




deepseek

- [deepseek 教程](#)
- [docker部署deepseek+open-webui](#)
- [deepseek+dify教程](#)
- [deepseek+llm-as-a-service](#)
- [llm-as-a-service+firecrawl](#)
- [deepseek+llm-as-a-service](#)
- [deepseek+llm-as-a-service](#)
- [llm-as-a-service&&deepseek+llm-as-a-service](#)
- [llm-as-a-service](#)

deepseek

deepseek



- ollama <https://ollama.com/download>
- open-webui  <https://github.com/open-webui/open-webui>
- dify <https://github.com/langgenius/dify>
- docker
- python
 - 3.11  open-webui 
- <https://iovhm.com/book/books/cee63/page/9872e>

ollama

hostbridge docker 



```
version: "3"
services:
  ollama:
    image: harbor.iovhm.com/hub/ollama/ollama:0.5.12
    container_name: ollama
    restart: always
    privileged: true
    ports:
      - "11434:11434"
    volumes:
      - ./ollama:/root/.ollama
    networks:
      - vpcub-bridge
```

ollama


```
# [][]
# []
# []
# []/bye []ctrl+d[]
# []run[]
ollama run deepseek-r1:1.5b

# []
ollama list

# []
ollama ps

# []
ollama show deepseek-r1:1.5b
```

API

```
ollama 1143 []postmain[]
```

```
[]
```

```
curl -X POST -H "Content-Type: application/json" \
  -d '{"model": "deepseek-r1:1.5b", "prompt": "[]"}' \
  http://localhost:11434/api/generate
```

open-webui

```
[]openai[]
```

<https://github.com/open-webui/open-webui>

<https://docs.openwebui.com/>

```
[]docker[]docker[]docker[]
```

```
# GPU[]
```

```
docker run -d -p 3000:8080 --gpus all --add-host=host.docker.internal:host-gateway -v open-  
webui:/app/backend/data --name open-webui --restart always ghcr.io/open-webui/open-webui:cuda
```

```
# CPU
```

```
docker run -d -p 3000:8080 --add-host=host.docker.internal:host-gateway -v open-  
webui:/app/backend/data --name open-webui --restart always ghcr.io/open-webui/open-webui:main
```

```
# pip
```

```
pip install open-webui -i https://pypi.tuna.tsinghua.edu.cn/simple
```

```
# python
```

```
# http://localhost:8080
```

```
open-webui serve
```

<https://iovhm.com/book/books/bbcbf/page/deepseek-docker>

docker deepseek open-

ollama ollama

docker-compose.yml

ollama open-webui

```
version: "3"
services:
  ollama:
    image: harbor.iovhm.com/hub/ollama/ollama:0.5.12
    container_name: ollama
    restart: always
    privileged: true
    ports:
      - "11434:11434"
    volumes:
      - ./ollama:/root/.ollama
    networks:
      - vpclub-bridge

# docker-compose --profile open-webui up -d
open-webui:
  # CPU
  image: harbor.iovhm.com/public/open-webui/open-webui:main
  # GPU
  # image: harbor.iovhm.com/public/open-webui/open-webui:main-gpu
  container_name: open-webui
  restart: always
  privileged: true
  ports:
    - "3000:8080"
  volumes:
    - ./open-webui:/app/backend/data
  environment:
    # ollama OLLAMA_BASE_URLS OLLAMA_BASE_URLS
```


点燃好奇心 无论你在哪里



开始使用

open-webui

- open-webui

openai


```
# open-webui

# openai api
- ENABLE_OPENAI_API=false

# RAG
- RAG_EMBEDDING_ENGINE=ollama

# 
- RAG_EMBEDDING_MODEL=nomic-embed-text:latest
```

openai

☰ **oi** 新对话 

🔗 工作空间

🔍 搜索

∨ 对话

用户 竞技场评估 函数 设置

⚙️ 通用

🔗 外部连接

📦 模型

📄 竞技场评估

📄 文档

🌐 联网搜索

📄 代码执行

🗨️ 界面

🔊 语音

🖼️ 图像

📡 Pipeline

🗄️ 数据库

关闭openai



OpenAI API

Ollama API



管理Ollama API连接

配置ollama地址



http://ollama:11434



访问 Ollama 时遇到问题? [点击这里获取帮助。](#)

