在阿里云快速启动LibreChat轻松玩转AI对话

什么是LibreChat

LibreChat有什么优势

如何部署LibreChat

前提条件

操作步骤

快速入门LibreChat

多AI模型

文件上传

快捷导出

语音输入

其他问题

- 1. 如何进入ECS
- 2. 如何添加转发地址
- 3. 如何添加第三方API模型

了解更多

参考资料

本文介绍了LibreChat的基本信息,并通过阿里云计算巢完成了LibreChat的快速部署,使用者不需要自己下载代码,不需要自己安装复杂的依赖,不需要了解底层技术,只需要在控制台图形界面点击几 下鼠标就可以快速部署并启动LibreChat,非技术同学也能轻松搞定。

什么是LibreChat

LibreChat是一个增强版的 ChatGPT , 它不仅集成了OpenAl、Assistants API、Azure、Groq 等多种 Al 模型, 还支持 GPT-4 Vision、Mistral、Bing 等前沿技术。

作为一个完全开源的自托管系统,LibreChat 提供了多用户支持、预设配置、消息搜索、DALL-E-3 图像生成等丰富功能,并且仍在不断开发中。更多信息,可访问链接。



LibreChat有什么优势

- **多模型支持**: LibreChat支持多种AI模型,包括OpenAI、Azure OpenAI、Anthropic Claude、 Google Vertex AI等。用户可以根据需求选择不同的模型,甚至在对话中切换模型。
- 本地和远程AI服务兼容:除了支持主流的云端AI服务,LibreChat还兼容多种本地AI服务,如 Ollama、koboldcpp等,为用户提供更多选择。
- **丰富的功能**:LibreChat提供了许多实用功能,如对话历史保存、消息搜索、自定义预设、多模态对 话(支持图像分析)、语音交互等。
- 插件系统: 支持各种插件, 如网络访问、DALL-E-3图像生成等, 进一步扩展了平台的能力。
- **多语言支持**:LibreChat的用户界面支持多种语言,包括中文、英语、德语、法语等,方便全球用户 使用。
- 安全性和隐私:提供多用户系统、安全认证和审核工具,保护用户数据和隐私。
- 灵活部署: 支持Docker部署、反向代理等多种部署方式, 适应不同的使用场景。

如何部署LibreChat

前提条件

• 已开通阿里云账号

• 确保账户余额超过100元

重要

使用阿里云资源,需要收费,请仔细阅读阿里云控制台的订单信息,然后再进行下一步操作。

操作步骤

1. 访问计算巢LibreChat社区版部署链接,按提示填写部署参数,确认参数后点击**下一步:确认订** 单:

									不业	≡ =	必填 1/6 全部 3/9
服务实例名称	请输入服务实例名称									服务实例名称	
	长度不超过64个字符,必须以粪;	文字母开头,可包含数字、英文字母、	短划线 (-) 和下划	助线 (_) ,	且需符合服务中定	义的资源的名称规范。				 付费类型配置 	
	40									付费类型	
- HERK	+26.1 (0570)									* 空间坐型	×
										* 实例密码	×
す 豊美型配置 へ										- 可用区配置	
										* 可用区ID	×
费类型	按量付费	包年包月								* 专有网络VPC实例ID	×
										交換机实例ID	
										 软件配置 	
源配置 ~										* 管理员密码	×
										 标签和资源组 	
实例类型	等选 选择vCPU	• 选择内存	•	搜索实例规	格	Q,				标签	
										资源组	
	架构 x86 计算										
	分类 通用算力型 共	享型									
	突倒规格族: ecs.u1,ecs.e										
	规格族 ②	实例规格	vCPU 🌐	内存 🔅	GPU/FPGA 🔿	处理器主频/睿频	参考价格 🔘 💠	处理器型号			
	 通用算力型 u1 	ecs.u1-c1m1.large	2 vCPU	2 GiB	-	2.5 GHz/-	¥ 0.33375/81	Intel(R) Xeon(R) Platinum	n.		
						0.5.000	¥ 0.351/8t	Intel(R) Xeon(R) Platinum			
	〇 通用算力型 u1	ecs.u1-c1m2.large	2 vCPU	4 GiB	-	2.0 GHZ/-	+ 0.00 (jag		n.		
	 通用算力型 u1 通用算力型 u1 	ecs.u1-c1m2.large ecs.u1-c1m4.large	2 vCPU 2 vCPU	4 GiB 8 GiB	-	2.5 GHz/-	¥ 0.45/时	Intel(R) Xeon(R) Platinum	n		
	 通用算力型 u1 通用算力型 u1 通用算力型 u1 	ecs.u1-c1m2.large ecs.u1-c1m4.large ecs.u1-c1m8.large	2 vCPU 2 vCPU 2 vCPU	4 GIB 8 GIB 16 GIB	-	2.5 GHz/- 2.5 GHz/- 2.5 GHz/-	¥ 0.45/#j ¥ 0.59625/#j	Intel(R) Xeon(R) Platinum Intel(R) Xeon(R) Platinum	n n		

