



概述

Casibase 是一个由 ChatGPT 驱动的开源 领域知识 数据库、即时通讯和论坛软件。

You need to enable JavaScript to run this app.

Casibase 特点

1. With a separate front-end and back-end architecture developed in Golang,

Casibase supports high concurrency, provides web-based management UI and supports multiple languages including Chinese, English, French, German, Spanish, Japanese, Korean, Russian, Indonesian, Malay, Portuguese, Italian, and Vietnamese.

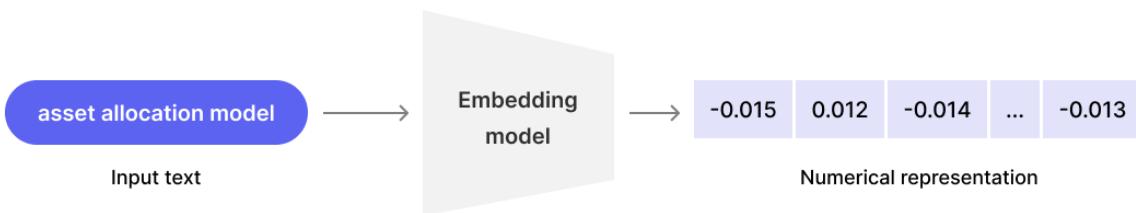
2. Casibase 支持第三方应用登录，如 GitHub、Google、QQ、微信等，并支持通过插件扩展第三方登录。
3. 基于嵌入和提示工程进行知识管理，Casibase 支持自定义嵌入方法和语言模型。
4. Casibase 通过数据库同步支持与现有系统集成，使用户可以平滑过渡到 Casibase。
5. Casibase 支持主流数据库：MySQL、PostgreSQL、SQL Server 等，并支持通过插件扩展新的数据库。
6. Casibase 是一个强大的资产管理工具，可以通过 RDP、VNC 和 SSH 协议轻松连接资产，并高效处理机器的远程连接。
7. Casibase 的安全日志审计功能允许你轻松跟踪和监控远程连接，详细记录连接开始时间、持续时间和其他相关详情，同时还能捕获和分析 Casdoor 操作的 API 日志，增强安全性和操作透明度。
8. Casibase 支持数据库管理。Casibase 支持数据库管理。Casibase 的数据库管理功能允许你轻松连接、管理和组织数据库，同时控制访问权限，简化数据库资源的用户管理和授权。
9. Casibase is an open-source container cloud platform based on Docker and Kubernetes. It is suitable for individuals or organizations to build their own dedicated container cloud environment. Based on the Casbin permission management engine, Casibase implements fine-grained access control policies. Users can easily create, orchestrate, and manage container applications on Casibase. The project focuses on optimizing Casibase's

application orchestration, service governance, and platform visualization core functions, improving platform usability and manageability, making it a leading lightweight container cloud platform.

工作原理

步骤 0（预备知识）

Casibase 的知识检索过程基于嵌入和提示工程，因此强烈建议您简要了解嵌入如何工作。嵌入的 [简介](#)。



步骤 1（导入知识）

要开始使用 Casibase，用户需要按照以下步骤导入知识并创建特定领域的知识数据库：

- 配置存储：**在 Casibase 仪表板中，用户首先应该配置存储设置。这涉及指定用于存储知识相关文件（如文档、图像或任何其他相关数据）的存储系统。用户可以根据其偏好和需求从多种存储选项中选择。这涉及指定用于存储知识相关文件（例如文档、图像或其它相关数据）的存储系统。用户可以根据他们的偏好和需求从多种存储选项中进行选择。
- 上传文件到存储：**一旦设置好存储，用户就可以将包含特定领域知识的文件上传到配置好的存储系统中。这些文件可以是各种格式，如文本文档、图像或结构化数据文

件（如 CSV 或 JSON）。这些文件可以采用各种格式，例如文本文档、图像，或 CSV、JSON 等结构化数据文件。

3. **选择知识生成的嵌入方法：**文件上传后，用户可以选择用于生成知识和相应向量的嵌入方法。嵌入是文本或视觉内容的数值表示，有助于高效的相似度搜索和数据分析。嵌入是文本或视觉内容的数字表示，有助于高效的相似性搜索和数据分析。

💡 提示

知识是如何嵌入的？

- 对于文本数据：用户可以选择各种嵌入方法，如 Word2Vec、GloVe 或 BERT，将文本知识转换为有意义的向量。
- 对于视觉数据：如果上传的文件包含图像或视觉内容，用户可以选择基于 CNN 的特征提取等图像嵌入技术来创建代表性向量。
- 更多方法即将推出...

通过遵循这些步骤，用户可以用相关信息和相应的嵌入来填充他们的领域知识数据库，这些将用于在 Casibase 中进行有效的搜索、聚类和知识检索。嵌入过程使系统能够理解不同知识片段之间的上下文和关系，实现更高效和有见地的知识管理和探索。嵌入过程使系统能够理解不同知识之间的上下文和关系，从而实现更高效和富有洞察力的知识管理与探索。

步骤 2（检索知识）

在导入你的 领域知识 后，Casibase 将其转换为 向量 并将这些向量存储在 向量数据库 中。这种向量表示启用了强大的功能，如 相似度搜索 和 相关信息的高效检索。你可以基于上下文或内容快速找到相关数据，实现高级查询并在你的领域知识中发现有价值的见解。这种向量表示使得诸如 相似性搜索 和 高效检索相关信息 之类的强大功能成为可

能。您可以根据上下文或内容迅速找到相关数据，从而实现高级查询并在您的领域知识中发现有价值的洞见。

步骤 3（构建提示）

Casibase 对存储的知识向量执行相似度搜索，以找到与用户查询最接近的匹配。使用搜索结果，它为 **语言模型** 创建一个 **提示模板** 来构建特定问题。这确保了基于 Casibase 中的领域知识提供准确和上下文相关的响应。利用搜索结果，它创建了一个 **提示模板** 来构建针对 **语言模型** 的具体问题。这确保了准确且符合上下文的响应，基于 Casibase 的领域知识提供全面答案。

步骤 4（实现目标）

在此阶段，通过使用 Casibase，您已成功获取所需的知识。在这个阶段，使用 Casibase，您已经成功获取了所需的知识。通过创新地将领域知识转换为向量并结合 ChatGPT 等强大的语言模型，Casibase 为您的查询提供准确和相关的响应。这使您能够高效地访问和利用存储在 Casibase 中的特定领域信息，轻松满足您的知识需求。这使您能够高效访问和利用存储在 Casibase 中的特定领域信息，从容满足您的知识需求。

步骤 5（可选的微调）

如果你发现结果不完全令人满意，你可以通过以下方式尝试获得更好的结果：

- 调整语言模型参数
- 提出多个问题
- 优化原始文件

通过利用这些微调选项，你可以提高在 Casibase 中的知识管理效率，确保系统更好地与你的目标保持一致，并提供更准确和有见地的信息。

① 提示

其他优化结果的方法（可能需要源代码更改）：

- 更新 `嵌入` 结果：通过调整领域知识的嵌入结果来改进知识表示。
- 修改 `提示` 模板：通过自定义提示，你可以从语言模型获得更精确的响应。
- 探索不同的 `语言模型`：尝试不同的模型，找到最适合你的响应生成需求的模型。

在线演示

只读站点（任何修改操作都会失败）

- 聊天机器人 (<https://ai.casibase.com>)
- 管理界面 (<https://ai-admin.casibase.com>)

可写站点（原始数据每 5 分钟恢复一次）

- 聊天机器人 (<https://demo.casibase.com>)
- 管理界面 (<https://demo-admin.casibase.com>)

全局管理员登录：

- 用户名：`admin`
- 密码：`123`

架构

Casibase 包含 2 个部分：

名称	描述	语言	源代码
前端	Casibase 应用程序的用户界面	JavaScript + React	https://github.com/casibase/casibase/tree/master/web
后端	Casibase 的服务器端逻辑和 API	Golang + Beego + MySQL	https://github.com/casibase/casibase



支持的模型

语言模型

模型	子类型	链接
OpenAI	gpt-4-32k-0613, gpt-4-32k-0314, gpt-4-32k, gpt-4-0613, gpt-4-0314, gpt-4, gpt-3.5-turbo-0613, gpt-3.5-turbo-0301, gpt-3.5-turbo-16k, gpt-3.5-turbo-16k-0613, gpt-3.5-turbo, text-davinci-003, text- davinci-002, text-curie-001, text- babbage-001, text-ada-001, text- davinci-001, davinci-instruct-beta, davinci, curie-instruct-beta, curie, ada, babbage	OpenAI
Hugging Face	meta-llama/Llama-2-7b, tiiuae/falcon-180B, bigscience/bloom, gpt2, baichuan-inc/ Baichuan2-13B-Chat, THUDM/chatglm2-6b	Hugging Face
Claude	claude-2, claude-v1, claude-v1-100k, claude- instant-v1, claude-instant-v1-100k, claude-v1.3, claude-v1.3-100k, claude-v1.2, claude-v1.0, claude-instant-v1.1, claude-instant-v1.1-100k, claude-instant-v1.0	Claude
OpenRouter	google/palm-2-codechat-bison, google/ palm-2-chat-bison, openai/gpt-3.5-turbo, openai/gpt-3.5-turbo-16k, openai/gpt-4,	OpenRouter

模型	子类型	链接
	openai/gpt-4-32k, anthropic/claude-2, anthropic/claude-instant-v1, meta-llama/llama-2-13b-chat, meta-llama/llama-2-70b-chat, palm-2-codechat-bison, palm-2-chat-bison, gpt-3.5-turbo, gpt-3.5-turbo-16k, gpt-4, gpt-4-32k, claude-2, claude-instant-v1, llama-2-13b-chat, llama-2-70b-chat	
Ernie	ERNIE-Bot, ERNIE-Bot-turbo, BLOOMZ-7B, Llama-2	Ernie
iFlytek	spark-v1.5, spark-v2.0	iFlytek
ChatGLM	chatglm2-6b	ChatGLM
MiniMax	abab5-chat	MiniMax
本地	custom-model	Local Computer

嵌入模型

模型	子类型	链接
OpenAI	AdaSimilarity, BabbageSimilarity, CurieSimilarity, DavinciSimilarity, AdaSearchDocument, AdaSearchQuery, BabbageSearchDocument, BabbageSearchQuery, CurieSearchDocument, CurieSearchQuery, DavinciSearchDocument, DavinciSearchQuery, AdaCodeSearchCode,	OpenAI

模型	子类型	链接
	AdaCodeSearchText, BabbageCodeSearchCode, BabbageCodeSearchText, AdaEmbeddingV2	
Hugging Face	sentence-transformers/all-MiniLM-L6-v2	Hugging Face
Cohere	embed-english-v2.0, embed-english-light-v2.0, embed-multilingual-v2.0	Cohere
Ernie	默认	Ernie
本地	custom-embedding	Local Computer

核心概念

作为 Casibase 的用户，您应当熟悉至少 4 个核心概念：[提供商](#)、[存储](#)、[聊天](#) 和 [向量](#)。

提供商

提供商是 Casibase 的支柱，提供基本服务并与外部系统集成。提供商类定义如下：

```
type Provider struct {
    Owner      string `xorm:"varchar(100) notnull pk" json:"owner"`
    Name       string `xorm:"varchar(100) notnull pk" json:"name"`
    CreatedTime string `xorm:"varchar(100)" json:"createdTime"`

    DisplayName string `xorm:"varchar(100)" json:"displayName"`
    Category    string `xorm:"varchar(100)" json:"category"`
    Type        string `xorm:"varchar(100)" json:"type"`
    ClientId   string `xorm:"varchar(100)" json:"clientId"`
    ClientSecret string `xorm:"varchar(2000)" json:"clientSecret"`
    ProviderUrl string `xorm:"varchar(200)" json:"providerUrl"`
}
```

提示

Casibase 中主要有两种类型的提供商：

- **存储提供商。** 存储提供商便于在 Casibase 中存储和检索数据。它支持多种存储选项，包括：

- AWS
 - Azure
 - 本地文件系统
- **AI 提供商。** AI 提供商负责处理 Casibase 中与 AI 相关的任务和服务。它支持多种 AI 模型和技术，包括：
 - OpenAI
 - ChatGLM
 - InternLM

向量

Casibase 中的向量代表不同类型数据的数值表示。这些向量使得信息处理和分析更加高效。部分可用的向量类型包括：

- 文本向量
- 图像向量
- (其他向量类型)

向量类定义如下：

```
type Vector struct {
    Owner      string      `xorm:"varchar(100) notnull pk"
                           json:"owner"`
    Name       string      `xorm:"varchar(100) notnull pk"
                           json:"name"`
    CreatedTime string     `xorm:"varchar(100)" json:"createdTime"`
    DisplayName string     `xorm:"varchar(100)" json:"displayName"`
}
```

聊天

聊天是用户与 Casibase 中 AI 模型之间交互沟通的核心。它们由三个基本组成部分构成：

- 问题：用户的输入或查询，用于寻求信息或帮助。
- 查询提示：经过格式化的用户问题版本，为 AI 模型的处理做准备。
- 答案：AI 生成的对用户问题的响应，提供相关信息或解决方案。

聊天类定义如下：

```
type Chat struct {
    Owner      string `xorm:"varchar(100) notnull pk"`
    json:"owner"`
    Name       string `xorm:"varchar(100) notnull pk"`
    json:"name"`
    CreatedTime string `xorm:"varchar(100)" json:"createdTime"`
    UpdatedTime string `xorm:"varchar(100)" json:"updatedTime"`

    DisplayName string `xorm:"varchar(100)" json:"displayName"`
    Category    string `xorm:"varchar(100)" json:"category"`
    Type        string `xorm:"varchar(100)" json:"type"`
    User1       string `xorm:"varchar(100)" json:"user1"`
    User2       string `xorm:"varchar(100)" json:"user2"`
    Users       []string `xorm:"varchar(100)" json:"users"`
    MessageCount int     `json:"messageCount"`
}
```

嵌入

嵌入是将各种类型的数据，如文本和图像，转换成密集向量表示的过程。此步骤对于促

进 Casibase 内高效的数据处理和分析至关重要。

💡 提示

- 通过嵌入，聊天中的问题和存储中的知识文件将转化为向量，并用于下一步的知识搜索。
- Casibase 默认的嵌入方法由 OpenAI 提供，每分钟最多调用三次。我们建议尽量减少知识文件之间的耦合，以便嵌入和后续处理。

Transactions

Transactions track AI token usage and costs through Casdoor integration. Each message generates a transaction record for billing and usage monitoring. See [Billing & Usage](#) for details.

Records

Records support data aggregation through a `Count` field, allowing consolidated logging without storing individual entries. See [Records](#) for more information.

Scans

Scans represent network and security scanning operations executed against infrastructure assets. The Scan object enables automated discovery, security auditing, and system assessment through integrated scan providers.

Each scan targets a specific asset (such as a virtual machine) or IP address, executes using a configured scan provider (like Nmap or OS Patch), and captures

results in both raw and structured formats. Scans integrate with the asset inventory to automatically resolve target addresses and maintain historical scanning records for security analysis.

服务器安装

要求

操作系统

支持所有主要操作系统，包括Windows、Linux和macOS。

环境

- Go 1.20+
- Node.js LTS (18)
- Yarn 1.x

信息

Casibase的使用分为两个步骤：

- 步骤1: [部署并运行Casdoor](#)
- 步骤2: 部署并运行Casibase（本文档）

我们强烈建议使用[Yarn 1.x](#)来运行和构建Casdoor和Casibase前端，使用NPM可能会导致UI样式问题，详见：[casdoor#294](#)

小心

对于中国用户，为了成功下载Go依赖包，您需要通过配置GOPROXY环境变量来使用Go代理。我们强烈推荐：<https://goproxy.cn/> 我们强烈推荐：

<https://goproxy.cn/>

数据库

Casibase 使用 XORM 与数据库通信。 Casibase使用XORM与数据库通信。 基于Xorm 驱动支持， Casibase目前支持以下数据库：

- MySQL
- MariaDB
- PostgreSQL
- CockroachDB
- SQL Server
- Oracle
- SQLite 3
- TiDB

guacd

Casibase使用guacamole-server提供远程桌面访问。 如果您想使用此功能， 需要先安装 guacamole-server。 如果您尚未安装guacamole-server，请参考[guacamole-server安装](#)。 如果您想使用此功能， 您需要先安装 guacamole-server。 如果您还未安装 guacamole-server，请参阅[guacamole-server 安装指南](#)。

您也可以使用以下命令在docker中运行guacd：

```
docker run -d --name guacd -p 4822:4822 guacamole/guacd
```

下载

Casibase can be installed using pre-built binaries or by building from source.

Pre-built Binaries

For production deployments, we recommend using pre-built binaries. Download the latest release from the [GitHub Releases page](#). Casibase provides binaries for:

- Linux: x86_64 (amd64) and ARM64
- Windows: x86_64 (amd64) and ARM64
- macOS: x86_64 (amd64) and ARM64

Extract the downloaded archive and you'll find the Casibase binary along with the web frontend and configuration files ready to use.

Build from Source

Casibase的源代码托管在GitHub上：<https://github.com/casibase/casibase>。Go后端代码和React前端代码都在同一个仓库中。 Go 后端代码和 React 前端代码都在同一个仓库中。

名称	描述	语言	源代码
前端	Casibase的Web前端UI	JavaScript + React	https://github.com/casibase/casibase/tree/master/web
后	Casibase的	Golang +	https://github.com/casibase/

名称	描述	语言	源代码
端	RESTful API后端	Beego + XORM	casibase

Casibase支持[Go Modules](#)。要下载代码，您只需通过git克隆代码：要下载代码，只需通过git克隆代码即可：

```
cd path/to/folder  
git clone https://github.com/casibase/casibase
```

配置

配置Casdoor

请参考[Casdoor-SSO](#)部分来配置Casdoor。

记住您在Casdoor配置中的[clientId](#)、[clientSecret](#)、[organization](#)、[application](#)等信息，我们稍后会用到。

配置数据库

Casibase支持mysql、mssql、sqlite3、postgres。Casibase默认使用mysql。
Casibase 默认使用 mysql。

MySQL

Casibase将在名为[casibase](#)的MySQL数据库中存储其用户、节点和主题信息。如果数据库不存在，需要手动创建。数据库连接字符串可以在以下位置指定：

<https://github.com/casibase/casibase/blob/master/conf/app.conf> 如果数据库不存在，则需要手动创建。数据库连接字符串可以在此指定：<https://github.com/casibase/casibase/blob/master/conf/app.conf>

```
driverName = mysql  
dataSourceName = root:123456@tcp(localhost:3306)/  
dbName = casibase
```

PostgreSQL

由于我们必须在使用xorm打开Postgres时选择一个数据库，因此您应该在运行Casibase之前手动准备一个数据库。

假设您已经准备好了[一个名为casibase的数据库](#)，那么您应该这样指定[app.conf](#)：

```
driverName = postgres  
dataSourceName = "user=postgres password=postgres host=localhost  
port=5432 sslmode=disable dbname=casibase"  
dbName =
```

① 信息

对于PostgreSQL，确保[dataSourceName](#)有非空的[dbName](#)，并像上面的例子一样将独立的[dbName](#)字段留空。

CockroachDB

您也可以使用[postgres](#)驱动程序使用Cockroachdb。它的配置与PostgreSQL相同。其配置与 PostgreSQL 相同。

```
driverName = postgres
```

① 信息

对于CockroachDB，不要忘记像上面的例子一样在`dataSourceName`中添加`serial_normalization=virtual_sequence`。否则，每当服务启动或重启时，您都会收到关于数据库已存在的错误。注意，这必须在数据库创建之前添加。否则，每当服务启动或重启时，您会收到有关数据库已存在的错误信息。请注意，此项必须在数据库创建之前添加。

Sqlite3

您应该这样指定`app.conf`:

```
driverName = sqlite
dataSourceName = "file:casibase.db?cache=shared"
dbName = casibase
```

自定义配置

Casibase支持自定义配置，您可以修改配置文件`conf/app.conf`来更改配置。

```
casdoorEndpoint = <您的Casdoor端点>
clientId = <您的Casdoor应用程序的客户端ID>
clientSecret = <您的Casdoor应用程序的客户端密钥>
casdoorOrganization = <您的Casdoor组织名称>
casdoorApplication = <您的Casdoor应用程序名称>
```

运行

⚠ 小心

Casibase需要Casdoor提供访问控制和一些后端服务，所以在运行**Casibase**之前，您必须确保**Casdoor**正常运行。

如何安装和运行Casdoor:

- [Casdoor安装](#)

Using Pre-built Binary

If you downloaded a pre-built binary, extract the archive and run Casibase directly. The binary already includes the compiled frontend.

For Linux/macOS:

```
# Example for Linux x86_64
tar -xzf casibase_Linux_x86_64.tar.gz
cd casibase
./casibase
```

For Windows (PowerShell):

```
# Example for Windows x86_64
Expand-Archive casibase_Windows_x86_64.zip -DestinationPath .
cd casibase
.\casibase.exe
```

The server will start on port 14000 by default.

开发模式

后端

Casibase的Go后端默认在端口14000上运行。您可以使用以下命令启动Go后端： 您可以使用以下命令启动 Go 后端：

```
go run main.go
```

服务器成功运行后，我们可以启动前端部分。

前端

Casibase的前端是一个非常经典的[Create-React-App \(CRA\)](#)项目。它默认在端口13001上运行。使用以下命令运行前端： 默认在端口 13001 运行。 使用以下命令运行前端：

```
cd web  
yarn install  
yarn start
```

Building from Source

If you're building from source, follow these steps to create a production build.

后端

将Casibase Go后端代码构建成可执行文件并启动。

对于Linux：

```
go build  
./casibase
```

对于Windows：

```
go build  
casibase.exe
```

前端

将Casibase前端代码构建成静态资源（.html、.js、.css文件）：

```
cd web  
yarn install  
yarn build
```

Nginx

💡 提示

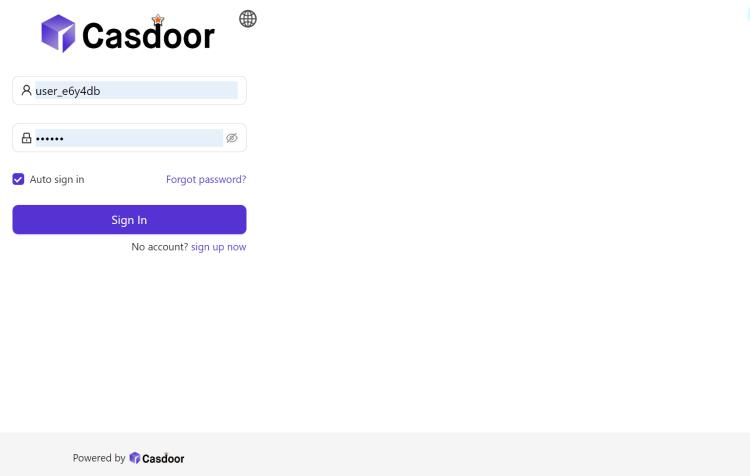
如果您使用nginx作为反向代理，需要在nginx配置文件中添加以下配置：

```
location / {  
    *** 您的配置 ***  
    proxy_set_header Upgrade $http_upgrade;  
    proxy_set_header Connection "upgrade";  
}
```

因为Casibase使用websocket与guacd通信。

预览

在浏览器中访问: `http://localhost:13001`。 在浏览器中访问:
`http://localhost:13001`。 使用您刚刚在Casdoor中注册的用户账户登录Casibase仪表板:



然后您将进入Casibase的主页:

Powered by Casibase

 提示

要使用其他端口, 请编辑 `conf/app.conf` 并修改 `httpport`, 然后重启 Go 后端。

(可选) 使用Docker尝试

要求

硬件

如果您想自己构建 Docker 镜像, 请确保您的机器至少有 **2GB** 的内存。Casibase 的前端是一个 React 的 NPM 项目。构建前端需要至少 **2GB** 的内存。内存少于 **2GB** 可能导致前端构建失败。

如果您只需要运行预构建的镜像, 请确保您的机器至少有**100MB**内存。

操作系统

支持所有操作系统 (Linux、Windows和macOS) 。

Docker

您可以在Linux中使用Docker (docker-engine版本 ≥ 17.05) 或在Windows和macOS中使用Docker Desktop。

- [Docker](#)

无论操作系统如何, 用户都必须确保他们拥有 docker-engine 版本 ≥ 17.05 。无论使用什么操作系统, 用户都必须确保有docker-engine版本 ≥ 17.05 。这是因为我们在 docker-compose.yml 中使用了多阶段构建功能, 该功能在17.05及以上版本中得到支持。更多信息, 请参见<https://docs.docker.com/develop/develop-images/multistage-build/>。更多信息, 请参见 <https://docs.docker.com/develop/develop-images/multistage-build/>。

如果您使用 docker-compose，请确保您有 docker-compose 版本 ≥ 2.2 。对于 Linux 用户，请注意 docker-compose 需要与 docker-engine 分开安装。对于 Linux 用户，请注意 docker-compose 需要与 docker-engine 分开安装。

获取镜像

我们提供了两个 DockerHub 镜像：

名称	描述	建议
casibase-all-in-one	镜像中包含 Casibase 和 MySQL 数据库	此镜像已包含一个测试数据库，仅用于测试目的
casibase	镜像中仅包含 Casibase	此镜像可以连接到您自己的数据库并用于生产环境

1. [casbin/casibase-all-in-one](#): 此镜像包含 casibase 可执行文件、MySQL 数据库和所有必要的配置。它是为想要快速尝试 Casibase 的新用户设计的。使用此镜像，您可以仅用一两个命令就立即启动 Casibase，无需任何复杂配置。但是，请注意，我们不建议在生产环境中使用此镜像。它旨在为想要快速体验 Casibase 的新用户而设计。使用此镜像，您可以通过一两条命令立即启动 Casibase，而无需任何复杂配置。但请注意，我们不建议在生产环境中使用此镜像。

选项-1：使用测试数据库

运行容器时将端口 [14000](#) 暴露给主机。如果本地主机上不存在该镜像，将自动拉取。如果本地主机上不存在该镜像，将会自动拉取。

```
docker run -p 14000:14000 casbin/casibase-all-in-one
```

在您的浏览器中访问 <http://localhost:14000>。 在浏览器中访问 <http://localhost:14000>。 使用默认的全局管理员账户登录Casibase仪表板: `built-in/admin`

```
admin  
123
```

选项-2：使用docker-compose尝试

在 `docker-compose.yml` 文件中创建一个 `conf/app.conf` 目录。 然后，从 Casibase 中复制 `app.conf`。 关于 `app.conf` 的更多详情，您可以查看[通过Ini文件](#)。

Below is a minimal but complete `docker-compose.yml` example that starts a MySQL database and the Casibase service. It configures Casibase to connect to the database using MySQL. Save this file as `docker-compose.yml` (next to a `conf` folder if you want to mount a custom `app.conf`).

```
services:  
  db:  
    image: mysql:8.0  
    restart: always  
    environment:  
      MYSQL_ROOT_PASSWORD: 123456  
      MYSQL_DATABASE: casibase  
    volumes:  
      - db_data:/var/lib/mysql  
    ports:  
      - "3306:3306" # optional: expose DB to host  
  
  casibase:  
    image: casbin/casibase:latest  
    restart: unless-stopped  
    depends_on:
```

What does the above compose file do:

- The Casibase container connects to the database using the Compose service name `db` (i.e. `db:3306`). When both services run in the same Docker network (default for compose), using the service name as host is the simplest and most reliable approach.
- The `dataSourceName` above uses the MySQL root account for simplicity. For production use please create a dedicated DB user and a strong password.
- Mounting `./conf/app.conf` into `/conf/app.conf` is optional. If you prefer environment variables, you can remove the mount and rely on the `driverName` and `dataSourceName` variables.
- If both `app.conf` and environment variables are provided, the environment variables take precedence and will override the corresponding settings in `app.conf`.

 注意

Casdoor: By default Casibase uses the hosted Casdoor instance at `https://door.casdoor.com` for user authentication. If you need to manage users, applications, or customize the authentication flow, you must deploy your own Casdoor instance and update Casibase's `app.conf` (or the equivalent environment variables) to point to your Casdoor server. You can look at [Casdoor configuration](#) for more details.

`RUNNING_IN_DOCKER`: By default `RUNNING_IN_DOCKER` is enabled in docker image. When enabled, Casibase replaces `localhost` with the Docker bridge address (for example, `host.docker.internal` or the equivalent bridge hostname) so that the container can reach services running on the host.

Bring up the services:

```
docker-compose up -d
```

Check logs (follow):

```
docker-compose logs -f casibase
```

在您的浏览器中访问 <http://localhost:14000>。 使用默认的全局管理帐户登录Casibas控制面板：`内置/管理`

admin

123

Stop and remove containers and volumes (data removed):

```
docker-compose down -v
```

选项-3：直接使用标准镜像



如果不方便将配置文件挂载到容器中，使用环境变量也是一种可能的解决方案。

example

```
docker run \
-e driverName=mysql \
-e dataSourceName='user:password@tcp(x.x.x.x:3306)/*' \
-p 14000:14000 \
```

创建 `conf/app.conf`。 创建 `conf/app.conf`。 您可以从Casibase的 `conf/app.conf` 复制。 关于 `app.conf` 的更多详情，您可以查看[通过Ini文件](#)。 关于 `app.conf` 的更多详情，您可以查看[通过Ini文件](#)。

然后运行

```
docker run -p 14000:14000 -v /folder/of/app.conf:/conf casbin/casibase:latest
```

总之，只需要将 `app.conf` 挂载到 `/conf/app.conf` 并启动容器即可。

在您的浏览器中访问 <http://localhost:14000>。 在浏览器中访问 <http://localhost:14000>。 使用默认的全局管理员账户登录 Casibase 仪表板：`built-in/admin`

```
admin  
123
```

(Optional) Try with K8s Helm

Introduction

This guide shows how to deploy Casibase on Kubernetes using Helm for easy and scalable management. Helm simplifies the deployment process and allows for easy configuration management.

Prerequisites

- A running Kubernetes cluster
- Helm v3 installed
- kubectl configured to connect to your cluster
- A MySQL/PostgreSQL database (recommended for production)
- A Casdoor instance for authentication

Configuration

Before installation, you must create an application configuration file. The Helm chart will not work with default values.

Pre step: Create app.conf File

Create an `app.conf` file with your specific settings:

You can view more details about the configuration options in the [Casibase Configuration Documentation](#).

Or check the latest configuration options in [app.conf example](#).

```
appname = casibase
httpport = 14000
runmode = prod
SessionOn = true
copyRequestBody = true

# Database Configuration - REQUIRED
driverName = mysql
dataSourceName = your-username:your-password@tcp(your-db-host:3306)-
dbName = your-database

# Casdoor Authentication - REQUIRED
casdoorEndpoint = https://door.casdoor.com
clientId = your-client-id
clientSecret = your-client-secret
casdoorOrganization = "your-organization"
casdoorApplication = "your-application"
redirectPath = /callback

# Optional Settings
redisEndpoint =
guacamoleEndpoint = 127.0.0.1:4822
isDemoMode = false
disablePreviewMode = false
logPostOnly = true
landingFolder =
cacheDir = "C:/casibase_cache"
appDir = ""
isLocalIpDb = false
audioStorageProvider = ""
providerDbName = ""
socks5Proxy = "127.0.0.1:10808"
publicDomain = ""
adminDomain = ""
enableExtraPages = false
shortcutPageItems = []
usageEndpoints = []
iframeUrl = ""
forceLanguage = ""
defaultLanguage = "en"
```

Optional: Using Secrets for Sensitive Data

For production environments, create a Kubernetes secret with your configuration:

```
# Create secret from app.conf file
kubectl create secret generic casibase-config --from-
file=app.conf=./app.conf
```

Installation Steps

Step 1: Prepare Configuration Files

Ensure you have created both:

- `app.conf` - Application configuration

Step 2: Install with Configuration File



Visit the [Casbin Helm Chart](#) to find the latest version.

Install Casibase by passing the app.conf file directly:

```
# Method 1: Pass app.conf content as appConfig parameter
helm install casibase oci://registry-1.docker.io/casbin/casibase-helm-
chart \
--version v1.549.0 \
--set-file appConfig=./app.conf
```

Alternative Installation with Secret

If using secrets for sensitive data:

```
# Create secret first
kubectl create secret generic casibase-config --from-
file=app.conf=./app.conf

# Install with secret reference (no additional values file needed)
helm install casibase oci://registry-1.docker.io/casbin/casibase-helm-
chart \
  --version v1.549.0 \
  --set appConfig="" \
  --set appConfigFromSecret=casibase-config
```

Step 3: Verify Installation

Check the deployment status:

```
kubectl get pods
kubectl get services
kubectl logs -l app.kubernetes.io/name=casibase
```

Step 4: Access Casibase

Once installed, Casibase will be accessible through the Kubernetes service on port 14000. If you enabled ingress, it will be available at your configured domain.

Configuration Options Reference

The following table shows the main configuration parameters available in the Helm chart:

Parameter	Description	Default Value
<code>replicaCount</code>	Number of Casibase replicas to run	<code>1</code>
<code>image.repository</code>	Docker image repository	<code>casbin</code>
<code>image.name</code>	Docker image name	<code>casibase</code>
<code>image.pullPolicy</code>	Image pull policy	<code>IfNotPresent</code>
<code>image.tag</code>	Image tag (defaults to chart appVersion)	<code>""</code>
<code>appConfig</code>	Application configuration (app.conf content)	See values.yaml
<code>appConfigFromSecret</code>	Mount app.conf from secret instead	<code>""</code>

Parameter	Description	Default Value
<code>service.type</code>	Kubernetes service type	<code>ClusterIP</code>
<code>service.port</code>	Service port	<code>14000</code>
<code>ingress.enabled</code>	Enable ingress	<code>false</code>
<code>ingress.hosts</code>	Ingress hosts configuration	<code>[]</code>
<code>resources</code>	CPU/ Memory resource requests and limits	<code>{}</code>
<code>autoscaling.enabled</code>	Enable horizontal pod autoscaler	<code>false</code>
<code>autoscaling.minReplicas</code>	Minimum number of replicas	<code>1</code>
<code>autoscaling.maxReplicas</code>	Maximum number of replicas	<code>100</code>

Parameter	Description	Default Value
<code>autoscaling.targetCPUUtilizationPercentage</code>	CPU utilization threshold	80
<code>nodeSelector</code>	Node labels for pod assignment	{}
<code>tolerations</code>	Toleration labels for pod assignment	[]
<code>affinity</code>	Affinity settings for pod assignment	{}

Advanced Configuration Options

For production deployments, consider these additional options:

```
# Autoscaling
autoscaling:
  enabled: true
  minReplicas: 2
  maxReplicas: 10
  targetCPUUtilizationPercentage: 70
```

Managing the Deployment

Upgrading Casibase

To upgrade your Casibase deployment to a new version:

```
helm upgrade casibase oci://registry-1.docker.io/casbin/casibase-helm-
chart --version <new-version>
```

To upgrade with custom values:

```
helm upgrade casibase oci://registry-1.docker.io/casbin/casibase-helm-
chart --version <new-version> \
-f custom-values.yaml
```

Checking Deployment Status

Monitor your deployment:

```
# Check pod status
kubectl get pods -l app.kubernetes.io/name=casibase

# Check service status
kubectl get svc -l app.kubernetes.io/name=casibase

# View logs
kubectl logs -l app.kubernetes.io/name=casibase

# Describe deployment
helm status casibase
```

Uninstalling Casibase

To completely remove Casibase from your cluster:

```
helm uninstall casibase
```

Troubleshooting

Common Issues

1. Pod not starting: Check logs with `kubectl logs <pod-name>`
2. Service not accessible: Verify service configuration and ingress setup
3. Database connection issues: Ensure database credentials and connectivity are correct
4. Configuration errors: Validate your `appConfig` syntax
5. Casdoor authentication failures: Verify Casdoor endpoint and credentials
6. Domain/URL issues: Check domain configuration and DNS resolution

Configuration-Related Issues

Problem: Casdoor authentication not working

- Verify `casdoorEndpoint` is accessible from the cluster
- Check `clientId` and `clientSecret` are correct
- Ensure `redirectPath` matches your Casdoor application configuration

Problem: Configuration syntax errors

```
# Validate YAML syntax before deployment
```

Getting Help

- Check pod events: `kubectl describe pod <pod-name>`
- View Helm release info: `helm status casibase`
- Review configuration: `helm get values casibase`

Conclusion

Using Helm to deploy Casibase on Kubernetes provides a robust, scalable solution for managing your knowledge base platform. The chart offers flexible configuration options to suit various deployment scenarios, from development environments to production clusters.

Key benefits of this approach:

- Easy deployment and updates through Helm commands
- Flexible configuration through values files
- Kubernetes-native scaling and management
- Production-ready with proper resource management and health checks

For more advanced configurations and troubleshooting, refer to the [Kubernetes documentation](#) and [Helm documentation](#).

Casibase Public API

Casibase frontend web UI is a [SPA \(Single-Page Application\)](#) developed in React. The React frontend consumes the Casibase RESTful API exposed by the Go backend code. This RESTful API is referred to as the [Casibase Public API](#). In other words, with HTTP calls, you can do everything just like how the Casibase web UI itself does. There's no other limitation. The API can be utilized by the following:

- Casibase's frontend
- Casibase client SDKs (e.g., casibase-java-sdk)
- Any other customized code from the application side

The full reference for the [Casibase Public API](#) can be found on Swagger: <https://ai-admin.casibase.com/swagger>. These Swagger docs are automatically generated using Beego's Bee tool. If you want to generate the Swagger docs by yourself, see: [How to generate the swagger file](#)

How to authenticate with [Casibase Public API](#)

Casibase Public API supports two application-level authentication methods: [Bearer Token](#) and [Basic Auth](#). The [Bearer Token](#) method is recommended as it is more secure.

SDK Authentication Example (Java)

To illustrate how authentication is handled in practice, here is an example from the Casibase Java SDK. The following code shows how the SDK constructs the credential for API requests. This process authenticates the SDK with application-level permissions, effectively acting as an admin.

```
// ...

protected void Service(Config config, AuthTypeEnum authType)
throws Exception {
    this.config = config;
    switch (authType){
        case BASIC:
            this.credential =
Credentials.basic(config.clientId, config.clientSecret);
            break;
        case BEARER:
            String token = config.clientId + ":" +
config.clientSecret;
            this.credential = "Bearer " +
DigestUtils.md5Hex(token);
            break;
        default:
            throw new Exception("Invalid auth type");
    }
}
```

The example above demonstrates how to prepare the credential for both authentication types:

- **BASIC:** It uses a helper (`okhttp3.Credentials`) to perform the standard Base64 encoding for Basic Authentication.
- **BEARER:** It constructs the token by taking the MD5 hash of

`clientId:clientSecret` (using `org.apache.commons.codec.digest.DigestUtils`) and prepending the result with "Bearer".

1. By **Bearer Token** (Recommended)

This method is more secure because it uses a static access token, which is a hashed value of your `clientId` and `clientSecret`.

How to get the access token?

The access token is calculated using the following formula: `md5(clientId + ":" + clientSecret)`

How to authenticate?

The access token must be provided in the `Authorization` header as a Bearer Token.

```
Authorization: Bearer <The access token>
```

2. By **Basic Auth**

This method uses the `clientId` and `clientSecret` directly for authentication. It is considered less secure because the `clientSecret` might be exposed. It is supported for convenience and compatibility purposes.

How to authenticate?

HTTP Basic Authentication: This is the standard way.

```
Authorization: Basic <The Base64 encoding of  
"clientId:clientSecret">
```

If you are not familiar with Base64 encoding, you can use a library for this, as [HTTP Basic Authentication](#) is a widely supported standard.

Where to find the Client ID and Secret?

Both authentication methods require a `clientId` and `clientSecret`. You can find these values for your application in the Casibase configuration file: [conf/app.conf](#).

System Info

Casibase provides real-time system monitoring to help administrators track resource usage and performance. The System Info page displays key metrics about your Casibase instance.

Accessing System Info

Global administrators can access the system monitoring dashboard through the SysInfo menu item in the navigation bar. This page automatically refreshes metrics every 3 seconds.

Available Metrics

The system monitoring dashboard displays key metrics in an organized layout that spans the full page width for better visibility.

CPU Usage: Displays the current CPU utilization percentage for each processor core. This helps identify if the system is experiencing high computational load.

Memory Usage: Shows the memory consumption of the Casibase process relative to total system memory. Monitor this to ensure sufficient memory is available for smooth operation.

Disk Usage: Tracks the storage space used by Casibase's data directory. This metric helps you monitor growth and plan for storage capacity.

Network Usage: Displays cumulative network I/O for the Casibase process,

including bytes sent, received, and total throughput. This helps track data transfer patterns and identify potential network bottlenecks.

API Performance Metrics

For deployments with Prometheus integration, the System Info page also shows:

- **API Latency:** Response time distribution across different API endpoints
- **API Throughput:** Request count per endpoint to identify high-traffic routes

These metrics help optimize performance and troubleshoot slow requests.

Container Cloud

Casibase is an open-source Container Cloud Platform built on the foundations of Docker and Kubernetes. It is designed for individuals and organizations to easily build, manage, and operate their own private cloud environments with a focus on simplicity and application-centric management.

The Challenge: Complexity in the Cloud-Native Era

In the world of modern software, Kubernetes has become the standard for running applications. However, its power comes with significant complexity. Deploying even a seemingly simple application, like a WordPress blog, requires orchestrating a multitude of distinct Kubernetes resources:

- Deployments to manage the application pods (the WordPress server itself).
- Services to expose the application to the network.
- PersistentVolumeClaims to request storage for the database and uploads.
- StatefulSets to manage the database pods (like MySQL).
- ConfigMaps and Secrets to handle configuration and sensitive data.

Managing these individual components manually is often called a "resource-centric" approach. This approach presents several challenges:

1. **High Learning Curve:** Users must have a deep understanding of various Kubernetes resources and how they interact.
2. **Operational Burden:** Manually creating, updating, and deleting these

resources is tedious and prone to human error.

3. **Lack of Atomicity:** There is no way to treat the entire "WordPress application" as a single, atomic unit. You cannot simply "install" or "uninstall" it with one action.
4. **Inconsistency:** Ensuring that the application is deployed identically across development, testing, and production environments is difficult.

The Casibase Approach: From Managing Resources to Managing Applications

Casibase fundamentally simplifies this process by shifting the focus from managing individual resources to managing the application as a whole. We believe you should be able to manage your applications without getting lost in the weeds of Kubernetes YAML configurations.