直接连接



直接连接功能允许用户连接至其自有的、兼容 OpenAI 的 API 端点。



新对话 工作空间 搜索 对话

用户 竞技场评估 函数 **设置**

通用

Content Extraction Engine 默认

PDF 图像处理 (使用 OCR)

Bypass Embedding and Retrieval

文本分切器 默认 (字符)

块大小 (Chunk Size) 1000 块重叠 (Chunk Overlap) 100

Embedding

语义向量模型引擎 Ollama

http://ollama:11434 API 密钥

语义向量模型 nomic-embed-text

警告: 如果您修改了语义向量模型, 则需要重新导入所有文档。

嵌入层批处理大小 (Embedding Batch Size) 1

完整上下文模式

混合搜索

Retrieval

Top K 3

RAG 提示词模板

```
### Task:
Respond to the user query using the provided context, incorporating inline citations in the format [source_id] **only when the <source_id> tag is explicitly provided** in the context.

### Guidelines:
- If you don't know the answer, clearly state that.
- If uncertain, ask the user for clarification.
- Respond in the same language as the user's query.
- If the context is unreadable or of poor quality, inform the user and provide the
```

保存

设置 已归档对话 AI 对话游乐场 **管理员面板** 登出 当前在线用户: 1 admin

ollama

```
# ollama
docker exec -it ollama /bin/bash

# ollama
ollama -v

# ollama
ollama run deepseek-r1:7b

# ollama
ollama pull nomic-embed-text
```


COLLECTION 威海智慧谷

威海智慧谷

FILE 威海智慧谷.md

威海智慧谷 - 威海智慧谷

#

+



代码解释器



Open WebUI · v0.5.18

deepseekdeepseekB

新对话

工作空间

搜索

对话

今天

宫廷玉液酒 🌈 @@ 网页流芳 ...

董列涛

deepseek-r1:1.5b +

宫廷玉液酒

deepseek-r1:1.5b Today at 23:46

思考用时 8 秒

“宫廷玉液酒”是一个源自中国传统文化的词汇,同时也因1996年春晚小品《打工奇遇》而广为人知。在1996年春晚的小品中,赵丽蓉和巩汉林等演员表演了一个关于“宫廷玉液酒”的搞笑片段。这一设定讽刺了商家的虚假宣传和哄抬物价的行为。小品中的经典台词“宫廷玉液酒,一百八一杯”成为了广为流传的网络梗。随着小品的流行,“宫廷玉液酒”在网络上成为了一个经典的暗号和调侃用语。它常被用来测试对方是否了解中国传统文化或春晚节目,甚至被用作网络社交中的“接头暗号”。

1 威海智慧谷.txt 2 宫廷玉液酒.txt

✎ 📄 🔊 🌐 📌 🔄

输入消息

+ 🌐 联网搜索 📄 代码解释器

🔊 🎧

deepseek-dify

dify-ollama AI



- <https://iovhm.com/book/books/cee63/page/9872e>
- <https://iovhm.com/book/books/k8s/page/harbordockerdocker>
- <https://github.com/docker/compose/releases>
- <https://iovhm.com/book/books/bbcbf/page/deepseek>
- <https://github.com/langgenius/dify>

ollama

```
version: "3"
services:
  ollama:
    image: harbor.iovhm.com/hub/ollama/ollama:0.5.12
    container_name: ollama
    restart: always
    privileged: true
    ports:
      - "11434:11434"
    volumes:
      - ./ollama:/root/.ollama
    # deploy:
    #   resources:
    #     reservations:
    #       devices:
    #         - driver: nvidia
    #           capabilities: [gpu]
    #           count: all
    networks:
      - vpclub-bridge

# ollama pull deepseek-r1
# ollama pull deepseek-r1
```

```
# ollama pull bge-m3
```

```
diffy docker docker
```

```
docker-compose.yml docker diffy
```

```
10
```

```
docker-compose up -d docker-compose --profile=xxx up -d
```

```
docker-compose.yml
```

- api
- worker
- web
- db
- redis
- sandbox
- plugin_daemon
- ssrf_proxy
- nginx
- weaviate

```
env
```

```
env.example env env80 env443
```

