



DeepSeek的技术解析和应用实践

杨海龙 博士

霞客湾创智园分享





杨海龙，江南大学人工智能与计算学院博士，研究领域包括编程语言设计、生成式模型和多智能体系统。曾华为编程语言实验室技术专家。现江汉大学特聘专家，华为仓颉社区KOL，苍穹（CangChain）AI智能体框架的作者。此外，他还是编程语言开放社区（PLOC）的创始理事以及DeepLang编程语言的主创。





第一部分：DeepSeek技术解析





DeepSeek是什么？

DeepSeek是杭州深度求索人工智能基础技术研究有限公司推出的一款创新大语言模型。公司成立于2023年7月17日，由知名私募巨头幻方量化孕育而生。DeepSeek致力于开发和应用先进的大语言模型技术。

深度小助手 聪明且低成本 聪明强大能干 中国本土AI



我是 DeepSeek，很高兴见到你！

我可以帮你写代码、读文件、写作各种创意内容，请把你的任务交给我吧~

给 DeepSeek 发送消息

深度思考 (R1)

联网搜索



17:52:49



DeepSeek开源模型



DeepSeek-R1 发布 2025/01/20

在开源 DeepSeek-R1-Zero 和 DeepSeek-R1 两个 660B 模型的同时，通过 DeepSeek-R1 的输出，蒸馏了 6 个小模型开源给社区，其中 32B 和 70B 模型在多项能力上实现了对标 OpenAI o1-mini 的效果。



DeepSeek-V3 发布 2024/12/26

DeepSeek-V3 为自研 MoE 模型，671B 参数，激活 37B，在 14.8T token 上进行了预训练。



DeepSeek为什么引起关注？

一、技术突破：为什么DeepSeek的模型值得关注？

1. 凭“MoE+MLA+MTP”成为效率狂魔

MoE：将一个巨大的AI模型切割成多个子模型

MLA：创造性地提高推理效率，多头潜在注意力机制，层次化索引。

MTP：机制提高预测能力，将“走一步看一步”的写作方式，变成了“先拟大纲再填充字词”

2. 蒸馏技术

数据蒸馏：有监督微调(SFT)，蒸馏出的数据对小模型进行直接微调

知识蒸馏：高效知识迁移策略，基于特征的蒸馏和特定任务蒸馏

二、开源生态：DeepSeek如何改变开发者社区？

1. 开放模型与工具链

全面开发：DeepSeek开发了多家训练代码，数据清洗Pipeline资源被覆盖工具包（如DeepSeek-Tuner），最大程度实现以下开发门槛：

标准化部署：提供模型压缩工具（如4-bit量化适配TensorRT-LLIT）

2. 社区驱动创新

开发者基于DeepSeek模型快速构建垂直应用

金融场景

教育场景

三、行业落地：DeepSeek推动的技术范式迁移

1. 从“通用模型”到“领域专家”

传统大数据（ISO/PT-3.5）依赖Prompt Engineering适配行业需求，而DeepSeek通过预测结果获取入微型模型，减少后端数据流失

2. 成本革命

通过模型结构高效管理模型，企业可基于本标准专业模型，推荐成本降低至CPT-4 PPI的1/50

如：未来将带来高使用DeepSeek-7B替代CPT-4，本次实现成本从0.05降至0.001，日均处理量提升10倍。

四、行业竞争格局：DeepSeek的“鲶鱼效应”