To achieve this, Casibase introduces a powerful, application-centric model built on two core concepts:

1. Templates: The Reusable Blueprint

A Template in Casibase is a complete, reusable blueprint for an application. It encapsulates all the necessary Kubernetes resource manifests required to deploy a service. Think of it as a "package" for a cloud application.

- **What it contains:** A template holds the base YAML configurations for all the components of an application (Deployments, Services, etc.), structured for use with Kustomize.
- **The Goal:** To make the underlying complexity transparent. Once a template for WordPress is created, anyone can use it to deploy WordPress without needing to know the details of its Kubernetes architecture.

2. Applications: The Live Instance

An Application is a live, running instance created from a Template. It represents a specific deployment of that template in your cluster.

- **Customization:** When you create an Application, you select a Template and then provide your own specific configurations, such as the number of replicas, a custom domain name, or a specific database password.
- **How it works:** These custom configurations are treated as "patches" or "overlays." Casibase uses Kustomize to intelligently merge your custom parameters with the base manifests from the template, generating the final, complete configuration.
- **Lifecycle Management:** The Application becomes the single unit you interact with. You can deploy, monitor, update, and delete the entire application with single clicks in the UI.

By adopting this model, Casibase transforms the complex task of cloud-native deployment into a streamlined, intuitive workflow. Instead of wrestling with `kubectl` and YAML files, you can manage the entire lifecycle of your applications through a clean web interface: select a template, fill in a few parameters, and click deploy.

初学者指南

添加存储提供商

了解如何将存储提供商集成到Casibase中

添加AI模型提供商

学习如何添加模型提供商来增强Casibase功能

添加嵌入提供商

探索如何将嵌入提供商集成到Casibase中

添加语音合成提供商

学习如何添加语音合成提供商来增强Casibase功能。

添加语音识别提供商

了解如何添加语音识别以增强Casibase功能。

添加存储

学习如何向您的Casibase知识库系统添加存储

与AI聊天

在您的Casibase知识库系统中实现AI聊天功能

添加存储提供商

本文档是为初学者设计的分步教程。它将指导您完成将存储提供者与Casibase（我们强大的知识库系统）集成的过程。它将引导您完成将存储提供商与 Casibase 我们强大的知识库系统集成的过程。

简介

向Casibase添加存储提供商可以让您高效地管理和存储数据，这是您的知识库系统的重要组成部分。

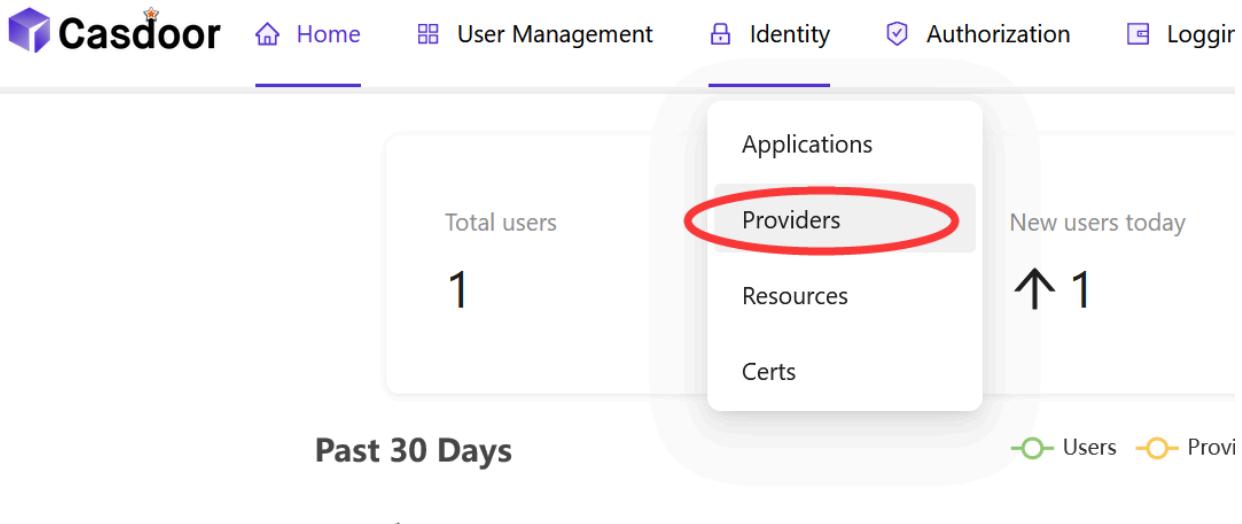
如果您是首次集成存储提供商，不必担心。我们已将整个过程分解为任何人都能遵循的简单步骤。

步骤1：部署Casdoor和Casibase

如果您还没有完成，请参考[部署Casdoor和Casibase](#)教程。

步骤2：添加新的存储提供商

存储提供用于存储数据。存储提供用于存储数据。可以通过点击主页上的 [Identity - Providers](#) 按钮在Casdoor中添加。



步骤2.1：添加存储提供商

点击 **Add** 按钮来添加存储提供商。

The screenshot shows the "Providers" management page. At the top, there are navigation links: Home, User Management, Identity, and Authorization. The "Identity" link is highlighted with a blue underline. Below the navigation, there's a table to manage providers:

Name	Organization	Created time	Di
provider_captcha_default	admin (Shared)	2023-09-10 19:31:50	Ca

A red circle highlights the "Add" button in the top-left corner of the table header.

步骤2.2：填写存储提供商信息

填写存储提供者信息并点击 **Save & Exit** 按钮。

New Provider [Save](#) [Save & Exit](#) [Cancel](#)

Name [?](#) : provider_storage_1

Display name [?](#) : Provider_storage_1

Organization [?](#) :

Category [?](#) : Storage

Type [?](#) : aws AWS S3

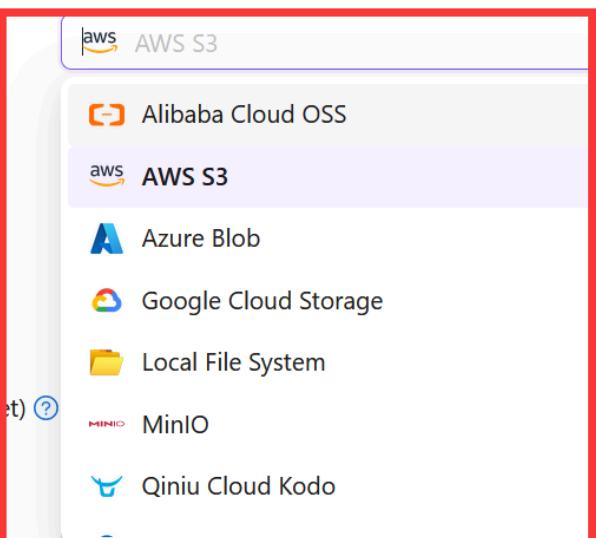
Client ID [?](#) :

Client secret [?](#) :

Endpoint [?](#) :

Endpoint (Intranet) [?](#) :

Bucket [?](#) :



Casdoor支持多种存储提供商，包括：

- [AWS S3](#)
- [Azure Blob](#)
- [Google Cloud Storage](#)
- [MinIO](#)

- [七牛云 Kodo](#)
- [阿里云OSS ...](#)

示例

添加阿里云OSS存储提供者



小心

- Client ID: 您的阿里云OSS账户的AccessKey ID。
- Client Secret: 您的阿里云OSS账户的AccessKey Secret。

*****是您的阿里云OSS账户信息的占位符。

Category [?](#) : Storage

Type [?](#) : Alibaba Cloud OSS

Client ID [?](#) : LTA***NLf

Client secret [?](#) : Vo6***pi8

Endpoint [?](#) : oss-cn-beijing.aliyuncs.com

Endpoint (Intranet) [?](#) :

Bucket [?](#) : xx-bucket-0

Path prefix [?](#) :

Domain [?](#) : https://xx-bucket-0.oss-cn-beijing.aliyuncs.com

Provider URL [?](#) : https://github.com/organizations/xxx/settings/applications/1234567

[Save](#) [Save & Exit](#) [Cancel](#)

步骤2.3：查看存储提供商

添加存储提供商后，您可以查看存储提供商信息。

Name	Organization	Created time	Display name	Category	Type	Client ID	Provider URL	Action
provider_storage_1	admin (Shared)	2023-09-10 21:23:02	Provider_storage_1	Storage	Alibaba Cloud OSS	[REDACTED]	https://github.com/organizations/xxx...	Edit Delete



提示

存储提供商来自Casdoor。 存储提供商来自Casdoor。 您可以在Casdoor中添加存储提供商，然后将其添加到Casibase中。

更多信息请参考步骤2：添加新的存储提供商。

The screenshot shows the Casbin web application interface. At the top, there is a navigation bar with links: Home, Chat, Stores (which is the active tab), Providers, Vectors, Chats, Messages, Tasks, and Help. Below the navigation bar, there is a form titled "Edit Store". The form has several input fields:

- Name: store_v6c22m
- Display name: New Store - v6c22m
- Storage provider: A dropdown menu with an option "Provider_storage_1 (provider_storage_1)" circled in red.
- Model provider: (empty)
- Embedding provider: (empty)
- File tree: (empty)

At the top left of the form, there is a "Save" button. The entire screenshot is framed by a green border.

存储示例

[Home](#)[Chat](#)[Stores](#)[Providers](#)[Vectors](#)[Chats](#)[M](#)[Edit Store](#)[Save](#)

Name:

my_store

Display name:

My_Store

Storage provider:

Provider_storage_1 (provider_storage_1)

Model provider:

Embedding provider:

File tree:

Save the configuration and navigate to the [Stores](#) page to see the file-tree of the storage provider.

The screenshot shows the casbin interface. On the left, a search results page displays a file tree under 'My_Store'. A red box highlights the search bar and the file tree. The tree includes categories like 'audio', 'document', and 'image', with specific files such as 'AC / DC - Highway To Hell.mp3', 'casdoor-knowledge.doc', and 'lena.jpg'. On the right, a sidebar features a 'New Chat' button, an AI icon, and a message input field.

Please input your search term

- My_Store
 - alibaba_oss
 - audio
 - AC / DC - Highway To Hell.mp3 (8.34 MB)
 - document
 - casdoor-knowledge.doc (18.0 KB)
 - casdoor-knowledge.docx (10.9 KB)
 - casdoor-knowledge.html (23.5 KB)
 - casdoor-knowledge.md (2.12 KB)
 - casdoor-knowledge.pdf (107 KB)
 - image
 - lena.jpg (105 KB)
 - lena.tiff (768 KB)
 - video
 - my_video.mkv (456 KB)

+ New Chat

AI

Type message here

现在您可以在Casibase中管理您的数据了。

在下一章中，我们将学习如何添加AI模型提供商到Casibase。

添加AI模型提供商

本文档是为初学者设计的分步教程。它将指导您完成将模型提供商与Casibase（我们强大的知识库系统）集成的过程。它将指导您完成将模型提供商与Casibase这套强大的知识库系统集成的过程。

简介

向Casibase添加模型提供商可以通过整合机器学习模型和AI功能来增强其功能。模型提供商允许您在知识库系统中分析和处理数据，使其更加智能和高效。模型提供商使您能够在知识库系统内分析和处理数据，从而使其更加智能高效。

如果您刚开始集成模型提供商，不用担心。我们已将该过程分解为任何人都能遵循的简单步骤。

步骤1：部署Casdoor和Casibase

在添加AI模型提供商之前，请确保已部署Casdoor和Casibase。在添加AI模型提供商之前，请确保您已经部署了Casdoor和Casibase。如果您还没有完成，请参考[部署Casdoor和Casibase教程](#)。

步骤2：添加新的模型提供商

模型提供商用于将LLM集成到Casibase中。您可以按照以下步骤添加：您可以按照以下步骤添加它们：

点击主页上的[Providers](#)按钮。



步骤2.1：添加模型提供商

点击 **Add** 按钮来添加模型提供商。

The table lists various model providers:

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUxnNltkKdIQyAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

步骤2.2：填写模型提供商详情

填写模型提供商详情并点击 **Save & Exit** 按钮。

[Home](#)[Chat](#)[Stores](#)[Providers](#)[Vectors](#)[Chats](#)[Me](#)[Edit Provider](#)[Save](#)

Name:

provider_openai_model

Display name:

OpenAI model

Category:

Model

Type:

OpenAI

Sub type:

text-davinci-003

Secret key:

Provider URL:

<https://platform.openai.com/account/api-keys>[Save](#)

提示

Casibase支持多种模型提供商，包括：

- [Hugging Face](#)
 - meta-llama/Llama-2-7b

- THUDM/chatglm2-6b
- baichuan-inc/Baichuan2-13B-chat
- gpt2
-
- OpenRouter
 - anthropic/claude-2
 - palm-2-chat-bison
 - palm-2-codechat-bison
 - openai/gpt-4
 -
- OpenAI
 - text-davinci-003
 - gpt-3.5-turbo
 - gpt-4
 -

⚠ 小心

- Category: 模型提供商的一级类别。例如, Model和Embedding。例如, Model和Embedding。
- 类型: 模型提供商的二级类别。例如, OpenAI和Hugging Face。
- SecretKey: 您的OpenAI账户的密钥。

示例

添加OpenAI模型提供商

The screenshot shows the 'Edit Provider' form on the casbin platform. A red circle highlights the 'Type' dropdown menu, which is currently set to 'OpenAI'. Other options in the dropdown include 'Hugging Face', 'OpenRouter', and 'Ernie'. The 'Save' button is visible at the top right of the form.

Field	Value
Name:	provider_openai_model
Display name:	OpenAI model
Category	Model
Type:	OpenAI
Sub type:	OpenAI
Secret key:	
Provider URL:	https://platform.openai.com/account/api-keys

⚠ 小心

某些模型不支持流式输出。已知支持流式输出的模型包括：已知支持流式输出的模型包括：

- gpt-3.5-turbo-0613

添加模型提供商后，您可以使用它来分析和处理Casibase中的数据，使用聊天机器人、问答等AI功能。

返回模型提供商列表页面：



The screenshot shows the Casibase interface with the 'Providers' tab selected. A single provider entry is listed in a table:

Name	Display name	Category	Type	Sub type	API key	Secret key	Provider URL	Action
provider_openai_model	OpenAI model	Model	OpenAI	text-davinci-003	***		https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>

A red box highlights the entire table row for the provider entry.

现在您已经添加了模型提供商，您可以使用它来分析和处理Casibase中的数据，使用聊天机器人、问答等AI功能。

在下一章中，我们将学习如何向Casibase添加嵌入提供商。

添加嵌入提供商

本文档是为初学者设计的分步教程。它将指导您完成将嵌入提供商与 Casibase（我们强大的知识库系统）集成的过程。它将引导您完成将嵌入提供商与 Casibase，我们强大的知识库系统集成的过程。

简介

嵌入是一种将单词和文档表示为向量的技术。嵌入提供商允许您在知识库系统中分析和处理数据，从而使系统更加智能高效。

有关嵌入的更多信息，请参考我们之前文档的[核心概念](#)部分。

在Casibase中，您可以按照以下步骤添加嵌入提供商：

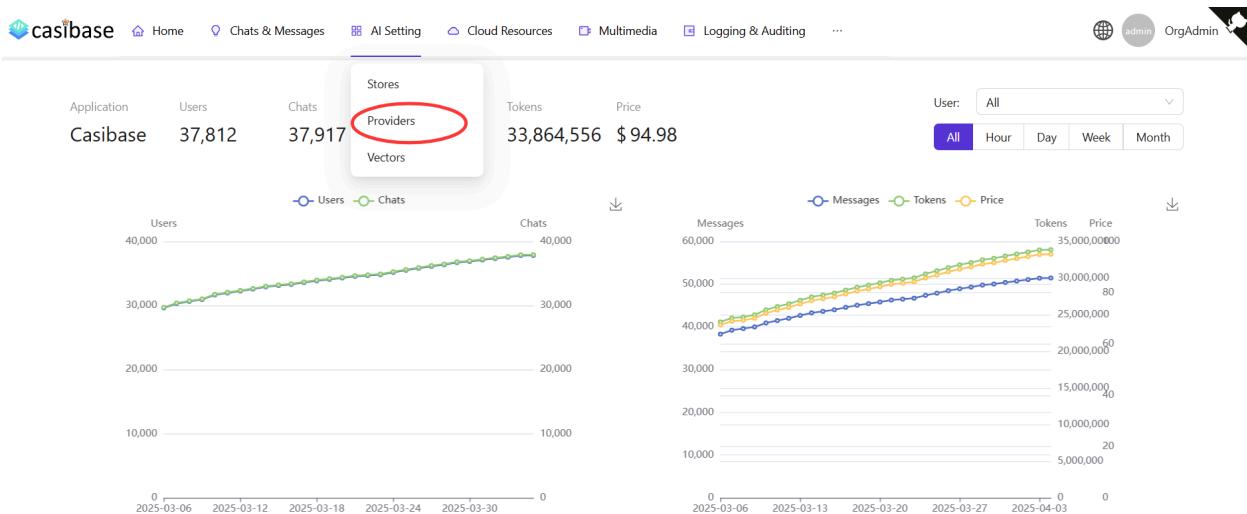
步骤1：部署Casdoor和Casibase

在您添加嵌入模型提供商之前，请确保已部署 Casdoor 和 Casibase。在添加嵌入模型提供商之前，请确保您已经部署了 Casdoor 和 Casibase。如果您还没有完成，请参考[部署 Casdoor 和 Casibase 教程](#)。

步骤2：添加新的嵌入提供商

嵌入提供商用于将嵌入功能集成到 Casibase 中。您可以按照以下步骤添加它们：

点击主页上的 [Providers](#) 按钮。



步骤2.1：添加嵌入提供商

点击 **Add** 按钮来添加嵌入提供商。

The table lists the following providers:

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUxnNltkKdIQyAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

步骤2.2：填写嵌入提供商详情

填写嵌入提供商详情并点击 **Save & Exit** 按钮。

[Home](#)[Chat](#)[Stores](#)[Providers](#)[Vectors](#)[CI](#)[Edit Provider](#)[Save](#)

Name:

embedding_openai_adasimilarity

Display name:

Embedding_OpenAI_AdaSimilarity

Category:

Embedding

Type:

OpenAI

Sub type:

AdaSimilarity

Secret key:

Provider URL:

<https://platform.openai.com/account/api-keys>[Save](#)

提示

与模型提供商部分相同， Casibase支持多种嵌入提供商，包括：

- [OpenAI](#)

- AdaSimilarity
- DavinciSimilarity
- AdaEmbedding2
-
- Hugging Face
 - sentence-transformers/paraphrase-MiniLM-L6-v2
 -

返回提供商列表页面：

Name	Display name	Category	Type	Sub type	API key	Secret key	Provider URL	Action
embedding_openai_adasimilarity	Embedding_OpenAI_AdaSimilarity	Embedding	OpenAI	1		***	https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>
model_openai_text_davinci_003	Model OpenAI text-davinci-003	Model	OpenAI	text-davinci-003		***	https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>

现在，您可以使用嵌入提供商将文本转换为向量。

添加嵌入提供商后，您可以使用它在Casibase中检索相似文档。有关更多信息，请参考我们之前文档的[核心概念](#)部分。有关更多信息，请参阅我们之前文档中的[核心概念](#)部分。

在下一章中，我们将学习如何将存储提供商、模型提供商和嵌入提供商与Casibase集成。

添加语音合成提供商

本文档是为初学者设计的分步教程。它将指导您完成将嵌入提供商与Casibase（我们强大的知识库系统）集成的过程。它将引导您完成将存储提供商与 Casibase 我们强大的知识库系统集成的过程。

简介

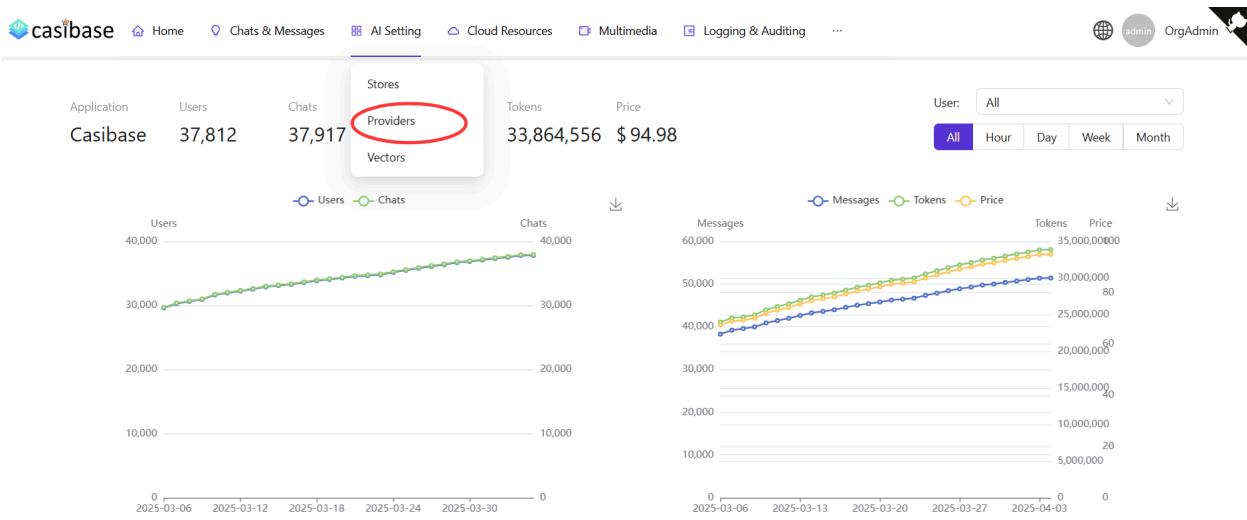
语音合成（TTS）是一项将文本转换为语音输出的技术。 TTS供应商允许您的Casibase 应用程序通过合成语音与用户交流，从而提升用户体验并增强知识库系统的可访问性。

在Casibase中，整合TTS提供商使您的AI应用能够口头回答查询，从而创造更具互动性和吸引力的用户体验。

添加新的语音合成提供商

文本转语音提供者用于将语音合成功能集成到Casibase中。 您可以按照以下步骤添加它们：

点击页面上的 **提供商** 按钮。



语音合成提供商

点击 **Add** 按钮来添加存储提供商。

The table lists various storage providers with the following columns:

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUxnNltkKdIQyAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

填写语音合成提供商的详细信息

填写嵌入提供商详情并点击 **保存并退出** 按钮。

casibase

Home Chats & Messages AI Setting Cloud Resources Multimedia Logging & Auditing ...

admin OrgAdmin

Edit Provider

Name: provider_tts_alibabacloud_cosyvoice

Display name: Provider TTS AlibabaCloud Cosyvoice

Category: Text-to-Speech

Type: Alibaba Cloud

Sub type: cosyvoice-v1

Flavor: longxiaochun

Secret key: ***

Provider URL: <https://bailian.console.aliyun.com/?apiKey=1#/api-key>

State: Active

Save Save & Exit

This screenshot shows the 'Edit Provider' page for a Text-to-Speech provider named 'provider_tts_alibabacloud_cosyvoice'. The provider is categorized as 'Text-to-Speech' and is of type 'Alibaba Cloud'. It uses the 'cosyvoice-v1' sub-type and flavor 'longxiaochun'. The secret key is set to '***'. The provider URL is listed as 'https://bailian.console.aliyun.com/?apiKey=1#/api-key'. The state is set to 'Active'. At the bottom, there are 'Save' and 'Save & Exit' buttons.

💡 提示

Casibase目前支持以下语音合成提供商:

- Alibaba Cloud
 - cosyvoice-v1 (具有多个语音选项)

Testing Your Text-to-Speech Provider

You can test your TTS provider by clicking the `Read it out` button. This will allow you to enter text and hear the synthesized speech output.

The screenshot shows the 'Edit Provider' page for a new provider named 'provider_r7fdnn'. The provider is categorized as 'Text-to-Speech' and is associated with the 'Alibaba Cloud' type and 'cosyvoice-v1' sub-type. A flavor note describes the voice as '龙小淳, 女, 中英双语。龙小淳的嗓音如丝般柔滑, 温暖中流淌着亲切与抚慰, 恰似春风吹过心田。'. The provider URL is set to <https://platform.openai.com/account/api-keys>. The provider is currently active. A red box highlights the 'Provider test:' input field containing 'Hello, I'm casibase AI.' and the 'Read it out' button.

casibase

Home Chats & Messages AI Setting Cloud Resources Multimedia Logging & Auditing ...

Edit Provider Save Save & Exit

Name: provider_r7fdnn

Display name: New Provider - r7fdnn

Category: Text-to-Speech

Type: Alibaba Cloud

Sub type: cosyvoice-v1

Flavor: 龙小淳, 女, 中英双语。龙小淳的嗓音如丝般柔滑, 温暖中流淌着亲切与抚慰, 恰似春风吹过心田。

Secret key: ***

Provider test: Hello, I'm casibase AI. Read it out

Provider URL: <https://platform.openai.com/account/api-keys>

State : Active

Save Save & Exit

This testing feature allows you to verify your TTS configuration before implementing it in your applications, ensuring the voice quality and settings meet your requirements.

Alibaba的语音选项

当使用 Alibaba Cloud 的 **cosyvoice-v1** 时, 您可以从各种语音选项中选择:

- 龙婉
- 龙橙
-

在存储中使用语音合成功能

添加语音合成提供商后, 您可以在商店设置中选择该提供商, 并决定是否启用TTS流式传输。

Edit Store

Name: store-built-in

Display name: Built-in Store

Title:

Avatar:

Storage provider: Built-in Storage Provider (provider-storage-built-in)

Image provider: Storage Aliyun OSS Casibase Casbin (provider_storage_casibase_casbin)

Split provider: Default

Model provider: Provider Model Azure GPT-4 (provider_model_azure_gpt4)

Embedding provider: Provider Embedding OpenAI V3 (provider_embedding_openai_v3)

Text-to-Speech provider: Provider TTS AlibabaCloud Cosyvoice (provider tts_alibabacloud_cosyvoice)

Enable TTS streaming:

Frequencies:

现在，您的商店可以将文本响应转换为语音，从而为用户提供更加互动的体验。

添加语音识别提供商

本文档是为初学者设计的分步教程。它将指导您完成将嵌入提供商与 Casibase（我们强大的知识库系统）集成的过程。它将引导您完成将存储提供商与 Casibase 我们强大的知识库系统集成的过程。

简介

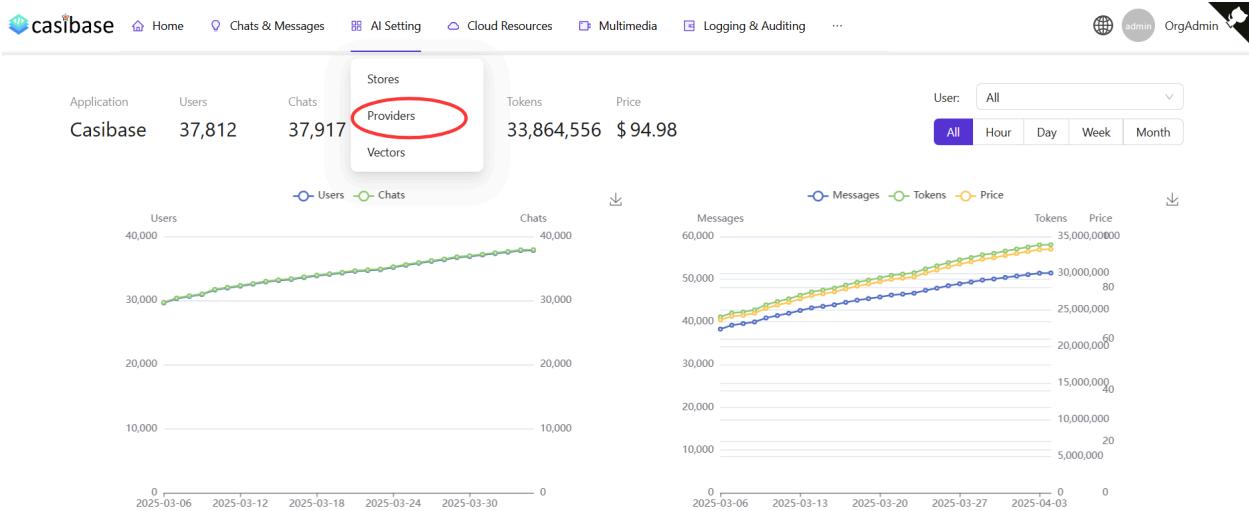
语音识别提供商 (STT) 是一种将口语转换成书面文字的技术。STT 提供商允许您的 Casibase 应用程序理解和处理用户口语输入，增强用户体验和您的知识库系统的访问能力。

在 Casibase 中，整合一个 STT 提供商使您的 AI 应用程序能够接收和处理语音查询，创建更多的互动和自然用户交互。

添加一个新的语音识别提供商

语音识别提供商用将语音识别功能集成到 Casibase 中。您可以按照以下步骤添加它们：

点击页面上的 [Providers](#) 按钮。



添加语音识别提供商

点击 **添加** 按钮来添加模型提供商。

The table lists the following providers:

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUxnNltkKdIQyAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

填写语音识别提供商详情

填写语音识别提供商详情并点击 **Save & Exit** 按钮。

Screenshot of the Casibase AI Setting interface showing the configuration of a new provider.

The provider details are as follows:

- Name: provider_njowpc
- Display name: New Provider - njowpc
- Category: Speech-to-Text
- Type: Alibaba Cloud
- Sub type: paraformer-realtime-v1
- Secret key: *** (highlighted with a red box)
- Provider URL: <https://platform.openai.com/account/api-keys>
- State: Active

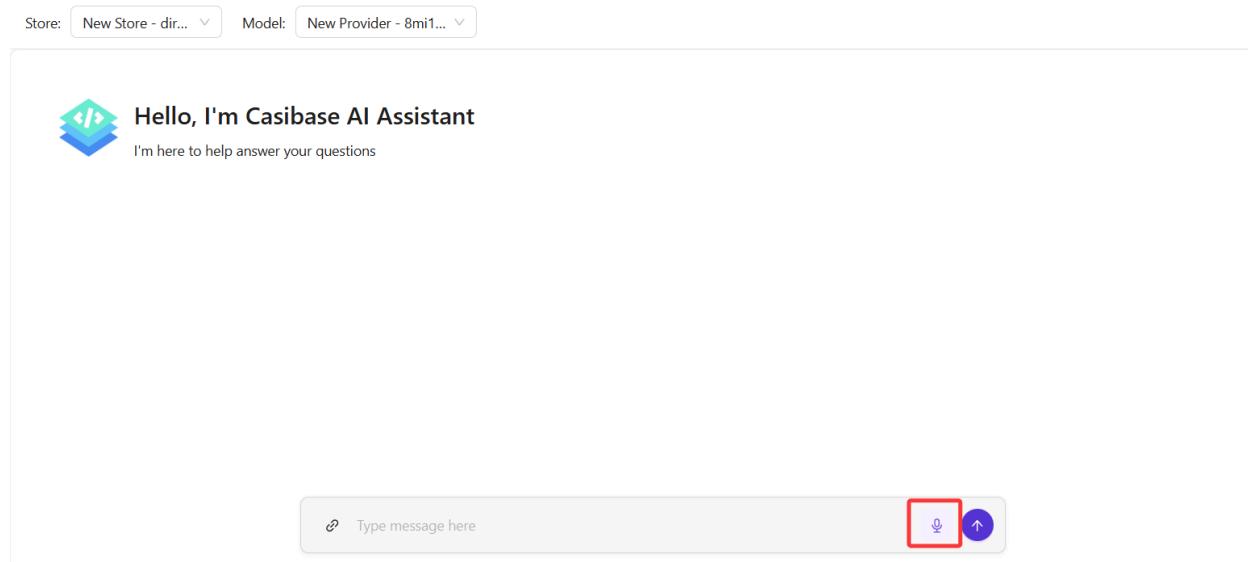
Buttons at the bottom: Save (gray), Save & Exit (purple).

Powered by  casibase

使用语音识别

当您在 Casibase 应用程序中点击语音识别按钮时，将发生以下过程：

1. 浏览器将请求访问您麦克风的权限
2. 一旦授予，系统将开始收听并自动将您的语音转换为文本
3. 在您完成发言后，识别的文本将自动作为消息发送



此功能允许与您的 Casisbase 应用程序进行无操作的互动，使它们更容易访问和使用。

 提示

CasiBase 目前支持以下语音识别提供商：

- [Alibaba Cloud](#)
 - paraformer-realtime-v1

添加存储

我们已经添加了存储提供商、模型提供商和嵌入提供商。现在我们需要配置一个存储来使用这些提供商。现在我们需要配置一个存储来使用这些提供商。

小心

本指南假定您已经部署了Casibase知识库系统。如果您还没有完成，请参考[部署 Casdoor](#)和[Casibase 教程](#)。

此外，本指南假定您已经添加了存储提供商、模型提供商和嵌入提供商。如果您没有，请关注[添加存储提供者](#), [添加一个 AI Model Provider](#), 和 [添加嵌入式提供商](#) 指南。

步骤1：添加新的存储

存储用于将存储、模型和嵌入提供商集成到Casibase中。您可以按照以下步骤添加：您可以按照以下步骤添加它们：

点击主页上的 **Stores** 按钮，然后点击 **Add** 按钮来添加存储。



Home Chat Stores Providers Vectors Chats Message

Stores	Add	Stores	Providers	Vectors	Chats	Message
Name		Display name		Storage provide		
my_store		My_Store		provider_storage		

步骤2：填写存储详情

选择您之前添加的存储提供商、模型提供商和嵌入提供商。

填写存储详情并点击 **Save & Exit** 按钮。

casbin

Home Chat Stores Providers Vectors Chats Messages Tasks Resources P

Edit Store Save

Name:	my_store
Display name:	My_Store
Storage provider:	Provider_storage_1 (provider_storage_1)
Model provider:	Model OpenAI text-davinci-003 (model_openai_text_davinci_003)
Embedding provider:	Embedding_OpenAI_Adasimilarity (embedding_openai_adasimilarity)
File tree:	<pre> └─ My_Store ├─ alibaba_oss │ ├─ audio │ │ └─ AC / DC - Highway To Hell.mp3 (8.34 MB) │ ├─ document │ │ └─ casdoor-knowledge.doc (18.0 KB) │ │ └─ casdoor-knowledge.docx (10.9 KB) │ └─ casdoor-knowledge.html (23.5 KB) └─ image └─ lena.jpg (105 KB) └─ lena.tiff (768 KB) └─ video └─ my_video.mkv (456 KB) </pre>

点击 **Save & Exit** 按钮并返回存储列表页面：

casbin

Home Chat Stores Providers Vectors Chats Messages Tasks Resources Permissions Logs

Jimmy

Stores Add

Name	Display name	Storage provider	Model provider	Embedding provider	Action
my_store	My_Store	provider_storage_1	model_openai_text_davinci_003	embedding_openai_adasimilarity	View Refresh Vectors Edit Delete

< 1 >

现在，您可以使用存储来存储知识库数据、将文本转换为向量，并与聊天机器人对话。

在下一节中，我们将学习如何在Casibase中与聊天机器人对话。

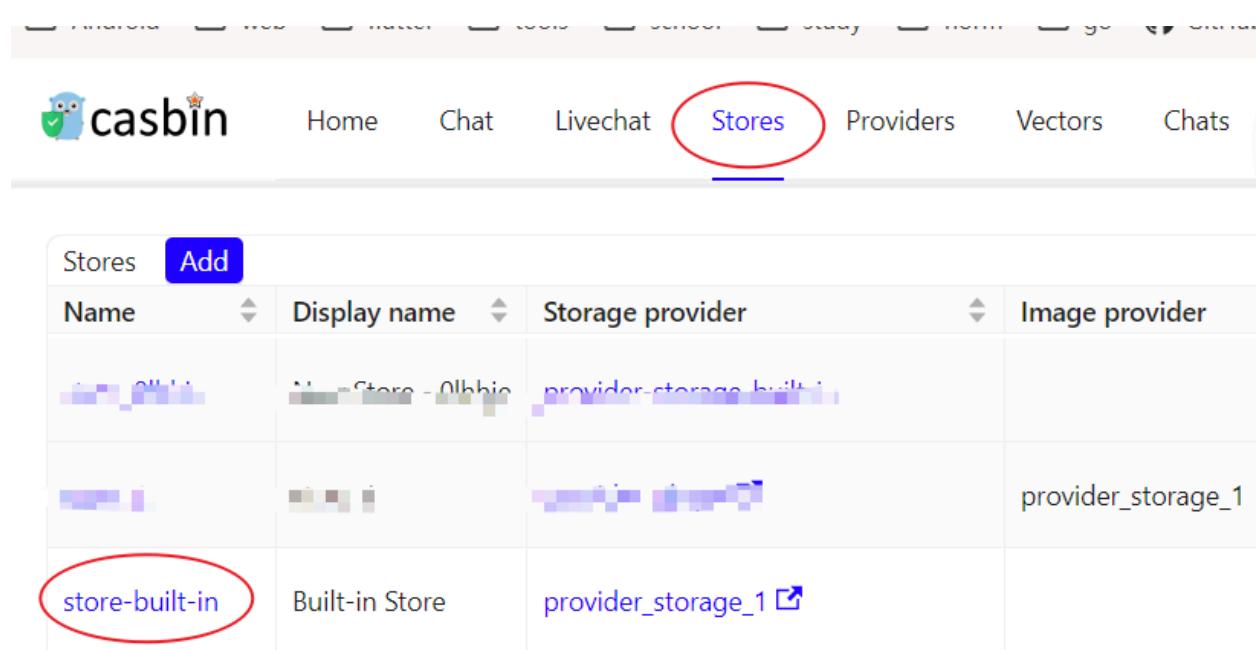
支持多存储

多存储模式为用户在每个不同的存储中提供不同的模型、建议等。

步骤1：启用多存储

首先，您应该在内置存储中启用多存储模式。

点击主页上的 Stores 按钮，然后点击 store-built-in 按钮进入内置存储。



The screenshot shows the Casibase interface with the 'Stores' tab selected. A red circle highlights the 'store-built-in' entry in the table below. The table has columns: Name, Display name, Storage provider, and Image provider. The 'store-built-in' entry has a blue border around its 'Name' field.

Name	Display name	Storage provider	Image provider
store-built-in	Built-in Store	provider_storage_1	

向下滚动找到 Can Select Store 字段，勾选它。



步骤2：添加可用存储

多存储模式仅提供可用的存储。要使存储可用，您需要配置其存储提供商、模型提供商和嵌入提供商。

步骤3：为对话选择存储

Casibase提供了一种非常方便的选择存储的方法。



Home Chat Livechat Stores Providers Vectors Chats Messages Usages Frameworks

The screenshot shows the casbin Chat interface. At the top, there's a navigation bar with links: Home, Chat (which is highlighted in blue), Livechat, Stores, Providers, Vectors, Chats, Messages, Usages, and Frameworks. Below the navigation bar, there's a purple header bar with the text "New Chat - 7". On the left side, there's a sidebar with a red border containing a button labeled "+ New Chat". Below this button, there's a list of storage options: "store_1" and "store-built-in". To the right of the sidebar, the main chat area displays a message from a user: "You are an expert in your field". Below this message, there's a response from a bot: "Thank you for recognizing my expertise. Whether it's related to my specific area of knowledge or expertise to provide insightful answers and solutions to any problems you may have." At the bottom of the main chat area, there are five small blue icons representing different actions: a square, a circle, a triangle pointing up, a triangle pointing down, and a triangle pointing right.

只需将鼠标悬停在"New Chat"上，然后您就可以从下面出现的列表中选择您想要使用的存储。

如果您点击"New Chat"按钮，系统将为您分配一个默认存储。

与AI聊天

本文档是为初学者设计的分步教程。它将指导您完成将嵌入提供商与Casibase（我们强大的知识库系统）集成的过程。本文档是为初学者设计的分步教程。它将指导您完成在Casibase知识库系统中实现AI聊天功能的过程。

简介

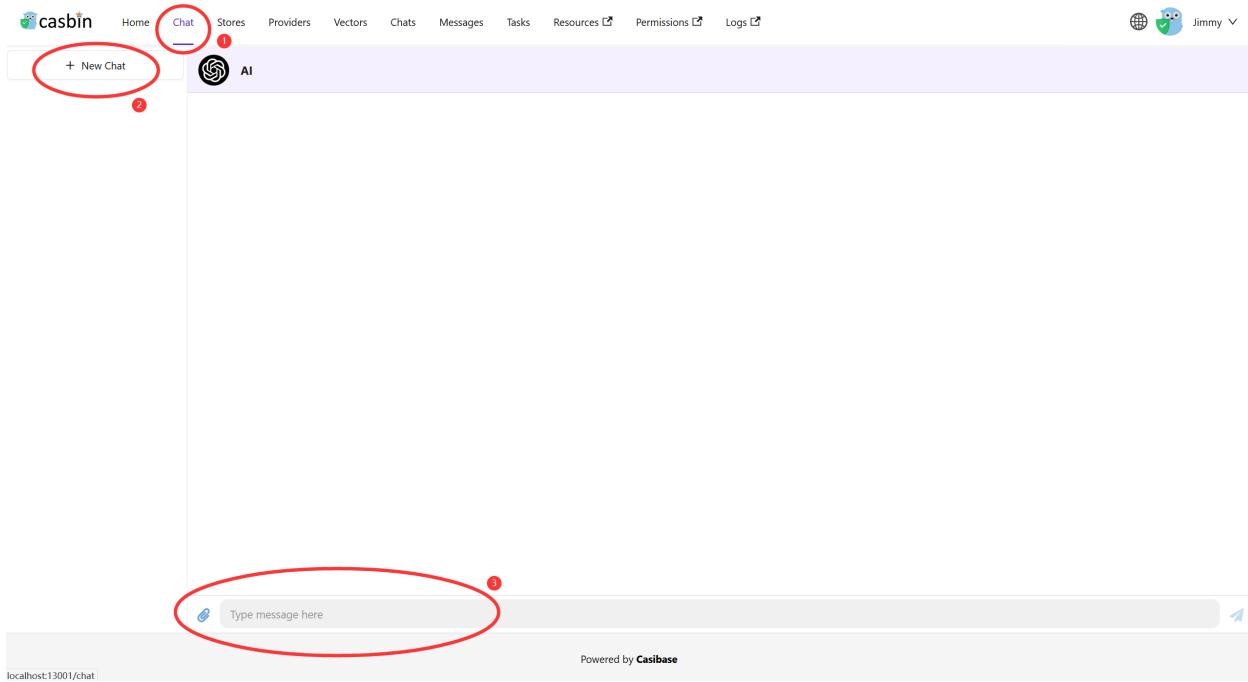
在之前的部分中，我们已经部署了Casdoor和Casibase，将存储提供商、模型提供商和嵌入提供商集成到Casibase中，并添加了一个存储来使用这些提供商。

有关存储的更多信息，请参考我们之前文档的[添加存储](#)部分。

现在，让我们在Casibase中实现AI聊天功能。

Step 1: Access the Chat Interface

When you log in to Casibase, you'll land directly on the chat interface. This gives you immediate access to start conversations with AI. If you need to create a new chat session, click the [New Chat](#) button.