<https://docs.dify.ai/zh-hans/getting-started/install-self-hosted/environments>

```
CONSOLE_API_URL=  
CONSOLE_WEB_URL=  
SERVICE_API_URL=  
APP_API_URL=  
APP_WEB_URL=  
FILES_URL=  
  
#  
EXPOSE_NGINX_PORT=80  
EXPOSE_NGINX_SSL_PORT=443  
  
# false, https://updates.dify.ai
```

```
# CloudFlare Worker
#
CHECK_UPDATE_URL=

#
VECTOR_STORE=weaviate

# Weaviate http://weaviate: 8080
WEAVIATE_ENDPOINT=http://weaviate: 8080

# Weaviate api-key
WEAVIATE_API_KEY=WVF5YThaHlkYwhGUSmCRgsX3tD5ngdN8pkih
```

ollama

dify

The screenshot shows the Dify AI interface. On the left, a sidebar menu has '模型供应商' (Model Providers) highlighted with a red box. The main area shows a list of model providers under the heading '模型供应商'. The providers listed are Ollama, 深度求索 (DeepSeek), OpenAI, Anthropic, AWS, Azure OpenAI, Azure AI Studio, and Cohere. A red text overlay is placed over the Ollama provider card, stating: '如果没有增加模型供应商，在这里会出来ollama 我已经安装了，所以就不出来了'.

llama

模型列表

系统模型设置

Ollama

LLM TEXT EMBEDDING

点击这里添加模型

4 个模型

点击这里查看已经增加的模型

添加模型

deepseek-r1:1.5b LLM CHAT 4K



deepseek-r1 LLM CHAT 4K



bge-m3 TEXT EMBEDDING 4K



nomiic-embed-text TEXT EMBEDDING 4K



2 LLM Text Embedding ollama

- deepseek-r1 LLM
- nomiic-embed-text
- bge-m3 100

添加 Ollama

模型类型 *

LLM
 Text Embedding

模型名称 *

deepseek-r1:7b

基础 URL *

http://ollama:11434

模型类型 *

对话 ×

模型上下文长度 *

4096

最大 token 上限 *

4096

是否支持 Vision

是
 否

最多支持 10 个模型

[如何集成 Ollama](#)



word markdown -> wps ai



选择数据源



导入已有文本



同步自 Notion 内容



同步自 Web 站点

上传文本文件



拖拽文件至此，或者 [选择文件](#)

已支持 TXT、MARKDOWN、MDX、PDF、HTML、XLSX、XLS、DOCX、CSV、MD、HTM，每个文件不超过 15MB。



威海智慧谷知识库.docx
DOCX · 0.02MB



下一步 →

[创建一个空知识库](#)



分段设置



通用

通用文本分块模式，检索和召回的块是相同的



父子分段

选择父子分段

使用父子模式时，子块用于检索，父块用作上下文

父块用作上下文



段落

此模式根据分隔符和最大块长度将文本拆分为段落，使用拆分文本作为检索的父块

分段标识符 [ⓘ]

分段最大长度

tokens



全文

整个文档用作父块并直接检索。请注意，出于性能原因，超过10000个标记的文本将被自动截断。

子块用于检索

分段标识符 [ⓘ]

分段最大长度

tokens



文本预处理规则

替换掉连续的空格、换行符和制表符

删除所有 URL 和电子邮件地址

预览块

重置



索引方式



高质量 推荐

调用嵌入模型处理文档以实现更精确的检索，可以帮助LLM生成高质量的答案。



经济

每个数据块使用10个关键词进行检索，不会消耗任何tokens，但会以降低检索准确性为代价。

⚠ 使用高质量模式进行嵌入后，无法切换回经济模式。

Embedding 模型

bge-m3

检索设置

[了解更多](#)关于检索方法，您可以随时在知识库设置中更改此设置。



向量检索

通过生成查询嵌入并查询与其向量表示最相似的文本分段



Rerank 模型

Top K

3

Score 阈值

0.5



全文检索

索引文档中的所有词汇，从而允许用户查询任意词汇，并返回包含这些词汇的文本片段



混合检索 推荐

同时执行全文检索和向量检索，并应用重排序步骤，从两类查询结果中选择匹配用户问题的最佳结果，用户可以选择设置权重或配置重新排序模型。



创建空白应用

选择应用类型

新手适用



聊天助手

简单配置即可构建基于 LLM 的对话机器人



Agent

具备推理与自主工具调用的智能助手



文本生成应用

用于文本生成任务的 AI 助手

进阶用户适用



BETA

Chatflow

支持记忆的复杂多轮对话 workflow



BETA

工作流

面向单轮自动化任务的编排 workflow

应用名称 & 图标

威海小智



描述 (可选)