1. 倒逼问题模型降低

DeepSeek的开发策略是底层厂商调整定位，例如，Anthropic和Claude 3 Sonnet API价格在DeepSeek开源后下降

2. 催化国产AI芯片生态

DeepSeek与华为开源，逐渐实现了高深度合作，优化模型在国产硬件的重要性，例如，DeepSeek不支持超30 LGI设备重出F100高

3. 推动AI技术产业化

中小企业研发阶段对现有开发模式进行优化，无需依赖巨头API，例如，非洲冰淇淋团队用DeepSeek-7B开发来优化项目导向，成本仅为50个方案的1/50



第二部分：DeepSeek应用实践





I.效率工具





LangBot大模型原生即时通信机器人平台

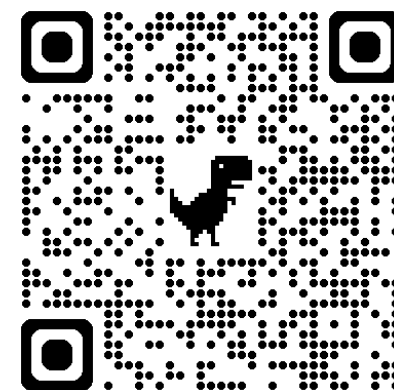
消息平台

平台	状态
QQ 个人号	✓
QQ 官方机器人	✓
企业微信	✓
个人微信	✓
微信公众号	✓
飞书	✓
钉钉	✓
Discord	✓
Telegram	✓
WhatsApp	🚧

🚧: 正在开发中

大模型

模型	状态
OpenAI	✓
DeepSeek	✓
Moonshot	✓
Anthropic	✓
xAI	✓
智谱AI	✓
Dify	✓
Ollama	✓
LMStudio	✓
GiteeAI	✓
SiliconFlow	✓
阿里云百炼	✓
火山方舟	✓



<https://github.com/RockChinQ/LangBot>



OfficeAI : DeepSeek与Excel处理

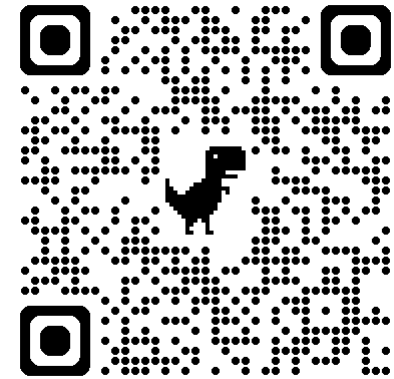
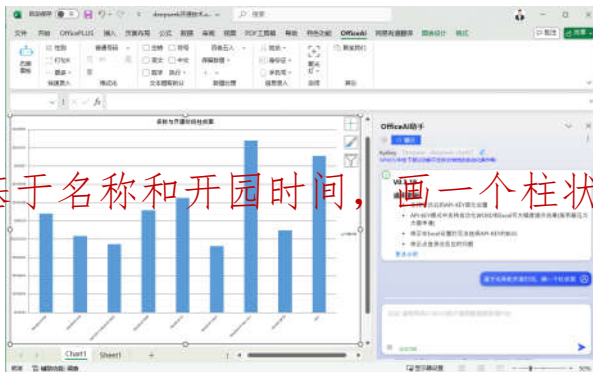
OfficeAI插件，通过VBS间接控制Excel