步骤2：发送消息

写一条消息并点击 Send 按钮来发送。



步骤3：知识库聊天

此外，您还可以在知识库中与聊天机器人对话。

知识库聊天有一些要求：

- 知识库必须有一个存储。
- 存储必须有一个模型提供商。
- 存储必须有一个嵌入提供商。

- 存储必须有一个存储提供商。
- 存储提供商必须有一个可读的文档（例如markdown文件、docx文件和pdf文件）。

一旦您满足了这些要求，您可以返回到 Stores 页面并点击 Refresh Vectors 按钮来嵌入知识库数据。

Name	Display name	Storage provider	Model provider	Embedding provider	Action
my_store	My_Store	provider_storage_1	model_openai_text_davinci_003	embedding_openai_adasimilarity	View Refresh Vectors (disabled)

当嵌入正在进行时，按钮将被禁用。

嵌入完成后，您可以点击导航栏中的 Vectors 按钮来查看向量。

结果：

Name	Display name	Store	File	Text	Data	Action
vector_7rss8s	Simplified development	my_store	alibaba_oss/document/casdoor-knowledge.pdf	Simplified development: Casdoor pro...	[{"-0.000106310275,0.02166452,0.02304..."]	Edit Delete
vector_gldg4u	Installation and Deployment: You can	my_store	alibaba_oss/document/casdoor-knowledge.pdf	Installation and Deployment: You ca...	[{"-0.002990207,0.018568026,-0.00580..."]	Edit Delete
vector_Owrasj	Privilege Control: With Casdoor	my_store	alibaba_oss/document/casdoor-knowledge.pdf	Privilege Control: With Casdoor, de...	[{"0.0054717776,0.017982274,0.0103428..."]	Edit Delete
vector_3tet51	Casdoor Knowledge Points	my_store	alibaba_oss/document/casdoor-knowledge.pdf	Casdoor Knowledge Points Casdoor is...	[{"-0.007692282,0.024387684,0.0001651..."]	Edit Delete

让我们在知识库中与聊天机器人对话。

AI

What's casdoor?

Answer: Casdoor is an OAuth2 and OIDC based authentication portal designed to help developers easily add user authentication and authorization features to their applications.

When the AI responds using your knowledge base, you'll see a "Knowledge sources" button showing how many document fragments were referenced. Click it to view each source with its relevance score, and click any source to navigate

directly to that document in your store.

与非知识库聊天的结果比较：

Casdoor is an online development platform for web and mobile applications that helps companies create and manage an end-to-end development process. It allows businesses to build, deploy, monitor, and manage apps quickly and cost effectively. It is equipped with powerful analytics tools to track the performance and usage of an app, as well as allowing companies to quickly identify and address any issues or problems.

⚠ 小心

嵌入速率与两个因素有关：

- 知识库中的文档：
 - 文档数量：文档越多，嵌入时间越长。
 - 文档大小：文档大小越大，嵌入时间越长。
- 嵌入提供商：
 - API速率限制：API速率限制越高，嵌入速度越快。
 - API并发：API并发越高，嵌入速度越快。

例如，如果您使用[OpenAI API](#)作为嵌入提供商，嵌入速率与[OpenAI API](#)的速率限制和并发有关。

结论

在本指南中，我们学习了如何在Casibase中实现AI聊天功能。

现在，您可以在Casibase中与聊天机器人对话了。尽情享受吧！好好享受！

有关Casibase的更多信息，可以在我门文档的[核心概念](#)部分找到。

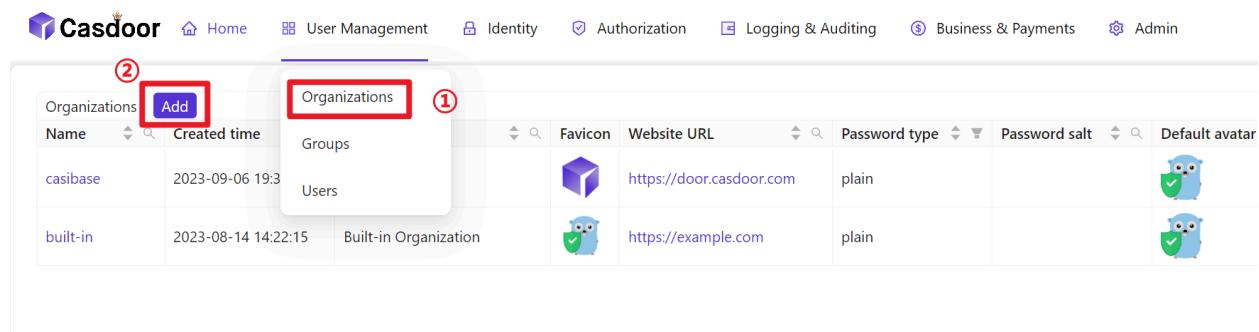
Casdoor-SSO

Casibase 使用 Casdoor 作为其身份认证和单点登录(SSO)提供商。请确保提前部署好 Casdoor。请确保提前部署它。

请参考 [Casdoor 服务器安装指南](#) 来安装和配置 Casdoor。

按照以下步骤为 Casibase 设置 Casdoor:

- 创建组织



Name	Created time	Favicon	Website URL	Password type	Password salt	Default avatar
casibase	2023-09-06 19:33		https://door.casdoor.com	plain		
built-in	2023-08-14 14:22:15		https://example.com	plain		

- 配置组织信息

casibase

Casibase

Save & Exit

- 创建新应用

Name	Created time	Display name	Providers	Organization	Providers
app-casibase	2023-09-06 19:38:54	Casibase		casibase	provider_captcha_default
app-built-in	2023-08-14 14:22:15	Casdoor		built-in	provider_captcha_default

- 配置应用信息（请记住应用名称、ClientID 和 ClientSecret）

The screenshot shows the 'Edit Application' form in the Casdoor dashboard. The 'Name' field is set to 'app-casibase' (1). The 'Display name' field is set to 'Casibase'. The 'Organization' field is set to 'casibase' (2). The 'Client ID' field is set to '548c8b9c7431d2621db1' (3). The 'Client secret' field contains the value '2bc7640d487fc4dea6f4b77f07f1bf4433e4ad40' (4). Other fields like 'Logo', 'URL', and 'Cert' are also visible.

- Create a Certificate: In the Casdoor dashboard, choose Cert → Add, keep Algorithm as RS256 (default), enter a name, and click Save.

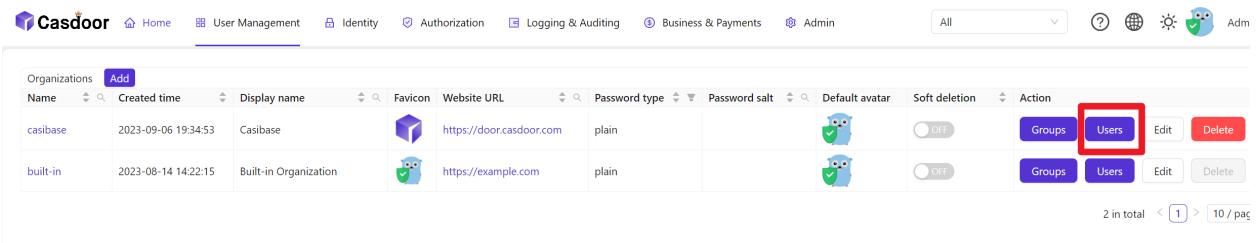
The screenshot shows the 'New Cert' form in the Casdoor dashboard. The 'Name' field is set to 'cert_casbin' (1). The 'Display name' field is set to 'New Cert - casbin'. The 'Scope' field is set to 'JWT'. The 'Type' field is set to 'x509'. The 'Crypto algorithm' field is set to 'RS256 (RSA + SHA256)' (2). The 'Bit size' field is set to '4096'. The 'Exire in years' field is set to '20'. The 'Certificate' section shows the certificate content starting with '-----BEGIN CERTIFICATE-----' and ends with '-----END RSA PRIVATE KEY-----'. The 'Cert' field at the bottom is set to 'cert_casbin'.

- Bind the Certificate to the Application: Open the *Config* tab of your newly

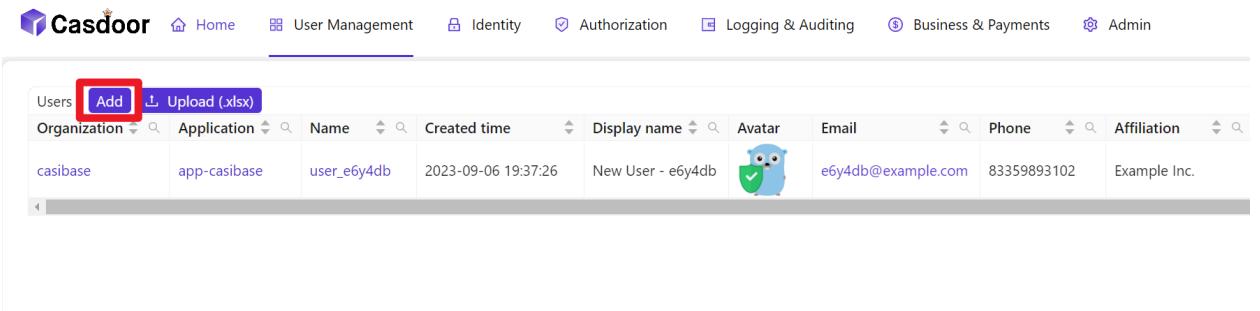
created Application, select the certificate you just created from the Cert dropdown, and click Save.



- 在新创建的组织中添加成员



Name	Created time	Display name	Favicon	Website URL	Password type	Password salt	Default avatar	Soft deletion	Action
casibase	2023-09-06 19:34:53	Casibase		https://door.casdoor.com	plain			<input checked="" type="radio"/>	Groups Users Edit Delete
built-in	2023-08-14 14:22:15	Built-in Organization		https://example.com	plain			<input checked="" type="radio"/>	Groups Users Edit Delete



Organization	Application	Name	Created time	Display name	Avatar	Email	Phone	Affiliation
casibase	app-casibase	user_e6y4db	2023-09-06 19:37:26	New User - e6y4db		e6y4db@example.com	83359893102	Example Inc.

- 配置成员信息（请记住用户名和密码）

The screenshot shows the Casdoor User Management interface for editing a user named 'user_e6y4db'. The interface includes fields for Name, Display name, Avatar, User type, Password, Email, Phone, and various profile details like Bio, Tag, Language, Gender, Birthday, Education, Score, Karma, and Ranking. A dropdown menu for 'Signup application' is open, showing 'app-casibase' selected. The 'Save & Exit' button is highlighted with a red box and the number ④.

Key elements highlighted with red boxes:

- ① Name: user_e6y4db
- ② Password: Modify password...
- ③ Signup application: app-casibase

User Data Synchronization

Casibase automatically keeps user information synchronized with Casdoor. When a user's account is accessed through the application, Casibase fetches the latest user data directly from Casdoor rather than relying on cached session data. This ensures that any updates to user profiles in Casdoor—such as changes to name, email, avatar, or permissions—are immediately reflected in Casibase without requiring users to log out and log back in.

This real-time synchronization happens transparently in the background whenever the application needs to access the current user's account information.

Billing Integration

Casibase integrates with Casdoor's transaction system for AI usage billing. When users interact with AI models, transaction records are automatically created in Casdoor to track token consumption and costs. This enables centralized billing management across all your applications. For details on how transactions work, see the [Billing & Usage](#) section.



>

Permissions

Permissions



Permissions Overview

Understanding Casibase's permission system



Frontend Permissions

Frontend permission utilities in Casibase



Backend Permissions

Backend permission utilities in Casibase

Permissions Overview

Casibase uses a role-based permission system to control access to features and resources. The system distinguishes between different user types, each with specific capabilities and access levels.

User Roles

Casibase supports several user roles:

System Admin - Full administrative access to all features, including system configuration, user management, and all resources. This is the highest privilege level.

Chat Admin - Administrative access focused on chat and knowledge base operations. Can manage stores, vectors, chats, activities, and view all users' data.

Chat User - Regular users who can interact with AI chat, access their assigned stores, and view their own usage data.

Video Normal User - Users with restricted video permissions. Can only modify videos in "Draft" state when no remarks are present.

Permission Hierarchy

The permission system follows this hierarchy (from highest to lowest):

1. System Admin (built-in admin or users with `isAdmin` flag)

2. Chat Admin (users with `type: "chat-admin"`)
3. Chat User (users with `type: "chat-user"`)
4. Anonymous/Guest users (limited read-only access)

Higher-level roles inherit the permissions of lower-level roles.

How Permissions Work

Permissions are checked throughout the application:

Frontend checks determine what UI elements users see and which client-side operations they can perform. These checks prevent unauthorized access to admin panels, configuration pages, and sensitive features.

Backend checks enforce permissions on API endpoints and data operations. Even if frontend checks are bypassed, backend validation ensures security at the server level.

This dual-layer approach provides defense in depth - the frontend offers a smooth user experience by hiding inaccessible features, while the backend guarantees security enforcement.

Authentication Integration

Casibase integrates with [Casdoor](#) for user authentication and SSO. User roles and permissions are managed through Casdoor and synchronized automatically with Casibase.

Implementation

Casibase centralizes permission logic to maintain consistency and simplify updates:

- Frontend utilities in `Setting.js` provide reusable permission checking functions
- Backend utilities in `util/permission.go` handle server-side authorization

For implementation details, see [Frontend Permissions](#) and [Backend Permissions](#).

Frontend Permissions

Casibase's React frontend uses centralized utility functions to check user permissions. These functions are defined in `web/src/Setting.js` and ensure consistent permission checking across all components.

Permission Utility Functions

`isAdminUser(account)`

Checks if a user is a system administrator.

```
if (Setting.isAdminUser(account)) {  
    // Show admin-only features  
}
```

Returns `true` for:

- Built-in admin account (`owner === "built-in"`)
- Users with the `isAdmin` flag set to true

`isChatAdminUser(account)`

Checks if a user has chat-admin privileges.

```
if (Setting.isChatAdminUser(account)) {  
    // Show chat admin features
```

Returns `true` for users with `type === "chat-admin"`.

canViewAllUsers(account)

Checks if a user can view data for all users (not just their own).

```
if (Setting.canViewAllUsers(account)) {  
    // Enable "All users" option in dropdown  
}
```

Returns `true` for:

- The admin account (`name === "admin"`)
- Chat admin users

This function is commonly used in usage pages and activity logs where admins need to see organization-wide data.

isLocalAdminUser(account)

Checks if a user has local admin privileges, which includes both system admins and chat admins.

```
const canManagePanes = Setting.isLocalAdminUser(account);
```

Returns `true` for:

- System admins
- Chat admin users

isLocalAndStoreAdminUser(account)

Similar to `isLocalAdminUser()` but also ensures the user can manage store content (not restricted to "non-store-admin" homepage).

```
if (Setting.isLocalAndStoreAdminUser(account)) {  
    // Allow file operations  
}
```

Practical Examples

Restricting Menu Items

```
// App.js - Building navigation menu  
if (!Setting.isAdminUser(this.state.account)) {  
    if (!Setting.isChatAdminUser(this.state.account)) {  
        return res; // Return limited menu for regular users  
    }  
}
```

Conditional UI Elements

```
// UsagePage.js - User selection dropdown  
<option key="all" value="All"  
       disabled={!Setting.canViewAllUsers(this.props.account)}>  
    All  
</option>
```

Permission-Based Features

```
// ProviderEditPage.js - Sensitive field access
<Input.Password
  value={this.state.provider.providerKey}
  disabled={!Setting.isAdminUser(this.props.account)}
  onChange={e => this.updateProviderField("providerKey",
e.target.value)}
/>
```

Benefits of Centralized Functions

Before these utilities existed, permission checks were scattered across components using inline conditions like `account.name === "admin" || account.type === "chat-admin"`. This made the code:

- **Fragile:** Changing role logic required updating multiple files
- **Inconsistent:** Different components might check permissions differently
- **Error-prone:** Easy to miss edge cases or make typos

Centralized utilities solve these issues by providing a single source of truth. Updates to permission logic happen in one place and automatically apply throughout the application.

Implementation Notes

All utility functions perform null/undefined checks on the account parameter, making them safe to call without prior validation. They return `false` for invalid or missing accounts, following a secure-by-default approach.

Backend Permissions

Casibase's Go backend centralizes permission checking through utility functions in `util/permission.go`. These functions provide consistent role validation across controllers, routers, and service layers.

Permission Functions

`IsAdminOrChatAdmin(user)`

Checks if a user is either a system admin or a chat admin.

```
if !util.IsAdminOrChatAdmin(user) {
    c.ResponseError("this operation requires admin privilege")
    return
}
```

Returns `true` for:

- Users with `IsAdmin` flag set
- Users with `Type == "chat-admin"`

This is the most commonly used permission check for operations that require elevated privileges.

`IsChatAdmin(user)`

Checks specifically for chat admin role.

```
if util.IsChatAdmin(user) {
    // Grant chat-admin specific permissions
}
```

Returns `true` only for users with `Type == "chat-admin"`.

IsVideoNormalUser(user)

Checks if a user has the video-normal-user role, which has restricted video editing permissions.

```
if util.IsVideoNormalUser(user) {
    if len(video.Remarks) > 0 || video.State != "Draft" {
        c.ResponseError("video can only be updated in Draft state")
        return
    }
}
```

Returns `true` for users with `Type == "video-normal-user"`.

Role Type Constants

The package defines constants for role types to eliminate hardcoded strings:

```
const (
    UserTypeChatAdmin      = "chat-admin"
    UserTypeVideoNormalUser = "video-normal-user"
)
```

These constants ensure type safety and prevent typos when checking user roles.

Usage in Controllers

Authorization Filter

The router middleware uses permission checks to protect admin endpoints:

```
// routers/authz_filter.go
func permissionFilter(ctx *context.Context) {
    user := GetSessionUser(ctx)

    if !util.IsAdminOrChatAdmin(user) {
        responseError(ctx, "this operation requires admin
privilege")
        return
    }
}
```

API Controllers

Controllers use these functions to enforce business logic permissions:

```
// controllers/account.go - User type assignment
if strings.Count(claims.Type, "-") <= 1 {
    if !util.IsAdminOrChatAdmin(&claims.User) {
        claims.Type = "chat-user"
    }
}

// controllers/video.go - Video update restrictions
if util.IsVideoNormalUser(user) {
    if len(video.Remarks) > 0 || video.State != "Draft" {
```

Service Layer

Object and service layers also benefit from centralized checks:

```
// object/util.go
func isAdmin(user *casdoorsdk.User) bool {
    return util.IsAdminOrChatAdmin(user)
}
```

Migration Impact

Before centralization, permission checks were duplicated across the codebase:

```
// Before: inline, repeated logic
isAdmin := user != nil && (user.IsAdmin || user.Type == "chat-
admin")
if !isAdmin {
    return error
}

// After: single, reusable function
if !util.IsAdminOrChatAdmin(user) {
    return error
}
```

The migration improved:

Maintainability: Permission logic changes happen in one place
Consistency: All parts of the system use identical checks
Security: Reduced risk of permission check errors

Safety Features

All permission functions include null safety checks:

```
func IsAdminOrChatAdmin(user *casdoorsdk.User) bool {
    if user == nil {
        return false
    }
    return user.IsAdmin || user.Type == UserTypeChatAdmin
}
```

This prevents nil pointer dereferences and ensures permission checks fail safely when user data is unavailable.



>

Developer Guide

Developer Guide



Generating Swagger Files

Generating Swagger Files

Generating Swagger Files

Overview

As we know, the beego framework provides support for generating swagger files to clarify the API via the command line tool called "bee". Casibase is also built based on beego. However, we found that the swagger files generated by bee failed to categorize the APIs with the "@Tag" label. So, we modified the original bee to implement this function.

How to write the comment

Most rules are exactly identical to the original bee comment formats. The only discrepancy is that the API shall be divided into different groups according to the "@Tag" label. Therefore, developers are obliged to ensure that this tag is correctly added. Here is an example:

```
// @Title Login
// @Tag Login API
// @Description login
// @Param oAuthParams     query    string  true      "oAuth
parameters"
// @Param body    body    RequestForm  true      "Login
information"
// @Success 200 {object} controllers.api_controller.Response The
Response object
// @router /login [post]
func (c *ApiController) Login() {
```

APIs with the same "@Tag" labels will be put into the same group.

How to generate the swagger file

0. Write comments for the API in the correct format.
1. Fetch this repository: <https://github.com/casbin/bee>.
2. Build the modified bee. For example, in the root directory of casbin/bee, run the following command:

```
go build -o mybee .
```

3. Copy mybee to the base directory of Casibase.
4. In that directory, run the following command:

```
mybee generate docs
```

5. (Optional) If you want to generate swagger document for specific tags or apis, here are some example commands:

```
mybee generate docs --tags "Adapter API"  
mybee generate docs --tags "Adapter API,Login API"  
mybee generate docs --apis "add-adapter"  
mybee generate docs --apis "add-adapter,delete-adapter"
```

Notably: We only accept a comma  as the separator when multiple tags/apis provided.

Then you will find that the new swagger files are generated.



>

部署

部署



部署 Casdoor 和 Casibase

了解如何部署 Casdoor 和 Casibase。

部署 Casdoor 和 Casibase

简介



提示

什么是 Casdoor?

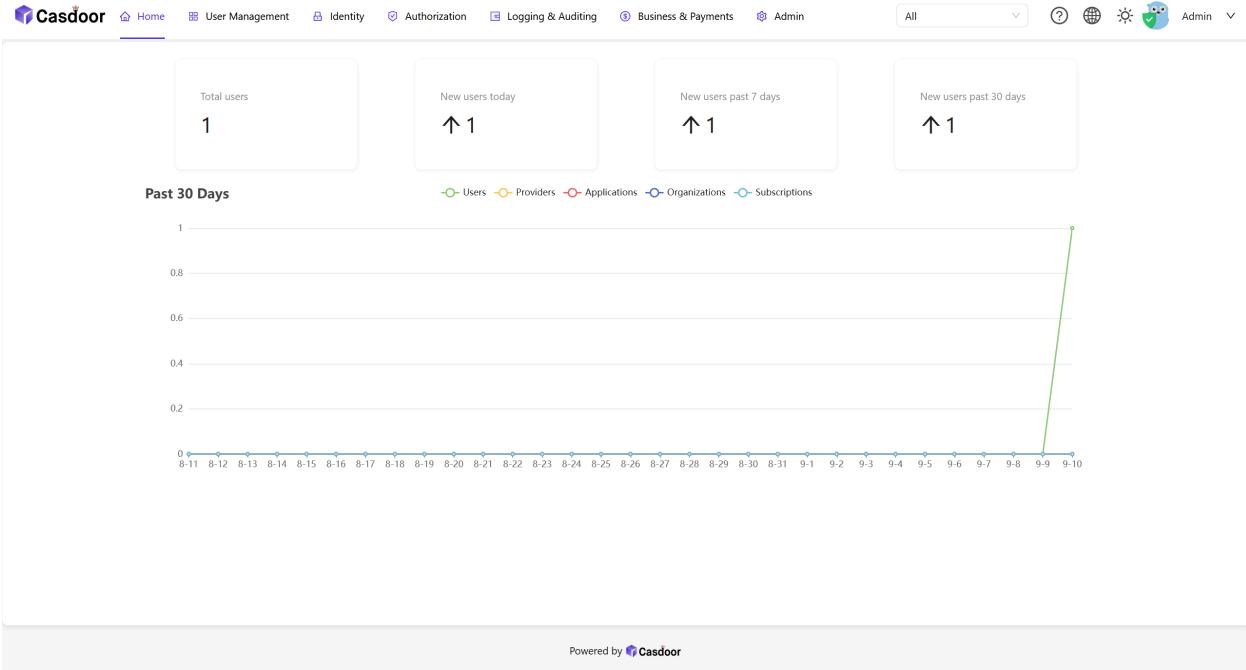
Casdoor 是一个强大的身份认证系统，提供安全可靠的登录体验。它是 Casibase 的前置要求，所以请确保先部署它。由于 Casibase 是前提条件，所以请务必先部署它。

访问 [Casdoor](#) 网站了解更多信息。

步骤 1：部署 Casdoor

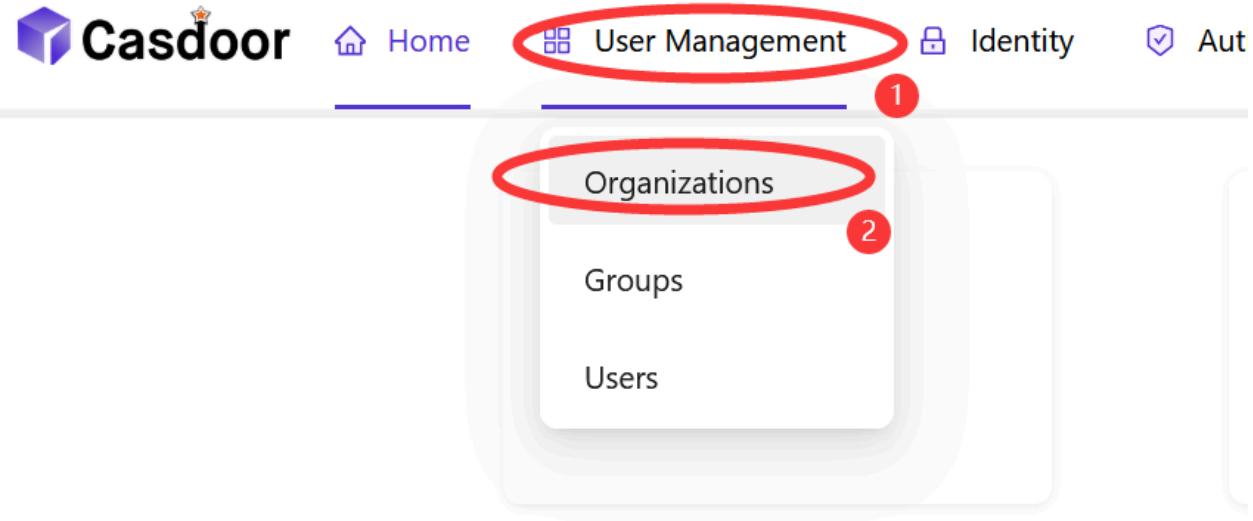
在 [Casdoor 部署指南](#) 中，你可以找到部署 Casdoor 的详细步骤。

部署完成 Casdoor 后，你将会看到如下效果：



步骤 2：在 Casdoor 中创建组织

在 Casdoor 中，你可以创建一个组织来管理你的用户和应用。你可以通过点击主页上的 **用户管理 - 组织** 按钮来创建组织。你可以通过点击主页上 **用户管理 - 组织** 按钮来创建一个组织。



Past 30 Days

步骤 2.1: 添加组织

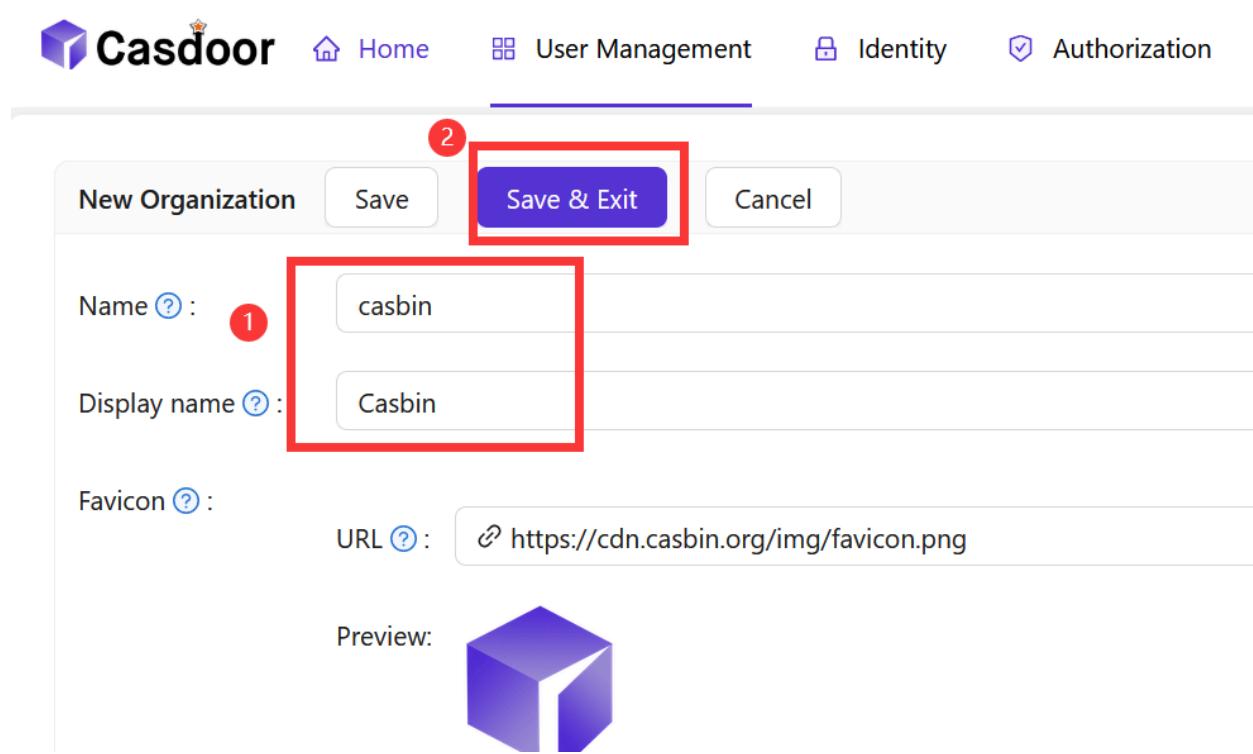
点击 **添加** 按钮来添加一个组织。

The screenshot shows the 'Organizations' management page. At the top, there is a search bar labeled 'Organizations' and a blue 'Add' button, both of which are circled in red. Below the search bar is a table with three columns: 'Name', 'Created time', and 'Display name'. A single row is visible, showing 'built-in' in the Name column, '2023-09-10 19:31:50' in the Created time column, and 'Built-in Organization' in the Display name column.

Name	Created time	Display name
built-in	2023-09-10 19:31:50	Built-in Organization

步骤 2.2：填写组织信息

填写组织信息并点击 **保存并退出** 按钮。



New Organization

Name ②: casbin

Display name ①: Casbin

Save Save & Exit Cancel

Favicon ②:

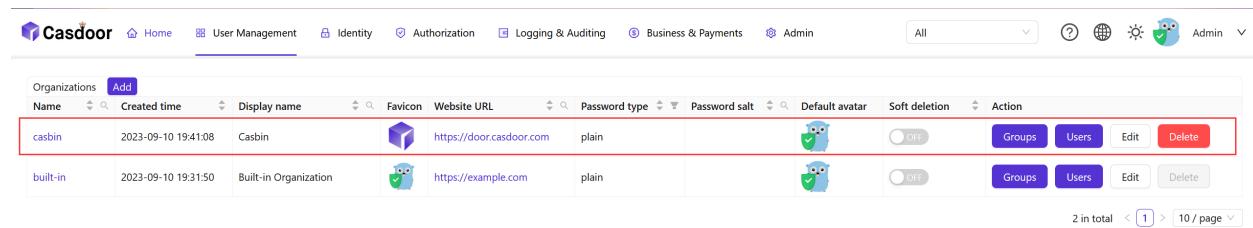
URL ①: <https://cdn.casbin.org/img/favicon.png>

Preview:



步骤 2.3：查看组织

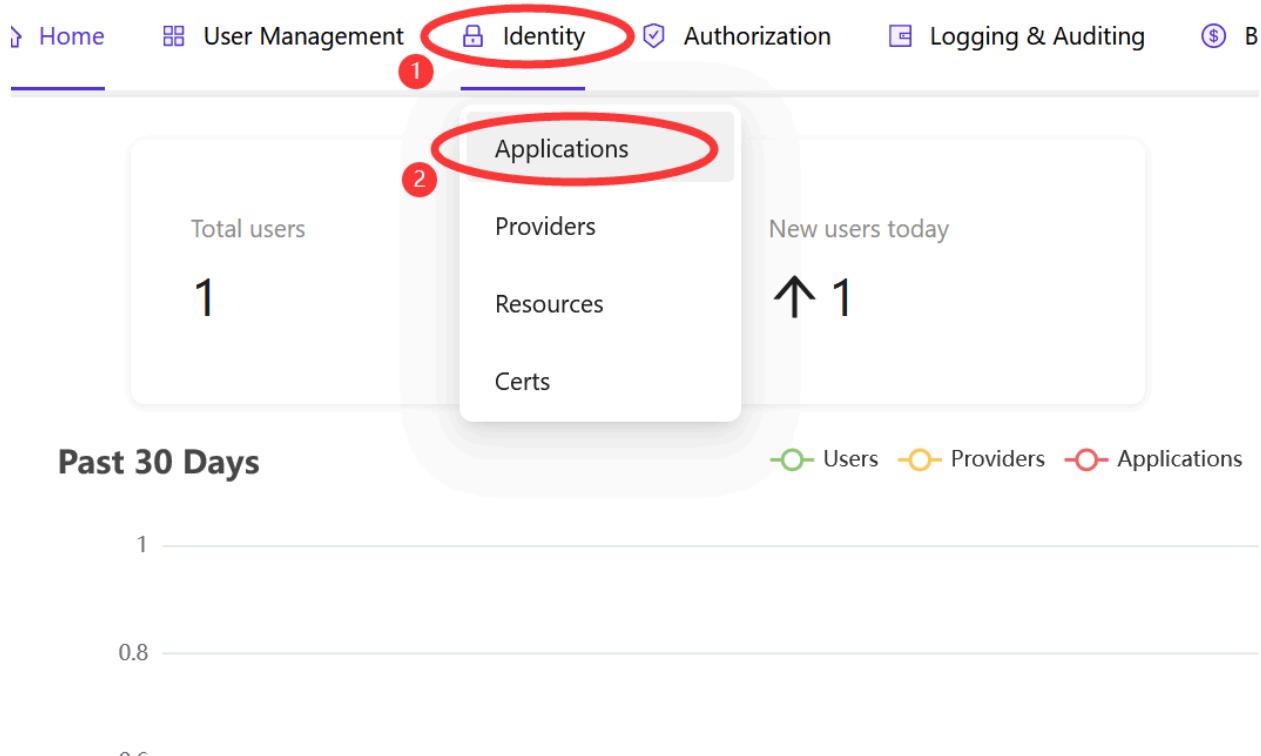
添加组织后，你可以查看组织信息。



Name	Created time	Display name	Favicon	Website URL	Password type	Password salt	Default avatar	Soft deletion	Action
casbin	2023-09-10 19:41:08	Casbin		https://door.casdoor.com	plain			<input checked="" type="radio"/> OFF	<button>Groups</button> <button>Users</button> <button>Edit</button> <button>Delete</button>
built-in	2023-09-10 19:31:50	Built-in Organization		https://example.com	plain			<input checked="" type="radio"/> OFF	<button>Groups</button> <button>Users</button> <button>Edit</button> <button>Delete</button>

步骤 3：在 Casdoor 中创建应用

在 Casdoor 中，你可以创建一个应用程序来管理你的用户和组织。你可以通过点击主页上 **身份 - 应用程序** 按钮来创建一个应用程序。



步骤 3.1：添加应用

点击 **添加** 按钮来添加一个应用。



步骤 3.2：填写应用信息

填写应用信息并点击 **保存并退出** 按钮。

Casdoor Home User Management Identity Authorization Logging & Auditing Business & Payments Ad

New Application Save Save & Exit Cancel

Name ? : app-casibase 5

Display name ? : Casibase 1

Logo ? : URL ? : https://cdn.casbin.org/img/casdoor-logo_1185x256.png

Preview: 

Home ? :

Description ? :

Organization ? : casbin 2

Tags ? :

Client ID ? : 2786e0cbadfb56287a9a 3

Client secret ? : 4f9957d3e679efdb3391eb42b38d274d46fa1232

Cert ? : cert-built-in

Redirect URLs ? :

Redirect URLs Add

Redirect URI 4

<http://localhost:14000/callback>

步骤 3.3：查看应用

添加应用后，你可以查看应用信息。

Casdoor Home User Management Identity Authorization Logging & Auditing Business & Payments Admin

All

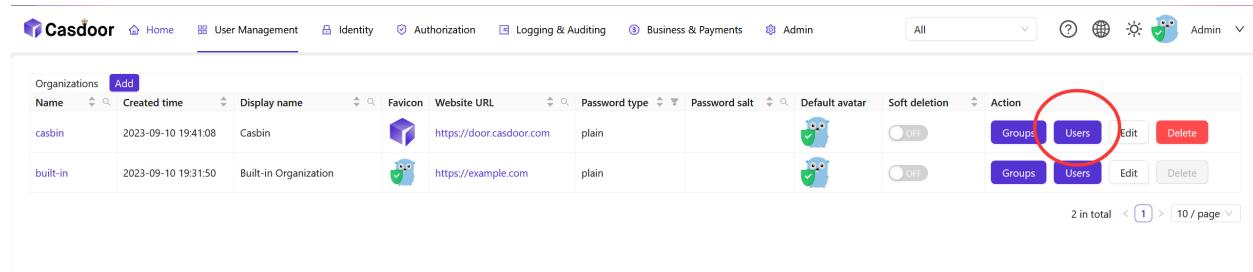
Applications Add

Name	Created time	Display name	Logo	Organization	Providers	Action
app-casibase	2023-09-10 19:44:08	Casibase		casbin	provider_captcha_default	<button>Edit</button> <button>Delete</button>
app-built-in	2023-09-10 19:31:50	Casdoor		built-in	provider_captcha_default	<button>Edit</button> <button>Delete</button>

2 in total < 1 > 10 / page

步骤 4：在 Casdoor 中为 Casibase 创建用户

在 Casdoor 中，你可以创建一个用户来登录 Casibase。你可以通过点击主页上的 **用户管理 - 组织 - 用户** 按钮来创建用户。你可以通过点击主页上 **用户管理 - 组织 - 用户** 按钮来创建一个用户。



The screenshot shows the Casdoor User Management interface. In the top navigation bar, the 'User Management' tab is selected. Below it, the 'Organizations' section lists two entries: 'casbin' and 'built-in'. For each organization, there are columns for Name, Created time, Display name, Favicon, Website URL, Password type, Password salt, Default avatar, Soft deletion, and Action. The 'Action' column contains three buttons: 'Groups', 'Users' (which is highlighted with a red circle), and 'Edit'. At the bottom right of the table, it says '2 in total' and '10 / page'.

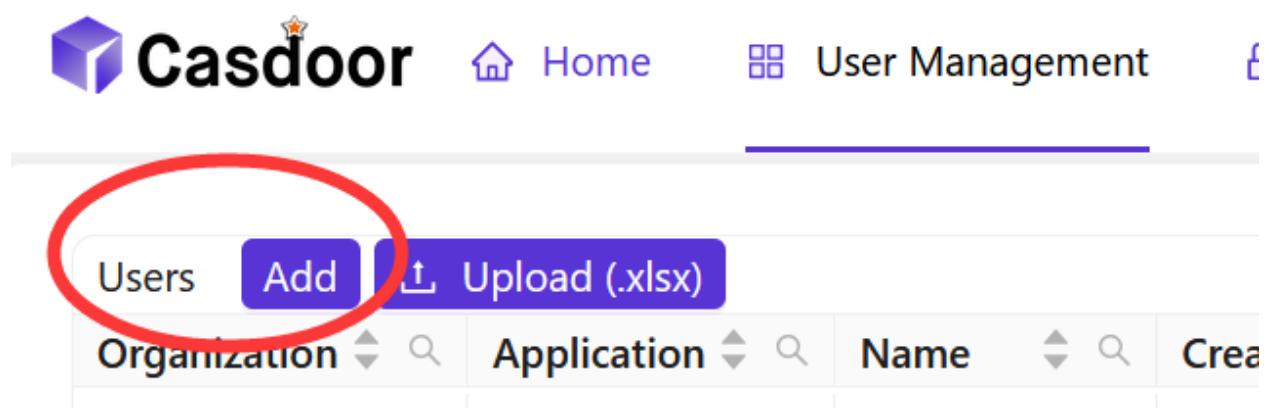


用户是组织中的成员，可以登录该组织中的应用。

访问 [Casdoor](#) 网站了解更多信息。

步骤 4.1：添加用户

点击 **添加** 按钮来添加一个用户。



The screenshot shows the Casdoor User Management interface. The top navigation bar has 'Home' and 'User Management' tabs. Below the navigation, a search bar includes fields for 'Organization', 'Application', 'Name', and 'Create'. A prominent red circle highlights the 'Add' button in the top left of the user list area.