输入应用的描述

没有想法? 试试我们的模板 →

取消

创建





编排

提示词 生成

在这里写你的提示词，输入 '{' 插入变量、输入 '/' 插入提示内容块

提示词，好的提示词能得到意想不到的效果

0

变量 添加

变量能使用户输入表单引入提示词或开场白，你可以试试在提示词中输入 {{input}}

知识库 召回设置 添加

您可以导入知识库作为上下文

模型选择和参数调整

deepseek-chat CHAT 发布

发布

调试与预览

保存并发布应用

调试窗口

和机器人聊天 发送

功能已开启

更多设置

管理





□□□□□□

□□AI□□□□□□□□□□□□□□

 代码改写 对代码进行修改，来实现纠错、注释、调优等。	 代码解释 对代码进行解释，来帮助理解代码内容。
 代码生成 让模型生成一段完成特定功能的代码。	 内容分类 对文本内容进行分析，并对齐进行自动归类
 结构化输出 将内容转化为 Json，来方便后续程序处理	 角色扮演 (自定义人设) 自定义人设，来与用户进行角色扮演。
 角色扮演 (情景续写) 提供一个场景，让模型模拟该场景下的任务对话	 散文写作 让模型根据提示词创作散文
 诗歌创作 让模型根据提示词，创作诗歌	 文案大纲生成 根据用户提供的主题，来生成文案大纲
 宣传标语生成 让模型生成贴合商品信息的宣传标语。	 模型提示词生成 根据用户需求，帮助生成高质量提示词
 中英翻译专家 中英文互译，对用户输入内容进行翻译	

[[[]]]curl [[[]]]postman[[[]]]deepseek api[[[]]]deepseek[[api key]]

```
curl --location 'https://api.deepseek.com/chat/completions' \  
--header 'Content-Type: application/json' \  
--header 'Authorization: Bearer sk-xxxxxxxxxxxxxxxxxxx' \  
--data '{  
  "model": "deepseek-chat",  
  "messages": [  
    {"role": "system", "content": "[[[]]]L. [[[]]] Markdo  
    {"role": "user", "content": "[[[]]] \' [[[]]] \' [[[]]]"}  
  ],  
  "stream": false  
}'
```

[[[]]]deepseek, [[[]]]

[[[]]]deepseek <https://www.deepseek.com/>

DeepSeek-R1 已发布并开源，性能对标 OpenAI o1 正式版，在网页端、APP 和 API 全面上线，点击查看详情。

deepseek

探索未至之境

API key

deepseek 开放平台

- 用量信息
- API keys
- 充值
- 账单

API keys

列表内是你的全部 API key，API key 仅在创建时可见可复制，请妥善保存。不要与他人共享你的 API key，或将其暴露在浏览器或其他客户端代码中。为了保护你的帐户安全，我们可能会自动禁用我们发现已公开泄露的 API key。我们未对 2024 年 4 月 25 日前创建的 API key 的使用情况进行追踪。

名称	Key	创建日期	最新使用日期	
iovhm-com	sk-cf29e*****db46	2025-03-06	2025-03-06	

创建 API key

10 5-10 key

deepseek 开放平台

- 用量信息
- API keys
- 充值
- 账单

充值

在线充值 对公汇款

支付金额

¥10 ¥20 ¥50 ¥100 ¥300 ¥500 自定义 查看价格

【错峰优惠活动】北京时间每日 00:30-08:30 为错峰时段，API 调用价格大幅下调：DeepSeek-V3 降至原价的 50%，DeepSeek-R1 降至 25%，在该时段调用享受更经济更流畅的服务体验。【查看价格详情】

支付方式

支付宝 ALIPAY

微信支付

去支付

deepseek

#

编排

提示词

- 3. **问题解答**：解答关于园区政策、安全规定、设备支持等常见问题。
- 4. **服务预约**：协助用户进行会议室预订、设备租赁等服务预约。
- 5. **反馈收集**：收集用户对园区服务的反馈和建议，帮助园区管理方优化服务。

知识库

1. **园区信息**：包括园区地图、设施介绍、服务项目等。
2. **政策法规**：园区相关的政策、规定及安全指南。
3. **技术支持**：常见技术问题的解决方案和操作指南。
4. **服务流程**：各类服务的预约流程、使用指南等。

