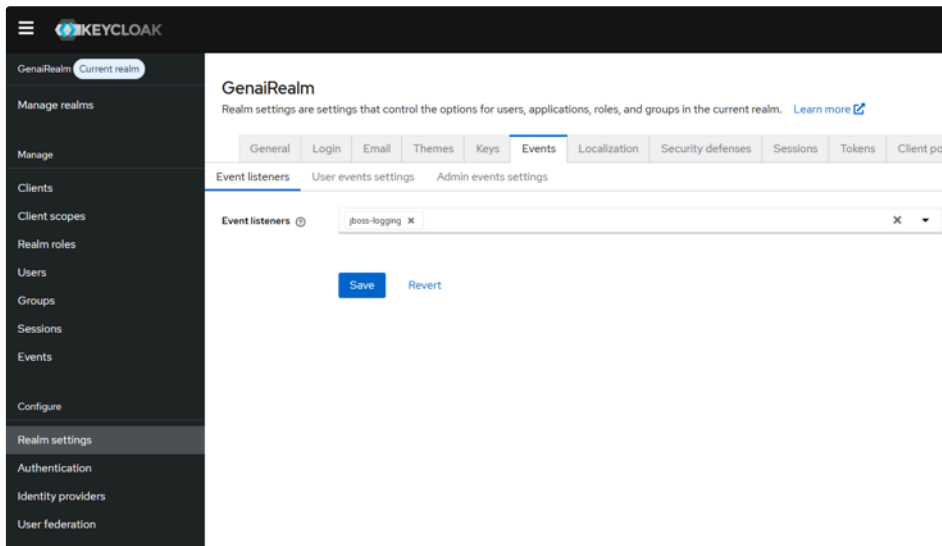
2. 左メニューから「Manage realms」を選択し、「GenaiRealm」のリンクをクリックします。



3. 左メニューから「Realm settings」を選択し、「Events」タブを選択します。



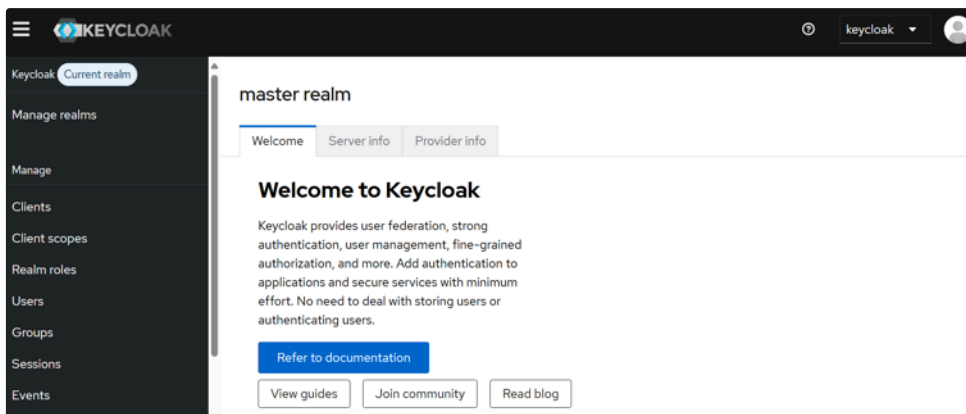
4. 「Event listeners」の「welcome-email-provider」の項目の×ボタンを押して削除後、Saveをクリックして設定を保存します。