步骤 4.2：填写用户信息

填写用户信息并点击 **保存并退出** 按钮。

New User Save Save & Exit Cancel

Organization ②: casbin 1

ID ②: d5bc730c-312c-406e-ae03-e6580d7590f4

Name ②: jimmy

Display name ②: Jimmy 2

Avatar ②: Preview:

User type ②: normal-user

Password ②: Modify password... 3

Email ②: t414w5@example.com

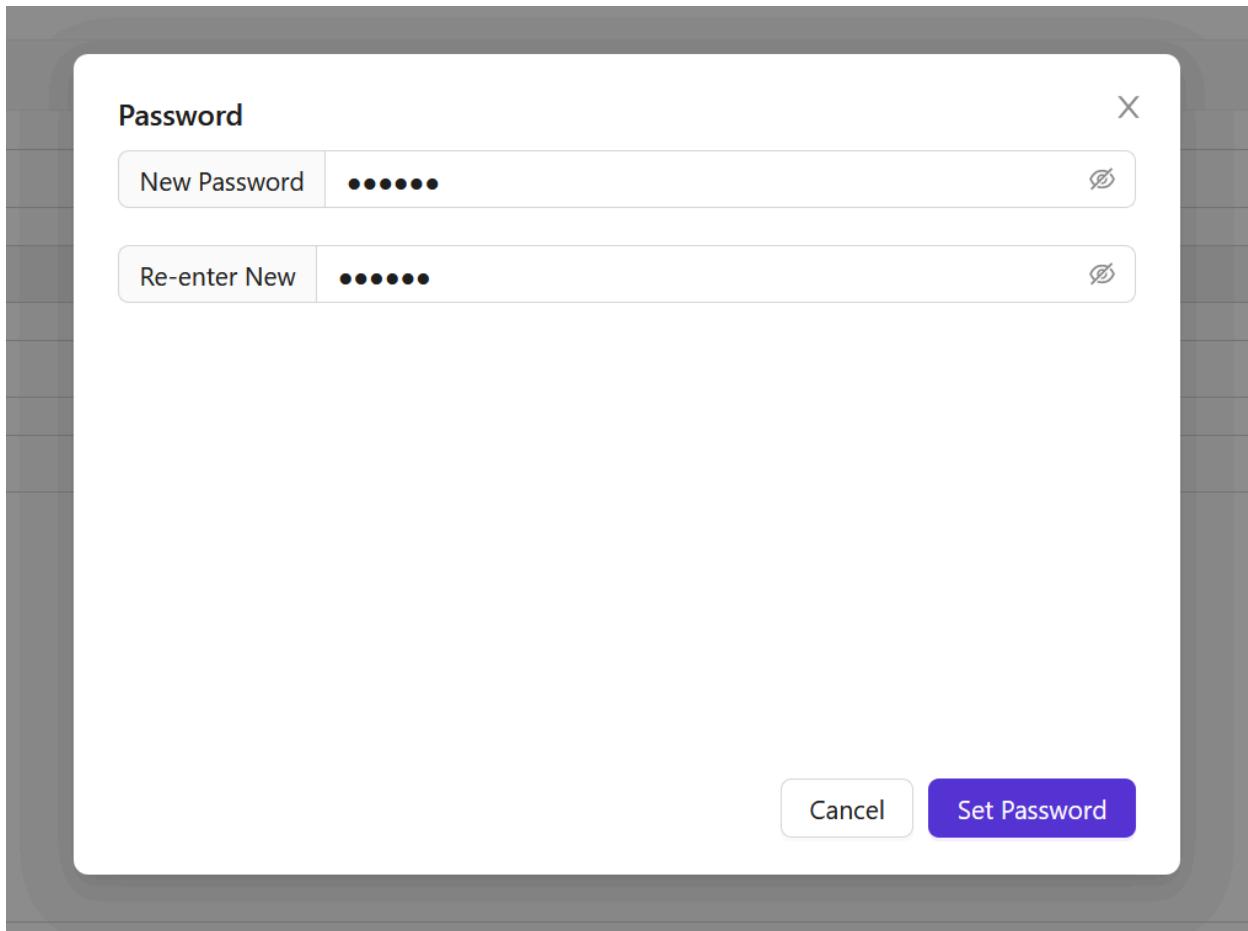
Phone ②: +1 71700415009

Country/Region ②: Please select country/region

Location ②:

- 密码

你可以通过点击 **修改密码** 按钮来设置用户密码。



- 管理员

你可以通过点击 **是否为管理员** 按钮来设置用户的管理员权限。

Permissions [?](#) :

Groups [?](#) :

3rd-party logins [?](#) :

Is admin [?](#) :



Is forbidden [?](#) :



Is deleted [?](#) :



Multi-factor authentication [?](#) :

Multi-factor methods

步骤 4.3：查看用户

添加用户后，你可以查看用户信息。

The screenshot shows the Casdoor User Management page. At the top, there are navigation links: Home, User Management, Identity, Authorization, Logging & Auditing, Business & Payments, Admin, and a search bar. Below the header is a table with a single row of data:

Organization	Application	Name	Created time	Display name	Avatar	Email	Phone	Affiliation	Country/Region	Tag	Is ac	Action
casbin	app-built-in	jimmy	2023-09-10 20:51:18	Jimmy		t414w5@example.com	71700415009	Example Inc.			Edit	Delete

At the bottom right of the table, there are pagination controls: '1 in total', a page number input field containing '1', and '10 / page' dropdown.

步骤 5：部署 Casibase

与 Casdoor 类似，你可以按照 [Casibase 部署指南](#) 来部署 Casibase。

部署完成 Casibase 后，你将会看到如下效果：



[Home](#) [Stores](#) [Providers](#) [Vectors](#) [Chats](#) [Messages](#) [Tasks](#) [Resources ↗](#) [Permissions ↗](#) [Logs ↗](#)

Jimmy ▾

Powered by **Casibase**



> How to Connect to Casibase

How to Connect to Casibase



Overview

Learn about different ways to connect to and integrate with Casibase.



Casibase SDKs

Learn how to integrate and use Casibase SDKs with your applications.



Using Casibase OpenAI API Compatible Interface

Learn how to connect external chat UIs to Casibase using OpenAI API compatibility.

Overview

Overview

In this section, we will show you how to connect your application to Casibase.

Casibase provides two main methods for integrating with your applications:

- [Casibase SDK](#) - For direct integration with Casibase's API
- [OpenAI API Compatibility](#) - For connecting existing OpenAI-compatible UIs and clients

Casibase SDK

What is Casibase SDK?

Casibase SDK provides a programmatic way to interact with Casibase services. It offers a convenient set of APIs that allow developers to manage tasks, knowledge bases, and other features of Casibase directly from their applications.

We recommend using the Casibase SDK for the following reasons:

1. It provides direct access to Casibase-specific functionality
2. It simplifies authentication and configuration
3. It handles error cases and provides a more developer-friendly experience

Currently, Casibase offers a Java SDK, with more language support planned for the future.

OpenAI API Compatibility

What is OpenAI API Compatibility?

Casibase supports the OpenAI API format, allowing you to connect any OpenAI-compatible chat UI or client application to Casibase. This makes it easy to use popular open-source chat interfaces with Casibase's backend.

We recommend using the OpenAI API compatibility for the following reasons:

1. It allows you to use your preferred chat UI with Casibase
2. It simplifies integration if you're already using OpenAI-compatible tools
3. It provides a standardized way to interact with Casibase's AI capabilities

This approach is particularly useful if you want to quickly integrate Casibase with existing applications that already support the OpenAI API format.

Casibase SDKs

Introduction

Casibase provides SDKs to help developers integrate with Casibase's APIs more easily. The SDKs offer a convenient way to interact with Casibase's services for tasks like managing AI conversations, knowledge bases, and more.

Currently, Casibase offers a Java SDK, with more language support planned for the future.

Backend SDK	Description	SDK code	Example code
Java SDK	For Java backends	casibase-java-sdk	-

How to use Casibase SDK?

1. Backend SDK configuration

When your application starts up, you need to initialize the Casibase SDK config by providing the required parameters.

Take casibase-java-sdk as an example:

```
CasibaseConfig config = new CasibaseConfig(  
    "https://demo-admin.casibase.com", // endpoint
```

All the parameters for initialization are explained as follows:

Parameter	Required	Description
endpoint	Yes	Casibase Server URL, like https://demo-admin.casibase.com or http://localhost:14000
clientId	Yes	Client ID for the Casibase application
clientSecret	Yes	Client secret for the Casibase application
organizationName	Yes	The name for the Casibase organization, e.g., casbin
applicationName	No	The name for the Casibase application, e.g., app-casibase

2. Available Services

Once you have initialized the configuration, you can create and use the available services. Currently, the only available service is `TaskService`.

```
TaskService taskService = new TaskService(config);
```

TaskService

`TaskService` supports basic task operations, such as:

- `getTask(String name)`: Get a single task by task name.
- `getTasks()`: Get all tasks under the `organizationName`.

- `addTask(Task task)`: Add a new task to the database.
- `updateTask(Task task)`: Update an existing task in the database.
- `deleteTask(Task task)`: Delete a task from the database.



> How to Connect to Casibase

> Using Casibase OpenAI API Compatible Interface

Using Casibase OpenAI API Compatible Interface

This document is a step-by-step tutorial designed for beginners. It will guide you through the process of connecting external chat UIs to Casibase using its OpenAI API compatibility feature.

Introduction

Casibase now supports integration with external chat UIs through OpenAI API compatibility. This feature allows you to use popular open-source chat interfaces with Casibase's backend, giving you more flexibility in how you interact with your knowledge base system.

If you're looking to use your preferred chat UI with Casibase, this guide will walk you through the simple setup process.

Step 1: Set Up Casibase with a Model Provider

Before connecting an external UI, ensure you have Casibase properly set up with a model provider. If you haven't done this yet, please refer to the [Add an AI Model Provider](#) tutorial.

Step 2: Get Your OpenAI-compatible API Key

When you create a model provider in Casibase, an API key is automatically generated. This key allows external applications to communicate with Casibase

using the OpenAI API format.

Step 2.1: Access Your API Key

Navigate to the **Providers** section and select your model provider. Only administrators can view and modify API keys.

Name:	provider_prm93r
Display name:	New Provider - prm93r
Category:	Model
Type:	OpenAI
Sub type:	text-davinci-003
Client secret	(?)
Temperature:	1.00
Top P:	1.00
Presence penalty:	0.00
Frequency penalty:	0.00
API key:	sk-UflKsbjYjzBvjeUFjbDpxuKg
Provider URL:	https://platform.openai.com/account/api-keys
State :	Active



If the API key field is empty, Casibase will automatically generate a new key when you save the provider.

Step 3: Configure Your External Chat UI

Once you have your API key, you can configure your external chat UI to connect to Casibase.

Step 3.1: Configure with chatgpt-web

For this example, we'll use [chatgpt-web](#), a popular open-source ChatGPT interface.

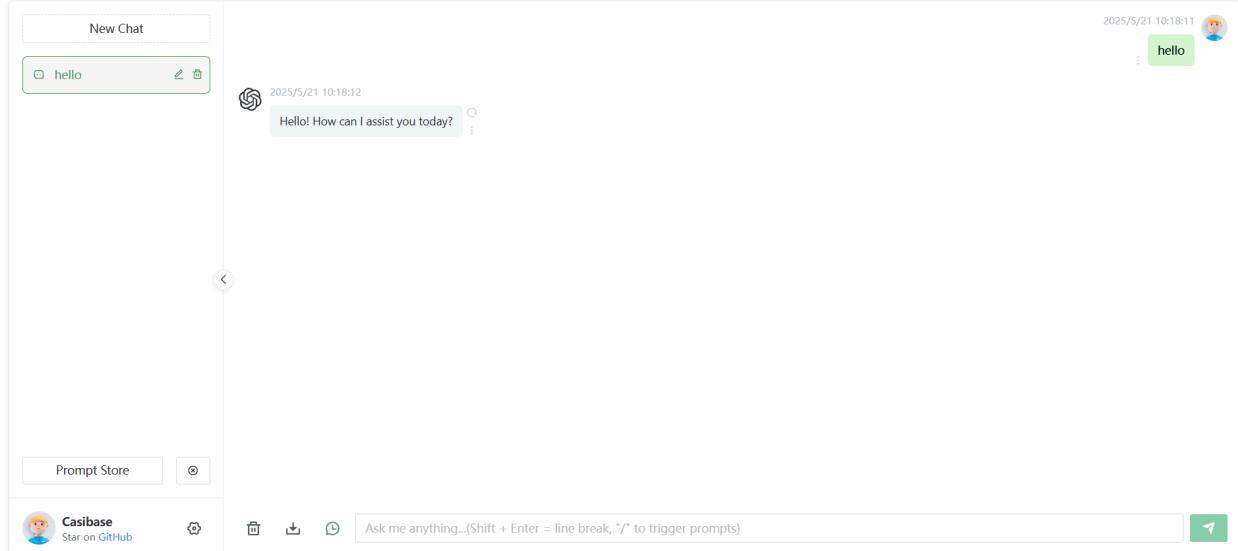
1. Locate the `service/.env` file in your chatgpt-web installation
2. Modify the following variables:
 - `OPENAI_API_KEY`: Set this to your Casibase-generated API key
 - `OPENAI_API_BASE_URL`: Set this to `http://your-casibase-backend:port/api`

```
# Example configuration
OPENAI_API_KEY=sk-UflKsbiYjzBvjeUFJbDpxuKg
OPENAI_API_BASE_URL=http://localhost:14000/api
```

Make sure your Casibase backend is accessible from the machine running your chat UI. Check firewall settings if you encounter connection issues.

Step 4: Test Your Integration

Start your chat UI application and test the connection. You should now be able to interact with Casibase through your preferred interface. If everything is set up correctly, you should see responses from Casibase in your chat UI.



Compatible Chat UIs

Casibase's OpenAI API compatibility has been tested with these popular chat interfaces:

- [chatgpt-web](#)

Other chat UIs that use the standard OpenAI API format should also work with Casibase.



>

提供商

提供商

概述

提供商概述

模型提供商

简介

嵌入提供商

简介

存储提供商

简介



文本转语音提供商

简介



语音识别提供商

简介



Model Providers

1 个项目



Scan Providers

6 个项目



Blockchain Providers

2 个项目



Private Cloud Providers

1 个项目



Public Cloud Providers

1 个项目

概述

Casibase 是一个开源的 AI 知识库系统，旨在为企业提供高效且灵活的知识管理与对话解决方案。One of its core features is Providers, which allow users to integrate multiple AI models, storage services, and infrastructure management capabilities. These integrations enhance the functionality and performance of the system.

Providers span several categories including AI Providers (Model and Embedding), Storage Providers, Cloud Providers, and Scan Providers, each serving distinct roles in the system. The provider list interface includes filtering capabilities, allowing you to quickly locate providers by category or type through dedicated filter controls.

1. 模型提供商

模型提供商是 Casibase 中用于集成和管理 AI 模型的组件。它允许用户将各种预训练的 AI 模型集成到系统中，以实现更智能的知识处理和对话生成。通过模型提供商，用户可以轻松地在不同的 AI 模型之间切换，根据具体需求选择最合适的模型。它允许用户将各种预训练的 AI 模型整合到系统中，从而实现更智能的知识处理和对话生成。借助模型提供者，用户可以轻松切换不同的 AI 模型，根据特定需求选择最合适的模型。

Casibase integrates with popular AI models through its provider system. Choose from cloud services like OpenAI and Alibaba Cloud, or connect to open-source models through Hugging Face and local deployments.

模型提供商类型

OpenAI — GPT-3.5-turbo, GPT-4, and reasoning-focused o1 models

Azure OpenAI — Access OpenAI models through Azure infrastructure

Alibaba Cloud — Qwen models plus DeepSeek v3, v3.1, v3.2, and R1

DeepSeek — Direct access to the v3.x series and R1 reasoning models

Claude — Anthropic's Claude 2 and Claude Instant models

Ernie — Baidu's ERNIE-Bot series

Hugging Face — Open-source models including Llama-2, ChatGLM2, and community contributions

2. 嵌入提供商

数据向量化

嵌入式提供者的主要作用是将各种类型的数据（例如，文本、图像等）进行转换 转换为稠密的向量表示。这种转换是 Casibase 中数据处理和分析的关键步骤，使得数据能够以更高效的方式存储、检索和分析。

知识检索

通过将知识库中的数据和用户的查询都转换为向量，嵌入提供商使系统能够基于向量相似度进行快速的知识检索。这大大提高了知识库检索的效率和准确性。这极大地提高了知识库检索的效率和准确性。

灵活的模型支持

嵌入提供商支持多种嵌入模型，用户可以根据需求选择最适合的模型。

3. 存储提供商

我们可以在 Casdoor 中配置存储提供者。 并将其用于 Casibase，该组件用于管理 Casibase 的数据存储和检索。 它允许用户将数据存储在不同的存储服务中，并通过统一的界面访问这些数据。 借助存储提供者，用户可以灵活选择存储服务，以确保数据安全和高效访问。 支持两种类型的存储：本地和云端。

4. 语音合成提供商

语音合成提供商是Casibase的一个组成部分，能够将文本答复转换成自然语音。 它使系统能够通过语音合成与用户进行交流，从而增强知识库系统的互动体验。

提供商支持

Currently, Casibase supports Alibaba Cloud's Text-to-Speech service, with various voice options available through the cosyvoice-v1 interface. 该系统设计为可扩展，以便未来集成更多的文本转语音服务提供商。

5. 语音识别提供商

语音识别提供商是Casibase的一个组成部分，可以将语音转化为书面文本。 它使该系统能够理解和处理语音查询，加强知识库系统的互动经验。

本地

我们支持将文件上传到本地系统。

云端

我们支持 AWS S3、Azure Blob Storage、MinIO、阿里云 OSS、腾讯云 COS，并且我们正在不断添加更多的云存储服务。

6. Cloud Providers

Cloud Providers enable Casibase to connect to and manage infrastructure across private and public cloud platforms. These providers allow centralized visibility and control of cloud resources.

Private Cloud Providers connect to containerized environments like Docker and Kubernetes, enabling container lifecycle management and orchestration directly from Casibase.

Public Cloud Providers scan and catalog cloud infrastructure resources across major platforms. They automatically discover virtual machines, storage volumes, network components, and other cloud assets, enriching each with detailed configuration information through multi-level API discovery.

7. Scan Providers

Scan Providers perform network discovery, security auditing, and system assessment operations. Casibase integrates specialized scanning tools that analyze infrastructure assets, detect vulnerabilities, and assess system

configurations.

Nmap Scan Provider conducts network reconnaissance by scanning IP addresses and port ranges. It identifies open ports, running services, service versions, and operating system details. The provider returns structured JSON results showing discovered hosts, accessible services, and potential security concerns.

OS Patch Provider evaluates system patch status and identifies missing security updates. It checks installed software versions against known vulnerabilities and recommends necessary patches to maintain system security.

Scan providers work with the Scan object to execute operations against target assets. Results are captured in both raw and structured formats, enabling detailed analysis through the web interface.

模型提供商

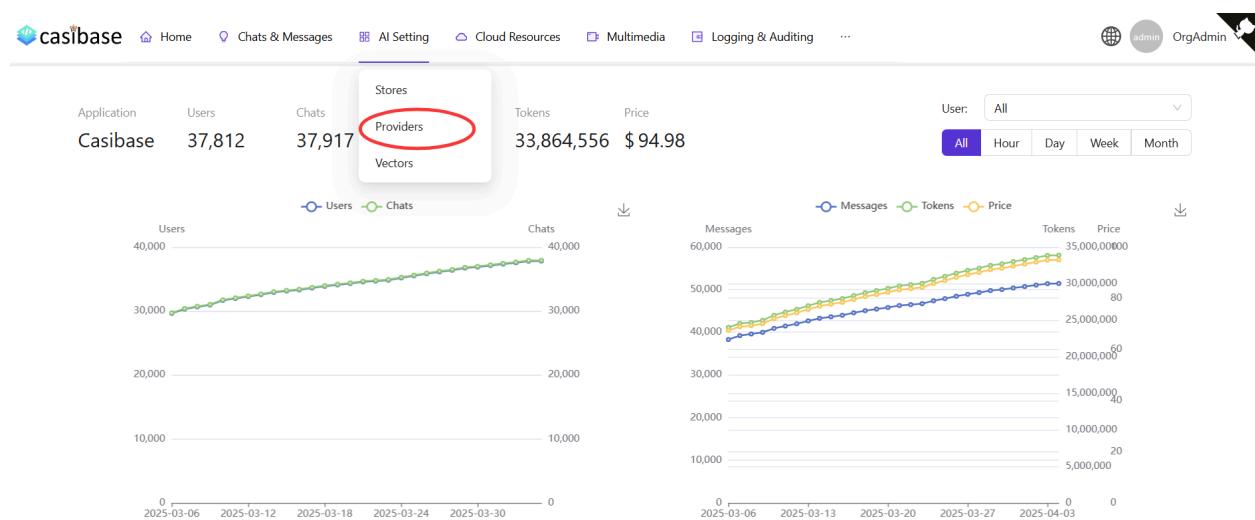
简介

Model providers connect AI capabilities to your Casibase instance. They power everything from chatbots to document analysis by integrating large language models (LLMs) directly into your knowledge base workflows.

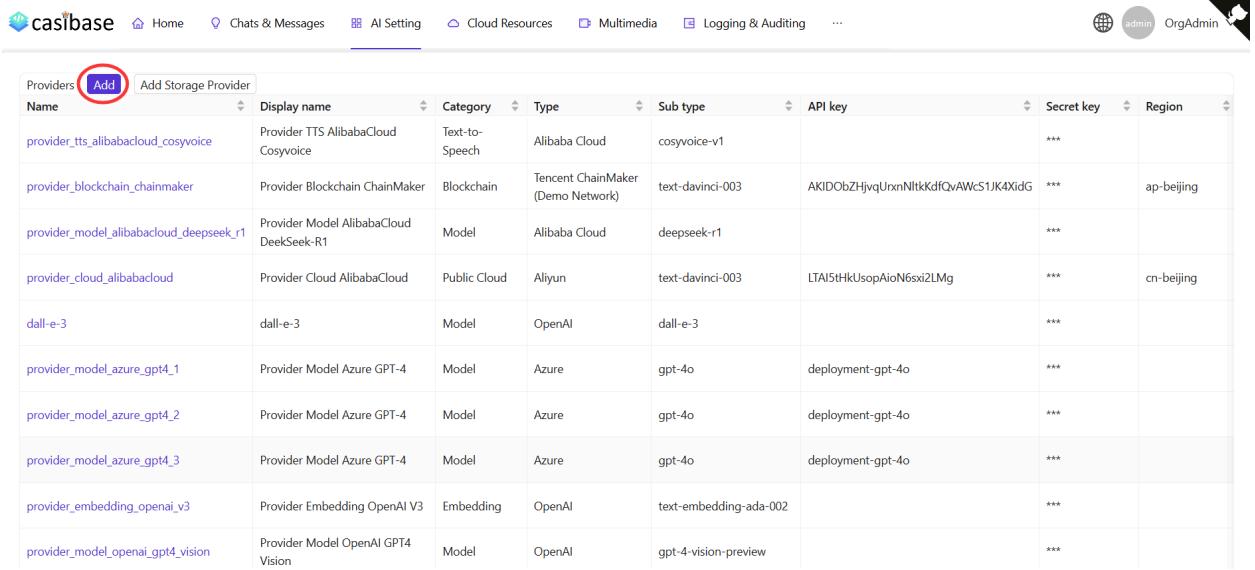
添加模型提供商

Adding a model provider takes just a few clicks:

点击主页上的**提供商**按钮。



点击**添加**按钮来添加模型提供商。



The screenshot shows the 'Storage Providers' section of the casibase AI Setting interface. The table has the following columns: Name, Display name, Category, Type, Sub type, API key, Secret key, and Region. The 'Add' button is located at the top left of the table area, highlighted with a red circle.

Providers	Add	Add Storage Provider						
Name		Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice		Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker		Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkKdfQvAWcs1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1		Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud		Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3		dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3		Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision		Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

填写模型提供商详情

填写模型提供商详情并点击 **保存并退出** 按钮。

[Home](#)[Chat](#)[Stores](#)[Providers](#)[Vectors](#)[Chats](#)[Me](#)[Edit Provider](#)[Save](#)

Name:

provider_openai_model

Display name:

OpenAI model

Category:

Model

Type:

OpenAI

Sub type:

text-davinci-003

Secret key:

Provider URL:

<https://platform.openai.com/account/api-keys>[Save](#)

Supported Providers

Provider	Models & Features
OpenAI	GPT-3.5-turbo, GPT-4, o1 series. Reasoning models show step-by-step thinking. Web search available for real-time information.
Azure OpenAI	Access OpenAI models through Azure infrastructure with official SDK support.
Alibaba Cloud	Qwen models and DeepSeek (v3, v3.1, v3.2, R1). Web search with inline results.
DeepSeek	DeepSeek v3.x series and R1 for reasoning-focused tasks.
Hugging Face	Open-source models including Llama-2, ChatGLM2, Baichuan2, and thousands more.
OpenRouter	Unified API access to Claude, PaLM, GPT-4, and other providers.

⚠ CONFIGURATION NOTES

When adding a provider, you'll need to set:

- **Category:** Choose `Model` for language models or `Embedding` for text vectorization
- **Type:** Select the specific provider (OpenAI, Hugging Face, etc.)

- **SecretKey:** Your API key or access token from the provider

Example: Adding OpenAI

The screenshot shows the Casbin provider configuration interface. At the top, there is a navigation bar with links: Home, Chat, Stores, Providers (which is highlighted), Vectors, Chats, and Message. Below the navigation bar, the main content area has a title "Edit Provider" and a "Save" button. The form fields are as follows:

Name:	provider_openai_model
Display name:	OpenAI model
Category:	Model
Type:	OpenAI
Sub type:	OpenAI Hugging Face OpenRouter Ernie
Secret key:	(Input field)
Provider URL:	https://platform.openai.com/account/api-keys

A red circle highlights the "Type" field, which is set to "OpenAI". A dropdown menu is open under "Sub type", showing options: OpenAI, Hugging Face, OpenRouter, and Ernie. The "Save" button is located at the bottom left of the form.

⚠ STREAMING SUPPORT

Not all models support streaming output (showing responses as they're generated). GPT-3.5-turbo and most modern models do, but if you're using older or custom models, check with your provider.

Once configured, your model provider powers Casibase's chat interface, document Q&A, and other AI features. You can switch between providers anytime to compare results or optimize for cost and performance.



The screenshot shows the Casibase interface with the 'Providers' tab selected. A single provider entry is listed in a table:

Name	Display name	Category	Type	Sub type	API key	Secret key	Provider URL	Action
provider_openai_model	OpenAI model	Model	OpenAI	text-davinci-003	***		https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>

The 'Action' column contains two buttons: 'Edit' and 'Delete'. The entire row is highlighted with a red border.

嵌入提供商

简介

嵌入是一种用于将单词和文档表示为向量的技术。嵌入提供者使您能够在知识库系统内分析和处理数据，从而使其更加智能高效。

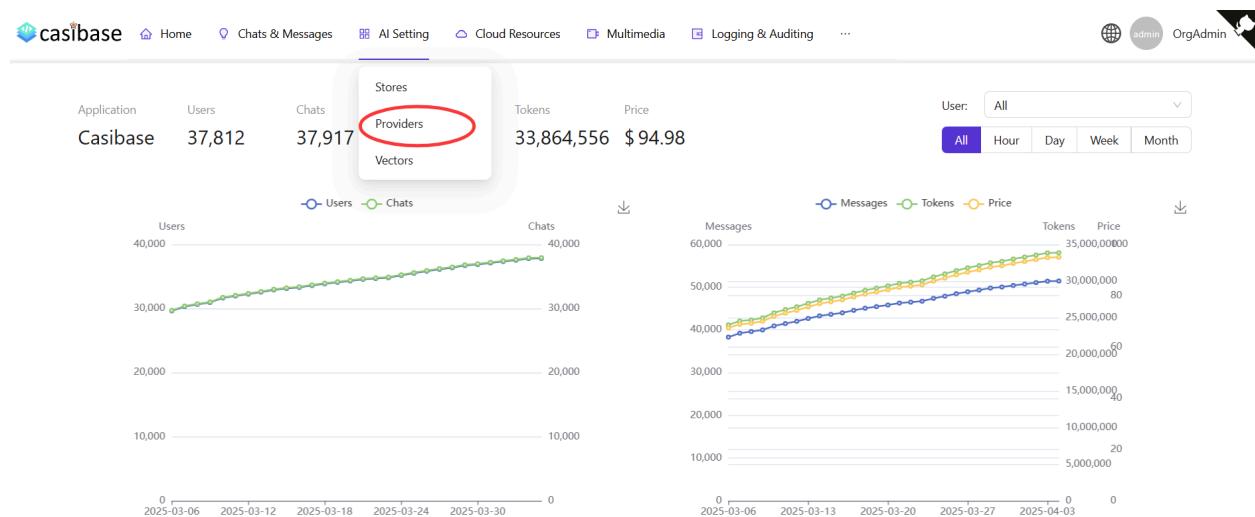
关于嵌入的更多信息，请参考我们之前文档中的[核心概念](#)部分。

在 Casibase 中，您可以按照以下步骤添加嵌入提供商：

添加新的嵌入提供商

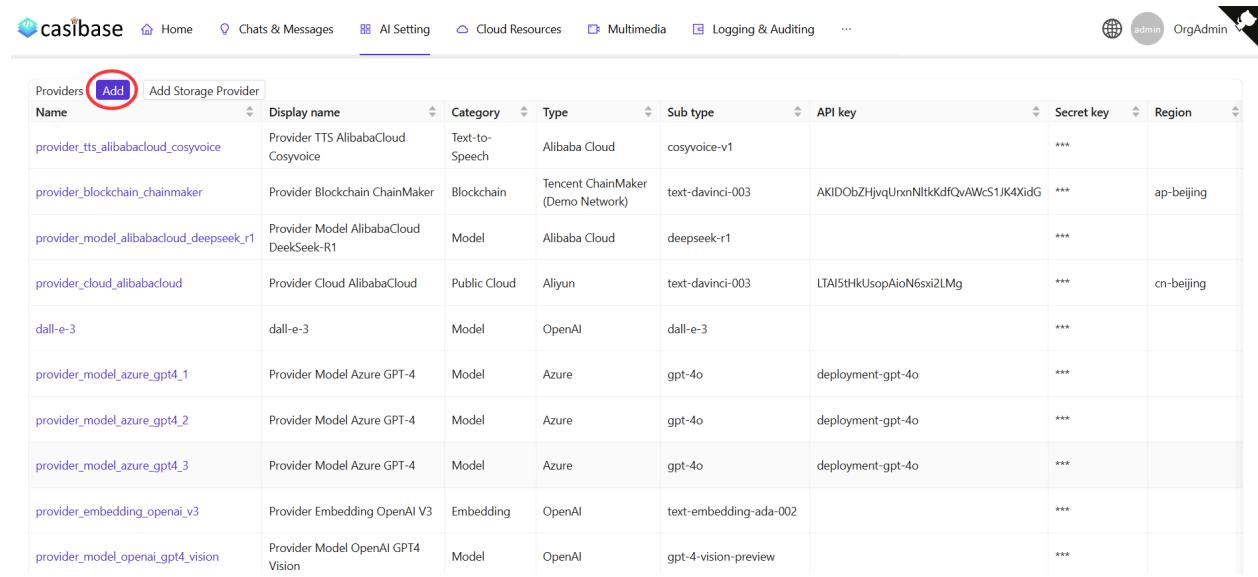
嵌入提供商用于将嵌入功能集成到 Casibase 中。您可以按照以下步骤添加它们：您可以按照以下步骤添加它们：

点击页面上的[提供商](#)按钮。



添加嵌入提供商

点击 **添加** 按钮来添加嵌入提供商。



Providers	Add	Add Storage Provider						
Name	Display name	Category	Type	Sub type	API key	Secret key	Region	
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***		
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkKdfQvAWcS1JK4XidG	***	ap-beijing	
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***		
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing	
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***		
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***		
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***		

填写嵌入提供商详情

填写嵌入提供商详情并点击 **保存并退出** 按钮。

[Home](#)[Chat](#)[Stores](#)[Providers](#)[Vectors](#)[CI](#)[Edit Provider](#)[Save](#)

Name:

embedding_openai_adasimilarity

Display name:

Embedding_OpenAI_AdaSimilarity

Category:

Embedding

Type:

OpenAI

Sub type:

AdaSimilarity

Secret key:

Provider URL:

<https://platform.openai.com/account/api-keys>[Save](#)

Casibase 支持多种嵌入提供商，包括：

- [OpenAI](#)

- AdaSimilarity
- DavinciSimilarity
- AdaEmbedding2
-
- Hugging Face
 - sentence-transformers/paraphrase-MiniLM-L6-v2
 -

返回提供商列表页面：

The screenshot shows the Casibin provider list interface. At the top, there are navigation links: Home, Chat, Stores, Providers (which is highlighted in blue), Vectors, Chats, Messages, Tasks, Resources, Permissions, and Logs. On the right, there is a user profile icon for 'Jimmy'.

Name	Display name	Category	Type	Sub type	API key	Secret key	Provider URL	Action
embedding_openai_adasimilarity	Embedding_OpenAI_AdaSimilarity	Embedding	OpenAI	1		***	https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>
model_openai_text_davinci_003	Model OpenAI text-davinci-003	Model	OpenAI	text-davinci-003		***	https://platform.openai.com/account/api-keys	<button>Edit</button> <button>Delete</button>

现在，您可以使用嵌入提供商将文本转换为向量。

添加嵌入提供商后，您可以使用它在 Casibase 中检索相似文档。更多信息请参考我们之前文档中的[核心概念](#)部分。有关更多信息，请参阅我们之前文档中的[核心概念](#)部分。

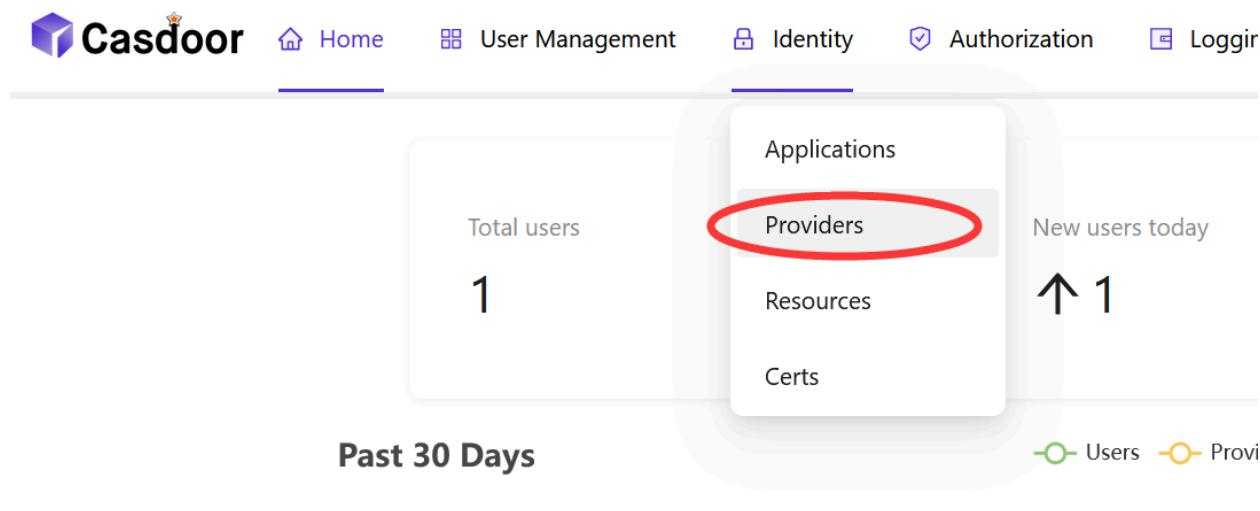
存储提供商

简介

向 Casibase 添加存储提供商可以让您高效地管理和存储数据，这是知识库系统的一个重要组件。

添加新的存储提供商

存储提供商用于存储数据。您可以通过点击主页上的 **身份 - 提供商** 按钮在 Casdoor 中添加它们。



点击 **添加** 按钮来添加存储提供商。



Providers

Add

Name

Organization

Created time

Di

provider_captcha_default

admin (Shared)

2023-09-10 19:31:50

Ca

填写存储提供商信息

填写存储提供商信息并点击 **保存并退出** 按钮。

New Provider

Save

Save & Exit

Cancel

Name ②:

provider_storage_1

Display name ②:

Provider_storage_1

Organization ②:

admin (Shared)

Category ②:

Storage

Type ②:

aws AWS S3

Client ID ②:

Alibaba Cloud OSS

aws AWS S3

Client secret ②:

Azure Blob

Endpoint ②:

Google Cloud Storage

Local File System

Endpoint (Intranet) ②:

MinIO

Bucket ②:

Qiniu Cloud Kodo

 提示

Casdoor 支持多种存储提供商，包括：

- [AWS S3](#)
- [Azure Blob](#)
- [Google Cloud Storage](#)
- [MinIO](#)

- [七牛云 Kodo](#)
- [阿里云 OSS ...](#)

示例

添加阿里云 OSS 存储提供商



小心

- 客户端 ID: 您的阿里云 OSS 账户的 AccessKey ID。
- 客户端密钥: 您的阿里云 OSS 账户的 AccessKey Secret。

***** 是您的阿里云 OSS 账户信息的占位符。

Category [?](#) : Storage

Type [?](#) : Alibaba Cloud OSS

Client ID [?](#) : LTA***NLf

Client secret [?](#) : Vo6***pi8

Endpoint [?](#) : oss-cn-beijing.aliyuncs.com

Endpoint (Intranet) [?](#) :

Bucket [?](#) : xx-bucket-0

Path prefix [?](#) :

Domain [?](#) : https://xx-bucket-0.oss-cn-beijing.aliyuncs.com

Provider URL [?](#) : https://github.com/organizations/xxx/settings/applications/1234567

[Save](#) [Save & Exit](#) [Cancel](#)

查看存储提供商

添加存储提供商后，您可以查看存储提供商信息。

Name	Organization	Created time	Display name	Category	Type	Client ID	Provider URL	Action
provider_storage_1	admin (Shared)	2023-09-10 21:23:02	Provider_storage_1	Storage	Alibaba Cloud OSS	[REDACTED]	https://github.com/organizations/xx...	Edit Delete

文本转语音提供商

简介

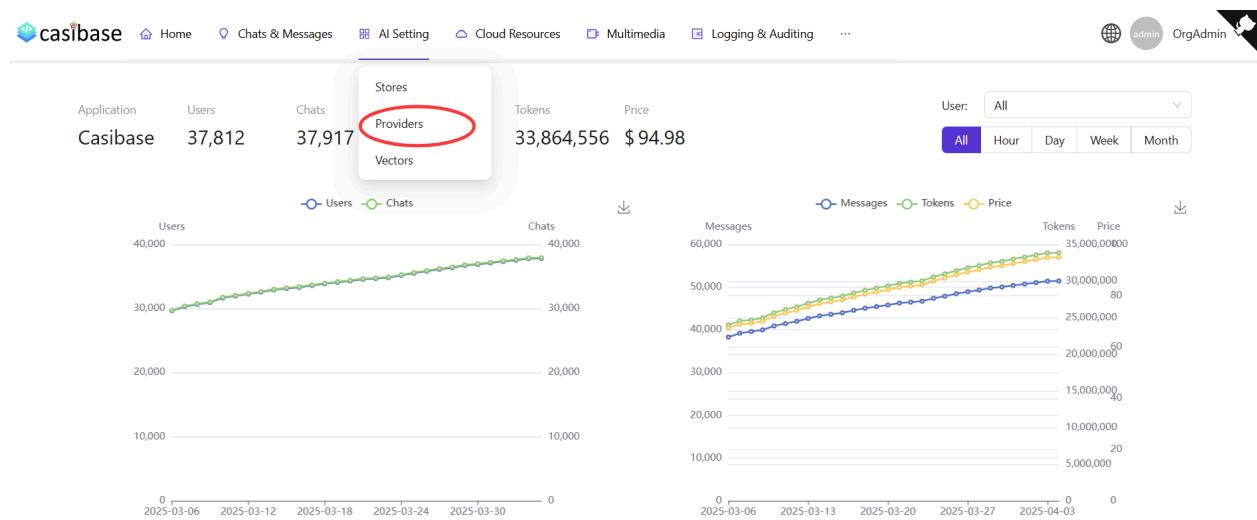
语音合成(TTS)是一种将文本转换成语音输出的技术。TTS 提供商使您的 Casibase 应用程序能够通过合成语音与用户交流，从而提高知识库系统的用户体验和可访问性。

在Casibase中，集成TTS提供商使您的AI应用能够口头响应查询，从而创造出更加互动且引人入胜的用户体验。

添加新的语音合成提供商

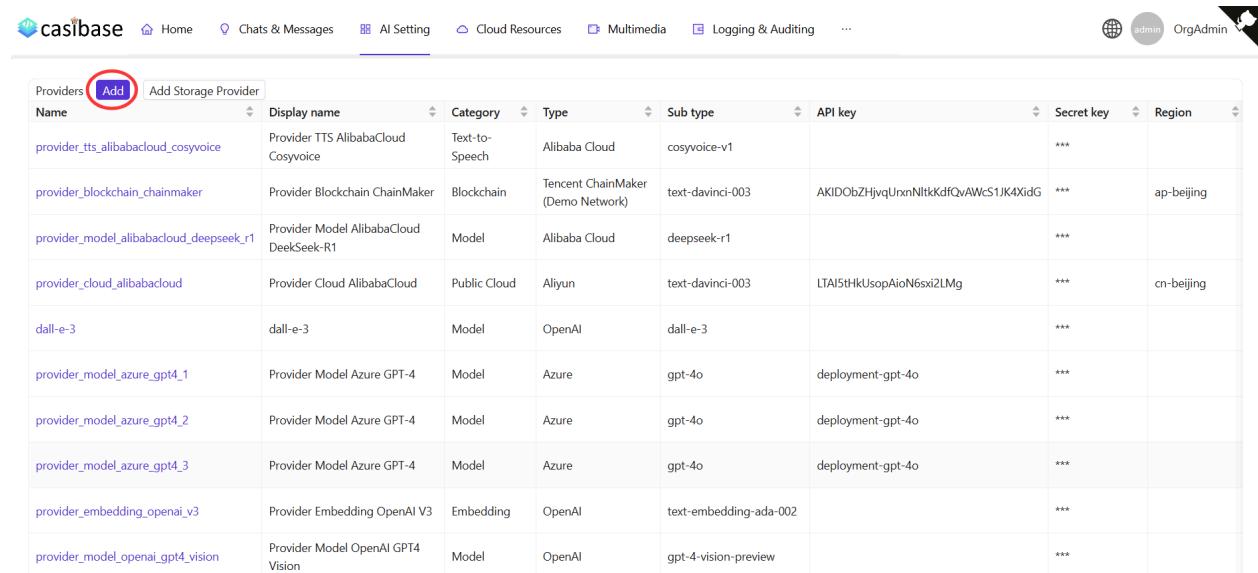
语音合成提供商用于将语音合成功能集成到 Casibase。您可以按照以下步骤添加它们：

点击页面上的 **提供商** 按钮。



添加语音合成提供商

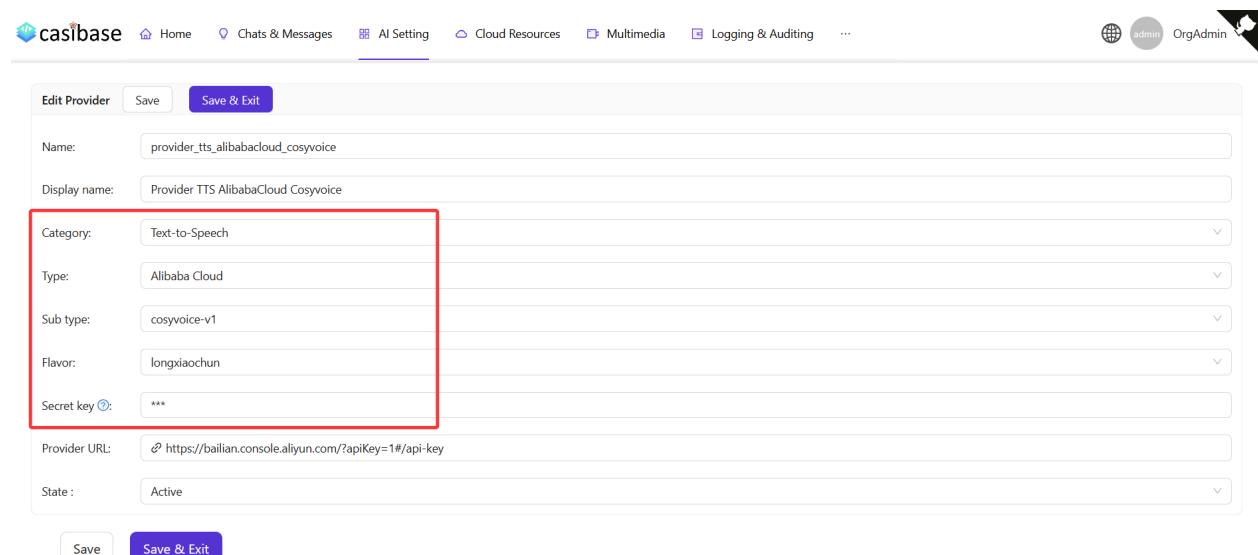
点击 **Add** 按钮来添加存储提供商。



Providers	Add	Add Storage Provider						
Name	Display name	Category	Type	Sub type	API key	Secret key	Region	
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***		
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkdfQvAWcS1JK4XidG	***	ap-beijing	
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***		
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing	
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***		
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***		
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***		
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***		

填写语音合成提供商详情

填写嵌入提供商详情并点击 **Save & Exit** 按钮。



Edit Provider Save Save & Exit

Name:	provider_tts_alibabacloud_cosyvoice
Display name:	Provider TTS AlibabaCloud Cosyvoice
Category:	Text-to-Speech
Type:	Alibaba Cloud
Sub type:	cosyvoice-v1
Flavor:	longxiaochun
Secret key:	***
Provider URL:	https://bailian.console.aliyun.com/?apiKey=1#api-key
State:	Active

Save Save & Exit

💡 提示

Casibase 目前支持以下语音合成提供商:

- Alibaba Cloud
 - cosyvoice-v1 (具有多个语音选项)

Testing Your Text-to-Speech Provider

You can test your TTS provider by clicking the `Read it out` button. This will allow you to enter text and hear the synthesized speech output.