交互示例

- **用户**：最近的咖啡厅在哪里？
- **助手**：贵宾，您好，园区内最近的咖啡厅位于A栋一楼，营业时间为早上8点到晚上8点。您可以通过园区导航系统找到具体位置，希望我的服务能帮助您。如有任何问题，欢迎随时咨询。
- **用户**：预订一个会议室
- **助手**：贵宾，您好，园区内可预订的会议室有A栋101、B栋202和C栋303，您可以通过园区APP或前台进行预订，希望我的服务能帮助您。如有任何问题，欢迎随时咨询。
- **用户**：...漏水...灯不亮了...门打不开了
- **助手**：贵宾，您好，感谢您的反馈，如果您需要报修，请在智慧谷app提交工单，收到您的工单后，我们会立即安排维修部门上门维修，希望我的服务能帮助您。如有任何问题，欢迎随时咨询。

提示词

- **查询信息**：请告诉我关于...
- **导航指引**：我需要找到...
- **问题解答**：关于...的问题，我需要帮助。

944

变量

变量能使用户输入表单引入提示词或开场白

知识库

调试与预览

你好，我是小智，是威海智慧谷的智能客服，我将竭诚为您服务，请问有什么可以帮到您的？

一卡通用不了捏

贵宾，您好，关于一卡通无法使用的问题，请您先确认以下几点：

1. **一卡通绑定**：请确保您已经按照正确的步骤在“智慧谷金领驿站”公众号中绑定了一卡通账号和车牌号。如果未绑定，请按照以下步骤操作：
 - 进入“智慧谷金领驿站”公众号聊天界面，点击“园区停车” - “车辆绑定”功能进行车牌号绑定。
 - 绑定完成后，点击“一卡通绑定”，填写一卡通账号和密码保存。
2. **月卡状态**：如果您使用的是月卡，请确认您的月卡是否在有效期内。月卡的有效期为购买次日生效后的一个月。例如，11月5日购买的月卡，有效期为11月6日0点至12月6日0点。
3. **车牌信息**：请确保您在绑定车辆信息时填写了正确的车牌颜色，错误的车牌颜色可能导致月卡不生效。

如果以上步骤都已确认无误，但一卡通仍然无法使用，建议您联系园区客服或通过智慧谷APP提交工单，我们将尽快为您解答。希望我的服务能帮助您。如有任何问题，欢迎随时咨询。

引用

威海智慧谷知识库.docx

耗时 24.78 秒 花瓣

月卡如何续费? 绑定失败怎么办? 客服联系方式?