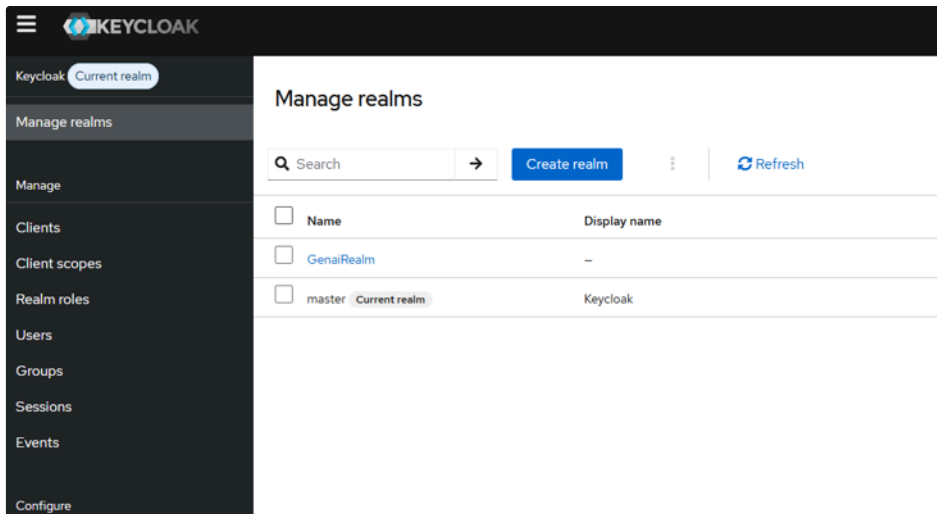
#### 4.16.2.1. 登録時のユーザへのメール連絡の再有効化

初期認証（メールサーバ連携有）のセットアップを実施した際のユーザ登録時のメール送信を無効化後に再度有効化したい場合の手順について説明します。

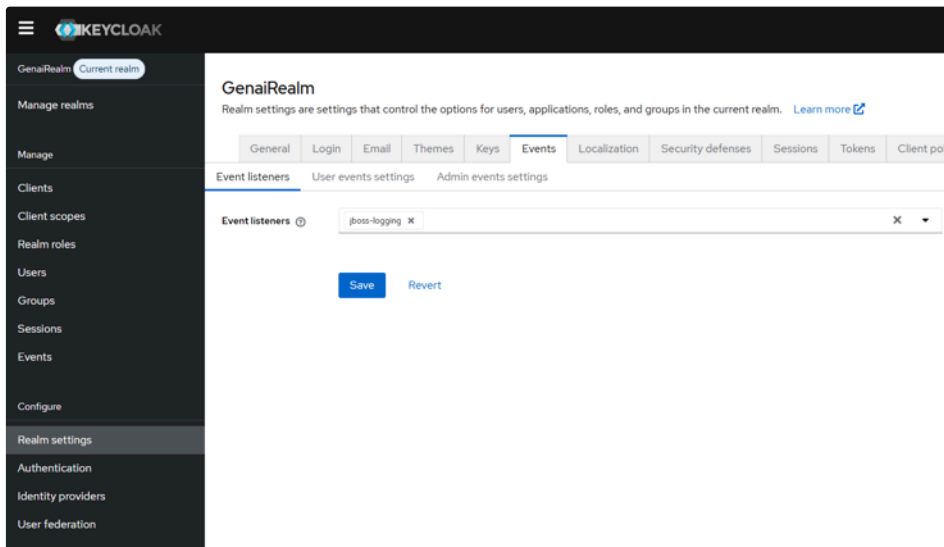
1. Keycloak管理画面にログインします。



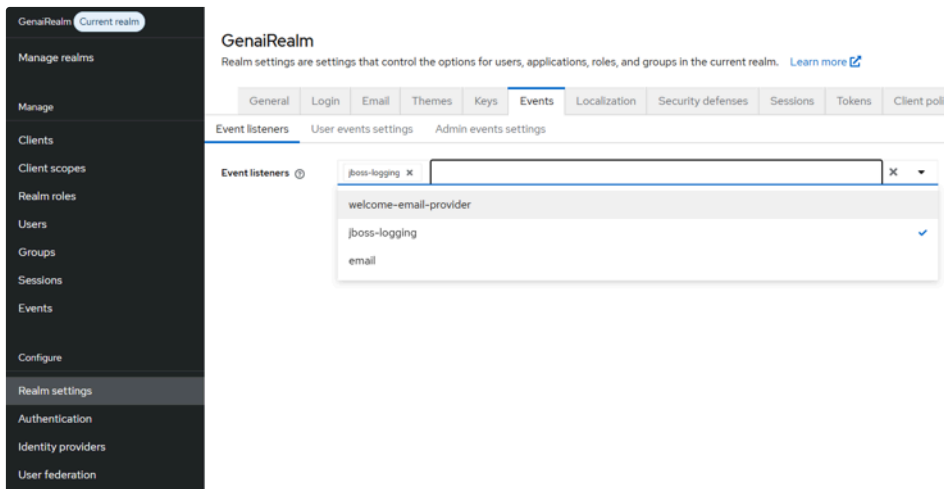
2. 左メニューから「Manage realms」を選択し、「GenaiRealm」のリンクをクリックします。



3. 左メニューから「Realm settings」を選択し、「Events」タブを選択します。



4. 「Event listeners」の「▼」ボタンを押下し、「welcome-email-provider」を選択後、Saveをクリックして設定を保存します。



## 4.17. 初期ユーザの削除

最後に、出荷時の初期ユーザを削除します。以下の手順に従い、ユーザを削除してください。

**i** 本手順は必ず一人以上の組織管理者ユーザが存在し、管理ポータル画面にログインできることを確認してから実施してください。

1. 初期ユーザ以外の組織管理者ユーザで管理ポータルにログインします。
2. 左メニューからユーザを選択し、初期ユーザ ([admin@example.com](mailto:admin@example.com)) の削除アイコンをクリックします。
3. 削除確認ダイアログが表示されますので、OKを押します。
4. 管理ポータルからログアウトします。
5. チャット画面にアクセスし、初期ユーザでログインできないことを確認します。

## 4.18. 管理者ユーザのパスワード変更

サーバの管理者ユーザadministratorの初期パスワードを変更します。管理者ユーザでログオンして以下のコマンドからパスワードを変更してください。

**i** サーバの管理者ユーザのユーザ名、パスワードは忘れないように保管することを推奨します。

```
1 | passwd
```