The screenshot shows the 'Edit Provider' page in the Casibase interface. The 'Type' field is set to 'Alibaba Cloud' and the 'Sub type' field is set to 'cosyvoice-v1'. In the 'Provider test' section, the input field contains 'Hello, I'm casibase AI.' and the 'Read it out' button is highlighted with a red box. Other fields include 'Name', 'Display name', 'Category', 'Flavor', 'Secret key', 'Provider URL', and 'State'.

Name:	provider_r7fdnn
Display name:	New Provider - r7fdnn
Category:	Text-to-Speech
Type:	Alibaba Cloud
Sub type:	cosyvoice-v1
Flavor:	龙小淳，女，中英双语。龙小淳的嗓音如丝般柔滑，温暖中流淌着亲切与抚慰，恰似春风吹过心田。
Secret key:	***
Provider test:	Hello, I'm casibase AI.
Provider URL:	https://platform.openai.com/account/api-keys
State:	Active

This testing feature allows you to verify your TTS configuration before implementing it in your applications, ensuring the voice quality and settings meet your requirements.

Alibaba 的语音选项

当使用 Alibaba Cloud 的 `cosyvoice-v1` 时，您可以从各种语音选项中选择:

- 龙婉

- 龙橙
-

在商店中使用语音合成功能

添加语音合成提供商后，您可以在您的商店设置中选择此提供商，然后选择是否启用TTS串流。

The screenshot shows the 'Edit Store' configuration page with the following fields:

- Name: store-built-in
- Display name: Built-in Store
- Title: (empty)
- Avatar: (empty)
- Storage provider: Built-in Storage Provider (provider-storage-built-in)
- Image provider: Storage Aliyun OSS Casibase Casbin (provider_storage_casibase_casbin)
- Split provider: Default
- Model provider: Provider Model Azure GPT-4 (provider_model_azure_gpt4)
- Embedding provider: Provider Embedding OpenAI V3 (provider_embedding_openai_v3)
- Text-to-Speech provider: Provider TTS AlibabaCloud Cosyvoice (provider_tts_alibabacloud_cosyvoice)
- Enable TTS streaming:

现在，您的商店可以将文本回复转换为语音，为用户提供更具互动性的体验。

语音识别提供商

简介

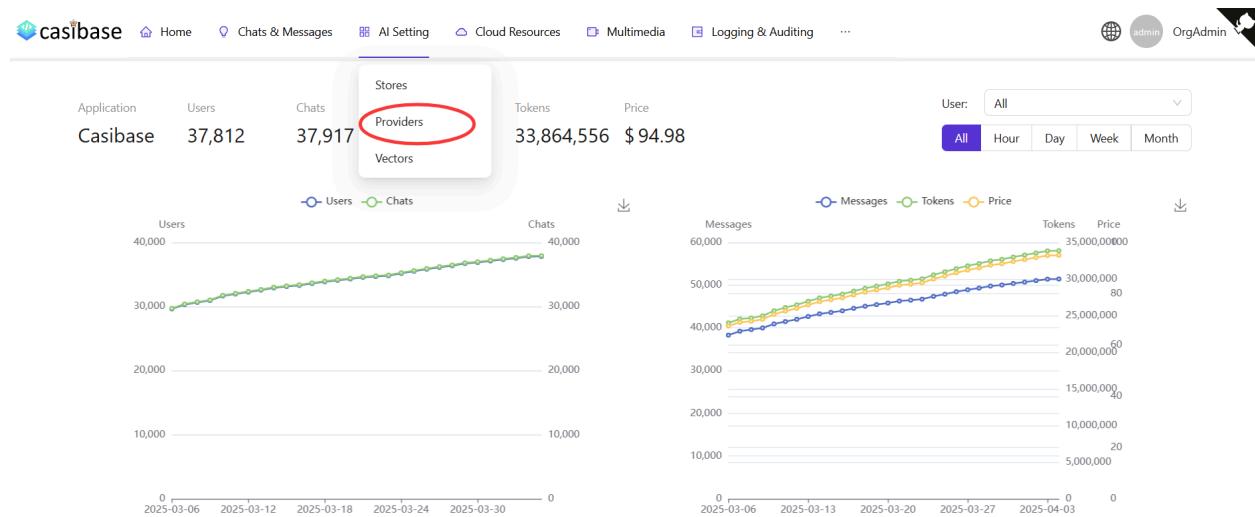
语音识别（STT）是一种将口语转换成书面文字的技术。STT 提供商允许您的 Casibase 应用程序理解和处理用户口语输入，增强用户体验和您的知识库系统的访问能力。

在 Casibase 中，整合一个 STT 提供商使您的 AI 应用程序能够接收和处理语音查询，创建更多的互动和自然用户交互。

添加一个新的语音识别提供商

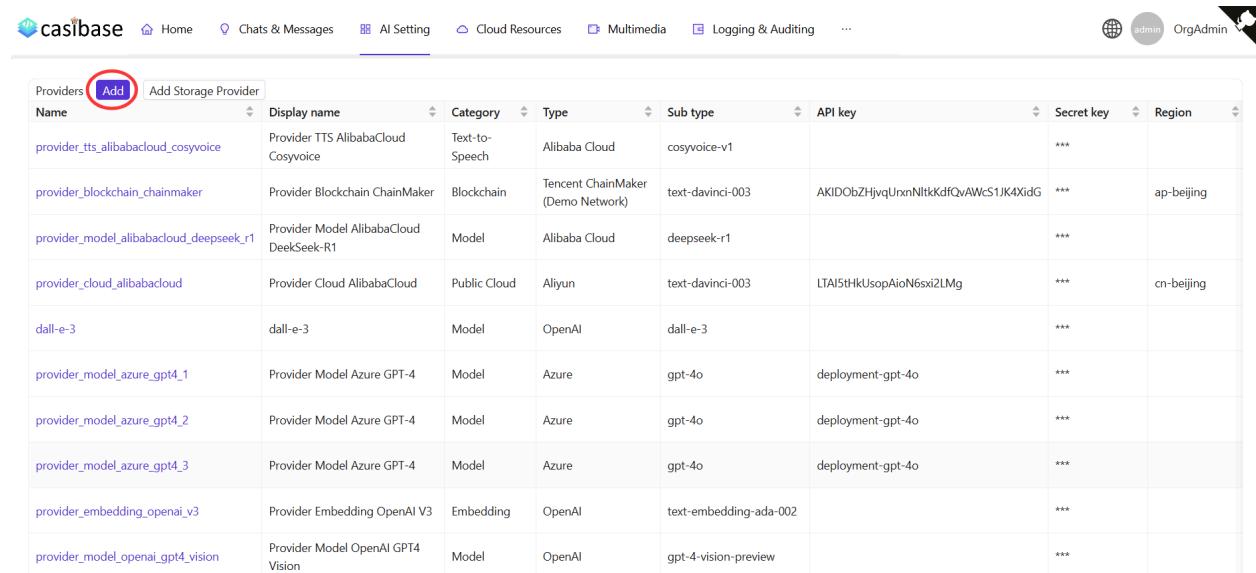
语音识别提供商用将语音识别能力整合到 Casibase 中。您可以按照以下步骤添加它们：

点击主页上的 Providers 按钮。



添加语音识别提供商

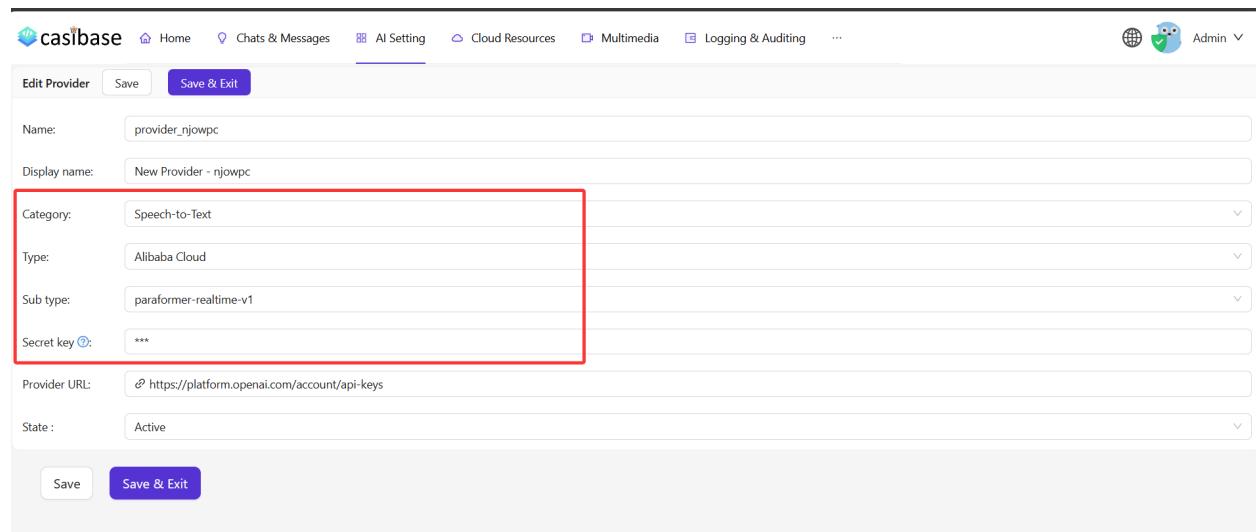
点击 **Add** 按钮来添加存储提供商。



Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkKdfQvAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

填写语音识别提供商详细信息

填写模型提供商详情并点击 **Save & Exit** 按钮。



casibase Home Chats & Messages AI Setting Cloud Resources Multimedia Logging & Auditing ... Admin

Edit Provider Save Save & Exit

Name: provider_njowpc

Display name: New Provider - njowpc

Category: Speech-to-Text

Type: Alibaba Cloud

Sub type: paraformer-realtime-v1

Secret key: ***

Provider URL: <https://platform.openai.com/account/api-keys>

State : Active

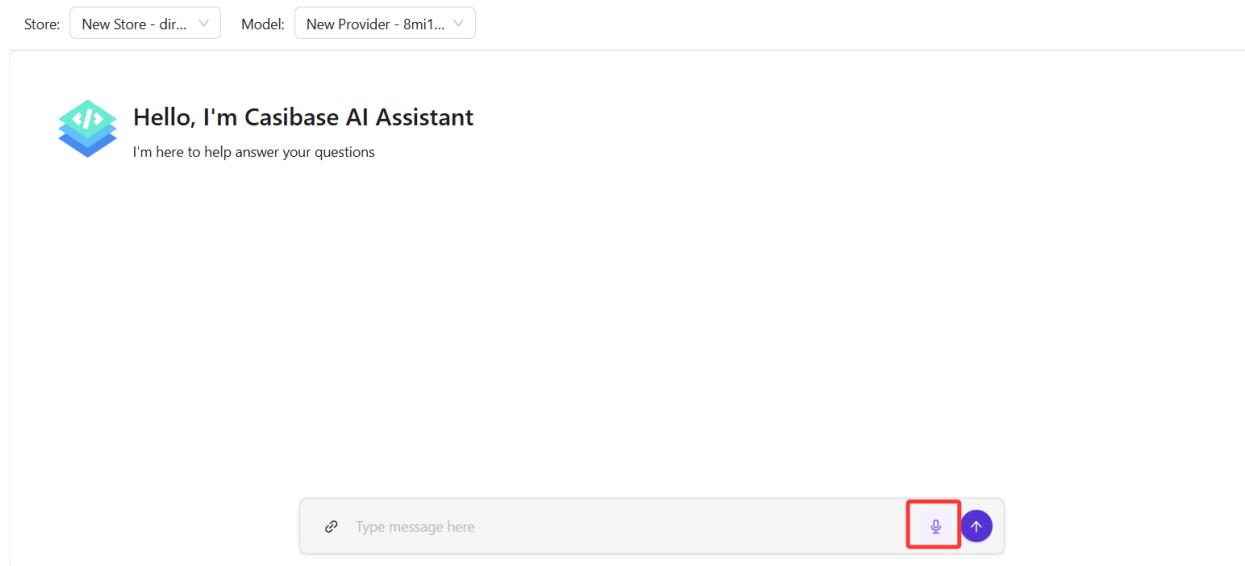
Save Save & Exit

使用语音识别

当您点击 Casibase 应用程序中的语音识别按钮时，将执行以下流程：

1. 浏览器将请求访问您麦克风的权限
2. 一旦授予，系统将开始收听并自动将您的语音转换为文本
3. 在您完成发言后，识别的文本将自动作为消息发送

此功能允许与您的 Casisbase 应用程序进行无操作的互动，使它们更容易访问和使用。



CasiBase目前支持以下语音识别提供商：

- Alibaba Cloud
 - paraformer-realtime-v1

Model Providers

Introduction

Model Providers enable AI capabilities in Casibase by integrating various large language models (LLMs) and AI services. These providers allow you to chat with AI, analyze documents, generate embeddings, and perform other intelligent tasks.

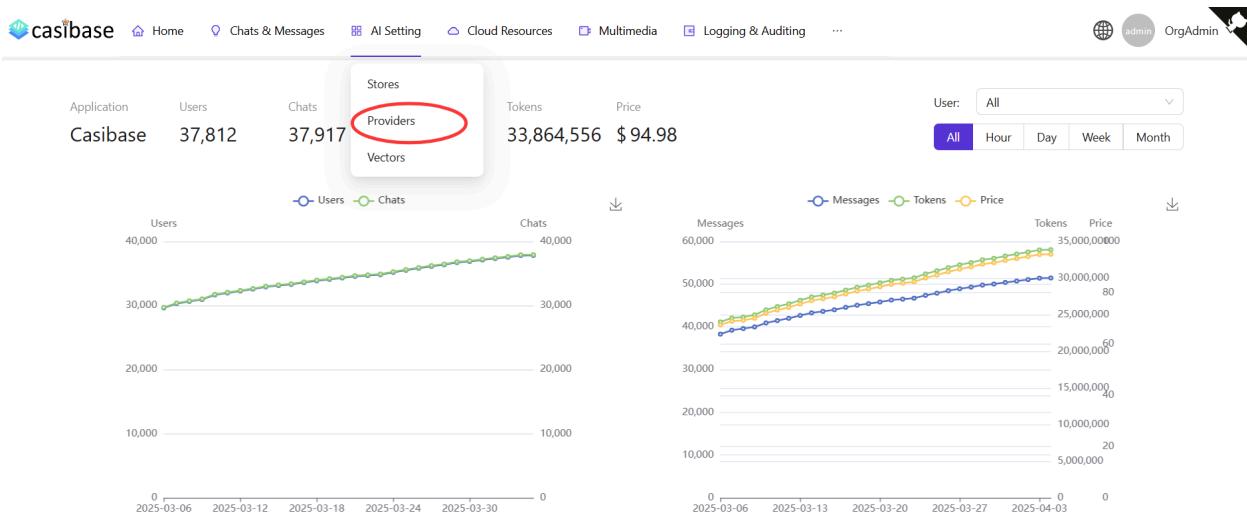
Casibase supports a wide range of model providers, from major cloud services like OpenAI and Azure OpenAI to local and custom model deployments. This flexibility lets you choose the right model for your use case, whether you prioritize performance, cost, privacy, or specific capabilities.

Refer to the [Core Concepts](#) section for more information about providers in general.

Add a New Model Provider

Model providers are used to integrate LLM capabilities into Casibase. You can add them by following these steps:

Click the [Providers](#) button on the page.



Add a Model Provider

Click the **Add** button to add a new model provider.

The table lists existing providers and includes an **Add** button at the top left:

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkKdIQvAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

Fill in Model Provider Details

Fill in the required configuration, setting the Category to "Model" and selecting the appropriate Type. Then click the **Save & Exit** button to save.

More information about specific model provider types can be found below:

Local Model Provider

Configure local and custom model providers with OpenAI-compatible APIs

Supported Providers

Provider	Description
OpenAI	GPT-3.5, GPT-4, o1 series with web search and step-by-step reasoning displays
Azure OpenAI	OpenAI models through Azure infrastructure with official SDK integration
Alibaba Cloud	Qwen models and DeepSeek (v3, v3.1, v3.2, R1) with web search support
DeepSeek	Direct access to DeepSeek reasoning models
Hugging Face	Open-source models hosted on Hugging Face
OpenRouter	Unified access to multiple AI models through one API
Local	Self-hosted models or custom providers with OpenAI-compatible APIs

After adding a model provider, you can use it for chatting, document analysis,

question answering, and other AI-powered features in Casibase.

Local Model Provider

The Local model provider in Casibase enables you to connect to self-hosted/local models or any custom model provider that implements an OpenAI Chat Completion-style interface. This gives you maximum flexibility to use models not directly supported by Casibase while maintaining full control over your infrastructure and data.

When to Use Local Provider

The Local provider is recommended in the following scenarios:

Custom Provider Integration:

Use the Local provider when you need to connect to a model provider that is not listed in the Casibase provider dropdown, but exposes an OpenAI-compatible Chat Completion API.

Many modern LLM services implement this standard interface for compatibility, making them easy to integrate via the Local provider.

Self-hosted or Local Models:

The Local provider is also suitable when running your own models in local or private environments. This includes popular frameworks such as vLLM, LocalAI, LM Studio, llama.cpp, or any custom deployment that supports the OpenAI Chat Completion format.

By using the Local provider, all data remains within your own infrastructure,

making it ideal for sensitive workloads, regulatory compliance, and development or testing environments.

Configuration

The following image shows the provider edit page when configured as a local type:

The screenshot shows a configuration interface for a provider. At the top left, there's a sidebar with a red border containing the following fields:

- Category: Model
- Type: Local
- Sub type: custom-model
- Compatible provider: deepseek-v3.2
- Input price / 1k tokens: 0
- Output price / 1k tokens: 0
- Currency: USD
- API key: ...

Below this sidebar is a large empty text area labeled "Provider test" with a "New Chat" button at the top right. At the bottom of the page, there are two input fields:

- Provider key: ...
- Provider URL: <https://aistudio.google.com/apikey>

A message input field with a placeholder "Type message here" and a send button is located at the bottom center.

Required Fields

When adding a Local model provider, configure these essential fields:

Category: Set to **Model** to indicate this is a model provider rather than embedding or other service types.

Type: Select **Local** from the dropdown.

Subtype: This field is automatically set to `custom-model` and cannot be changed. It identifies the provider as a custom implementation.

Compatible provider: Enter the model name you want to use. You can either select from the dropdown suggestions or type any model name directly.

Provider URL: The HTTP(S) endpoint where your model service is running. This is the base URL that Casibase will use to make requests. For example:

- `http://localhost:8000/v1` for a local OpenAI-compatible server
- `http://192.168.1.100:8000/v1` for vLLM
- `https://my-model-service.company.com/api/v1` for a custom deployment
- `https://cloud.infini-ai.com/maas/v1`

The endpoint should implement the `/chat/completions` path that accepts OpenAI format requests.

API key: If your model service requires authentication, provide the API key or token here. Leave empty if your service doesn't require authentication (common for local deployments). The key is securely stored and sent in the Authorization header as `Bearer <key>`.

Pricing Configuration

These fields help track usage costs when using paid services or for internal billing:

Input price / 1k tokens: Cost per 1,000 input tokens. Enter the numeric value (e.g., `0.001` for \$0.001 per 1k tokens). Set to `0` for free models.

Output price / 1k tokens: Cost per 1,000 output tokens.

Currency: The currency for pricing. This is used for cost tracking and reporting.

Configuration Example

OpenAI-compatible service:

- Compatible provider: Any model you prefer
- Provider URL: Service endpoint URL
- API key: Your service API key
- Input/Output price: According to the service pricing

Using the Provider

After saving your Local model provider, you can use it just like any other provider in Casibase. Select it when creating chats, configuring stores for RAG, or any other feature that requires a model provider.

When the provider is in use, Casibase sends requests to your configured Provider URL using the OpenAI Chat Completion format. Your service should respond with compatible JSON responses.

Troubleshooting

Connection refused: Verify the Provider URL is correct and the service is running. Check firewalls and network connectivity.

Authentication errors: Ensure the API key is correct if your service requires authentication. Some services use different authentication methods - verify your service supports Bearer token authentication.

Unexpected responses: Confirm your service implements the OpenAI Chat

Completion API format correctly. Check the service logs for details about request/response formats.

Model Not Found: If the service returns a “Model Not Found” error, verify that the Compatible provider field in Casibase is configured correctly and that the model has been loaded and is available in the deployment environment.

Scan Providers

Introduction

Scan Providers enable automated network discovery, security auditing, and system assessment in Casibase. These providers integrate specialized scanning tools to analyze infrastructure assets and identify security issues.

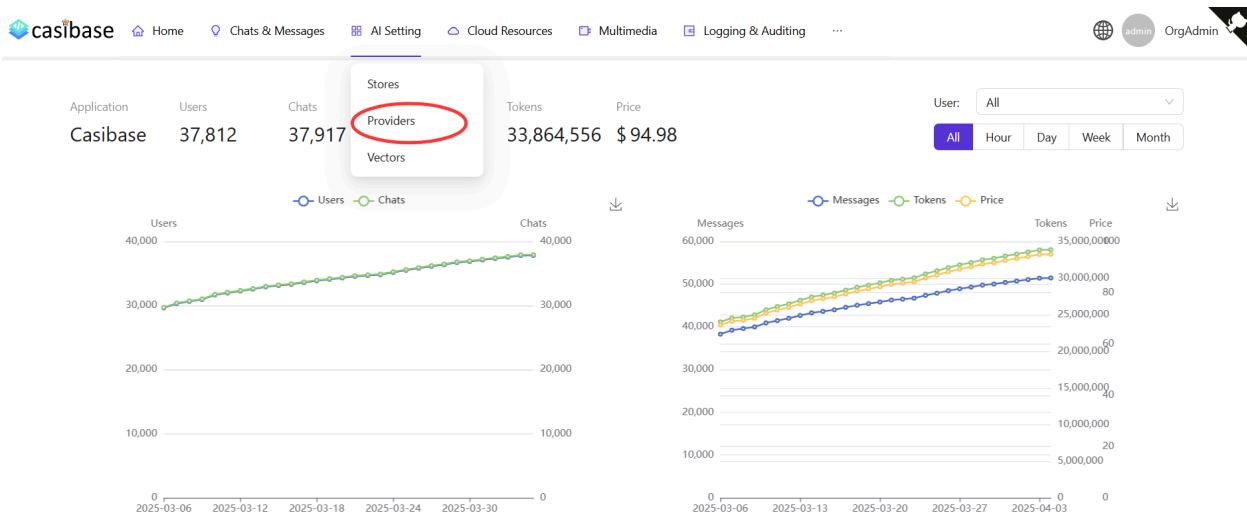
By configuring scan providers, you can perform network reconnaissance, detect open ports and running services, assess patch status, and identify missing security updates across your infrastructure. The results are captured in both raw and structured formats for detailed analysis.

Refer to the [Core Concepts](#) section of our previous documentation for more information about providers.

Add a New Scan Provider

Scan providers are used to integrate network and security scanning capabilities into Casibase. You can add them by following these steps:

Click the [Providers](#) button on the page.



Add a Scan Provider

Click the **Add** button to add a new scan provider.

The table lists existing providers and has a header row with columns for Name, Display name, Category, Type, Sub type, API key, Secret key, and Region. A red circle highlights the **Add** button in the top-left corner of the table header.

Name	Display name	Category	Type	Sub type	API key	Secret key	Region
provider tts_alibabacloud_cosyvoice	Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker	Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUxnNltkKdIQvAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1	Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud	Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6xi2LMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3	Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3	Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision	Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

Fill in Scan Provider Details

Fill in the required configuration, setting the Category to "Scan" and selecting the appropriate Type. Then click the **Save & Exit** button to save.

More information about each scan provider type can be found below:

Nmap Scan Provider

Network discovery and security auditing with Nmap

OS Patch Scan Provider

System patch assessment and security update detection

Nuclei Scan Provider

Vulnerability scanning with Nuclei

ZAP Scan Provider

Web application security testing with OWASP ZAP

Subfinder Scan Provider

Subdomain discovery with Subfinder

httpx Scan Provider

HTTP service probing with httpx

提示

Casibase supports multiple scan provider types:

- **Nmap**: Network discovery and security auditing through port scanning, service detection, and OS fingerprinting
- **OS Patch**: System security assessment by checking for missing patches and available updates
- **Nuclei**: Template-based vulnerability scanning for detecting security issues and misconfigurations
- **ZAP**: Web application security testing using OWASP Zed Attack Proxy
- **Subfinder**: Passive subdomain discovery for attack surface mapping
- **httpx**: HTTP service probing for web server enumeration and technology detection

Now you can use scan providers to assess your infrastructure security and maintain visibility into your network assets.

After adding a scan provider, you can create scans to execute network discovery or security assessments against your infrastructure assets, with results stored for historical analysis.

Nmap Scan Provider

The Nmap Scan Provider integrates Network Mapper into Casibase for network discovery and security auditing. It scans target systems to identify open ports, running services, and operating system details.

Configuration

Create an Nmap Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "Nmap". Configure the display name and owner as needed. The provider operates locally within Casibase and requires no external credentials or API keys.

Network Scanning

Nmap performs comprehensive network analysis by probing target systems. Port discovery identifies which ports are open, closed, or filtered. Service detection determines what software is running on each port and attempts to identify version numbers. OS fingerprinting analyzes network responses to determine the target's operating system and version.

Using the Provider

From the Scans page, create a new scan and select the Nmap provider. Choose your target by selecting an asset from your inventory or entering an IP address manually for ad-hoc scanning.

Command Templates

The provider edit page includes command templates for common scan types.

Select a template from the dropdown to quickly configure your scan:

- Ping Scan: Quick host discovery without port scanning
- Quick Scan: Fast scan of common ports with service detection
- Intense Scan: Comprehensive scan with OS detection and script scanning
- Port Scan: Scan specific port ranges
- OS Detection: Identify target operating system
- Version Detection: Detect service versions on open ports
- Stealth Scan: Lower visibility SYN scan
- UDP Scan: Scan UDP ports instead of TCP

Custom Commands

You can customize the Nmap command to meet specific requirements. The command field supports the `%s` placeholder for the target IP address. For example, `-sV -p 1-1000 %s` scans ports 1-1000 with version detection. The command input validates input to prevent command injection by rejecting shell metacharacters such as `;`, `&`, `|`, `$`, and backticks.

The test widget allows you to verify functionality before running scans. Enter a target, optionally modify the command, click "Scan", and view results immediately.

Scan Results

Nmap returns results as structured JSON containing host information and discovered services. A typical response includes the target IP, host state, and an

array of port details:

```
{  
  "hosts": [  
    {  
      "ip": "192.168.1.1",  
      "state": "up",  
      "ports": [  
        {  
          "port": 22,  
          "state": "open",  
          "service": "ssh",  
          "version": "OpenSSH 8.2"  
        }  
      ]  
    }  
  ]  
}
```

The web interface renders this data in organized tables showing port numbers, states, service names, and detected versions. This structured presentation makes it easy to quickly assess what services are exposed on your network.

Network Considerations

Ensure Casibase has network connectivity to your target systems. Firewalls may need configuration to permit scanning traffic from the Casibase server. Network scanning can trigger intrusion detection systems, so coordinate with your security team before running scans against production infrastructure.

Schedule scans during maintenance windows when possible to minimize any impact on running systems. While Nmap is designed to be non-intrusive, scanning can still generate significant network traffic depending on the scope and scan type configured.

OS Patch Scan Provider

The OS Patch Scan Provider assesses system security by checking for missing patches and available updates. It analyzes installed software versions and compares them against known security vulnerabilities to identify systems requiring updates.

Configuration

Create an OS Patch Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "OS Patch". Configure the display name and owner. Like other scan providers, it requires no external credentials or API keys.

Patch Assessment

The provider evaluates system patch status by connecting to target systems and querying package managers for update information. It identifies the current patch level, finds missing security updates, and compares installed software versions against available updates. The assessment generates prioritized recommendations for patches that should be installed.

The scan results include three views: **All Patches** (default), showing both available and installed patches in one unified view, **Available Patches** for updates ready to install, and **Installed Patches** for currently applied updates. The All Patches view displays available patches first, making it easy to identify pending updates at a glance.

Running Scans

Execute OS Patch scans against Linux-based virtual machines and servers. Target systems must be accessible from Casibase and have package management tools properly configured.

From the Scans page, create a new scan and select the OS Patch provider. Choose your target asset from the inventory or enter system details manually. The provider edit page also includes a test widget for verifying connectivity and functionality before running production scans.

Remote Scanning

For distributed deployments, Casibase supports remote patch scanning across multiple machines. Deploy Casibase instances on each asset machine with a shared database. When a scan is created for a specific asset (matched by hostname to asset.displayName), the Casibase instance on that machine automatically detects and executes the scan. Results are saved to the shared database, allowing all instances to view scan progress and results in real-time.

Installing Patches

The scan interface includes patch installation capabilities directly from the scan results. Click the install button next to any available patch to initiate installation on the target system. The interface displays installation progress and status, updating in real-time as the patch is applied. Installation requires appropriate permissions on the target system.

Scan Results

Results provide patch assessment data formatted for review. The structured output includes available security updates, identifies package version discrepancies, and provides update priority recommendations. Installation impact analysis helps you understand what changes each update will make to the system.

This information appears in both structured format for quick scanning and raw output for detailed analysis. The web interface organizes the data to highlight critical security patches that require immediate attention versus routine updates.

System Requirements

Target systems should have network connectivity allowing Casibase to reach them. Package manager tools must be installed and functional on the target. The scan queries the package database to retrieve information about installed software and available updates.

Regular scanning helps maintain visibility into your security posture across infrastructure. Compare results over time to track how patch levels change and ensure critical updates are applied promptly. The historical record preserved in Casibase makes it easy to demonstrate compliance with patch management policies.

Nuclei Scan Provider

The Nuclei Scan Provider integrates the Nuclei vulnerability scanner into Casibase for automated security testing. Nuclei uses template-based scanning to detect vulnerabilities, misconfigurations, and exposed services across web applications and infrastructure.

Configuration

Create a Nuclei Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "Nuclei". The provider operates locally and requires the Nuclei binary to be installed on the Casibase server.

Vulnerability Scanning

Nuclei executes template-based vulnerability checks against target systems. Templates define specific vulnerability signatures, allowing for accurate detection of security issues ranging from common misconfigurations to critical CVEs. The scanner can test web applications, APIs, and network services using its extensive template library.

Using the Provider

From the Scans page, create a new scan and select the Nuclei provider. Enter your target URL or IP address for scanning.

Command Templates

The provider includes command templates for common scanning scenarios:

- Quick Scan: Fast scan with high-severity templates only
- CVE Scan: Focus on known CVE vulnerabilities
- Web Scan: Target web applications with HTTP templates
- All Templates: Comprehensive scan using all available templates

Custom Commands

Customize the Nuclei command to match your requirements. The command field supports the `%s` placeholder for the target. For example, `-u %s -severity critical,high` scans only for critical and high-severity issues. Input validation prevents command injection by blocking dangerous characters.

Use the test widget to verify scanning before running production scans. Enter a target, adjust the command if needed, and review the results.

Scan Results

Nuclei returns structured JSON output containing vulnerability findings. Each result includes:

```
{  
  "template-id": "cve-2021-12345",  
  "info": {  
    "name": "Vulnerability Name",  
    "severity": "high",  
    "description": "Detailed description"  
}
```

The web interface displays findings in organized tables showing template IDs, severity levels, matched locations, and vulnerability details. Color coding by severity helps prioritize remediation efforts.

Security Considerations

Run Nuclei scans only against systems you have authorization to test.

Vulnerability scanning can trigger security alerts and may impact system performance. Coordinate with your security team before scanning production systems.

Keep Nuclei templates updated regularly to detect the latest vulnerabilities. The scanner's effectiveness depends on having current vulnerability signatures in its template database.

ZAP Scan Provider

The ZAP Scan Provider integrates OWASP Zed Attack Proxy into Casibase for automated web application security testing. ZAP actively tests web applications to identify security vulnerabilities including injection flaws, broken authentication, and cross-site scripting.

Configuration

Create a ZAP Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "ZAP". The provider requires the ZAP daemon to be installed and running on the Casibase server.

Web Application Security Testing

ZAP performs active security testing by sending crafted requests to web applications and analyzing responses for vulnerability indicators. The scanner tests for OWASP Top 10 vulnerabilities and other common security issues. It can spider web applications to discover pages, then systematically test each endpoint for security flaws.

Using the Provider

From the Scans page, create a new scan and select the ZAP provider. Enter the target web application URL for testing.

Command Templates

The provider includes templates for different scanning modes:

- **Quick Scan:** Fast baseline scan with minimal configuration
- **Baseline Scan:** Standard security baseline assessment
- **Full Scan:** Comprehensive active scanning
- **AJAX Spider:** Advanced crawling for JavaScript-heavy applications

Custom Commands

Customize ZAP commands to target specific testing scenarios. The command field accepts the `%s` placeholder for the target URL. For example, `-quickurl %s -j` runs a quick scan with JSON output. Security validation prevents command injection attacks.

Test your configuration using the test widget before running production scans. Enter a target URL, modify the command as needed, and review results immediately.

Scan Results

ZAP returns detailed JSON output containing identified vulnerabilities:

```
{  
  "site": "https://example.com",  
  "alerts": [  
    {  
      "pluginid": "10021",  
      "alert": "X-Content-Type-Options Header Missing",  
      "severity": "INFO",  
      "uri": "/index.html",  
      "line": 1,  
      "col": 1  
    }  
  ]  
}
```

The web interface organizes findings by risk level (High, Medium, Low, Informational) with color-coded indicators. Each alert includes the affected URL, vulnerability description, solution recommendations, and CWE/WASC references for further research.

Security Considerations

Active security testing can modify application data and trigger security controls. Always obtain proper authorization before scanning web applications. Run scans against test environments when possible, and coordinate with application owners for production testing.

ZAP generates significant HTTP traffic during scans. Schedule testing during maintenance windows to avoid impacting users. Configure scan policies appropriately to balance thoroughness with testing time and resource consumption.

Subfinder Scan Provider

The Subfinder Scan Provider integrates subdomain discovery capabilities into Casibase through the Subfinder tool. Subfinder passively discovers subdomains using certificate transparency logs, search engines, and other public sources to map an organization's attack surface.

Configuration

Create a Subfinder Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "Subfinder". The provider operates locally and requires the Subfinder binary on the Casibase server.

Subdomain Discovery

Subfinder performs passive subdomain enumeration by querying various data sources without directly probing target infrastructure. This approach discovers subdomains quietly through certificate transparency logs, DNS databases, and search engine results. The tool aggregates results from multiple sources to provide comprehensive subdomain mapping.

Using the Provider

From the Scans page, create a new scan and select the Subfinder provider. Enter the target domain name for subdomain discovery.

Command Templates

The provider includes templates for different discovery modes:

- **Basic Scan:** Standard subdomain discovery
- **Silent Mode:** Minimal output for scripting
- **Recursive Scan:** Discover subdomains of found subdomains
- **All Sources:** Query all available data sources
- **Passive Only:** Strictly passive discovery

Custom Commands

Customize Subfinder commands for specific requirements. The command field supports the `%s` placeholder for the domain. For example, `-d %s -all -json` queries all sources with JSON output. Input validation prevents command injection by blocking shell metacharacters.

Use the test widget to verify discovery before running production scans. Enter a domain name, adjust the command if needed, and review discovered subdomains.

Scan Results

Subfinder returns JSON output containing discovered subdomains:

```
{  
  "host": "mail.example.com",  
  "source": "crtsh"  
}
```

The web interface displays results in sortable tables showing each discovered

subdomain and its data source. Color-coded source tags help identify which services provided each finding. Summary statistics show total subdomains discovered and source breakdown.

Reconnaissance Considerations

Subdomain discovery is typically legal since it uses only publicly available information. However, check local regulations and organizational policies before conducting reconnaissance. Discovered subdomains may reveal infrastructure details useful for security assessments or unauthorized access attempts.

Use discovery results responsibly. Subdomain mapping helps organizations understand their external attack surface, but the same information could assist malicious actors. Protect subdomain lists appropriately and use findings to improve security posture.

httpx Scan Provider

The httpx Scan Provider integrates HTTP service probing capabilities into Casibase through the httpx tool. httpx efficiently probes HTTP and HTTPS services to gather detailed information about web servers, technologies, and configurations.

Configuration

Create an httpx Scan Provider by navigating to Providers and adding a new provider with Category set to "Scan" and Type set to "httpx". The provider operates locally and requires the httpx binary on the Casibase server.

HTTP Service Probing

httpx probes HTTP and HTTPS services to extract detailed service information. The tool detects web servers, identifies technologies through fingerprinting, extracts titles and response headers, and determines TLS certificate details. It efficiently handles large-scale probing while respecting rate limits and connection constraints.

Using the Provider

From the Scans page, create a new scan and select the httpx provider. Enter target URLs or hosts for HTTP probing.

Command Templates

The provider includes templates for common probing scenarios:

- **Basic Probe:** Standard HTTP/HTTPS detection
- **Full Scan:** Comprehensive information gathering
- **Technology Detection:** Identify web technologies
- **Title Extraction:** Extract page titles
- **Status Only:** Quick service availability check

Custom Commands

Customize httpx commands for specific requirements. The command field supports the `%s` placeholder for targets. For example, `-u %s -tech-detect -json` enables technology detection with JSON output. Input validation prevents command injection attacks.

Use the test widget to verify probing before running production scans. Enter target URLs or hosts, adjust the command if needed, and review service information.

Scan Results

httpx returns structured JSON output containing service details:

```
{  
  "url": "https://example.com",  
  "status-code": 200,  
  "content-length": 1234,
```

The web interface displays results in organized tables showing URLs, status codes, response sizes, page titles, web servers, and detected technologies. Response time metrics help identify performance characteristics. TLS information displays certificate details for HTTPS services.

Network Considerations

Ensure Casibase has network connectivity to target services. Firewalls may need configuration to permit HTTP/HTTPS probing from the Casibase server. Rate limiting prevents overwhelming target services, but scanning can still generate significant traffic depending on scope.

HTTP probing is generally non-intrusive but may trigger monitoring alerts. Coordinate with service owners before scanning production infrastructure. Schedule scans appropriately to minimize any impact on service performance.

BlockChain Providers

Introduction

Blockchain technology provides an immutable and transparent ledger for data integrity verification. In Casibase, blockchain providers serve as a crucial security layer by uploading data to blockchain networks, ensuring that your knowledge base data cannot be tampered with or altered maliciously.

By leveraging blockchain's decentralized and cryptographic properties, Casibase can guarantee data authenticity and provide audit trails for all data modifications. This is particularly important for organizations that need to maintain data integrity compliance or require verifiable proof of data authenticity.