和机器人聊天

功能已开启

AI

AI

AI dify API



威海智慧谷智能客服问答

编辑信息 复制 导出 DSL

威海小智 运行中

公开访问 URL

http://192.168.0.11/chat/nCl66e2YMD8rpt94

启动 嵌入 定制化 设置

后端服务 API 运行中

API 访问凭据

http://192.168.0.11/v1

API 密钥 查阅 API 文档

生成

调试

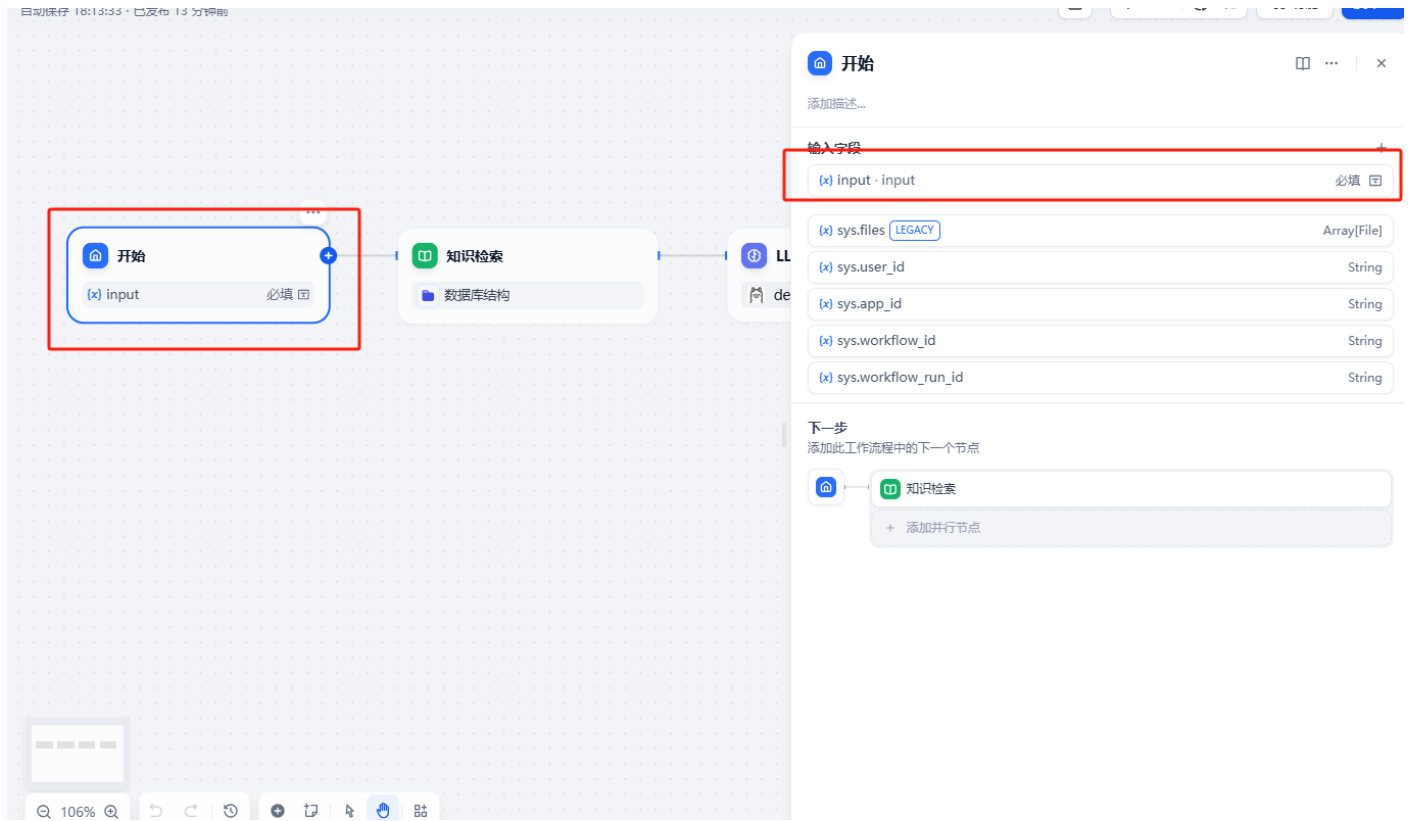
...等常见问题。
...服务预约。
...办园区管理方优化服务。

...楼，营业时间为早上8点到晚上8点。您可以通过园区导
...问题，欢迎随时咨询。

...1、B栋202和C栋303，您可以通过园区APP或前台进
...行咨询。

...，请在智慧谷app提交工单，收到您的工单后，我们会
...有任何问题，欢迎随时咨询。






```
- **[]**[]20[]
- **[]**[]select name,idcard from users where age>=20 and sex=1
```

- [][],[]sql[]

[][]vp_xxx

```
# []vp_xxx_user []
## []
- id[]ID[]
- name[]
- email[]
- idcard[]

# []vp_xxx_order []
## []
- order_id[]ID[]
- user_id[]ID[]
- amount[]
- img_url[]

# []vp_xxx_customer []
## []
- id[]ID[]
- name[]
- email[]
- idcard[]
```

[][]

 结束

 ... | 

添加描述...

输出变量

+

output

 LLM / text String



 结束

 LLM / (x) text



Test Run#45

输入 结果 详情 追踪

- > 开始 13.713 ms ✓
- > 知识检索 63.368 ms ✓
- > LLM 351 tokens · 17.733 s ✓
- > 结束 6.265 ms ✓



Test Run#40

输入 结果 详情 追踪

SQL

```
1 select name, idcard from vp_xxxx_customer where age > 20;
```

复制





firecrawl



#####S#####VUEhtml#####chrome#####



- #####S
- #####LLM#####markdown#####HTML
- #####main#####



- <https://docs.firecrawl.dev/introduction>
- ~~firecrawl~~#####M#####<https://docs.firecrawl.dev/v1-welcome>
- github#####<https://github.com/mendableai/firecrawl>