1、接入Deepseek



2、chat的方式处理当前数据

基于名称和开园时间，画一个柱状图

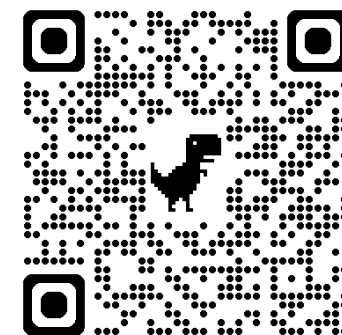


<https://www.office-ai.cn/>



ChatExcel: 另一个Excel的效率工具

- 自动处理Excel
- 智能数据分析
- 数据可视化
- 文档转Excel



<https://chatexcel.com>

DeepSeek与PPT生成

1、先使用 DeepSeek 生成大纲和内容

我需要写一个ppt，主标题是“半导体智能制造的调研报告”，输出PPT的内容

我需要写一个ppt，主标题是“半导体智能制造的调研报告”，输出PPT的内容



17:52:49

2、豆包AI PPT生成





II.多智能体系统



智能体是什么？

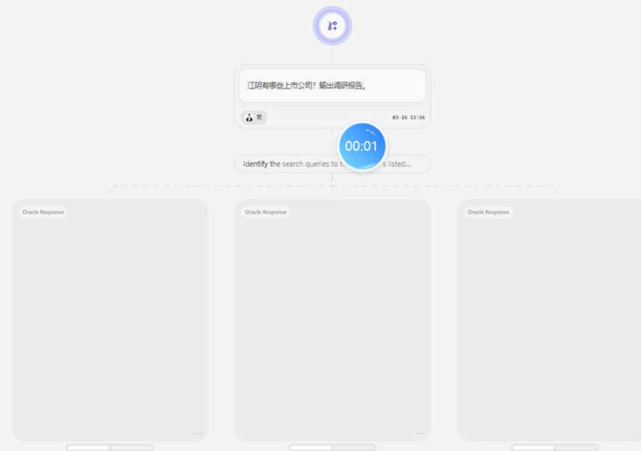
基于大语言模型（LLM）的智能体是指利用大语言模型作为核心组件，能够理解、生成自然语言并执行任务的智能系统。这类智能体通过语言模型的能力与环境交互，处理复杂任务，并展现出一定的自主性和适应性。



Flowith通用智能体：调研报告

flowith

江阴有哪些上市公司？输出调研报告。



Identify the search queries to find Jiangyin's listed companies. Formulate search queries using...



Flowith通用智能体：调研报告

Flowith: 一款创新的AI交互式搜索和对话工具，基于节点式交互方式，支持多种AI模型和图像生成技术，提供插件系统和社区功能。类似工具Manus, OpenManus, MGX, Dify, AutoGen, Crew.AI

Good afternoon

Ready to bring your ideas to life?

江阴有哪些上市公司?输出调研报告

ORACLE MODE

江阴有哪些上市公司？输出调研报告。

开始

常规模式

DeepSeek V3

KNOWLEDGE BASE



Summarization

好的，这是一份关于江阴上市公司信息的调研报告：

江阴上市公司调研报告

江阴市，作为中国A股上市公司数量最多的县级市之一，被誉为“A股第一县”。本报告旨在梳理江阴市的上市公司情况，包括上市公司名称、股票代码、所属行业及简要介绍。

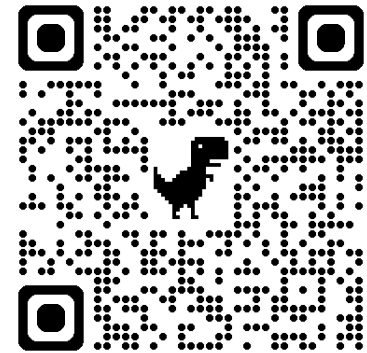
上市公司数量： 江阴共有61家上市公司

- 截至2023年底，江阴共有61家上市公司，其中包括A股、港股及其他类型的境外上市公司。
- 其中，A股上市公司数量为33家。

上市公司列表 (部分):

以下表格列出了部分江阴的上市公司信息。由于数据来源不统一，部分信息可能存在差异，最新数据请以交易所公告为准。

公司名称	股票代码	所属行业	简要介绍
海澜之家集团 股份有限公司	600398.SH	纺织服装	国内知名服装品牌，主营服装及相关产品的设计、生产和销售。
江苏三房巷聚 材股份有限公 司	600370.SH	石化新材 料	主要从事PTA、聚酯切片等石化产品的生产和销售。



<https://flowith.io/blank>

<https://flowith.io/oracle-play/32f343bd-2130-44ae-8df7-fb4c25a74a48?oracle-id=9dcabdf5-6f34-4621-8f42-19cb553179b3>

MGX通用多智能体：调研报告

MGX通用多智能体：调研报告

多角色的调研团队

Team 5

Mike Team Leader

Emma Product Manager

David Data Analyst

Bob Architect

Alex Engineer

团队负责人 产品经理 数据分析师 架构师 工程师

Mike Emma David Bob Alex

当前聊天文件

docs

intelligent_manufactu...

code.ipynb

competitors_chart.png

market_size_chart.png

tech_applications_chart....