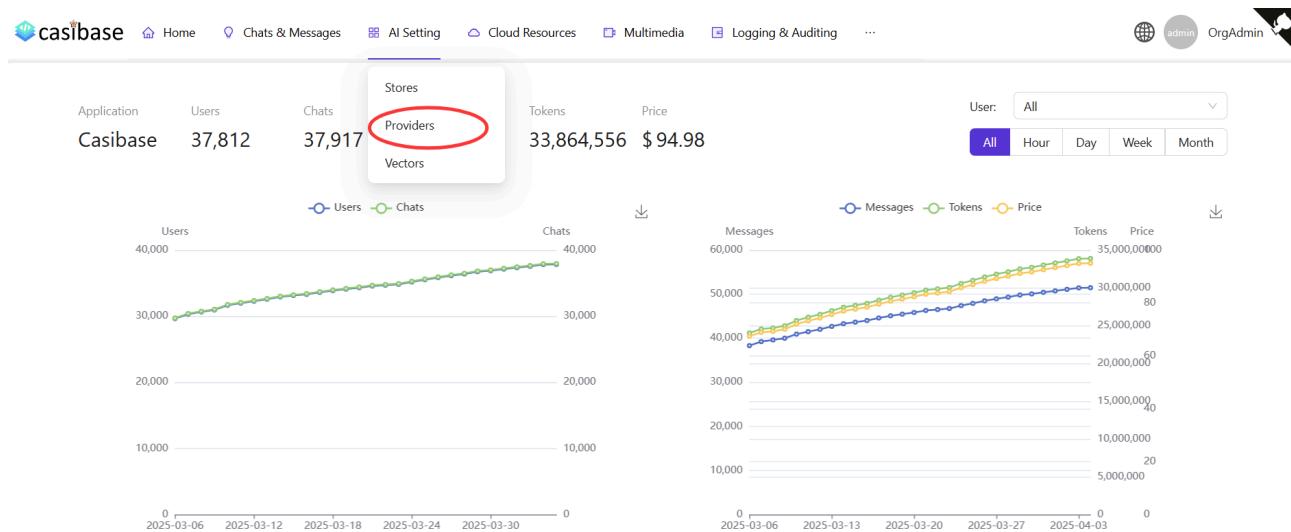
Refer to the [Core Concepts](#) section of our previous documentation for more information about providers.

In Casibase, you can add a blockchain provider by following these steps:

Add a New Blockchain Provider

Blockchain providers are used to integrate blockchain data integrity features into Casibase. You can add them by following these steps:

Click the [Providers](#) button on the page.



Add a Blockchain Provider

Click the [Add](#) button to add a blockchain provider.

Providers	Add	Add Storage Provider						
Name		Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice		Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker		Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjvqUrxnNtkKdfQvAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1		Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud		Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6sxilMg	***	cn-beijing
dall-e-3		dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3		Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision		Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

Fill in Blockchain Provider Details

Fill in the required configuration according to the blockchain you use, then click the **Save & Exit** button to save.

More information about the configuration can be found below:

Chainmaker Configuration

ChainMaker is a high-performance, enterprise-grade blockchain platform developed under the leadership of China. It is designed to provide secure, controllable, and sc...

Ethereum Configuration

Ethereum is a decentralized blockchain platform that enables smart contracts and decentralized applications (dApps). It is one of the most popular blockchain platforms...

Edit Provider		Save	Save & Exit
Name	eth_wm_test		
Display name	eth_wm		
Category	Blockchain		
Type	Ethereum		
Private key	***		
Contract Address	D4600b1B04b4FD07194476C35825175B30F0f9Ec		
Invoke method	save		
Browser URL	http://127.0.0.1:5051/txpage?blocknumber={bh}		
Provider URL	http://192.168.31.234:8545		
Is default	<input checked="" type="checkbox"/>		
State	Active		



提示

Casibase supports various blockchain networks for data integrity verification, including:

- [ChainMaker](#)
 - [ChainMaker](#)
 - [Tencent Chainmaker](#)
- [Ethereum](#)
 - Private networks (compatible with Ethereum JSON-RPC)
 - [Geth](#)
 - [Ganache](#)
 - Other Ethereum JSON-RPC compatible chains

Now, you can use the blockchain provider to ensure data integrity and prevent tampering.

After adding a blockchain provider, you can use it in Casibase to create immutable data records. The data records will be committed to the blockchain, providing proof of authenticity and preventing any unauthorized tampering.

Chainmaker Configuration

ChainMaker is a high-performance, enterprise-grade blockchain platform developed under the leadership of China. It is designed to provide secure, controllable, and scalable blockchain infrastructure for industries such as finance, government, and supply chain. ChainMaker supports multiple consensus mechanisms, smart contracts, privacy protection, and other features, meeting the needs of large-scale commercial applications. Its open-source, modular architecture allows developers to flexibly customize and extend functionalities according to actual business scenarios.

In this chapter, you will learn how to configure and use ChainMaker, including setting up storage providers and other operations, to help you quickly get started and apply the features of the ChainMaker platform.

1. Configuration field description

When configuring a ChainMaker provider in Casibase, you need to fill in several key fields. Each field has a specific meaning and is required for the correct integration with the ChainMaker blockchain. The following list explains the purpose of each field:

- `Name`: The unique identifier for this blockchain provider.
- `Display name`: The display name shown in the UI for this provider.
- `Category`: The type of service, here it should be `Blockchain`.
- `Type`: The blockchain type, here it should be `ChainMaker`.
- `orgId`: The organization ID in the ChainMaker network.

- `ChainId`: The chain ID of the ChainMaker blockchain.
- `AuthType`: The account mode. Currently, only `permissionedwithcert` is supported.
- `User key`: The user's private key for authentication in the ChainMaker server.
- `User cert`: The user's certificate for authentication in the ChainMaker server.
- `Sign key`: The user's private key for signing transactions in the ChainMaker server.
- `Sign cert`: The user's certificate for signing transactions in the ChainMaker server.
- `Node address`: The address of the ChainMaker node to connect to.
- `Contract name`: The name of the smart contract to interact with.
- `Invoke method`: The method name to invoke on the contract.
- `Browser URL`: The URL for viewing the blockchain in a browser.
- `Chainmaker endpoint`: The API endpoint for the ChainMaker service. See:
<https://github.com/casibase/chainmaker-server>

Please make sure to fill in each field accurately according to your ChainMaker deployment information. This will ensure that Casibase can successfully connect and interact with your ChainMaker blockchain.

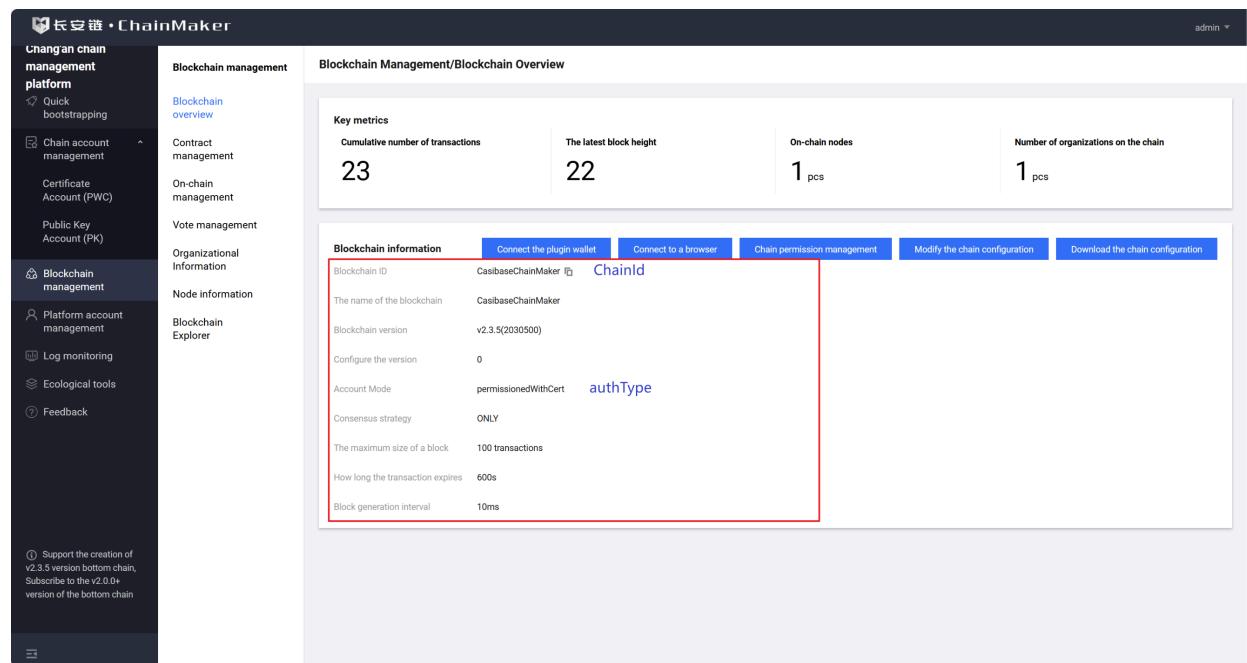
Next, we will use "Deploying ChainMaker via the Management Console" as an example for our introduction. If you have not yet deployed ChainMaker, please follow the [ChainMaker documentation](#) for deployment.

2. Configure ChainMaker

2.1 chainId, orgId and authType configuration

Obtaining Blockchain Information from the Web Panel

To retrieve blockchain information such as `chainId`, `orgId`, and `authType`, log in to the **ChainMaker Management Console** (Web panel). Navigate to the relevant blockchain management section, where these configuration details are displayed. Copy the required values and use them when configuring the ChainMaker provider in Casibase.



The screenshot shows the Chang'an chain management platform web interface. The left sidebar has sections for 'Chang'an chain management platform' (Quick bootstrapping, Chain account management, Certificate Account (PWC), Public Key Account (PK)), 'Blockchain management' (Blockchain overview, Contract management, On-chain management, Vote management, Organizational Information, Node information, Blockchain Explorer), and 'Platform account management', 'Log monitoring', 'Ecological tools', and 'Feedback'. A note at the bottom left says: 'Support the creation of v2.3.5 version bottom chain, Subscribe to the v2.0.0+ version of the bottom chain'. The main content area is titled 'Blockchain Management/Blockchain Overview' and displays 'Key metrics': Cumulative number of transactions (23), The latest block height (22), On-chain nodes (1 pcs), and Number of organizations on the chain (1 pcs). Below this is a 'Blockchain information' section with fields: Blockchain ID (CasibaseChainMaker), Name of the blockchain (CasibaseChainMaker), Blockchain version (v2.3.5(2030500)), Configure the version (0), Account Mode (permissionedWithCert), Consensus strategy (ONLY), Maximum size of a block (100 transactions), How long the transaction expires (600s), and Block generation interval (10ms). Buttons for 'Connect the plugin wallet', 'Connect to a browser', 'Chain permission management', 'Modify the chain configuration', and 'Download the chain configuration' are also present. The 'ChainId' and 'authType' fields are highlighted with red boxes.

Blockchain information	
Blockchain ID	CasibaseChainMaker
Name of the blockchain	CasibaseChainMaker
Blockchain version	v2.3.5(2030500)
Configure the version	0
Account Mode	permissionedWithCert
Consensus strategy	ONLY
The maximum size of a block	100 transactions
How long the transaction expires	600s
Block generation interval	10ms

Blockchain management

- Blockchain overview
- Contract management
- On-chain management
- Vote management
- Organizational information**
- Node information
- Blockchain Explorer

Blockchain management/organization information

Organization ID	The name of the organization	Creation time	Number of nodes
TestCMorg1 orgId	cmtestorg1	2025-06-10 10:18:33	1

Please enter the organization name to search Q

10 Article/page 1 / 1 page

2.2 Contract name and Invoke method configuration

Blockchain management

- Blockchain overview
- Contract management**
- On-chain management
- Vote management
- Organizational Information
- Node information
- Blockchain Explorer

Blockchain management/contract management

The name of the contract	Current version	Affiliation	Created by	Updated	Voting status	On-chain status	operate
casicbase Contract name	1.0.0	TestCMorg1	cmtestorg1	2025-06-10 13:21:08	normal	normal	freeze logout upgrade edit

Please enter the contract name to search Q

20 Article/page 1 / 1 page

Click the edit button in the "operate" column:

Blockchain management

- Blockchain overview
- Contract management**
- On-chain management
- Vote management
- Organizational Information
- Node information
- Blockchain Explorer

Edit the contract

The name of the contract ① casicbase

state The contract is initialized and deployed

Current version 1.0.0

The type of virtual machine ① WASM

Contract Invocation Method (Optional) ①

save	Invoke method	invoke	file_hash,file_name,time	increase
find_by_file_hash	Inquire	file_hash	Delete	

Are you sure Cancel

2.3 key, certificate, and nodeAddr configuration

The screenshot shows the 'Blockchain Management/Blockchain Overview' page. On the left, there's a sidebar with links: Blockchain management (Blockchain overview, Contract management, On-chain management, Vote management, Organizational Information), Node information, and Blockchain Explorer. The main area has a header 'Blockchain Management/Blockchain Overview'. Below it is a 'Key metrics' section with four cards: Cumulative number of transactions (23), The latest block height (22), On-chain nodes (1 piece), and Number of organizations on the chain (1 piece). Underneath is a 'Blockchain information' section with various parameters like Blockchain ID (CasibaseChainMaker), Version (v2.3.5(2030500)), and Account Mode (permissionedWithCert). A prominent blue button at the bottom right of this section is labeled 'Download the chain configuration', which is also highlighted with a red rectangular border.

Click the "download the chain configuration" button and extract the archive:

2.3.1 node addr

You can find the node_addr configuration in ~\CasibaseChainMaker\sdk_configs\sdk_config.yml.

Alternatively, you can check nodeAddr elsewhere, but note that the displayed port is the p2p port. The one we use should be the rpc port, which is the p2p port + 1000 by default.

The screenshot shows a sidebar with navigation options like Blockchain management, Blockchain overview, Contract management, On-chain management, Vote management, Organizational Information, Node information (which is selected and highlighted with a red box), and Blockchain Explorer.

The main content area is titled "Blockchain management/node information". It displays a table with columns: Node name, Affiliation, Node type, Node ID, Node address, Ledger synchronization type, and operate. One row is shown for "cmtestnode1" with affiliation "cmtestorg1", node type "Consensus nodes", node ID "Qmdcq5NhATkgqEi7q3Tvx...", node address "39.107.236.48:11301", ledger sync type "FULL", and operate status "operate".

A search bar at the top right says "Please enter the node name to search". At the bottom, there are pagination controls: "10 Article/page", "1 / 1 page", and navigation arrows.

A blue box contains the text: "The correct port is the current value plus 1000; i.e.: 12301".

2.3.2 user certificate and key

You can find the corresponding user certificate and key in `~\CasibaseChainMaker\ sdk_configs\crypto-config\TestCMorg1\user\cmtestuser1`. Fill in the provider fields accordingly.

cmtestuser1.sign.crt	2025/6/12 23:49	安全证书	1 K
cmtestuser1.sign.key	2025/6/12 23:49	KEY 文件	1 K
cmtestuser1.tls.crt	2025/6/12 23:49	安全证书	1 K
cmtestuser1.tls.key	2025/6/12 23:49	KEY 文件	1 K

2.4 Browser URL and ChainMaker endpoint

The Browser URL refers to the URL of the ChainMaker management console, which allows for quick access from Casibase in the future:

| <https://manage.casvisor.com/chains/CasibaseChainMaker/nodes?chainMode=permissionedWithCert> [Browser URL](#)

The screenshot shows the Casibase interface with the 'Logging & Auditing' tab selected. The main area displays a table of records with columns: Organization, ID, Name, Client IP, Created time, Sessions, Action, Block, and Action. The 'Sessions' column shows a list of sessions with their details. A red box highlights the 'Sessions' header, and a red arrow points to the 'Commit' button in the session log. A red callout box contains the text: 'If you have already committed, then the block will appear here, click on it to jump to the admin console'.

Organization	ID	Name	Client IP	Created time	Sessions	Action	Block	Action
casbin	9457	36fd22c8-1771-4083-9bc5-7cc401ed3a40	124.64.124.134	2025-06-13 00:39:	2025-06-13 00:38:57 / provider_blockchain update-provider	signin	Commit	<button>View</button> <button>Delete</button>
casbin	9456	f0bc2228-c10a-420e-90b9-535318658a9a	⋮1		2025-06-13 00:30:03 / provider_blockchain update-provider		Commit	<button>View</button> <button>Delete</button>
casbin	9455	438f014d-b808-40f6-bf52-62289d70d6f5	⋮1		2025-06-13 00:30:03 / provider_blockchain update-provider		Commit	<button>View</button> <button>Delete</button>
casbin	9454	18f40ece-d988-461b-9f4a-5d9cb80ff192	⋮1	2025-06-13 00:27:52	provider_blockchain update-provider		Commit	<button>View</button> <button>Delete</button>

The ChainMaker endpoint refers to the IP and port of the ChainMaker server that Casibase needs to connect to, for example: 127.0.0.1:13900

Ethereum Configuration

Ethereum is a decentralized blockchain platform that enables smart contracts and decentralized applications (dApps). It is one of the most popular blockchain platforms, supporting a wide range of applications from DeFi to NFTs. Ethereum uses a proof-of-stake consensus mechanism and provides a robust ecosystem for developers to build and deploy smart contracts.

① 信息

In this chapter, you will learn how to configure and use Ethereum, including setting up blockchain providers and other operations, to help you quickly get started and apply the features of the Ethereum platform.

1. Configuration field description

When configuring an Ethereum provider in Casibase, you need to fill in several key fields. Each field has a specific meaning and is required for the correct integration with the Ethereum blockchain.

Field Descriptions:

- **Name**: The unique identifier for this blockchain provider.
- **Display name**: The display name shown in the UI for this provider.
- **Category**: The type of service, here it should be **Blockchain**.
- **Type**: The blockchain type, here it should be **Ethereum**.
- **Private key**: The private key of the Ethereum account used for signing

transactions.

- **Contract Address**: The address of the smart contract to interact with on the Ethereum blockchain.
- **Invoke method**: The method name to invoke on the smart contract.
- **Browser URL**: The URL for viewing the blockchain in a browser, with block number template support.
 - Format: `http://127.0.0.1:5051/txpage?blocknumber={bh}` where `{bh}` will be replaced with the actual block number when visiting the block.
- **Provider URL**: The JSON-RPC endpoint URL for connecting to the Ethereum network (e.g., Geth, Ganache, or other node).

Please make sure to fill in each field accurately according to your Ethereum deployment information. This will ensure that Casibase can successfully connect and interact with your Ethereum blockchain.

2. Configure Ethereum

Example

Edit Provider		Save	Save & Exit
Name ⓘ :	eth_win_test		
Display name ⓘ :	eth_win		
Category ⓘ :	Blockchain		
Type ⓘ :	 Ethereum		
Private key ⓘ :	***		
Contract Address ⓘ :	D4600b1B04b4FD07194476C35825175B30F0f9Ec		
Invoke method ⓘ :	save		
Browser URL ⓘ :	http://127.0.0.1:5051/txpage?blocknumber={bh}		
Provider URL ⓘ :	http://192.168.31.234:8545		
Is default ⓘ :	<input checked="" type="checkbox"/>		
State ⓘ :	Active		

2.1 Provider URL Configuration

The Provider URL is the JSON-RPC endpoint that Casibase will use to communicate with the Ethereum network. This is the first and most important configuration as it establishes the connection to your Ethereum network.

Example Provider URL

```
http://127.0.0.1:8545
```

You can use:

- [geth](#): A popular Ethereum client that provides a JSON-RPC interface.
- [ganache](#): A personal blockchain for Ethereum development that can be used for testing and development purposes.
- other: Ethereum JSON-RPC compatible chains. Any other Ethereum-compatible chain that supports the JSON-RPC interface.

Example: Geth Dev Mode

To quickly experiment with Ethereum using Geth, you can start Geth in developer mode. This mode launches a local Ethereum node with instant mining and pre-funded accounts, making it ideal for testing and development.

```
geth --dev --http --http.api eth,web3,net --http.corsdomain  
"https://remix.ethereum.org"
```

This command starts a local Ethereum node with HTTP JSON-RPC enabled and sets the CORS domain to allow cross-origin requests from <https://remix.ethereum.org>. This configuration is suitable for online contract deployment and interaction using Remix web-based tools. You can use the default account (private key can be found in the `geth` console at first launch) and the endpoint (`http://127.0.0.1:8545`) as your Provider URL in Casibase for immediate testing and development.

```

WARN [07-19|02:42:46.376]      stored on a ramdisk, and will be lost if your machine is restarted.
WARN [07-19|02:42:46.376] 4. Mining is enabled by default. However, the client will only seal blocks if trans-
actions
WARN [07-19|02:42:46.376]      are pending in the mempool. The miner's minimum accepted gas price is 1.
WARN [07-19|02:42:46.376] 5. Networking is disabled; there is no listen-address, the maximum number of peers
is set
WARN [07-19|02:42:46.376]      to 0, and discovery is disabled.
WARN [07-19|02:42:46.376]
WARN [07-19|02:42:46.376]
WARN [07-19|02:42:46.376] Running in ephemeral mode. The following account has been prefunded in the genesis
:
WARN [07-19|02:42:46.376]      Account
WARN [07-19|02:42:46.376] -----
WARN [07-19|02:42:46.376]      0x71562b71999873db5b286df957af199ec94617f7 (10^49 ETH)
WARN [07-19|02:42:46.376]      Private Key
WARN [07-19|02:42:46.376] -----
WARN [07-19|02:42:46.376]      0xb71c71a67e1177ad4e901695e1b4b9ee17ae16c6668d313eac2f96
WARN [07-19|02:42:46.376]
INFO [07-19|02:42:46.376] Starting peer-to-peer node           instance=Geth/v1.16.1-stable-12b4131f/linux
-amd64/goi.24.4
WARN [07-19|02:42:46.376] P2P server will be useless, neither dialing nor listening
INFO [07-19|02:42:46.381] IPC endpoint opened           url=/tmp/geth.ipc
INFO [07-19|02:42:46.381] HTTP server started           endpoint=127.0.0.1:8545 auth=false prefix=
cors=https://remix.ethereum.org vhosts=localhost
INFO [07-19|02:42:46.382] New local node record           seq=1,752,864,166,381 id=5a498da1b5df4f0c i
p=127.0.0.1 udp=0 tcp=0
INFO [07-19|02:42:46.382] Started P2P networking           self=enode://2947b9f976fea97f00cf1be7e58b88
995a40f02daacb1eb6052fd298e7acb9e52e7481686d3f6101762a7a48e5b639e1540db8d958baff182b2bfdafb8a79e04@127.0.0.1:0
INFO [07-19|02:42:46.382] Started log indexer

```

Understanding the Console Output:

When you run Geth in dev mode, the console will display important information as shown in the image above:

- Private Key:** The console shows the private key of the pre-funded account that you can use for testing. This key is automatically generated and displayed in the console output.
- HTTP Endpoint:** The console confirms that the HTTP JSON-RPC server is running on `http://127.0.0.1:8545`. This is the endpoint address you should use as your Provider URL in Casibase.
- Account Address:** The corresponding Ethereum address for the generated private key is also displayed.

Important: Copy and save these values immediately as they are essential for configuring your Casibase provider. The private key will be needed for the `Private key` field, and the HTTP endpoint will be your `Provider URL`.

For more details, see the [Geth Dev Mode documentation](#).

2.2 Private key Configuration

The private key is essential for signing transactions on the Ethereum blockchain.



The private key should be provided without the hexadecimal prefix `0x`.

Example Private Key

```
# Example private key in geth dev mode (without 0x prefix)
b71c71a67e1177ad4e901695e1b4b9ee17ae16c6668d313eac2f96dbcda3f291
```



This is just an example private key for demonstration purposes. Never use this key in production or for real funds!

You can obtain your private key from various sources:

- **Initially generated by Geth:** When you start Geth in `dev` mode, it generates a pre-funded account with a private key displayed in the console.
- **Ethereum Clients:** Generate a new account using Ethereum clients like Geth.



Casibase will use `***` to replace the private key on the frontend after the submission.

2.3 Invoke Method and Contract Address Configuration

Invoke Method Configuration

The invoke method is the specific function name in the smart contract that you want to call.

In Casibase, your smart contract should implement specific methods to ensure compatibility:

- `save`: This method is used to store data in the contract. It should accept parameters as a tuple (struct).

```
struct DataItem {
    string key;
    string field;
    string value;
}

// Define event, returns key, field, value in order
event DataSaved(string key, string field, string value);

// Save struct data and emit an event for tracking
function save(DataItem memory _data) public {
    emit DataSaved(_data.key, _data.field, _data.value);
}
```

Method name `save` can be customized, but it should accept a struct as an argument.

You can refer to the [Example](#) to see how to implement the `save` method in your smart contract.

Make sure your contract includes these methods to enable seamless integration with Casibase.

Contract Address Configuration

```
# Example contract address (without 0x prefix)
c36fED2CE2E1Bb14b330465f4498D4892C8ee194
```

The contract address is the deployed smart contract's address on the Ethereum blockchain. You can obtain the contract address after deploying a smart contract.

Example for Contract Deployment Reference

To deploy a smart contract on Ethereum, you can refer to the [Casibase/contract-storage-eth](#). This repository provides sample Solidity contracts and deployment scripts using Go and Remix.

Getting Started with the Repository:

1. **Get Example Code:** Clone or download the repository to access sample Solidity contracts and deployment scripts.
2. **Contract Compilation:** Pre-compiled contract artifacts (ABI and bytecode) are available in the [releases](#) section of the repository.
3. **Setup for Go Script Deployment:** If using the Go deployment script, download the contract artifacts from releases and place them in the `build/` folder within the Go script's working directory.
4. **Deployment Options:** You can deploy the contract using either:
 - **Go Script:** Use the provided Go deployment script in the repository for programmatic deployment (requires contract artifacts in `build/` folder)
 - **Remix IDE:** Deploy contracts online using [Remix](#) with the contract source code

After deployment using either method, you can obtain the contract address from the deployment output.

Use the go script in the reference to deploy the contract

```
$ go run deploy.go
Starting contract deployment...
Connected to Ethereum node: http://192.168.31.234:8545
Deploying from address: 0x71562b71999873DB5b286dF957af199Ec94617F7
Loaded bytecode from: build/SaveContract.bin
Loaded ABI from: build/SaveContract.abi
Gas price: 9 wei
Gas limit: 0
Deploying contract...
Transaction sent: 0xf377a667d3216a1a45b3c3d0944745ea1cbe8ab17745f6e95c44d4f7a5a3fd8f
Contract address: 0xc36fED2CE2E1Bb14b330465f4498D4892C8ee194
Waiting for transaction confirmation...
Contract deployed successfully!
Gas used: 611787
Block number: 198

Running contract test...
Calling save function with: key=test_key_123, field=test_field, value=test_value_456
Save transaction: 0xb060530b6de01bd0537595b86dd0ebcac9007ebab3f24c096c42857ac6fdb3f2
Save function called successfully!
Retrieved data - Key: test_key_123, Field: test_field, Value: test_value_456
Log data - Key: test_key_123, Field: test_field, Value: test_value_456
```

Or you can use the block explorer to find the contract address.

Transaction Dashboard

TOTAL NO. OF TRANSACTIONS 1	BLOCK NUMBER 198	BLOCK HASH 0x58ddd0285c82aca 9676f47e96b2f6f08 7ddc83c4155f5b88 6373a115f466cb5	Transaction Status SUCCESSFUL
--------------------------------	---------------------	---	----------------------------------

Transaction Overview

Transaction Hash	0xf377a667d3216a1a45b3c3d0944745ea1cbe8ab17745f6e95c44d4f7a5a3fd8f
Transaction Gas	617810
Transaction Gas Price	9
Transaction Nonce	197
Transaction To	0xc36fED2CE2E1Bb14b330465f4498D4892C8ee194 [CONTRACT CREATION]
Transaction From	0x71562b71999873DB5b286dF957af199Ec94617F7
Transaction Value [wei]	0
Transaction Status	SUCCESSFUL

2.4 Browser URL Configuration

The Browser URL lets you view specific blockchain blocks and transactions in a web browser. By using a template with the `{bh}` placeholder, Casibase can automatically redirect you to the corresponding block details in your chosen blockchain explorer.

`http://127.0.0.1:5051/txpage?blocknumber={bh}`

Organization	ID	Name	Client IP	Created time	Provider	User	Method	Request URI	Action	Block	Action
built-in	115	0af1c434-a708-4238-a55a-5ae322b2f3f2	127.0.0.1	2025-07-19 03:08:20	eth_win_test	admin	POST	/api/signin?code=f33196946c	signin	200	<button>Query</button> <button>View</button> <button>Delete</button>
built-in	111	aea40549-4bfb-41a4-a9cd-51fb4004fe49	::1	2025-07-14 00:20:36	eth_win_test	admin	POST	/api/update-provider	update-provider		<button>Commit</button> <button>View</button> <button>Delete</button>
built-in	110	88cfef21e-b4e3-4c8b-8b37-8ca5941fed55	::1	2025-07-14 00:19:24	eth_win_test	admin	POST	/api/delete-provider	delete-provider		<button>Commit</button> <button>View</button> <button>Delete</button>
built-in	109	cf93e75c-501d-4aa7-a350-70509a90ba4	::1	2025-07-14 00:19:21	eth_win_test	admin	POST	/api/delete-provider	delete-provider	25	<button>Query</button> <button>View</button> <button>Delete</button>
built-in	108	357977eb-4d44-44d5-8714-13f7aff5a2fe	::1	2025-07-14 00:19:02	eth_win_test	admin	POST	/api/update-provider	update-provider	24	<button>Query</button> <button>View</button> <button>Delete</button>
built-in	107	66b29d6d-2b16-4d6b-9349-6be11748bea7	::1	2025-07-14 00:03:25	eth_win_test	admin	POST	/api/signin?code=77a6534407	signin	23	<button>Query</button> <button>View</button> <button>Delete</button>

💡 TEMPLATE FOR BROWSER URL

When you use the `{bh}` placeholder in the Browser URL template, Casibase will replace it with the actual block number and allow you to jump directly to the relevant block information in your blockchain explorer.

Example: Ganache CLI Block Explorer

To quickly view Ethereum blocks and transactions, you can use the open-source blockchain explorer [casibase/ganache-cli-block-explorer](#). This tool provides a simple web interface for browsing blocks, transactions, and contract events on your local Ethereum node.

The screenshot shows the Ganache Block-Explorer interface. On the left, there's a sidebar with a logo, the title "GANACHE BLOCK-EXPLORER", and two menu items: "Dashboard" and "Menu". The main area is titled "Transaction Dashboard" and contains four cards: "TOTAL NO. OF TRANSACTIONS" (1), "BLOCK NUMBER" (200), "BLOCK HASH" (0x65594c64b90b5805e0c876786b8319f5838179d6d0c314e53e3272b588fc83be), and "Transaction Status" (green). Below this is a section titled "Transaction Overview" with a table:

Transaction Hash	0xd05c3fb8aa26d168ae6d3cdb88948a6487282813afb738f2ec7773d86ae63a6a
Transaction Gas	336909
Transaction Gas Price	9
Transaction Nonce	199
Transaction To	0xD460001B04b4FD07194476C35825175B30F09Ec
Transaction From	0x71562b71999873DB5b286dF957af199Ec94617F7
Transaction Value [wei]	0
Transaction Status	SUCCESSFUL

Note: This explorer is based on [vivekganesan01/ganache-cli-block-explorer](#) and includes additional features contributed by Casibase.

After installation and startup, you can access block details directly from the above address as the Browser URL in Casibase.

Private Cloud Providers

Introduction

In Casibase, Private Cloud Providers act as a bridge, allowing you to connect to and manage various cloud-native resources, such as Docker and Kubernetes (K8s), directly from the Casibase interface. Their core objective is to provide a centralized dashboard for monitoring and operating your containerized services, integrating their management seamlessly into your Casibase workflow.

By configuring a provider, you enable Casibase to communicate with your private cloud or on-premises data center. This provides an ideal solution for organizations that want a unified interface to manage both their knowledge base and the infrastructure it runs on, enhancing operational efficiency and control.

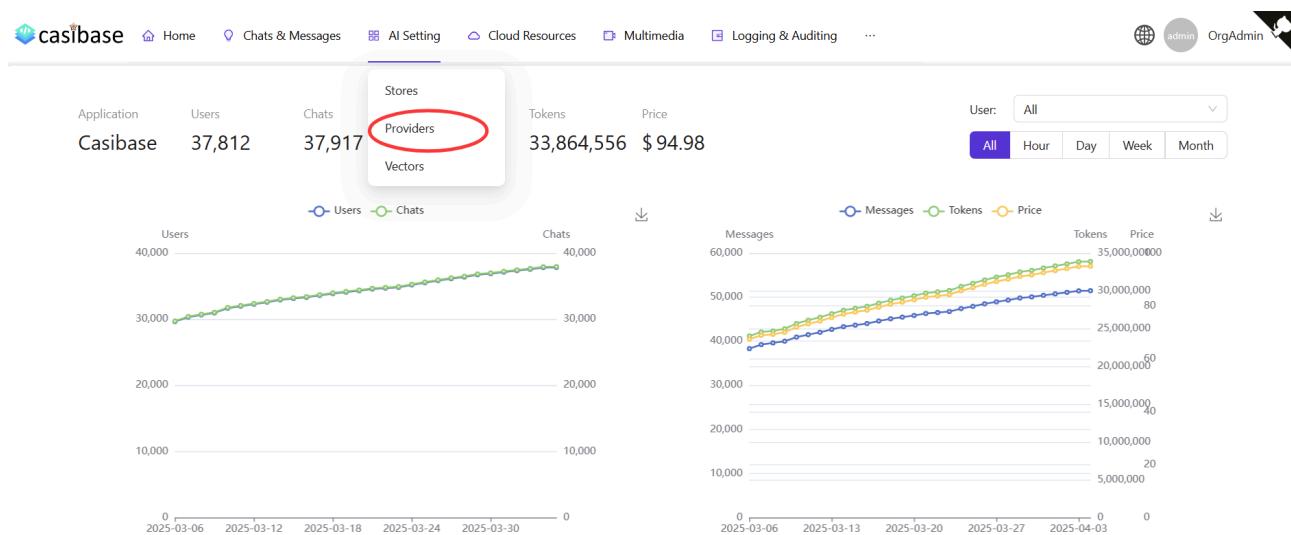
Refer to the [Core Concepts](#) section of our previous documentation for more information about providers.

In Casibase, you can add a private cloud provider by following these steps:

Add a New Private Cloud Provider

Private cloud providers are used to integrate cloud-native management features into Casibase. You can add them by following these steps:

Click the [Providers](#) button on the page.



Add a Private Cloud Provider

Click the [Add](#) button to add a new private cloud provider.

Providers	Add	Add Storage Provider						
Name		Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice		Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker		Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjvqUrxnNtkKdfQvAWcS1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1		Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud		Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHkUsopAioN6sxilMg	***	cn-beijing
dall-e-3	dall-e-3	Model	OpenAI	dall-e-3			***	
provider_model_azure_gpt4_1		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3		Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision		Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

Fill in Private Cloud Provider Details

Fill in the required configuration according to the cloud-native platform you use, then click the **Save & Exit** button to save.

More information about the configuration can be found below:

Kubernetes Configuration

Kubernetes (K8s) is an open-source container orchestration platform for automating the deployment, scaling, and management of containerized applications. It has bec...

提示

Casibase supports several mainstream cloud-native technologies and platforms, including:

- Docker:** To connect to a Docker host and manage the lifecycle of its containers (e.g., start, stop, view status) directly within Casibase.
- Kubernetes (K8s):** To connect to a Kubernetes cluster and manage its resources, such as Pods and Deployments, providing a high-level orchestration view within Casibase.

Now, you can use this private cloud provider to monitor and manage services in your cloud-native environment.

After adding a private cloud provider, you can use Casibase as a control panel to oversee your containerized applications, simplifying management and providing a unified view of your services alongside your knowledge base.

Kubernetes Configuration

Kubernetes (K8s) is an open-source container orchestration platform for automating the deployment, scaling, and management of containerized applications. It has become the de facto standard for managing applications in modern, cloud-native environments. By providing a robust framework for running distributed systems resiliently, Kubernetes simplifies complex operational tasks.

In this chapter, you will learn how to configure and use a Kubernetes provider in Casibase. This will allow you to connect Casibase to your Kubernetes cluster, enabling you to monitor and manage your cloud resources directly from the Casibase interface.

1. Configuration Field Description

When configuring a Kubernetes provider in Casibase, you need to fill in several key fields. Each field has a specific meaning and is required for the correct integration with your Kubernetes cluster. The following list explains the purpose of each field:

- `Name`: The unique identifier for this private cloud provider.
- `Display name`: The display name shown in the UI for this provider.
- `Category`: The type of service; here it should be `Private Cloud`.
- `Type`: The cloud-native platform type; here it should be `Kubernetes`.
- `Config text`: The raw text content of your `kubeconfig` file, which contains the credentials and endpoint information needed to connect to your Kubernetes cluster.

Please make sure to fill in each field accurately. The `Config text` is crucial for

establishing a successful connection.

2. Configure Kubernetes

The primary method for connecting Casibase to your Kubernetes cluster is by using your `kubeconfig` file.

2.1 Obtain Your Kubeconfig File

Before proceeding, you must ensure that the `kubeconfig` file you intend to use can successfully connect to your Kubernetes cluster. A reliable way to get the raw configuration is to run the following command in your terminal:

```
kubectl config view --raw > kubeconfig.yaml
```

This command will save the complete, flattened configuration into a file named `kubeconfig.yaml` in your current directory. You can then open this file to copy its contents.

You can test your configuration file with a command that checks for pods across all namespaces. This is a more reliable test to confirm connectivity.

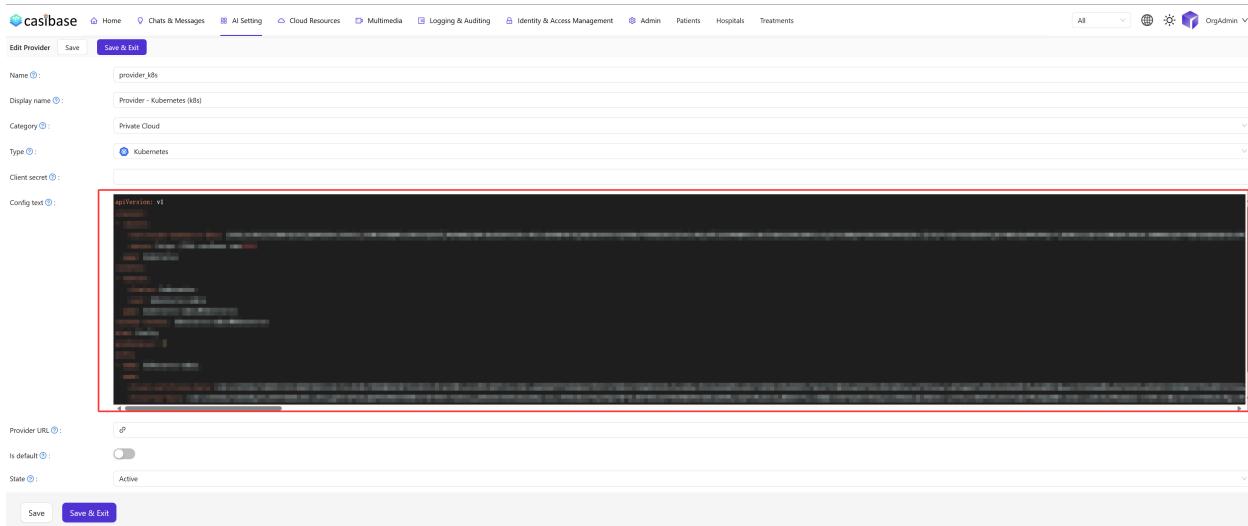
```
kubectl --kubeconfig=./kubeconfig.yaml get pods -A
```

If this command runs successfully (even if it just lists pods from system namespaces), you are ready to proceed.

2.2 Fill in the Provider Details

Copy the entire content of your valid `kubeconfig.yaml` file and paste it into the

`Config text` field in the provider configuration form.



3. Verify the Connection

After you have filled in the details and saved the provider, you can verify if the connection was successful.

Navigate to the Cloud Resources > Applications section within Casibase. Here, you will see a list of your configured providers. Check the status of the Kubernetes provider you just added.

- **Active:** If the status is `Active`, Casibase has successfully connected to your Kubernetes cluster.
- **Inactive:** If the status is `Inactive`, there was an issue with the connection. Please double-check the content of your `Config text` and ensure that there is network connectivity between Casibase and your Kubernetes cluster's API server.

Applications								Status: Active		Actions		
	Name	Display name	Created time	Description	Template	Status	Namespace	Action	Action	Action		
<input type="checkbox"/>	application_grekm	New Application - grekm	2025-08-02 21:59:58		template_2	Running	casibase-application_grekm	<button>Edit</button>	<button>Undeploy</button>	<button>Delete</button>		
<input type="checkbox"/>	application_jijkl	New Application - jijkl	2025-08-02 09:43:55		template_2	Running	casibase-application_jijkl	<button>Edit</button>	<button>Undeploy</button>	<button>Delete</button>		
<input type="checkbox"/>	application_x04b7i	New Application - x04b7i	2025-08-01 23:44:27		template_2	Not Deployed	casibase-application-x04b7i	<button>Edit</button>	<button>Deploy</button>	<button>Delete</button>		

3 in total | [1](#) > 10 / page

Once the connection is active, you can begin to monitor and manage your Kubernetes resources through Casibase.

Public Cloud Providers

Introduction

Public Cloud Providers enable Casibase to connect to and scan cloud infrastructure across major cloud platforms. The system automatically discovers and catalogs cloud resources, providing centralized visibility and management of your cloud assets.

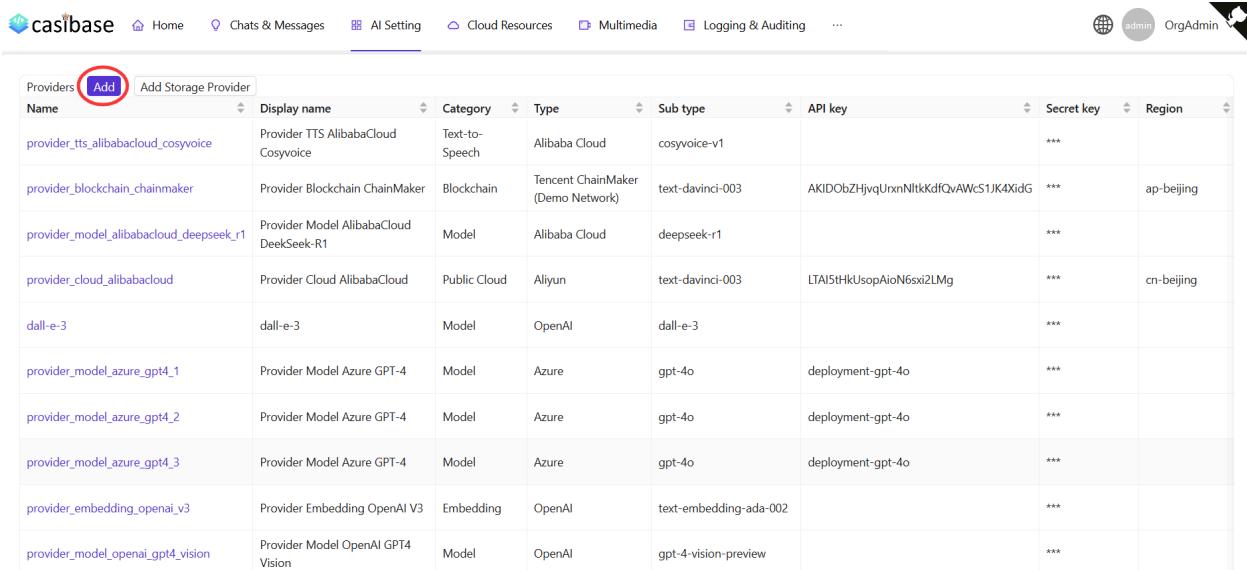
When configured, Casibase can scan resources like virtual machines, storage volumes, and network components, enriching each asset with detailed configuration information through multi-level API calls.