~~docker~~-compose build ~~docker~~-compose up -d

playwright-service

#####S#####

#####.env#####

```
#  
NUM_WORKERS_PER_QUEUE=8  
PORT=3002  
HOST=0. 0. 0. 0  
REDIS_URL=redis://redis: 6379  
REDIS_RATE_LIMIT_URL=redis://redis: 6379  
PLAYWRIGHT_MICROSERVICE_URL=http://playwright-service: 3000/html  
  
#  
USE_DB_AUTHENTICATION=false
```



```
-H 'Authorization: Bearer YOUR_API_KEY' \  
-d '{  
  "url": "https://firecrawl.dev"  
}'  
  
``  
  
- [ ]  
  
```shell
```

```
curl -X POST http://localhost:3002/v1/search \
-H 'Content-Type: application/json' \
-H 'Authorization: Bearer YOUR_API_KEY' \
-d '{
 "query": "AI tools",
 "limit": 5,
 "scrapeOptions": {
 "formats": ["markdown"]
 }
}'
```

- [ ]

```
curl -X POST http://localhost:3002/v1/extract \
-H 'Content-Type: application/json' \
-H 'Authorization: Bearer YOUR_API_KEY' \
-d '{
 "url": "https://example.com",
 "extract": {
 "schema": {
 "type": "object",
 "properties": {
 "title": {"type": "string"},
 "price": {"type": "number"}
 }
 }
 }
}'
```

```
}'
```

- `{}{}``

```
curl -X POST http://localhost:3002/v1/batch/scrape \
 -H 'Content-Type: application/json' \
 -H 'Authorization: Bearer YOUR_API_KEY' \
 -d '{
 "urls": ["https://example1.com", "https://example2.com"],
 "options": {
 "formats": ["markdown"]
 }
 }'
}'
```

# deepseek

deepseek, A! A! wps

- Role:
- Goals: XML, python, curl
- Constrains: " Swagger "
- OutputFormat: " Swagger "
- Example:
- Optional:

deepseek

#

##

##

- <context></context>XML

-

- " " "

##

- \*\*

- \*\* A B B

- 

##

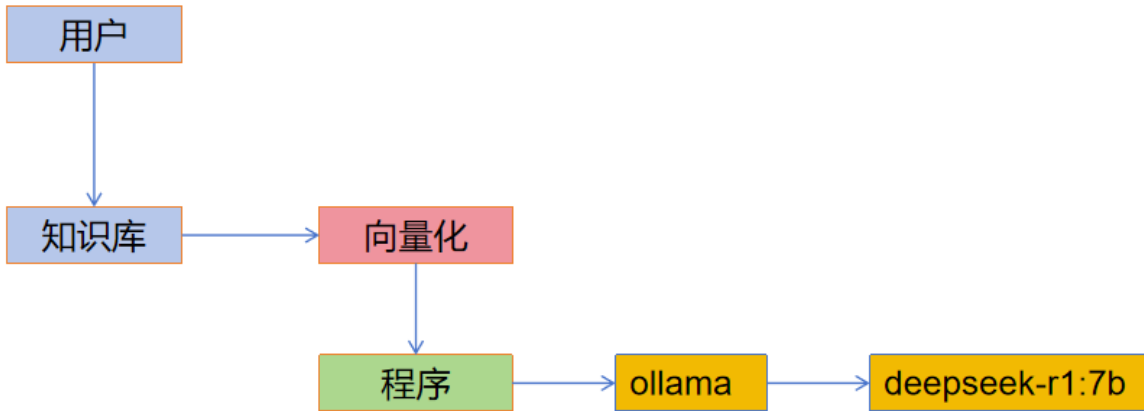


# deepseek

## deepseek

LLM

用户编写，采集知识库并向量化



程序从向量话程序RAG检索，并传递给运算平台云上

- 

- 

dify OpenaI http dify

- 

ollama

- 

deepseek



```
import requests
import json

bas_url = "http://localhost:3002/v1/scrape"

headers = {"Content-Type": "application/json"}

req_data = {
 "url": "http://www.eweihai.gov.cn/art/2025/3/10/art_159136_5310185.html",
 "formats": ["markdown", "links"],
 "includeTags": [".page-bd.article-bd"],
 "onlyMainContent": True,
}

response = requests.post(bas_url, headers=headers, data=json.dumps(req_data))
print(response.json())
```



ollama

```
import openai

base_url = "http://192.168.0.11:11434/v1"
api_key = "sk-"

response = client.chat.completions.create(
```

```

model="deepseek-r1:1.5b",
messages=[
 {"role": "system", "content": "You are a helpful assistant."},
 {"role": "user", "content": "Hello!"},
],
stream=False,
)

print(response.choices[0].message.content)