Global resource

My Resource

目前智能制造行业形成了四大主要产业主体：

1. 传统企业：装备/自动化企业、软件企业及制造企业
2. ICT企业：以阿里、华为为代表的互联网企业
3. 初创企业：如三一重工孵化的树根互联、TCL孵化的格创东智等
4. 研究机构：提供技术支持和研发创新

2.2 头部企业分析

2023年智能制造领域的主要企业：

- 第一梯队：比亚迪、航天科技、华为
- 第二梯队：海尔智家、宁德时代、中国商飞
- 第三梯队：阿里巴巴、大疆创新、中国船舶、京东方

2.3 区域分布特征

- 企业主要集中在东部和中南地区
- 广东、江苏分布最为集中
- 上海地区次之
- 西部地区分布相对较少

2.4 市场竞争特点

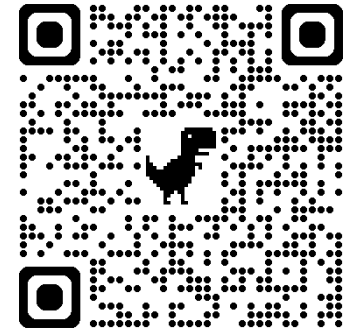
1. 行业整体集中度不高
2. 细分领域各有龙头
3. 尚未出现行业整体绝对龙头
4. 技术专利分散，前十企业专利申请量占比14%-18%

David: Convert the enhanced report with charts to PDF format and save it to the specified path.

Reply task:

@ to chat, # to select

Add deepse...



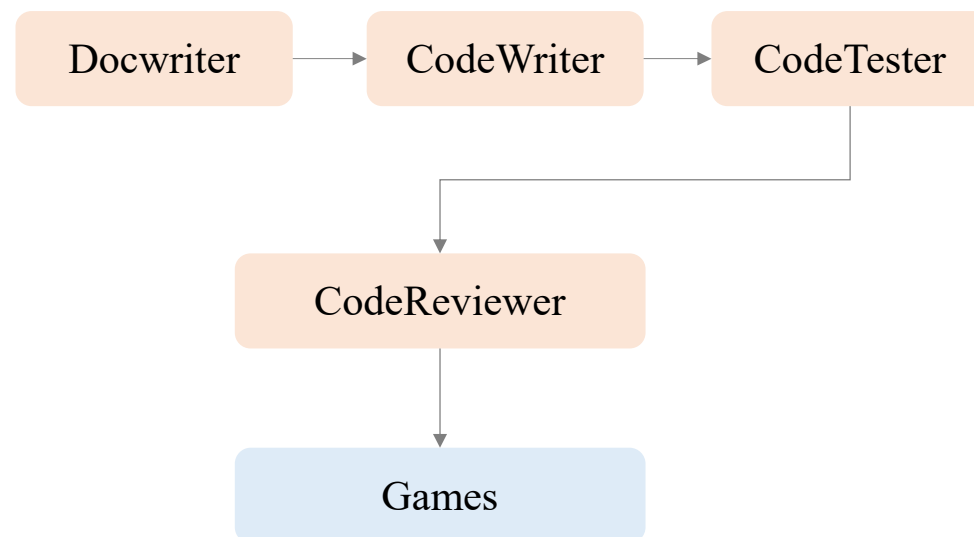
<https://mgx.dev/>

</> 项目1：仓颉代码生成

```
src > agent > planner.q > PlannerPromptTemplate > formatHashMap<String, String>: String
45 // Planner主要是用来挑选下一步的Agent角色
46 public class PlannerAgent <: Role {
47     public init(tools!: ArrayList<BaseTool> = ArrayList<BaseTool>())
48         super("Planner", "The manager of multi agent to decide which
49         println("PlannerAgent init done.")
50     }
51
52     public func generate(input: Dict): String {
53         let output = super.predict(input["text"])[0]["text"]
54         output
55     }
56
57     public func shouldContinue(res: JsonObject): Bool {
58         match (res.get("message")) {
59             case None => return true
60             case Some(_) => return false
61         }
62     }
63 }
```

```
OpenAI init done.
=====
agent_name: Doc_Writer
query:编写一个2048游戏的文档，其中包括游戏规则、功能需求和实现细节。
=====
Executing role: Doc_Writer
OpenAI init done.
2048_Game_Documentation.txt is create successfully in current directory.
OpenAI init done.
=====
OpenAI init done.
=====
OpenAI init done.
=====
agent_name: Code_Writer
query:根据2048游戏的文档，编写2048游戏的代码。
=====
Executing role: Code_Writer
OpenAI init done.
2048_game.py is create successfully in current directory.
```