Supported Cloud Platform

Casibase currently supports Alibaba Cloud for complete asset discovery including ECS instances, disks, VPCs, and other resources. Additional cloud platforms are planned for future releases.

Add a Public Cloud Provider

Navigate to the Providers page and click  to create a new provider.



Providers	Add	Add Storage Provider						
Name		Display name	Category	Type	Sub type	API key	Secret key	Region
provider_tts_alibabacloud_cosyvoice		Provider TTS AlibabaCloud Cosyvoice	Text-to-Speech	Alibaba Cloud	cosyvoice-v1		***	
provider_blockchain_chainmaker		Provider Blockchain ChainMaker	Blockchain	Tencent ChainMaker (Demo Network)	text-davinci-003	AKIDObZHjqUrxnNltkKdfQvAWcs1JK4XidG	***	ap-beijing
provider_model_alibabacloud_deepseek_r1		Provider Model AlibabaCloud DeepSeek-R1	Model	Alibaba Cloud	deepseek-r1		***	
provider_cloud_alibabacloud		Provider Cloud AlibabaCloud	Public Cloud	Aliyun	text-davinci-003	LTAI5tHKUsopAioN6xi2LMg	***	cn-beijing
dall-e-3		dall-e-3	Model	OpenAI	dall-e-3		***	
provider_model_azure_gpt4_1		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_2		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_model_azure_gpt4_3		Provider Model Azure GPT-4	Model	Azure	gpt-4o	deployment-gpt-4o	***	
provider_embedding_openai_v3		Provider Embedding OpenAI V3	Embedding	OpenAI	text-embedding-ada-002		***	
provider_model_openai_gpt4_vision		Provider Model OpenAI GPT4 Vision	Model	OpenAI	gpt-4-vision-preview		***	

Configuration

Select "Aliyun" (Alibaba Cloud) as the provider type and fill in the required credentials:

- Type: Select "Aliyun" (Alibaba Cloud)
- Client ID: Your Alibaba Cloud Access Key ID
- Client Secret: Your Alibaba Cloud Access Key Secret
- Region: The cloud region to scan (e.g., cn-hangzhou)

Click **Save & Exit** to complete the configuration.

Asset Scanning

After adding a public cloud provider, Casibase can scan and catalog your cloud resources. The scanning process:

1. Connects to the cloud platform using provided credentials
2. Discovers resources using the SearchResources API

3. Enriches resource details through secondary API calls
4. Stores asset information with full configuration details

Scanned Resources

For Alibaba Cloud, the following resource types are automatically discovered:

- ECS instances with CPU, memory, IP addresses, and OS details
- Disk volumes with size, type, and attachment information
- VPC networks with CIDR blocks and routing configuration
- Security groups and network interfaces

Viewing Assets

After scanning, view discovered assets in the Assets page. Each asset displays:

- Resource ID and display name
- Resource type and region
- Current state
- Tags and metadata
- Detailed properties specific to the resource type

More configuration details:

Alibaba Cloud Configuration

Configure Alibaba Cloud asset scanning

Alibaba Cloud Configuration

Overview

The Alibaba Cloud provider serves dual purposes in Casibase. For model integration, it connects you to Qwen's language models and DeepSeek's reasoning-focused models (v3, v3.1, v3.2, and R1) through Alibaba's Model Studio platform. When chatting with these models, you can enable web search to have the AI fetch current information from the internet—results appear inline with citations.

For infrastructure management, the same provider scans your Alibaba Cloud resources, discovering ECS instances, disks, VPCs, and other assets across your deployment.

Prerequisites

Before configuring the Alibaba Cloud provider, you need:

- An Alibaba Cloud account
- Access Key ID and Access Key Secret with appropriate permissions
- Access to at least one region where resources are deployed

Required Permissions

The Access Key should have permissions to call:

- `resourcecenter:SearchResources` - For resource discovery
- `ecs:DescribeInstances` - For ECS instance details
- `ecs:DescribeDisks` - For disk details
- `vpc:DescribeVpcs` - For VPC details

Configuration Steps

Create Access Key

1. Log in to Alibaba Cloud console
2. Navigate to AccessKey Management
3. Create a new Access Key pair
4. Save both the Access Key ID and Access Key Secret securely

Add Provider in Casibase

1. Open Casibase and navigate to Providers
2. Click the `Add` button
3. Fill in the configuration:
 - Name: A descriptive name for this provider
 - Type: Select "Aliyun" (this is the internal identifier for Alibaba Cloud)
 - Category: Select "Public Cloud"
 - Client ID: Paste your Access Key ID
 - Client Secret: Paste your Access Key Secret

- Region: Enter the target region (e.g., cn-hangzhou, cn-beijing)

4. Click **Save & Exit**

Asset Discovery

Once configured, Casibase performs two-level asset discovery:

Level 1: Resource Discovery

Uses the SearchResources API to scan all resource types in the specified region. This provides basic information about each resource including ID, type, region, and tags.

Level 2: Detailed Enrichment

For specific resource types, Casibase makes additional API calls to gather comprehensive details:

ECS Instances

- Instance specifications (CPU cores, memory size)
- Operating system information
- Public and private IP addresses
- Billing method and instance status

Disk

- Disk category (cloud, cloud_efficiency, cloud_ssd)
- Size and attachment status
- Associated instance information

VPCs

- CIDR block configuration
- Associated VSwitches
- Routing table information

Resource Properties

Each discovered asset includes standard fields plus resource-specific properties. The properties are stored as key-value pairs and displayed in the asset detail view.

Common properties include instance types, network configurations, storage specifications, and billing information.

Scanning Multiple Regions

To scan resources in multiple regions, create separate providers for each region. This approach allows granular control over which regions are scanned and monitored.

Troubleshooting

No resources found: Verify the Access Key has correct permissions and the specified region contains resources.

Authentication errors: Check that the Access Key ID and Secret are correctly entered and the key is active.

Missing details: Ensure the Access Key has permissions for the detailed API calls (DescribeInstances, DescribeDisks, DescribeVpcs).



>

存储

存储

概述

存储概述

存储配置

在添加存储提供商、模型提供商和嵌入提供商后，我们可以配置存储

概述

1. 存储功能概述

在 Casibase 中，存储功能是其核心模块之一，它允许用户集成存储、建模和嵌入服务提供商，用于知识库数据存储、文本向量转换以及与聊天机器人的交互。通过存储功能，用户可以构建一个高效、灵活且强大的 AI 知识管理系统。借助 Stores 功能，用户可以构建一个高效、灵活且功能强大的 AI 知识管理系统。

2. 存储的优势

2.1 多模型集成

Casibase 的存储功能支持多个主流 AI 语言模型，包括 OpenAI（如 GPT-3.5、GPT-4）、Azure OpenAI、HuggingFace、Google Gemini 等。这种多模型支持允许用户根据具体需求选择最合适的 AI 模型，在性能、成本和功能之间找到平衡。这种多模型支持使用户能够根据特定需求选择最合适的 AI 模型，并在性能、成本和功能之间取得平衡。

2.2 多种存储和嵌入选项

用户可以自由选择存储和嵌入服务提供商，以满足不同的数据存储和处理需求。用户可以自由选择存储和嵌入服务提供商，以满足不同的数据存储和处理需求。这种灵活性使用户能够根据其技术栈和业务需求配置最合适的存储和嵌入解决方案。

2.3 多存储模式

Casibase 支持多存储模式，允许用户在不同的存储中使用不同的模型、存储和嵌入服务，为不同场景和用户提供定制化服务。这个功能使用户能够根据不同的业务需求灵活配置和切换存储。此功能使用户能够根据不同的业务需求灵活配置和切换 Stores。

2.4 Cross-Store Vector Sharing

Stores in Casibase can be configured to use vectors from other stores through the **Vector stores** field. This allows you to create a main store that searches across multiple specialized knowledge bases, or let different stores share their knowledge with each other. Instead of duplicating content, stores can dynamically access relevant information from other stores while maintaining their own separate vector databases.

2.5 Streamlined Management

The interface adapts to how you work. File-focused workflows can use the "Hide chat" toggle to clear away AI provider columns from the store list. Each store can also include example questions that appear when users start chatting, helping them understand what to ask without reading documentation.

2.6 Direct File Navigation

When reviewing knowledge sources in chat responses, you can click on any referenced file to open the store's file viewer with that specific document automatically selected and displayed. This seamless navigation lets you verify AI responses against source materials without manually searching through your document tree.

3. 总结

Casibase 的存储功能为用户提供了一个强大的知识管理工具，通过集成多个 AI 模型、存储和嵌入服务，使他们能够灵活地构建和管理知识库。其多存储模式和企业级功能进一步增强了系统的灵活性和安全性，适用于各种应用场景。其多 Stores 模型和企业级功能进一步增强了系统的灵活性和安全性，适用于各种应用场景。

Casibase 是一个开源的 AI 知识库系统，旨在为企业提供高效灵活的知识管理和对话解决方案。其核心功能之一是 Providers，允许用户整合多种 AI 模型和存储服务，以提升系统的功能和性能。Providers 分为三大类：Model Providers、Embedding Providers 和 Storage Providers，分别负责处理 AI 模型和数据存储。

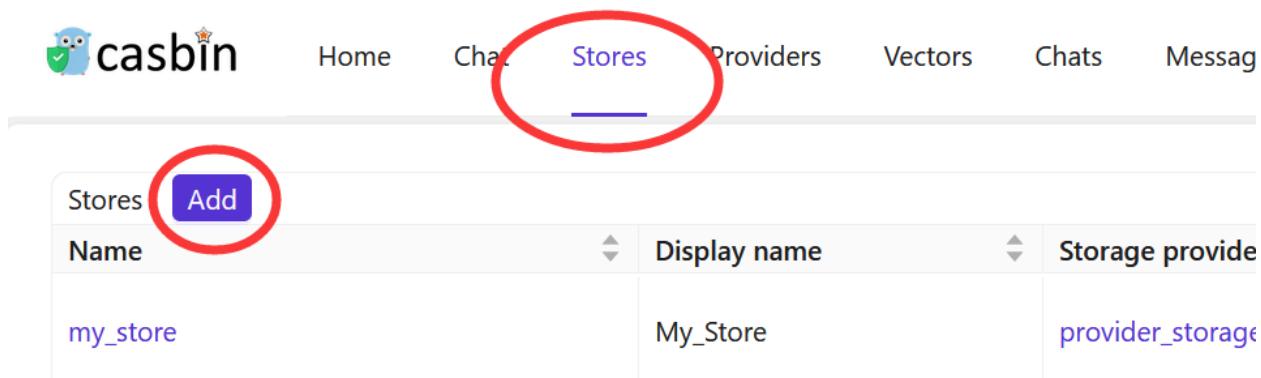
存储配置

在添加存储提供商、模型提供商和嵌入提供商后，我们可以配置存储

1. 添加新存储

存储用于将存储、模型和嵌入提供商集成到 Casibase 中。你可以按照以下步骤添加它们：您可以按照以下步骤添加它们：

点击主页上的 存储 按钮，然后点击 添加 按钮来添加存储。



Name	Display name	Storage provider
my_store	My_Store	provider_storage

填写存储详情

填写存储详情并点击 保存并退出 按钮。

casbin Home Chat Stores Providers Vectors Chats Messages Tasks Resources Permissions Logs Jimmy

Edit Store Save

Name: store_v6c22m

Display name: New Store - v6c22m

Storage provider:

Model provider:

Embedding provider:

File tree:



storage provider is empty

Go to Store

Save

Powered by Casibase

请选择之前添加的存储提供商、模型提供商、嵌入服务提供商、文本转语音服务提供商和语音转文本服务提供商。

Edit Store Save

Name:	<input type="text" value="my_store"/>
Display name:	<input type="text" value="My_Store"/>
Storage provider:	<input type="text" value="Provider_storage_1 (provider_storage_1)"/>
Model provider:	<input type="text" value="Model OpenAI text-davinci-003 (model_openai_text_davinci_003)"/>
Embedding provider:	<input type="text" value="Embedding_OpenAI_Adasimilarity (embedding_openai_adasimilarity)"/>
File tree:	<div style="border: 1px solid #ccc; padding: 5px;"> [] My_Store <ul style="list-style-type: none"> [] alibaba_oss <ul style="list-style-type: none"> [] audio <ul style="list-style-type: none"> [] AC / DC - Highway To Hell.mp3 (8.34 MB) [] document <ul style="list-style-type: none"> [] casdoor-knowledge.doc (18.0 KB) [] casdoor-knowledge.docx (10.9 KB) [] casdoor-knowledge.html (23.5 KB) [] casdoor-knowledge.md (2.12 KB) [] casdoor-knowledge.pdf (107 KB) [] image <ul style="list-style-type: none"> [] lena.jpg (105 KB) [] lena.tiff (768 KB) [] video <ul style="list-style-type: none"> [] my_video.mkv (456 KB) </div>

点击 保存并退出 按钮并返回存储列表页面：

Stores Add

Name	Display name	Storage provider	Model provider	Embedding provider	Action
my_store	My_Store	provider_storage_1	model_openai_text_davinci_003	embedding_openai_adasimilarity	View Refresh Vectors Edit Delete

The store list shows many columns for AI configurations like model providers, embedding providers, and voice settings. If you're mainly using stores for file management, toggle "Hide chat" at the top of the list to simplify the view and show only the essentials: name, storage provider, and status.

现在，你可以使用该存储来存储知识库数据、转换文本为向量，并与聊天机器人对话。

When configuring your store, consider adding example questions that appear when users start a new chat. These suggestions help users discover what they can ask and get them started quickly. You can also configure which other stores this one should search through using the **Vector stores** field - useful when you want one store to pull knowledge from multiple sources.

Vector Stores

Sometimes you need one store to search through knowledge from multiple other stores. For example, you might have separate stores for different topics or departments, but want a main store that can answer questions by searching across all of them. The **Vector stores** field makes this possible by letting a store use vectors from other stores in addition to its own.

To configure this, navigate to the store edit page and find the **Vector stores** field. You can select one or more stores from the dropdown list. When you chat with this store, it will automatically search through both its own vectors and the vectors from all the stores you selected. This way, you can create a centralized knowledge hub without duplicating content across multiple stores.



The store always uses its own vectors automatically. You only need to specify additional stores in the **Vector stores** field when you want to search across multiple knowledge bases.

在下一节中，我们将学习如何在 Casibase 中与聊天机器人对话。

2.File Management

Automatic Vector Generation

When you upload files to a store, Casibase automatically generates vector embeddings in the background. Each file goes through several states that you can track in the file list:

- Pending: File uploaded, waiting for embedding to start
- Processing: Currently generating vectors from the file content
- Finished: Vector generation completed successfully
- Error: Something went wrong during embedding (check the error text for details)

This happens without any manual intervention - just upload your files and let Casibase handle the rest. The system splits your documents into chunks, generates embeddings for each chunk, and stores them for fast retrieval during chat.

File Status Tracking

Navigate to the Files page to see all uploaded files across your stores. Each file shows its current status and token count. Files in the "Error" state display detailed error messages so you can troubleshoot issues like unsupported file formats or provider configuration problems.

When you delete a file from storage, Casibase automatically removes its associated vectors from the database. This keeps your vector database clean and prevents outdated information from appearing in chat responses.

Re-embedding Individual Files

Sometimes you need to regenerate vectors for a specific file - perhaps you've changed the embedding model or fixed a previous error. The "Refresh Vectors" button in the file list lets you re-embed individual files without affecting other content in the store.

Click the button next to any file to delete its existing vectors and generate fresh ones. The file status updates in real-time as processing completes, and the token count reflects the newly generated embeddings. This targeted approach saves time compared to refreshing the entire store when you only need to update a few files.

3. Store Isolation for Users

Casibase lets you restrict users to specific stores through Casdoor's Homepage field. When a user's Homepage matches a store name, they become bound to that store and work within it exclusively - perfect for multi-tenant setups where teams need separate knowledge bases.

Bound users see their assigned store locked in the top bar selector. They can view and query only that store's data, and API calls to other stores fail with an error. They also cannot add, delete, or rename stores, keeping their workspace stable and isolated.

To bind a user, edit their Casdoor profile and set the Homepage field to the exact store name (case-sensitive). The binding activates on their next login. Users with an empty Homepage or one that doesn't match any store keep full access to all stores, so you can mix restricted and unrestricted users easily.

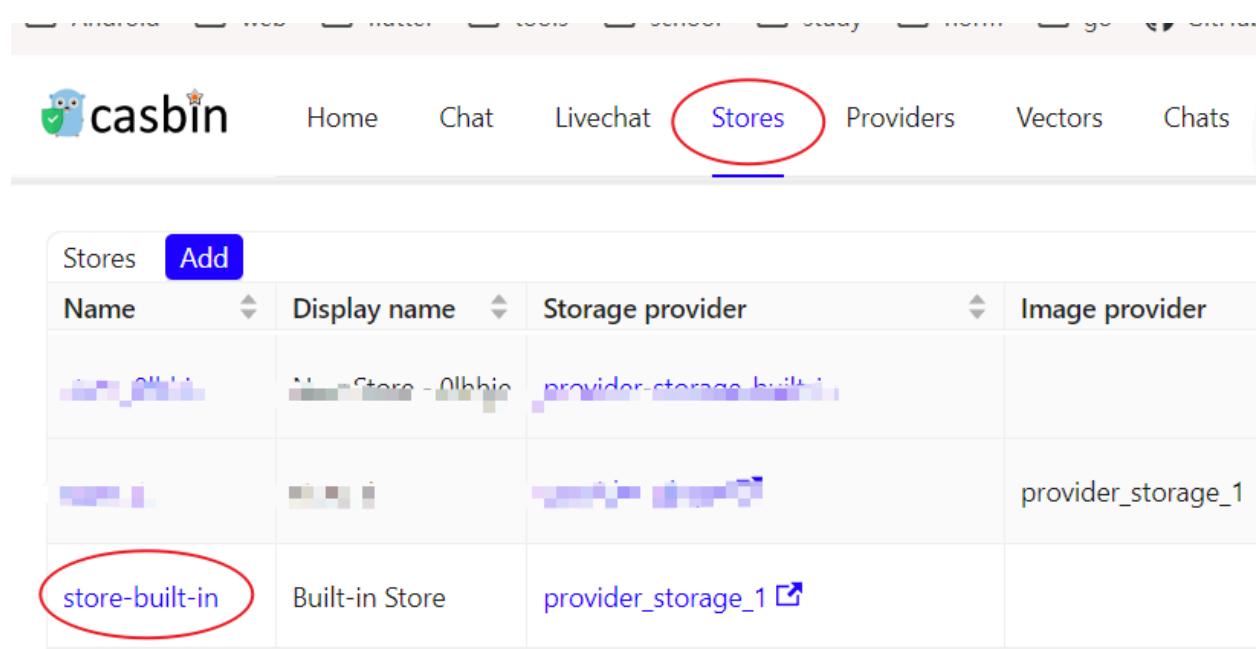
4. Support Multi-store

多存储模式为用户在每个不同的存储中提供不同的模型、建议等功能。

启用多存储

首先，你需要在内置存储中启用多存储模式。

点击主页上的 `存储` 按钮，然后点击 `store-built-in` 按钮进入内置存储。



Name	Display name	Storage provider	Image provider
store-built-in	Built-in Store	provider_storage_1	

向下滚动找到 `可选择存储` 字段，勾选它。

The screenshot shows a user interface for managing files. At the top, there is a message "No data". Below it, a "Suggestion count: 3" field is shown with a blue theme color swatch. A red oval highlights the "Can Select Store: ". The main area is a "File tree" with a search bar that says "Please input your search term". The tree structure includes a "Built-in Store" folder containing "casibase", which further contains "user_iaycd9" and "chat_1gs4l9" folders. "user_iaycd9" contains several files like "20240508_070749-1.png" and "20240508_070937-1.txt". "chat_1gs4l9" contains a single file "20240509_202051.jpg".

添加可用存储

多店模式仅提供可用的店铺。要使存储可用，您需要配置其存储提供商、模型提供商和嵌入提供商。

选择对话存储

Casibase 提供了一个非常方便的方法来选择存储。



Home Chat Livechat Stores Providers Vectors Chats Messages Usages Frameworks

+ New Chat

New Chat - 7

You are an expert in your field

Thank you for recognizing my expertise. Whether it's related to my specific area of knowledge or expertise to provide insightful answers and solutions to any problems you may have.

ReplyForwardShareSavePrint

store_1

store-built-in

New Chat - 8

New Chat - 5

New Chat - 8

New Chat - 7

只需将鼠标悬停在"新建对话"上，然后你就可以从下方出现的列表中选择你想要使用的存储。

如果你点击"新建对话"按钮，系统将为你分配一个默认存储。



>

Forms

Forms



Forms Overview

Introduction to Forms in Casibase



Form Configuration

How to configure and customize forms in Casibase

Forms Overview

Introduction

Forms in Casibase provide a powerful way to customize the display of list pages throughout the application. By configuring forms, you can control which columns are visible, their order, and their width in various list views such as Records, Providers, Stores, and more.

The Forms feature was ported from Casdoor and allows administrators to tailor the user interface to their specific needs, improving usability and focusing on the most relevant information.

Key Features

Customizable Column Visibility

Control which columns appear in list pages by toggling the visibility setting for each form item. This helps reduce clutter and focus on the most important data.

Adjustable Column Width

Set custom widths for each column to optimize the display based on the content type and your screen size preferences.

Column Ordering

Arrange columns in the order that makes the most sense for your workflow by reordering form items.

Multiple Form Categories

Forms support different categories to serve various purposes:

- **Table:** Traditional table-based forms for displaying structured data
- **iFrame:** Forms that can embed external content
- **List Page:** Forms specifically designed to customize list page columns

Form Structure

Each form in Casibase consists of:

- **Organization:** The organization that owns the form
- **Name:** Unique identifier for the form
- **Display Name:** Human-readable name shown in the interface
- **Position:** Placement or order of the form
- **Category:** Type of form (Table, iFrame, or List Page)
- **Type:** Specific form type (e.g., "records" for the Records list page)
- **URL:** Associated URL or endpoint (for Table and iFrame categories)
- **Form Items:** Collection of columns or fields to display

Form Items

Form items define the individual columns in a list page. Each form item includes:

- **Name:** Internal column identifier (e.g., "organization", "name", "createdTime")
- **Label:** Display label shown in the column header
- **Type:** Data type (currently "Text" for list pages)
- **Visible:** Whether the column is displayed
- **Width:** Column width in pixels

Use Cases

Forms are particularly useful for:

1. **Customizing Record Views:** Tailor the Records list page to show only relevant columns for your use case
2. **Simplifying Complex Tables:** Hide technical columns that aren't needed by all users
3. **Optimizing Screen Space:** Adjust column widths to fit more information on screen
4. **Role-Based Views:** Create different forms for different user roles or workflows

Getting Started

To start using Forms in Casibase:

1. Navigate to the Forms section in the Casibase admin interface
2. Create a new form or edit an existing one

3. Configure the form category, type, and items
4. Preview your changes in real-time
5. Save and apply the form to the corresponding list page

For detailed instructions on configuring forms, see the [Form Configuration](#) guide.

Form Configuration

Accessing Forms

Forms can be accessed through the Casibase admin interface:

1. Log in to your Casibase admin dashboard
2. Navigate to the Forms section from the main menu
3. You'll see a list of existing forms organized by category and type

Creating a New Form

Step 1: Basic Information

To create a new form:

1. Click the Add button on the Forms list page
2. Fill in the basic form information:
 - Organization: Select the organization (typically your organization name)
 - Name: Enter a unique identifier for the form
 - Display Name: Provide a human-readable name
 - Position: Set the position or order (optional)

Step 2: Select Form Category

Choose the appropriate category for your form:

- **Table:** For traditional table-based forms
- **iFrame:** For embedding external content
- **List Page:** For customizing list page columns (recommended for most use cases)

Step 3: Configure Form Type

If you selected **List Page** as the category:

1. Choose the **Type** from the dropdown menu
 - Currently supported: **Records**
 - More types will be added in future versions
2. The form will automatically populate with default columns for the selected type

Step 4: Customize Form Items

For each form item (column), you can configure:

Name

The internal identifier for the column. This corresponds to the data field being displayed.

Available columns for Records include:

- `organization`: Organization name
- `id`: Record ID
- `name`: Record name
- `clientIp`: Client IP address

- `createdTime`: Creation timestamp
- `provider`: AI provider name
- `provider2`: Secondary provider
- `user`: Associated user
- `method`: HTTP method
- `requestUri`: Request URI
- `language`: Language
- `query`: Query parameters
- `region`: Geographic region
- `city`: City
- `unit`: Unit information
- `section`: Section
- `response`: Response data
- `object`: Related object
- `errorText`: Error messages
- `isTriggered`: Trigger status
- `action`: Action column
- `block`: Block information
- `block2`: Secondary block information

Visible

Toggle to show or hide the column in the list page. Hidden columns are still available in the form configuration but won't appear in the UI.

Width

Set the column width in pixels. This helps optimize the display based on the content length and your screen size.

提示

- Use narrower widths (90-120px) for short fields like IDs, dates, and status indicators
- Use wider widths (200-300px) for longer text fields like names, descriptions, and URLs
- Adjust widths based on your typical content length to avoid truncation

Step 5: Reorder Columns

Arrange columns in your preferred order:

1. Use the Up arrow button to move a column up
2. Use the Down arrow button to move a column down
3. The leftmost columns will appear first in the list page

Step 6: Add or Remove Columns

- Add Column: Click the Add button to create a new custom column
- Remove Column: Click the Delete button next to a column to remove it
- Reset to Default: Click Reset to Default to restore the original column configuration

Step 7: Preview and Save

1. View the Preview section at the bottom of the form editor
2. The preview shows how your form will appear in the actual list page
3. Click on the preview to open the full list page in a new window
4. Once satisfied, click Save & Exit to apply your changes

Editing Existing Forms

To modify an existing form:

1. Navigate to the Forms list page
2. Click on the form name or the edit button
3. Make your desired changes
4. Click **Save & Exit** to apply the updates

Changes take effect immediately for all users viewing the corresponding list page.

Form Categories in Detail

List Page Forms

List Page forms are the most commonly used type in Casibase. They allow you to:

- Customize which columns appear in list views
- Control column order and width
- Show/hide columns based on user needs
- Create optimized views for different workflows



When configuring List Page forms:

- The **Action** column is always displayed at the end, regardless of form item configuration
- At least one column should be visible for the list page to be functional

- Column names must match the actual data fields available in the backend

Table Forms

Table forms are used for structured data display in table format. Configuration is similar to List Page forms but may have different available fields based on the data source.

iFrame Forms

iFrame forms allow you to embed external content or applications within the Casibase interface:

1. Set the URL field to the external content address
2. Configure display settings as needed
3. The content will be displayed in an embedded frame



>

向量

向量

概述

向量概述

向量生成

向量的生成需要与存储配合使用，这意味着你需要在理解向量之前配置存储。

概述

在 Casibase 中，向量是其核心优势之一。向量技术在知识表示和检索中起着关键作用，通过结合将文本和图像等数据转换为密集向量的 stores 功能，Casibase 实现了高效的相似性搜索和数据分析。

关于向量的定义，请参阅我们之前文档中的 [核心概念](#) 部分。

Casibase 中向量技术的应用

知识嵌入

用户可以上传各种格式的文件（例如 TXT、Markdown、Docx、PDF 等）并选择嵌入方法（例如 Word2Vec、GloVe、BERT 等）生成知识及相应的向量。这些向量存储在向量数据库中，以便快速检索和查询。

相似度搜索

Casibase 将知识转换为向量并存储在向量数据库中。这种向量表示支持强大的相似度搜索功能，使用户能够基于上下文或内容快速找到相关信息。这种向量表示支持强大的相似性搜索功能，使用户能够根据上下文或内容快速找到相关信息。

When you chat with a knowledge base, the system automatically finds and ranks the most relevant document fragments based on semantic similarity. Each chat response includes a knowledge sources button that shows exactly which parts of your documents were used to generate the answer, along with relevance scores displayed as percentages. This transparency helps you understand and verify the AI's reasoning while providing quick access to source materials.

向量生成

向量的生成需要与存储配合使用，这意味着你需要在理解向量之前配置存储。

向量实际上是嵌入的结果，即将各种数据（例如文本和图像）转换为密集向量表示的过程。此步骤对于促进 Casibase 内高效数据处理和分析至关重要。通过嵌入，聊天中的问题和存储中的知识文件将被转换为用于下一步知识搜索的向量。

1. Automatic Embedding

Casibase automatically generates vectors when you upload files to a store. After configuring a store with storage, model, and embedding providers, simply upload your files through the file tree interface. The system immediately queues them for embedding and processes them in the background.

Each file progresses through states you can monitor in the Files page: Pending, Processing, Finished, or Error. The system calculates token counts during embedding and displays them alongside each file. If embedding fails for any file, the error details appear in the file list so you can quickly identify and fix issues like missing provider configurations or unsupported file formats.

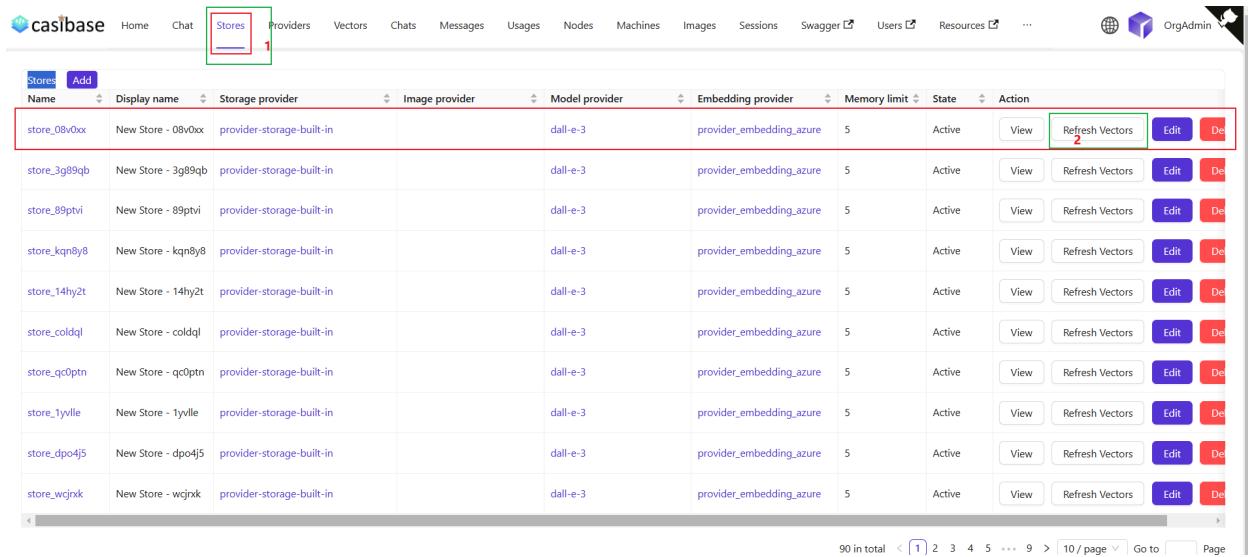
When you delete files from storage, their vectors are automatically removed from the database. This keeps your knowledge base synchronized - you won't get outdated responses from files that no longer exist.

2. 刷新向量

在 stores 菜单下的每个存储数据中，都设置了 "刷新向量" 操作按钮。After

configuring stores with storage providers, navigate to the [Stores](#) page to view the file tree for the storage providers.

By clicking on the Refresh Vectors button for a particular store, it will generate the corresponding vectors for all the files in the file tree for that store by embedding them. 下图显示了页面和操作。



Name	Display name	Storage provider	Image provider	Model provider	Embedding provider	Memory limit	State	Action
store_08v0xx	New Store - 08v0xx	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_3g89qb	New Store - 3g89qb	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_89ptvi	New Store - 89ptvi	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_kqn8y8	New Store - kqn8y8	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_14hy2t	New Store - 14hy2t	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_coldql	New Store - coldql	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_qc0ptn	New Store - qc0ptn	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_1yville	New Store - 1yville	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_dpo4j5	New Store - dpo4j5	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete
store_wcjnxk	New Store - wcjnxk	provider-storage-built-in		dall-e-3	provider_embedding_azure	5	Active	View Refresh Vectors Edit Delete

When you refresh vectors, the system automatically removes all existing vectors for that store before generating new ones. This ensures your vector database stays synchronized with your current files - if you've deleted files from storage, their vectors will be removed as well. The refresh always creates a clean, up-to-date vector set based on your current file tree.

Re-embedding Individual Files

For targeted updates, use the "Refresh Vectors" button in the Files page next to any specific file. This regenerates vectors for just that file while leaving others untouched - useful when you've changed embedding settings or need to fix errors on specific documents without reprocessing your entire knowledge base.

3. 查看向量

After that, we can view the specific vectors generated by that store in the vector menu.

Name	Store	Provider	File	Index	Text	Size	Data	Dimension	Action
vector_opfabu	store_08v0xx	provider_embedding_azure	casdoor-website/static/ads.txt	0	google.com, pub-3509678918753247, DIRECT, f08c47fec0942fa0	1536	[-0.011380682,-0.019052684,-0.017577048,-0.0364077...]	1536	<button>Edit</button> <button>Delete</button>
vector_mqhdg8	store_08v0xx	provider_embedding_azure	casdoor-website/docs/webhooks/overview.md	0	--- title: Overview description: Adding Webhooks in Casdoor	1536	[-0.0041976073,-0.0032014789,-0.0019043133.0.00293...]	1536	<button>Edit</button> <button>Delete</button>
vector_vagn2q	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/roles.md	0	--- title: User Roles description: Roles assigned to users k...	1536	[-0.005397057,0.0042889416,-0.0072895316,-0.035272...]	1536	<button>Edit</button> <button>Delete</button>
vector_elpwyu	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/permissions.md	0	--- title: Permissions description: User Permissions keyword...	1536	[0.0022299571.0.0068149953.0,0.028915738,-0.0544376...]	1536	<button>Edit</button> <button>Delete</button>
vector_ev350m	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/overview.md	2	- 'Baidu' - 'Casdoor' - 'Infoflow' - 'Apple' - 'AzureAD'	1536	[-0.009141525,0.021197738,-0.019946484,-0.02071195...]	1536	<button>Edit</button> <button>Delete</button>
vector_pj5q7q	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/overview.md	3	When migrating users from an external database to Casdoor, t...	1536	[-0.004490083.0.010098264,0.0035382845,-0.0212970...]	1536	<button>Edit</button> <button>Delete</button>
vector_7exw5a	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/overview.md	0	--- title: Overview description: Managing Users in Casdoor k...	1536	[0.014196097,0.017991772,0.009771907,-0.03401034...]	1536	<button>Edit</button> <button>Delete</button>
vector_8l70kd	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/overview.md	1	- 'Gender' - 'Birthday' - 'Education' - 'Score' - 'Karma' -	1536	[-0.008841777,0.034623653.0,0.8250997,-0.03711953...]	1536	<button>Edit</button> <button>Delete</button>
vector_tp40In	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/multi-factor-authentication.md	12	2. Then enter the code into the "Enter your code" field and ...	1536	[-0.00106068328,-0.0078019355,-0.025427068,-0.01174...]	1536	<button>Edit</button> <button>Delete</button>
vector_vqx3g6	store_08v0xx	provider_embedding_azure	casdoor-website/docs/user/multi-factor-authentication.md	13	3. Above the "Enable" button, copy your recovery codes and s...	1536	[-0.021876315.0.013569352,-0.0122898435,-0.0218238...]	1536	<button>Edit</button> <button>Delete</button>

我们可以看到，上一步刷新向量中的存储文件已被转换为向量在此显示。

我的向量编辑页面显示了具体信息，如存储名称、嵌入模型名称、进行嵌入的文件名、文件大小、维度、向量数据等。



>

文本切分器

文本切分器

概述

文本切分器概述

概述

文本切分器 是构建大型语言模型(LLM)应用程序的关键组件。它们的主要作用是将长文本分割成多个较短的段落，从而便于后续任务，例如文本嵌入、增强检索生成(RAG)和问答系统。

在大型语言模型中，文本分割主要基于以下几个原因：

- 提高效率和准确性：通过将大段文本分解为较小的段落，可以优化大型语言模型生成的嵌入向量的相关性和准确性。分块有助于确保嵌入内容在保留语义相关性的同时，噪音最小。例如，在语义搜索中，当为文档库建立索引时，每个文档都包含特定主题的有价值信息。采用有效的分块策略确保搜索结果能够准确捕捉用户查询的本质。
- 限制上下文窗口大小：在使用类似 GPT-4 的模型时，可处理的 token 数量是有限的。例如，GPT-4 的上下文窗口大小限制为 32K 个 token。虽然这一限制通常不会构成问题，但从一开始就考虑分块大小非常重要。如果文本块过大，可能会丢失信息或无法将所有内容嵌入上下文中，从而影响模型的性能和输出。
- 处理长文档：虽然长文档的嵌入向量可以捕捉整体上下文，但它们可能会忽略与特定主题相关的重要细节，导致输出不精确或不完整。分块使得对信息的提取和嵌入有更好的控制，从而降低信息丢失的风险。

Casibase目前提供多种文本切分方法，允许用户针对不同文本场景采用不同的处理策略。

默认文本切分器

默认文本切分器旨在根据标记数量和文本结构高效地分段文本。其分割策略包括：

- 逐行读取和段落识别：文本按行读取，通过连续空行准确确定段落分隔。它还通过标记敏感地识别自然断点，确保文本分割逻辑严谨且精确。
- 代码块的特殊处理：被```符号包围的代码块将被单独处理。代码块内的行数决定了它是否可以独立作为一个段落。这一机制既保持了代码块的完整性，又有效防止单个文本段落超过标记限制。
- 维护句子完整性：在整个分割过程中，严格保持句子完整，确保句子不会被分割。这一特性保证每个文本段落都包含一个完整的信息单元。不論文本多么复杂，分割均精确地在句子边界进行，有效避免因句子被拆分而引起的歧义和信息丢失。

问答切分器

问答切分器专注于对问答格式文本进行精确分割，并提供以下核心优势：

- 问答单元的准确切分：它采用逐行扫描机制，智能识别问答文本的结构。通过确定每行是否以“Q:”或“A:”开头，它能精确定位问题与答案之间的边界，确保每个问答对都被完整分割。这确保了每个问答单元的独立性和完整性，为后续的问答处理和分析提供了清晰的数据。
- 清晰且逻辑性强的实现：代码简洁直观，易于理解和维护。通过管理当前问答对的状态以及指示是否正在收集答案的标志，文本分割过程得到了清晰控制，确保每个问答单元正确配对。



聊天

聊天

概述

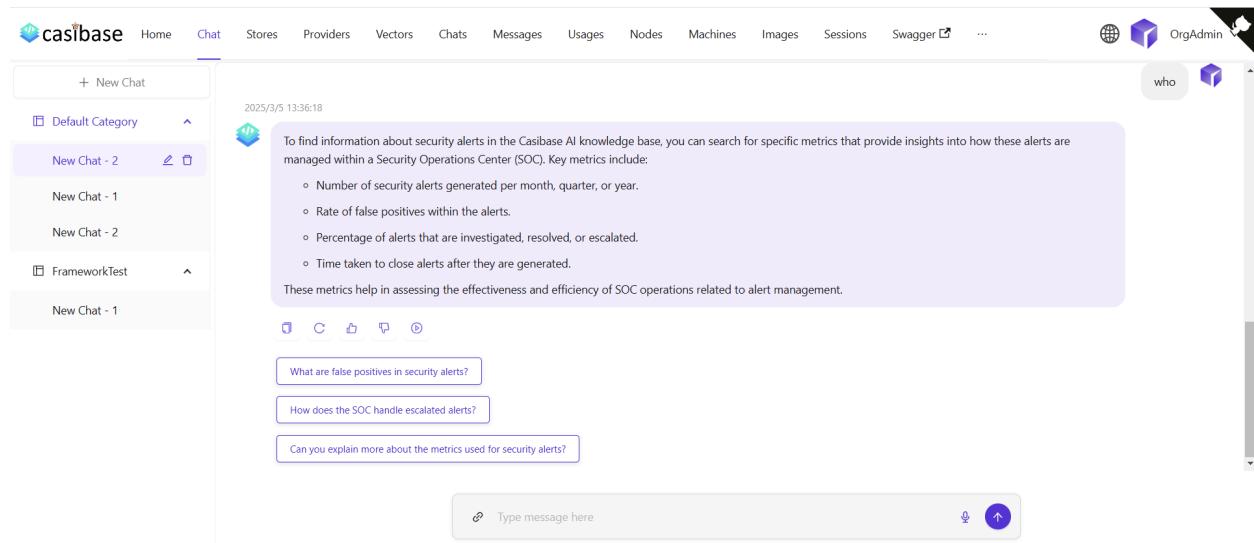
聊天概述

概述

在本节中，我们介绍Casibase最核心的部分：聊天及其管理。

1. 聊天

When you log in to Casibase, the chat interface is your starting point. This makes it easy to immediately begin conversations with AI without navigating through menus. 一旦我们配置好存储，我们就可以与AI进行对话。如下图所示：



The chat interface adapts to different AI capabilities. When using OpenAI or Alibaba Cloud providers, you'll see a web search toggle that lets the AI fetch current information from the internet. Search results appear inline with the response, complete with source URLs and timestamps for verification. For OpenAI's reasoning models (o1 and o1-mini), the chat displays the model's step-by-step thinking process in expandable sections, letting you follow along as it works through complex problems.