[]
response = client.chat.completions.create(
model="deepseek-r1:1.5b",
messages=[
{"role": "system", "content": "You are a helpful assistant."},
{"role": "user", "content": "Hello!"},
],
stream=True,
)

for chunk in response:
[]
if chunk.choices and chunk.choices[0].delta.content:
print(chunk.choices[0].delta.content)

```

# Deepseek

- [ ]
- [ ]



deepseek-v3 qwen-plus deepseek-chat deepseek-v3

模型名称	耗时	模型类型	效果	结果
同义千文qwq-plus	32.02	深度推理	较好	贵宾，您好，办理停车月卡的流程如下：请登录智慧谷收费系统（网址： <a href="https://wvpark.com:802">https://wvpark.com:802</a> ），依次进入服务大厅、个人中心、变更车辆登记，填写相关信息并提交申请。审核通过后即可录入车辆信息。需要注意的是，每家企业仅限一辆车免费通行，且每年只能变更一次免费车牌。如果需要办理付费月卡，每月费用为120元。希望我的服务能帮助您。如有任何问题，欢迎随时咨询。
deepseek-r1	28.88	深度推理	较好	贵宾，您好，办理停车月卡请登录智慧谷收费系统（网址： <a href="https://wvpark.com:802">https://wvpark.com:802</a> ），依次进入【服务大厅】-【个人中心】-【变更车辆登记】，填写相关信息并提交申请。待园区审核通过后，即可完成车辆信息录入并享受月卡权益。请注意，每家园区企业仅限申请1辆免费通行车辆，且每年仅允许变更1次车牌。停车月卡费用为每月120元。希望我的服务能帮助您。如有任何问题，欢迎随时咨询。
deepseek-v3	5.5	chat	较好	贵宾，您好，办理停车月卡可以登录智慧谷收费系统，依次进入服务大厅、个人中心、变更车辆登记，按照规定填写相关信息，并点击确定以提交申请。每月费用为120元。待园区相关部门审核并确认后，您即可录入车辆信息，享受停车月卡服务。请注意，每家园区企业仅限申请一辆车辆享受免费通行，且每年仅允许变更一次免费车牌。智慧谷收费系统的网址为： <a href="https://wvpark.com:802">https://wvpark.com:802</a> 。希望我的服务能帮助您。如有任何问题，欢迎随时咨询。
qwen-plus	6.52	chat	较好	办理停车月卡，请登录智慧谷收费系统，进入服务大厅、个人中心、变更车辆登记页面，按照要求填写相关信息并提交申请。审核通过后，您就可以享受园区内停车月卡的服务了。智慧谷收费系统的网址是： <a href="https://wvpark.com:802">https://wvpark.com:802</a> 。需要注意的是，每家园区企业仅限申请一辆车享受免费通行，且每年仅允许变更一次免费车牌。如果需要办理收费的停车月卡，每月费用为120元。
qwen-turbo	3.63	chat	差	您可以登录智慧谷收费系统，进入服务大厅，然后选择个人中心，接着点击变更车辆登记。按照要求填写相关信息后，点击确定提交申请。经过园区相关部门审核并确认后，就可以录入车辆信息，开始享受停车月卡的权益了。
deepseek-chat(应该就是deepseek-v3)	8.47	chat	较好	贵宾，您好，办理停车月卡需要登录智慧谷收费系统（网址： <a href="https://wvpark.com:802">https://wvpark.com:802</a> ），依次进入服务大厅、个人中心、变更车辆登记，填写相关信息并提交申请。审核通过后即可录入车辆信息。每月费用为120元，请注意每家园区企业仅限申请一辆免费通行车辆，且每年仅允许变更一次车牌。希望我的服务能帮助您。如有任何问题，欢迎随时咨询。



- Role: [ ]
- Goals: [ ]
- Constrains: [ ]
- OutputFormat: [ ]
- Example: [ ]
- Optional: [ ]



- Role: [ ]
- Goals: [ ], [python][curl][<context></context>XML [ ]
- Constrains: [ ] " [Swagger [ ]"
- OutputFormat: [ ] " [Swagger [ ]"

## mock [ ]

- Role: [API [ ]
- Goals: [ ]
- Constraints: [ ]
- Output Format: [ ] Markdown [ ]