团队设计并实现了一个多智能体（Multi-Agent）的系统，包含以下四个关键角色：Docwriter、CodeWriter、CodeTester和CodeReviewer。





项目2：多智能体邮件处理系统

基于GLM构建，使用MindIE套件，运行在昇腾310芯片上。

原始邮件

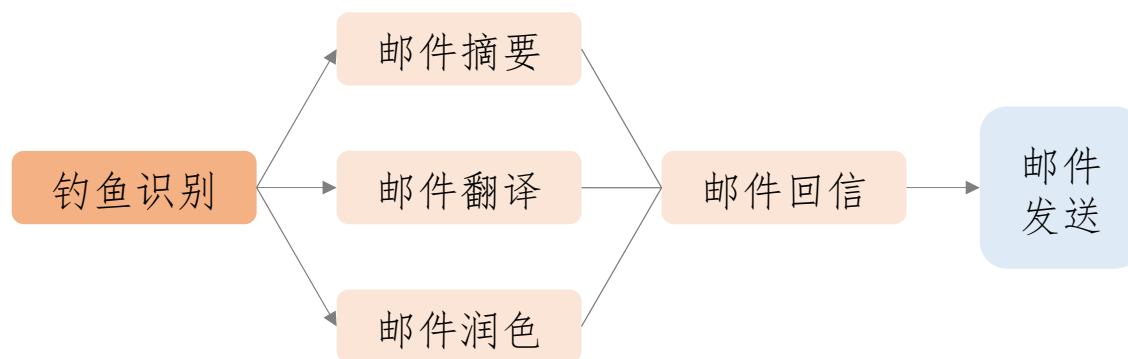
请输入原始邮件内容

请选择功能

☐ 邮件摘要 ☐ 邮件翻译 ☐ 邮件回信 ☐ 钓鱼邮件识别 ☐ 邮件润色

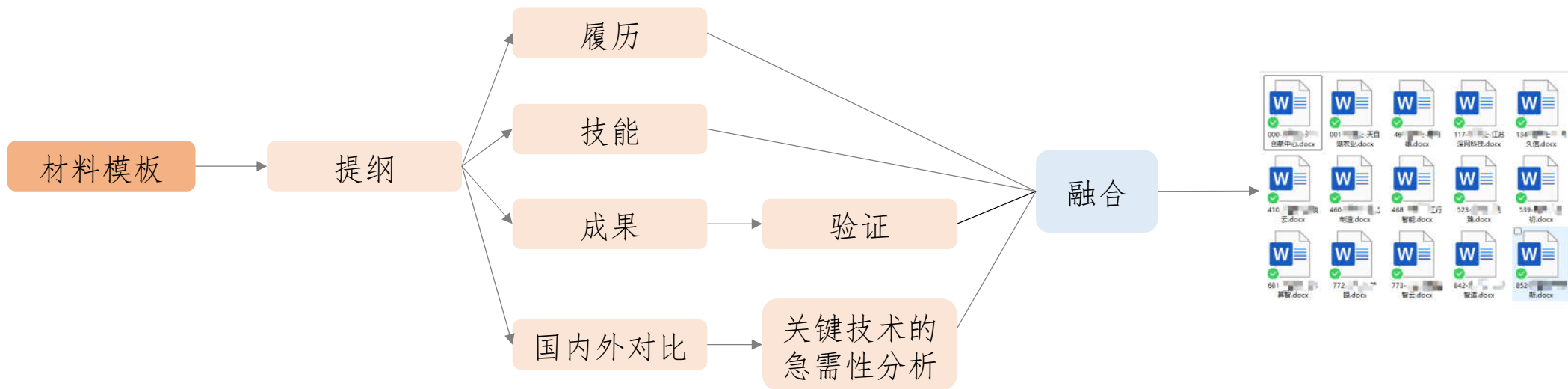
清空 提交

邮件处理执行图



项目3: xx人才文档系统

人才材料撰写执行图





III. 知识库





FastGPT: 在线知识库

FastGPT 是一个基于大语言模型的知识库问答系统，支持开箱即用的数据处理和模型调用。

The image displays the FastGPT web interface. On the left is a large text input area with a '调试预览' (Debug Preview) button. The right sidebar contains several sections:

- 企业知识库 (Enterprise Knowledge Base):** A table listing documents with columns for name, processing mode, data volume, creation/update time, status, and actions. The table shows four documents, all with a status of '已就绪' (Ready).
- 全部应用 (All Applications):** A section for managing applications, currently showing '企业知识库' (Enterprise Knowledge Base) with a description '招投标, 客户关系' (Tendering, Customer Relationship).
- AI 配置 (AI Configuration):** A section for configuring the AI model, currently set to 'Deepseek-v3'.
- 我的知识库 (My Knowledge Base):** A section for managing personal knowledge bases, currently showing '企业知识库' (Enterprise Knowledge Base) with a description '招投标资料' (Tendering Materials).

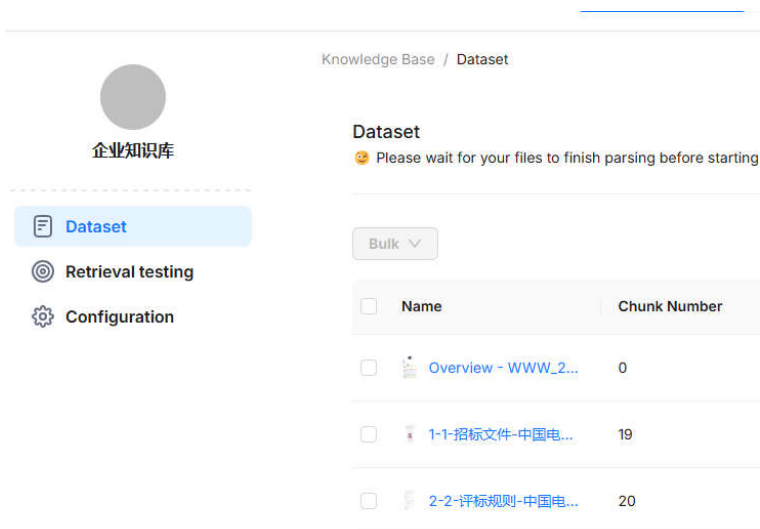
At the bottom left, a timestamp '17:52:49' is visible.



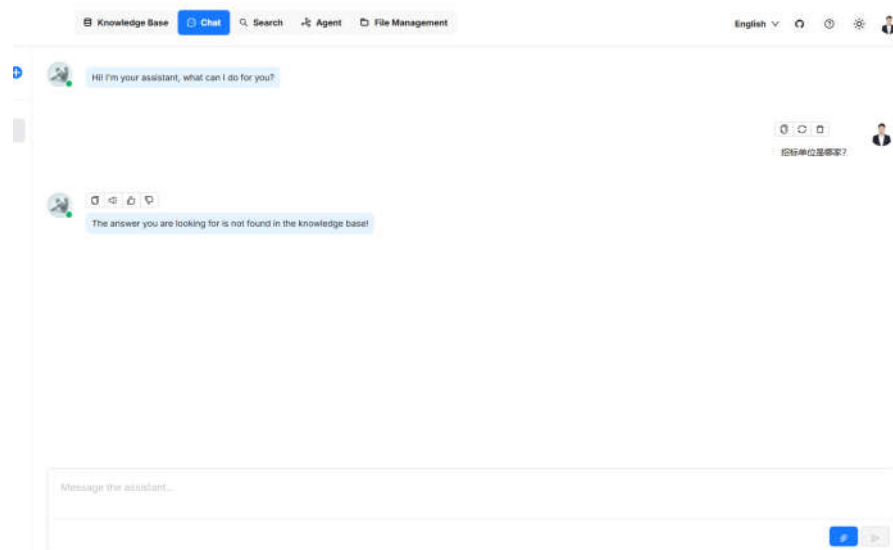
RagFlow

RAGFlow 是一个端到端的 RAG 解决方案，专注于深度文档理解技术，能够处理多种格式的文档（如 PDF、Word、Excel 等），并支持高质量的文本切片和向量化处理。

1、上传文档



2、chat 问答





Ollama部署DeepSeek

1、下载Ollama



Get up and running with large language models.

Run Llama 3.3, DeepSeek-R1, Phi-4, Mistral, Gemma 3, and other models, locally.