Source Attribution

When the AI provides answers, you can trace where the information came from through two types of sources:

Knowledge sources appear when chatting with a store that has embedded documents. Click the button showing the number of knowledge sources to open a drawer that displays each relevant document fragment with its relevance score. You can click on any source to view the complete vector entry, or click on the file name/icon to jump directly to that file in the store's file viewer with the specific document pre-selected.

Web sources appear when web search is enabled. These show the websites the AI referenced when answering your question, complete with titles, URLs, and brief descriptions. The drawer displays search results with timestamps and allows you to visit the original sources.

2. 聊天管理

我们可以从聊天菜单管理我们的聊天会话。

The screenshot shows the casibase Chats management interface. At the top, there are navigation links: Home, Chat, Stores, Providers, Vectors, Chats (highlighted with a red box), Messages, Usages, Nodes, Machines, Images, Sessions, Swagger, and more. On the right, there are icons for OrgAdmin and a user profile.

The main area displays a table of chat sessions. The columns include: Name, Updated time, User, Client IP, Count, Token count, Price, and Messages. A red box highlights the 'Chats' tab in the navigation bar and the 'Count' column header in the table.

Three rows of chat session data are shown:

- chat_j916c0**: Updated 2025-03-05 13:37:02 by user u-0b9800aa. Client IP: 119.164.218.30 (中国 山东 济南). Edge 133.0.0, Windows 10. Count: 1, Token count: 1006, Price: \$0.002705.
- chat_v67r4z**: Updated 2025-03-05 13:36:48 by user u-649ef853. Client IP: 101.129.8.189 (中国 台湾, N/A). Chrome 133.0.0, Mac OS X 10.15.7. Count: 1, Token count: 1004, Price: \$0.00269.
- chat_252ftr**: Updated 2025-03-05 13:36:18 by user admin. Client IP: ::1. Count: 12, Token count: 7107, Price: \$0.0203981.

A red box highlights the third row (chat_252ftr). To the right of this row, a modal window shows a message from the AI: "Hello! How can I assist you today?". Below the message are icons for reply, like, dislike, and share. The modal has a date header "2025/" and a footer "2025/ Delete".

该页面允许用户查看已创建的聊天信息，用户也可以点击编辑来查看或编辑它们。它们显示以下信息： Regular users can access and view their own chat sessions, while administrators have full access to all chats in the system for management purposes. 它们显示以下信息：

- 名称：创建的聊天名称。
- 更新时间：聊天更新的时间。
- 用户：聊天所属的用户。
- 客户端IP：聊天的客户端IP。
- 计数：此聊天的输入和输出数量。
- 令牌计数：此聊天使用的总令牌数。
- 价格：此聊天花费的总价格。
- 消息：显示聊天的消息内容。
- 存储：显示聊天所属的存储。
- Model Provider：The AI provider used for this chat (e.g., OpenAI, Claude).
- 类别：显示聊天所属的类别。

Token usage and costs are automatically tracked for billing purposes. See the [Billing & Usage](#) section for more details on how transactions are created and

managed.



>

消息

消息



概述

消息概述

概述

在本节中，我们介绍Casibase中消息功能。

消息

The messages module manages all the messages in our sessions. Users can view and track their own message history, while administrators can access all messages across the system. Each message displays detailed metadata including creation time, associated chat, parent message references, token usage, pricing information, and the AI model provider that generated the response.

The model provider field identifies which AI service (OpenAI, Claude, etc.) processed each message. Token costs are automatically calculated and recorded as transactions for billing. Learn more about cost tracking in the [Billing & Usage](#) section.

Name	Created time	User	Chat	Reply to	Author	Token count	Text token count	Price	Reasoning text	Text	Action
message_aoudc5	2025-03-05 14:19:25.446	u-0492894e	chat_fh3ywr	Welcome	AI	1013	9	\$0.0027575		Hello! How can I assist you today?	<button>Edit</button> <button>Delete</button>
message_vj91k5	2025-03-05 14:18:14.061	u-fb61ee56	chat_cbzfz2	Welcome	AI	1011	9	\$0.0027425		Hello! How can I assist you today?	<button>Edit</button> <button>Delete</button>
message_9b7xud	2025-03-05 14:16:51.497	u-3fc996d	chat_hwq271	message_0nv1sh	AI	994	9	\$0.00269		Hello! How can I assist you today?	<button>Edit</button> <button>Delete</button>
message_0nv1sh	2025-03-05 14:16:51.433	u-3fc996d	chat_hwq271		u-3fc996d	1	1	\$0.0000001		hi	<button>Edit</button> <button>Delete</button>
message_q3b6j6	2025-03-05 14:16:25.266	u-3fc996d	chat_hwq271	Welcome	AI	1007	9	\$0.0027125		Hello! How can I assist you today?	<button>Edit</button> <button>Delete</button>



Billing & Usage

Billing & Usage



Overview

Billing and Usage Tracking



Transactions

Transaction tracking and billing integration

Overview

Casibase tracks AI token usage and costs automatically through its integration with Casdoor. Every conversation with an AI model generates transaction records that capture consumption details for billing and analytics.

How It Works

When you send a message to an AI model, Casibase validates your balance before generating the response. The system estimates the token cost, checks if you have sufficient funds, and only proceeds if validation succeeds. This prevents situations where you receive an AI response but can't pay for it.

Chat → Balance Check → Message → Transaction

The validation happens in two steps. First, Casibase estimates the cost by asking the model provider how many tokens your question will likely consume. It then validates this estimated cost against your balance using a dry run transaction. If you have insufficient funds, you'll see an error immediately—no AI processing occurs. If validation succeeds, the AI generates your answer and the actual transaction records the real cost.

Each chat session maintains a running total of tokens used and costs incurred. Individual messages record their own token consumption, and transactions are created in Casdoor after the AI completes its response. This ensures accurate billing records and enables usage monitoring across your organization.

Token Tracking

Token usage is tracked at multiple levels:

- Chats display total tokens used and cumulative price for the entire conversation
- Messages show tokens consumed for each AI response
- Transactions record the billing details in Casdoor for centralized accounting

The model provider field identifies which AI service (OpenAI, Claude, etc.) handled each interaction, making it easy to understand usage patterns across different providers.

Price Calculation

Costs are calculated using proper decimal handling to prevent floating-point precision loss. The system automatically applies the correct pricing model based on the provider and model being used, ensuring accurate billing for all token usage.

Transactions

Transactions connect Casibase usage to Casdoor's billing system. Each AI-generated message automatically creates a transaction record that captures the associated costs and metadata.

Transaction Structure

Transactions use a structured format to track chat interactions:

- **Type:** Chat identifier - links the transaction to a specific conversation
- **Subtype:** Message identifier - tracks which message generated the cost
- **Provider:** Model provider name (e.g., "OpenAI", "Claude")
- **Amount:** Token cost calculated for the message

This structure provides complete traceability from billing records back to the original conversations and messages. Organizations can use this data for cost analysis, usage reporting, and budget tracking across teams.

Balance Validation

Before any AI processing begins, Casibase validates that you have sufficient balance to cover the estimated cost. The system asks the model provider to calculate how many tokens your question will likely consume, then checks this amount against your balance using a dry run transaction—a temporary validation that doesn't actually deduct funds.

If the validation fails, you'll receive an error message right away and the AI won't process your request. This saves you from waiting for an AI response only to discover you can't afford it. Once validation succeeds, the AI generates your answer and a real transaction records the actual cost, which may differ slightly from the estimate based on the final response length.

Automatic Creation

Transactions are created after the AI completes your answer, ensuring that billing records stay synchronized with actual usage. You don't need to manually track or create transaction records - the system handles this automatically for every AI interaction.

Integration with Casdoor

Transaction records are stored in Casdoor, enabling centralized billing management across all your applications. This integration allows organizations to:

- Track AI usage costs alongside other application expenses
- Generate unified billing reports
- Set up usage alerts and budgets
- Analyze spending patterns across different AI providers

The transaction data flows seamlessly from Casibase to Casdoor, providing real-time visibility into AI usage costs without requiring manual data entry or reconciliation.



>

Records

Records



Records

Data records and aggregation

Records

Records in Casibase support efficient data logging and analytics through a built-in aggregation mechanism. Instead of storing thousands of individual entries, you can create consolidated records that represent multiple data points.

Count Field

Records include a `Count` field that defaults to `1` for all new records. This field indicates how many actual data points the record represents.

When uploading consolidated data, set the count to match the number of underlying records. For example, if you have 100 similar log entries, create one record with `count: 100` rather than inserting 100 separate records. This approach significantly reduces storage requirements while maintaining accurate statistics.

API Behavior

The count field works consistently across all record operations:

- Default value: `1` for all new records created via API or UI
- Backward compatibility: Zero values display as `1` in the frontend
- Type: Integer field supporting positive values

When using the `add-record` or `add-records` API endpoints, include the count value to specify how many data points the record aggregates.

Use Cases

Records with aggregation are particularly useful for:

- **Log consolidation:** Summarize multiple similar log entries into a single record
- **Analytics:** Maintain accurate counts for reporting without storing raw data
- **System monitoring:** Track event frequencies without detailed individual records
- **Data archival:** Compress historical data while preserving statistical accuracy

The count field enables efficient data management by separating the logical representation (one record) from the actual count (potentially many occurrences).



>

Scans

Scans



Scans

Network and security scanning in Casibase

Scans

Scans in Casibase enable network discovery, security auditing, and system assessment across your infrastructure. The scan feature provides automated scanning capabilities for assets like virtual machines and network hosts, delivering structured results for analysis.

What is a Scan?

A Scan represents a scanning operation executed against a target asset. Each scan instance records the target, provider used, execution time, and results. Scans can be triggered manually or configured to run on demand against assets in your infrastructure.

Scan Workflow

Start by navigating to the Scans page and creating a new scan instance. You'll need to configure the target selection, choose a scan provider like Nmap or OS Patch, and set any necessary scan parameters.

When selecting your target, you can use **Asset Mode** to pick an existing asset from your inventory, or **Manual Mode** to enter an IP address or hostname directly. For Virtual Machine assets, the system automatically uses the public IP address from the asset properties.

Click the "Start" button to initiate the scan. The scan executes asynchronously, transitioning through states from Pending to Running to Completed. The Runner

field tracks which system instance is actively executing the scan, useful in distributed deployments. If errors occur, the **ErrorText** field captures diagnostic information to help troubleshoot issues.

The scan provider performs the configured operation against your target and returns results. These results appear in three formats: a **Structured View** with organized tables and formatted data, the **Raw JSON** showing parsed data structures, and the **Raw Text** showing the complete unprocessed output from the scan tool. All formats are stored in the database for historical reference, allowing you to review previous scans and track changes over time.

To reset a scan and clear its results, use the **Clear** button available on scan pages. This removes the scan state and results while preserving the scan configuration, letting you run a fresh scan with the same settings.

From the Scans list page, you can quickly review scan results without opening the full editor. The **Result** column provides a popup view displaying the same structured, JSON, and raw text formats available in the scan detail page.

Scan Providers

Casibase supports multiple scan provider types, each optimized for specific scanning tasks.

The **Nmap Scan Provider** performs network discovery and security auditing by scanning ports, detecting services, and identifying system information. It handles port scanning with customizable ranges, service version detection, operating system fingerprinting, and network topology mapping. Results come back as structured JSON containing host information, open ports, detected services, and system details, which the web interface renders in organized tables.

The **OS Patch Provider** checks system patch status and identifies missing security

updates. The provider uses efficient local cache queries to list installed patches, ensuring fast scan performance without querying online update services. When listing available patches for installation, the system does query Windows Update online to identify new updates.

This hybrid approach balances speed and accuracy, providing quick status checks while enabling comprehensive update discovery when needed. Results include patch status information, available updates, and security recommendations in a structured format for quick review.

Scan Configuration

When configuring a scan in the provider edit page, you can test the provider functionality directly. The scan configuration widget automatically selects the first available provider to streamline setup, and lets you switch between Asset and Manual Input modes for target selection, execute test scans to verify provider configuration, and view scan output before saving configurations. The system saves both the scan configuration and provider settings to the database before initiating scan execution, ensuring all settings are persisted.

For Manual Input mode, the target matching logic intelligently routes scans to the appropriate runner instance. When you specify a hostname as the target, the system matches it against each runner instance's hostname to ensure the scan executes on the correct machine. For IP address targets (excluding localhost), the scan routes to the instance whose network interfaces include that IP address, checking both private and public IPs.

Localhost and loopback addresses like 127.0.0.1 can be claimed by any instance, allowing flexible local scanning. This ensures distributed scan deployments work correctly, with each instance claiming scans intended for its specific machine.

API Integration

Scans support programmatic access through REST APIs. Use `GET /get-scans` to retrieve all scans with pagination, or `GET /get-scan` to fetch a specific scan by ID. The `POST /add-scan` endpoint creates a new scan, while `POST /update-scan` modifies scan configuration and `POST /delete-scan` removes a scan.

The `POST /scan-asset` API performs on-demand scanning and returns results without creating a persistent scan record, making it useful for quick assessments.

Working with Scans

For cloud-based virtual machines, Asset Mode automatically resolves the correct public IP address, ensuring scans reach the intended target even as infrastructure changes. Schedule regular scans to maintain visibility into your security posture, but consider the impact on target systems when determining frequency.

When analyzing results, review both the structured and raw formats. The structured view provides quick insights, while raw output offers complete details for in-depth analysis. Choose Nmap for network discovery and port scanning, or OS Patch for security update assessment depending on what you need to learn about your systems.



> Container Cloud

Container Cloud

Overview

Container Cloud Overview

Template

In Casibase, a Template is a reusable base configuration for an application. It contains the core Kubernetes manifest files, typically structured for use with Kustomize. Y...

Applications

1 个项目

Overview

Once you have successfully connected Casibase to your private cloud providers (like Kubernetes), this section will guide you on how to manage cloud-native resources directly through the Casibase interface.

Casibase provides a powerful system based on Docker and Kubernetes, designed for individuals and organizations to build their own dedicated container cloud environment. Built on the Casbin permission management engine, it implements fine-grained access control policies for secure and controllable private cloud operations.

Core Concepts

Casibase utilizes a streamlined two-part system for managing container deployments:

1. Templates: Reusable Application Blueprints

Templates are pre-configured Kubernetes manifests that serve as blueprints for your applications. Each template contains:

- **Base Configuration:** Complete Kubernetes resources (Deployments, Services, ConfigMaps, etc.) required to run an application
- **Customizable Parameters:** Configurable fields that can be modified during deployment
- **Version Management:** Template versioning for consistent deployments across environments

2. Applications: Live Application Instances

Applications are running instances created from templates. They represent actual workloads deployed to your Kubernetes cluster:

- **Parameter Customization:** Override template defaults with specific configurations (replicas, image versions, resource limits)
- **Namespace Isolation:** Each application runs in its own dedicated namespace for security and organization
- **Lifecycle Management:** Complete application lifecycle control from deployment to termination

Key Features

Declarative Application Orchestration: Transform from resource-level management to application-level management, simplifying complex multi-resource deployments into simple "select template → configure → deploy" workflows.

Service Governance Integration: Built-in support for service mesh and gateway templates (Istio, Linkerd, Nginx Ingress) enabling one-click deployment of microservice governance capabilities including service discovery, circuit breaking, and rate limiting.

Enhanced Platform Visualization:

- **Application Dashboard:** Monitor and manage all your deployed applications with real-time status updates
- **Resource Insights:** Deep visibility into underlying Kubernetes resources, logs, and events for each application

Kustomize-Powered Flexibility: Leverages Kubernetes-native Kustomize for configuration management, ensuring consistency across development, testing, and production environments while maintaining the ability to customize deployments per environment.

This approach helps you standardize your infrastructure, ensure deployment consistency, and streamline the process of launching and managing containerized services. It eliminates the complexity of manual Kubernetes resource orchestration while maintaining full control over your applications.

Please proceed to the following sections to learn more about managing templates and applications:

- [Kubernetes Templates](#)
- [Kubernetes Applications](#)

Template

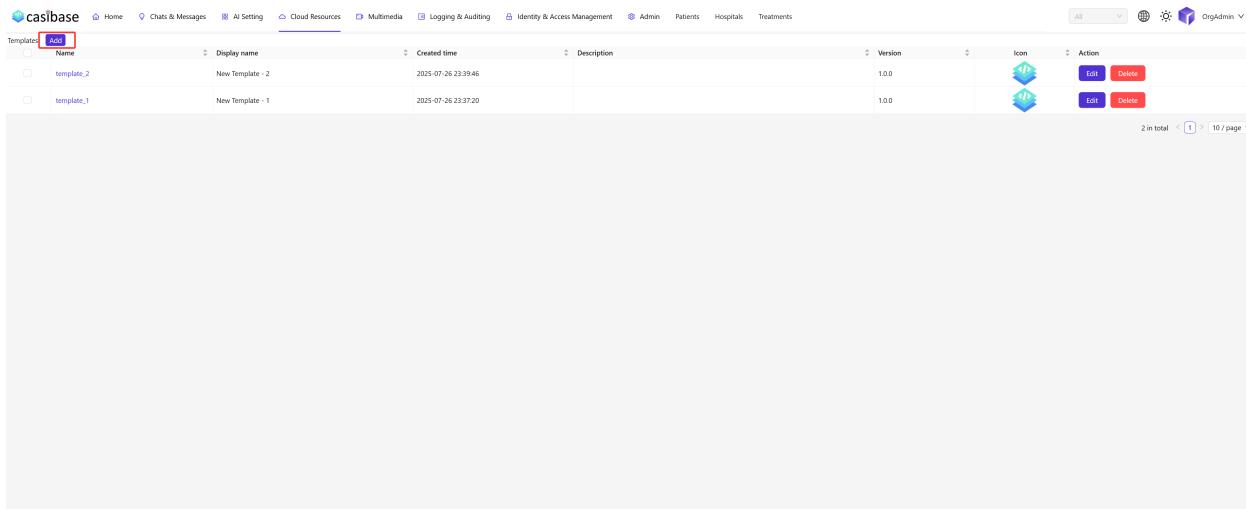
In Casibase, a **Template** is a reusable base configuration for an application. It contains the core Kubernetes manifest files, typically structured for use with Kustomize. You define a template once, and it can then be used as a blueprint to create multiple, customized application instances.

This model allows you to standardize your deployment patterns, ensuring consistency and simplifying the process of launching new services.

This chapter will guide you through creating and managing templates in Casibase.

Create a New Template

First, navigate to the **Cloud Resources > Templates** section and click the  button to open the creation page.



Name	Display name	Created time	Description	Version	Icon	Action
template_2	New Template - 2	2025-07-26 23:39:46		1.0.0		Edit Delete
template_1	New Template - 1	2025-07-26 23:37:20		1.0.0		Edit Delete

You will need to fill in the following fields, which correspond to the template's

properties:

- **Name**: A unique identifier for the template (e.g., `my-app-template`). This is a required field.
- **Display name**: A user-friendly name that will be shown in the UI (e.g., `My App Template`).
- **Description**: A brief description of what this template is for.
- **Version**: The version of the template (e.g., `1.0.0`).
- **Icon**: A URL to an icon image that represents the template in the UI.
- **Manifest**: The raw YAML text of your Kubernetes manifests. This content serves as the base for Kustomize deployments.

The screenshot shows the Casibase application interface for editing a template. At the top, there's a navigation bar with links like Home, Chats & Messages, AI Setting, Cloud Resources, Multimedia, Logging & Auditing, Identity & Access Management, Admin, Patients, Hospitals, and Treatments. Below the navigation is a toolbar with 'Edit Template', 'Save', and 'Save & Exit' buttons. The main area has several input fields: 'Name' (template_2), 'Display name' (New Template - 2), 'Description' (empty), 'Version' (1.0.0), 'Icon' (a URL to a Casibase logo image), and a large 'Manifest' text area. The 'Manifest' area contains the following YAML code, which is highlighted with a red border:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
        - name: nginx
          image: nginx
          ports:
            - containerPort: 80
```

After saving, your template will be available in the selection list when you create a new application.

Application

An Application in Casibase is a specific, deployable instance created from a Template. Before you can create an application, you must first have at least one template defined. When you create an application, you select a base template and can then apply specific customizations before deploying it to your Kubernetes cluster.

This chapter will guide you through creating, deploying, and managing applications in Casibase.

Create a New Application

Navigate to the Cloud Resources > Applications section and click the [Add](#) button to open the creation page.

The screenshot shows the Casibase application management interface. At the top, there's a navigation bar with links like Home, Chats & Messages, AI Setting, Cloud Resources, Multimedia, Logging & Auditing, Identity & Access Management, Admin, Patients, Hospitals, and Treatments. Below the navigation is a search bar and a filter dropdown set to 'All'. The main area is titled 'Applications' and features a large 'Add' button with a plus sign. A table lists three existing applications:

Name	Display name	Created time	Description	Template	Status	Namespace	Action
application_grpkxm	New Application - grpkxm	2025-08-02 21:59:58		template_2	Running	casibase-application-grpkxm	Edit Undeploy Delete
application_3k8d1	New Application - 3k8d1	2025-08-02 09:43:55		template_2	Running	casibase-application-3k8d1	Edit Undeploy Delete
application_x04b7i	New Application - x04b7i	2025-08-01 23:44:27		template_2	Not Deployed	casibase-application-x04b7i	Edit Deploy Delete

At the bottom right of the table, it says '3 in total < 1 > 10 / page'.

The key fields for an application are:

- **Name**: A unique name for your application instance (e.g., `my-app-prod`). This is a required field.
- **Display name**: A user-friendly name that will be shown in the UI (e.g., `My App (Production)`).
- **Description**: A brief description of this specific application instance.
- **Template**: Select a pre-existing template from the dropdown list. This will be the base for your application.
- **Parameters**: This field is used for customization. Here you can provide specific Kustomize patches or other variable substitutions in YAML format to override or extend the base **Manifest** from the selected template.

Note: Fields like **Status** and **Namespace** are managed by the system. The **Namespace** is automatically generated based on the application name upon creation and cannot be modified by the user. The **Status** is updated based on its deployment state (e.g., `Not Deployed`, `Running`, `Pending`).

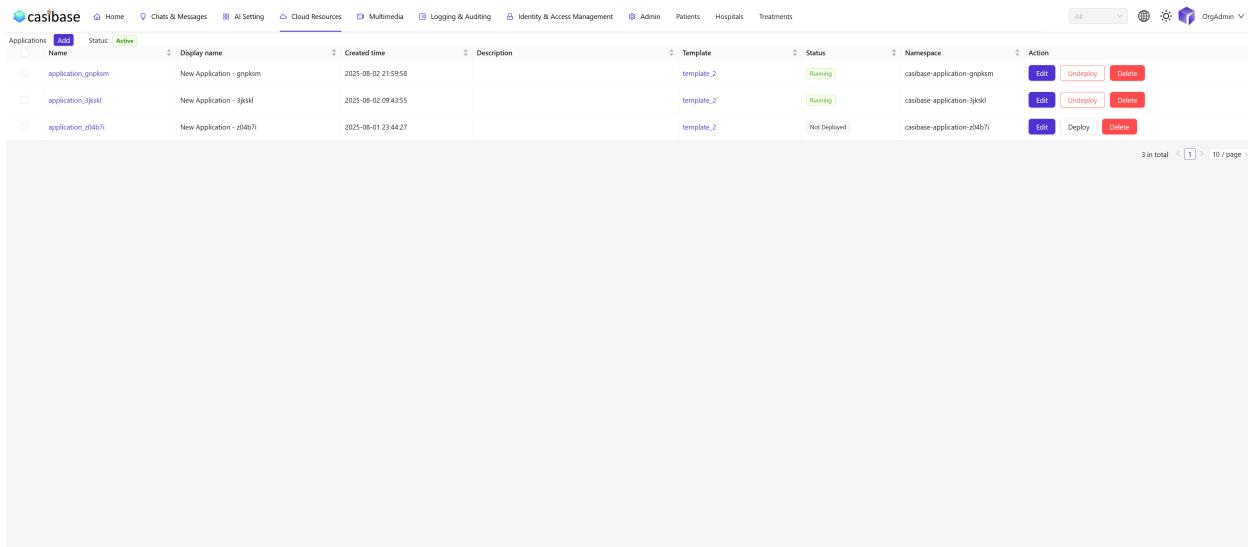
The screenshot shows the 'Edit Application' form in the casibase interface. The form includes fields for Name, Display name, Description, Template, Status, Namespace, and Parameters. The Parameters field contains a redacted YAML code block, indicated by a large black rectangle with a red border. At the bottom of the form are 'Save' and 'Save & Exit' buttons.

Deploying and Monitoring an Application

After creating an application, it will appear in the applications list. From here, you

can manage its lifecycle.

- **Deploy:** Click the `Deploy` button to apply the application's configuration to your Kubernetes cluster. Casibase will use Kustomize to merge the base template's `Manifest` with your application's `Parameters` and run `kubectl apply`.
- **Undeploy:** The `Undeploy` button will remove the application's resources from your Kubernetes cluster.



The screenshot shows the Casibase application management interface. At the top, there is a navigation bar with links like Home, Chats & Messages, AI Setting, Cloud Resources, Multimedia, Logging & Auditing, Identity & Access Management, Admin, Patients, Hospitals, and Treatments. On the far right, there are user profile icons and a dropdown for OrgAdmin.

The main area is a table titled "Applications". The columns are: Name, Display name, Created time, Description, Template, Status, Namespace, and Action. There are three entries in the table:

Name	Display name	Created time	Description	Template	Status	Namespace	Action
application_grpkm	New Application - grpkm	2025-08-02 21:59:58		template_2	Running	casibase-application-grpkm	<button>Edit</button> <button>Undeploy</button> <button>Delete</button>
application_3jekl	New Application - 3jekl	2025-08-02 09:43:55		template_2	Running	casibase-application-3jekl	<button>Edit</button> <button>Undeploy</button> <button>Delete</button>
application_z04b7i	New Application - z04b7i	2025-08-01 23:44:27		template_2	Not Deployed	casibase-application-z04b7i	<button>Edit</button> <button>Deploy</button> <button>Delete</button>

At the bottom right of the table, it says "3 in total" and has a page number "1" with a "10 / page" dropdown.

By using this template-and-application model, you can effectively standardize and scale your Kubernetes deployments through the Casibase interface.



Creating Databases with KubeBlocks

Overview

KubeBlocks is an open-source Kubernetes operator and toolset designed to simplify the complexity of running and managing data infrastructure, such as databases, message queues, and streaming systems, on Kubernetes. It provides a declarative approach to deploying and managing stateful applications, allowing you to manage them as easily as stateless ones.

This guide will walk you through the process of creating a database cluster using KubeBlocks.

Why KubeBlocks?

- **Simplified Management:** Automates the database lifecycle, including deployment, upgrades, scaling, and monitoring.
- **Production-Ready:** Supports high availability, backup and restore, and robust monitoring.
- **Versatile:** Supports a wide range of databases, including MySQL, PostgreSQL, MongoDB, Redis, and more.

Installing KubeBlocks

You can install KubeBlocks using [Helm](#). For more installation options, refer to the [KubeBlocks Installation Guide](#).

This guide will use Helm for the installation.

Step 1: Deploy Snapshot Controller

KubeBlocks requires the Snapshot Controller to manage volume snapshots. First, check if it is already installed in your cluster.

```
kubectl get crd volumesnapshotclasses.snapshot.storage.k8s.io  
kubectl get crd volumesnapshots.snapshot.storage.k8s.io  
kubectl get crd volumesnapshotcontents.snapshot.storage.k8s.io
```

If it is not installed, you can deploy it using the following commands:

```
helm repo add piraeus-charts https://piraeus.io/helm-charts/  
helm repo update  
# Update the namespace to an appropriate value for your  
environment (e.g. kb-system)  
helm install snapshot-controller piraeus-charts/snapshot-  
controller -n kb-system --create-namespace
```

Then, verify the installation:

```
kubectl get pods -n kb-system | grep snapshot-controller
```

The Snapshot Controller should be in the [Running](#) state.

Step 2: Get the Latest KubeBlocks Version

Get the latest stable version tag (e.g. v1.0.1):

```
curl -s "https://api.github.com/repos/apecloud/kubeblocks/releases?per_page=100&page=1" | jq -r '.[] | select(.prerelease == false) | .tag_name' | sort -V -r | head -n 1

# Example output:
# v1.0.1
```

Step 3: Create KubeBlocks CRDs

Create the Custom Resource Definitions (CRDs) required by KubeBlocks.

```
# Replace <VERSION> with the version you selected
kubectl create -f https://github.com/apecloud/kubeblocks/releases/download/<VERSION>/kubeblocks_crds.yaml

# Example: If the version is v1.0.1
kubectl create -f https://github.com/apecloud/kubeblocks/releases/download/v1.0.1/kubeblocks_crds.yaml
```

Step 4: Install KubeBlocks with Helm

1. Add the KubeBlocks Helm repository:

```
helm repo add kubeblocks https://apecloud.github.io/helm-charts
helm repo update
```

2. Install KubeBlocks: This command installs the KubeBlocks chart into the kb-

system namespace.

```
helm install kubeblocks kubeblocks/kubeblocks --namespace kb-system --create-namespace
```

Creating a Database

Once KubeBlocks is installed, you can create a database cluster using `kubectl`.

Currently supported databases include:

- MySQL
- PostgreSQL
- MongoDB
- Redis
- Kafka
- Milvus
- Qdrant
- RabbitMQ
- Elasticsearch

Example: Creating a Demo MySQL Cluster

1. Create a file named `my-mysql-cluster.yaml`:

```
apiVersion: apps.kubeblocks.io/v1
kind: Cluster
metadata:
```

2. Apply the manifest to create the cluster:

```
kubectl apply -f my-mysql-cluster.yaml
```

More Information

For more detailed information, advanced configurations, and troubleshooting with KubeBlocks, please refer to the [official KubeBlocks documentation](#).



>

节点

节点

概述

Casibase节点概述

远程桌面协议

Casibase节点RDP

VNC

Casibase节点VNC

概述

Casibase帮助您管理节点，并通过RDP、VNC、SSH和Telnet远程连接到您的节点。

协议：

- 安全外壳协议
- 远程桌面协议
- 虚拟网络计算
- 远程终端协议

每个节点都具有以下基本属性：

- 组织：节点所属的组织。
- 名称：唯一的节点名称。
- 描述：节点的描述。
- IP：域名或IP地址。
- 协议：协议的端口号。
- 端口：节点的端口号。
- 用户名：连接到节点的用户名，如root、administrator、sa等。
- 密码：连接到节点的密码。
- 操作系统：节点的操作系统，包括Windows和Linux，用于分类节点。
- 标签：节点的标签，用于分类节点。

在本章中，您将学习如何开始连接到您的节点。

让我们一起探索！

远程桌面协议

Casibase支持通过RDP协议连接到您的节点:

RDP连接

1. 启动Guacamole服务器

```
docker run --name guacd -d -p 4822:4822 guacamole/guacd
```

2. 添加一个新节点, 将协议设置为 rdp

Organization	Name	Created time	Description	Protocol	IP	Port	Username	Language	Auto query	Is perm	Action		
casbin	node_eqiwer	2025-03-09 23:37:34		VNC	127.0.0.1	5900	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>	<button>Edit</button>	<button>Delete</button>
casbin	node_apacdj	2025-03-09 23:32:12		VNC	127.0.0.1	5900	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>	<button>Edit</button>	<button>Delete</button>
casbin	node_qf773r	2025-02-25 11:12:14+03:30		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>	<button>Edit</button>	<button>Delete</button>
casbin	node_zbj7av	2025-02-21 17:18:08		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>	<button>Edit</button>	<button>Delete</button>
casbin	node_cy3c9s	2025-02-14 11:59:43		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>	<button>Edit</button>	<button>Delete</button>

The screenshot shows the 'Edit Node' form for a 'host-base' node. The node is associated with the organization 'casbin'. It has the IP address 47.93.49.234 and port 3389. The OS is Windows. The 'Auto query' and 'Is permanent' options are enabled. A table at the bottom lists services, with one entry for 'RemoteApp'.

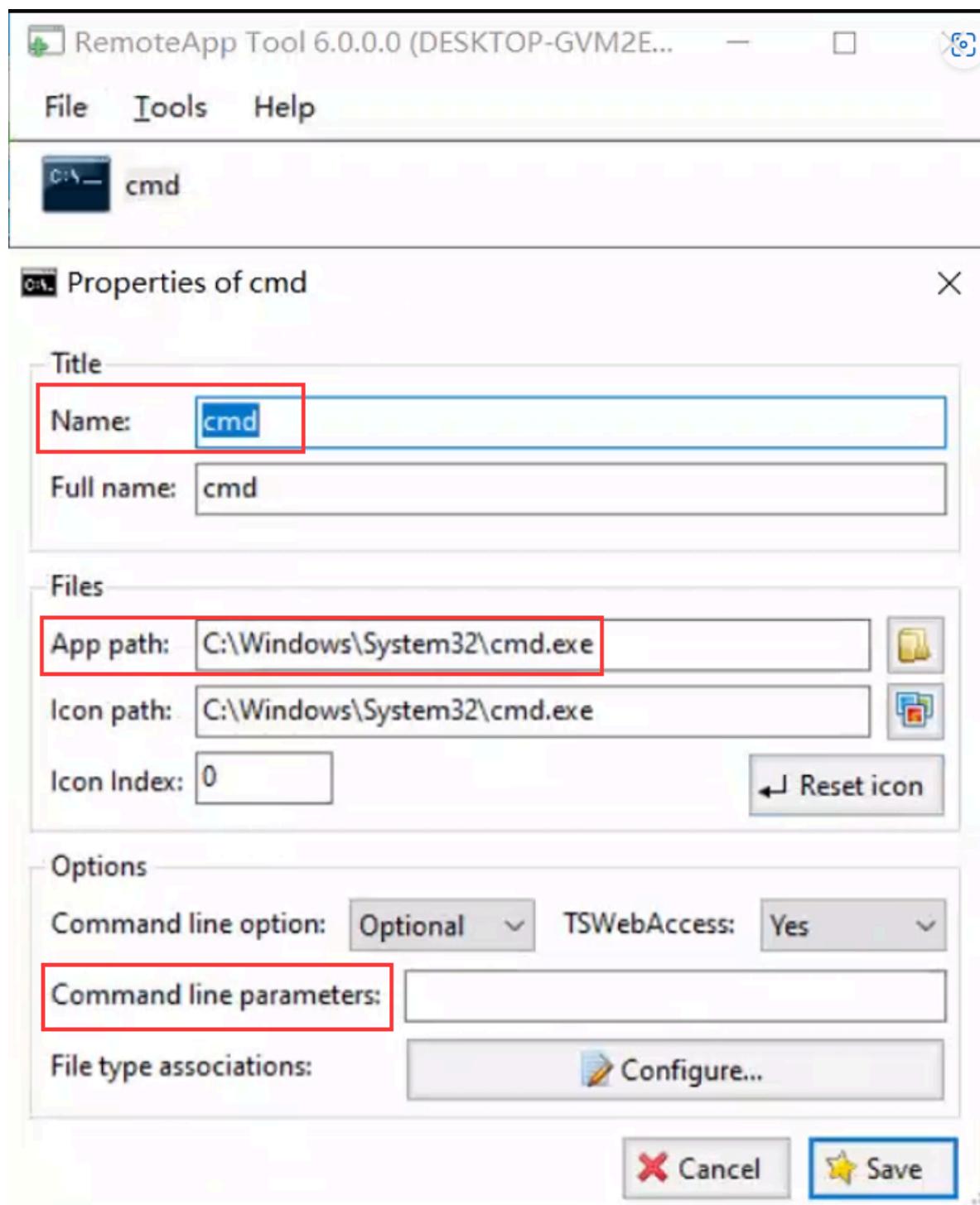
No.	Name	Path	Port	Process ID	Expected status	Status	Message	Action
1	RemoteApp							

3. 点击 **连接** 按钮连接到您的节点

远程应用程序

我们支持在Windows节点上使用远程应用程序，您可以在**节点编辑**页面添加远程应用程序，然后通过点击**连接**按钮连接到您的远程应用程序。

1. 在服务器端配置您的远程应用程序。\\您可以使用 [RemoteApp工具](#) 注册应用。



2. 根据服务器端配置，在节点编辑页面中配置远程应用信息。“remoteAppName”、“remoteAppDir”和“remoteAppArgs”是必需的。



refer to [Configuring Guacamole — Apache Guacamole Manual v1.5.3](#)

3. 连接到您的远程应用程序。

VNC

VNC连接

VNC连接与RDP连接类似。

1. 启动Guacamole服务器

```
docker run --name guacd -d -p 4822:4822 guacamole/guacd
```

2. 添加一个新节点，将协议设置为vnc

Organization	Name	Created time	Description	Protocol	IP	Port	Username	Language	Auto query	Is perm	Action
casbin	node_eqiwer	2025-03-09 23:37:34		VNC	127.0.0.1	5900	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>
casbin	node_apacdj	2025-03-09 23:32:12		VNC	127.0.0.1	5900	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>
casbin	node_qf773r	2025-02-25 11:12:14+03:30		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>
casbin	node_zbj7av	2025-02-21 17:18:08		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>
casbin	node_cy3c9s	2025-02-14 11:59:43		RDP	127.0.0.1	3389	Administrator	en	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<button>Connect</button>

3. 点击connect按钮连接到您的节点。



>

Graphs

Graphs



Graphs

Graph Visualization in Casibase

Graphs

Overview

Graphs in Casibase provide visual representation of relationships and connections between entities. The interactive graph interface allows you to explore complex data structures through an intuitive node-and-link visualization.

Asset Graph Transformation

Casibase can automatically generate graphs from cloud infrastructure assets. When you have assets defined (such as virtual machines, containers, or network resources), use the asset-to-graph transformation feature to visualize their relationships. The system analyzes asset properties to create nodes and establishes links based on resource dependencies and connections. This provides instant visibility into your infrastructure topology without manual graph definition.

Creating a Graph

Navigate to the Graphs section and click **Add** to create a new graph. Configure the graph with:

- **Name:** Unique identifier for the graph
- **Display Name:** Human-readable name shown in the UI
- **Category:** Select the graph category (Assets, Chats, etc.)
- **Data:** JSON structure defining nodes and links

Graph Categories

Casibase supports different graph categories for various data types:

Assets: Visualize cloud infrastructure and network topology. Define nodes and links manually or generate them automatically from asset data.

Chats: Analyze chat conversations with word cloud visualization. Select a store and time range to filter conversations, then generate word clouds from message content. The system automatically processes chat messages, removes stop words, and creates visualizations showing frequently used terms.

Graph Data Structure

Graphs use a JSON format with two main components:

Nodes: Individual entities in the graph

```
{  
  "id": "node1",  
  "name": "Node Name",  
  "label": "Display Label",  
  "icon": "icon-url"  
}
```

Links: Connections between nodes

```
{  
  "source": "node1",  
  "target": "node2",  
  "label": "Connection Type"  
}
```

Interactive Features

The graph visualization provides several interactive capabilities:

Node Selection: Click any node to view its details in a floating panel. Selected nodes are highlighted with a distinct visual style. The detail panel displays node metadata including all custom properties defined in the asset or node configuration, and shows related scan results when available. For assets with associated scans, you can review the scan history and results directly from the node detail view.

Canvas Navigation: Drag anywhere on the canvas to pan the view. The graph supports full directional movement for exploring large networks.

Visual Feedback: Nodes display custom icons and labels. Links connect at node edges rather than centers for cleaner visualization. When assets have associated scan results, the node displays a notification badge showing the scan count. The badge appears as a red pill-shaped indicator with white text, similar to iOS app notifications, making it immediately obvious which nodes have scan data available without needing to click through each one.

Layout Controls

Graphs support multiple layout algorithms to organize nodes:

- **Force-directed layout:** Nodes automatically arrange based on their connections, creating organic-looking graphs where connected nodes pull together while maintaining spacing.
- **Word Cloud layout:** Available for Chats category graphs, displays words sized by frequency for visual analysis of conversation topics.

- **None layout:** Positions nodes based on provided coordinates, useful when you want manual control over the exact placement.

Chat Analysis with Word Clouds

For Chats category graphs, Casibase provides specialized visualization through word clouds. Configure these graphs by:

1. **Selecting a Store:** Choose which chat store to analyze
2. **Setting Time Range:** Define start and end times to focus on specific periods
3. **Viewing Chat Data:** Review the filtered chats and messages in table format
4. **Generating Visualizations:** The system automatically creates word clouds from message content

The word cloud generation process extracts text from all messages in the selected time range, removes common stop words (in English and Chinese), and calculates word frequencies. Words appear larger based on how often they occur in the conversations, making it easy to identify key topics and themes.

Density Control

For graphs with many nodes, use the density slider to adjust node spacing. Higher density brings nodes closer together for a compact view, while lower density spreads them out for better readability. This is particularly useful when working with auto-generated graphs from cloud infrastructure.

Graph Editing

In edit mode, you can modify the graph structure and preview changes:

- Edit the JSON data to add, remove, or modify nodes and links

- Preview shows real-time visualization with a bordered canvas
- Changes are saved when you click `Save & Exit`

Node Properties

Each node can have custom properties:

- **Icon:** URL to an icon image displayed on the node
- **Label:** Text displayed below the node
- **Properties:** Additional metadata shown in the detail panel

When viewing node details, the panel shows comprehensive information about the selected node. For asset-derived graphs, this includes all metadata from the original asset such as IP addresses, resource IDs, locations, and other cloud-specific properties. Property values are displayed in an organized format with proper label formatting. Long text values are automatically truncated with ellipsis for readability.

For assets with scan results, the detail panel includes a dedicated section showing related scans. You can view scan execution history, status information, and access the full scan results without leaving the graph view. This integration provides immediate visibility into the security and operational status of your infrastructure assets as you explore the topology.

Error Handling

Graphs include built-in error handling. If data cannot be loaded or visualization fails, an error message displays in place of the graph, allowing you to identify and fix configuration issues.

Customization

The graph visualization adapts to the Casibase theme, using consistent colors and styling throughout the interface. Selected nodes use theme colors for visual emphasis while maintaining clarity.

Best Practices

Keep graph structures focused and organized. For complex networks, consider creating multiple smaller graphs rather than one large visualization. This improves performance and makes relationships easier to understand.

Use descriptive node labels and meaningful link labels to make the graph self-explanatory. Clear naming helps users quickly grasp the relationships without additional documentation.