Download ↓

Available for macOS,
Linux, and Windows

2、运行Deepseek R1

```
ollama run deepseek-r1:1.5b
```

```
PS C:\Windows\System32> ollama run deepseek-r1:1.5b
```

3、进行Chat

你好！请问有什么我可以帮助你的吗？无论是问题、建议还是其他想法，我都乐意听到哦！ 😊

>>> 你是谁？
<think>

您好！我是由中国的深度求索（DeepSeek）公司开发的智能助手DeepSeek-R1。如您有任何任何问题或需要anything帮助，请随时告诉我！ 😊

>>> Send a message (/? for help)

DeepSeek-R1-Distill-Qwen-1.5B

```
ollama run deepseek-r1:1.5b
```

DeepSeek-R1-Distill-Qwen-7B

```
ollama run deepseek-r1:7b
```

DeepSeek-R1-Distill-Llama-8B

```
ollama run deepseek-r1:8b
```

DeepSeek-R1-Distill-Qwen-14B

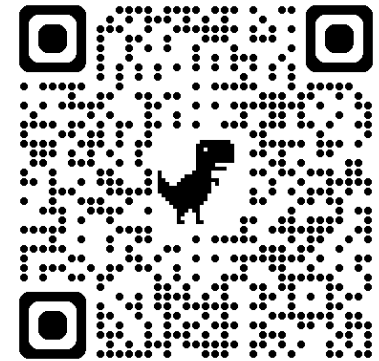
```
ollama run deepseek-r1:14b
```

DeepSeek-R1-Distill-Qwen-32B

```
ollama run deepseek-r1:32b
```

DeepSeek-R1-Distill-Llama-70B

```
ollama run deepseek-r1:70b
```



<https://ollama.com/>

AnythingLLM构建本地私有知识库

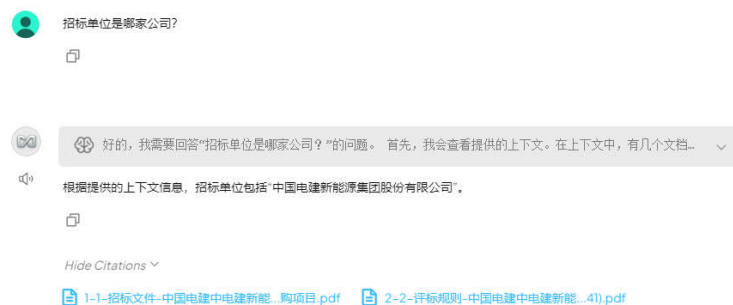
1、设置本地知识库



2、上传文档到知识库



4、问答Chat



3、基于上传文档的问答

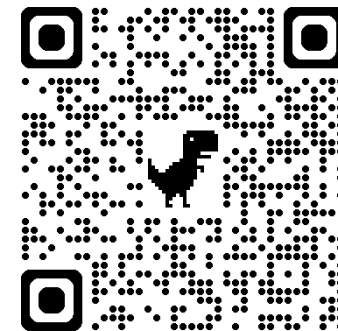
介绍一下江南大学。



Send a message



<https://anythingllm.com>





项目4：公/私域融合知识库

商业多智能体项目界面



首页 > 知识库管理

首页

市场池

后台管理

1. 取

2. 3

3. 5

4. 证

更新

统计

7. 与

8. Cha

9. 更 录入

10. 储存

11. 用户

17:52:49

上传文件

公域

私域

交互

请选择文件格式

重置

<input type="checkbox"/>	文件名	文件格式	上传时间	企业名称	上传者	操作
<input type="checkbox"/>	www.ennewpower.com_productCenter_info.html_1734348492616.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productApplication_info.html_1734348482654.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productCenter_1734348500504.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productCenter_info.html_1734348368621.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productCenter_info.html_1734348495143.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productBMS_info.html_1734348468187.md	文档	2024-12-16		admin	下载 删除
<input type="checkbox"/>	www.ennewpower.com_productApplication_info.html_1734348488040.md	文档	2024-12-16		admin	下载 删除

26



THANKS



AGI&FBHC



扫一扫上面的二维码图案，加我为朋友。