2. 确认订单完成后同意服务协议并点击**立即创建**进入部署阶段。

创建服务实	例									4
LibreChat beta / 202410 LibreChat E-	社区版 三方服务 009172301 - 个开源的聊天应用程序、旨在提供用户	*发好的交流伴袖。它支持多种功能,包括实时消息传递、文件共享以	及集成各种第三方服务。Lit	breChat 以隐私和安	全为重点、确保/	用户的数据不补	皮溢用或收集。	服务商名称 服务商主页 服务商描述	计算巢社区服务 https://computenest.com 计算巢社区服务	isole.aliyun.con
张 务实例信息 ~									生成	OpenAPI 请求
±10	ap-southeast-1		付费类型	Ł	PostPaid					
¥买资源时长周期 □個素問	-		购买资源	明长 L						
用区ID	ap-southeast-1a		专有网络	VPC实例ID	vpc-t	(internet)				
E换机实例ID	vsw-t		管理员密	调						
蔷			资源组		-					
格预览。										
预估算费用仅供参考。 时长费用: x0.773	最终金额以实际产生的费用为准。									导出文件
资源类型	资源用途	规格	地域 费	数量 计费7	方式	时长	原价	优惠		預估算费用
ECS实例 (ECS)	InstanceGroup	实例: ecs.u1-c1m1.large (通用算力型 u1, 2 vCPU, 2 GiB) 系统盘: cloud essd 200G/B	亚太东南 1 (新加坡) 1	按量(す費	-	ecs.u1-c1m1.large: ¥0.82	7/小时 ecs.u1	-c1r ecs.u1-c1m1.large 波量费用:	¥0.773/Jult
書簡報户的余額充足 一步 立即前除	充储支付									
「「「「「「「「」」」 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	/ 创建超条车侧									
PERMIT PARA MENAN	0340.08.77 24.03									
- 创建服	务实例									
					A 10-4					
					♥ 提3	と成切				
						- 71 - 25 - 00	面地的建			
					ž	175-00 BL 48	1977,2530			

3. 等待部署完成后就可以开始使用服务,进入服务实例详情点击LibreChat链接。

计算巢	计算刷 / 服务实例				
欢迎页	服务实例				
服务目录		-			
我的服务	我的服务实例用户的服务实	例			
服务实例	本账号创建的服务实例				
服务审核	Ø 您有待处理的运维项				
服务分销 服务部署物	私力如果服务 今红等服务	经运进服务 过田服务	往下如黑肥条		
运维公告	14.有即有服务 主元首服务	10脸堆服另 临时服务	3次(1、中小銀10K35)		
学习中心	标签筛选 > 服务实例名称 >	输入搜索内容	Q. 试用服务实例		
iQ111	服务实例ID/名称 划	大志 〒 板笠	服务ID/名称	服务版本 服务商	1
应用管理	si-877di si-877di	· 部署中 29%	● service-35cb216b9cf14efe LibreChat社区版	a57d beta 计算巢社区服务	
计算媒 / 服务实例 / 服务实例详细 ← si-985	(si-98	(Charlostonial and			
調盟 咨询 重件	the adv ST Adv with ST A	4 *** C + 66 70 *1 40 C +			
04,45 JU.MA 9717	监控 运班管理 运转	推坝 日志管理 并级历史	费用统计		
事件	新拓 723班属1番 7233	^{田坝} 日志管理 并级历史	费用统计服务资源		
事件 计划内运维事件 非预	<u> </u>	组织 日志管理 开致历史 实例安全事件	费用统计 服务资源 ECS		
事件 近応 単行 计划内运维事件 非預 0 0	<u>編我2</u> 返班留理 2015 期运施事件 实例费用事件 O	度現 日志管理 开致历史 实例安全事件 0	费用统计 服务资源 ECS 1		
	<u> </u>	理境 日志管理 开致历史 实例安全事件 0	费用统计 服务资源 ECS 1		
事件 计划内运维事件 単類 O O 立即使用 ServerAddress ③ http	<u> 編長2</u> 205日第2章 2051 2051 2051 2051 2051 2051 2051 2051	理項 日志管理 开致历史 实例安全事件 0	费用统计 服务资源 ECS 1		
事件	<u> 崩投</u> 取 近 推 留 建 文 列 费 用 事件 の の の の	<u>推填 日志管理 开致历史</u> 实例安全事件 O	费用统计 服务资源 ECS 1		
事件 分類内語維事件 単類 の の の の の の の の の の の の の の の の の の の	<u> 編長2</u> 返班電理 2015 期运地事件	度項 日志管理 开致历史 实例安全事件 O	费用统计 服务资源 ECS 1 服务资源	si-9	
事件	<u> </u>	<u>理項 日志管理 开致历史</u> 实例安全事件 0	费用统计 服务资源 ECS 1 服务实例名称 创建时间	si-9 2024年10月10日 10:09:18	
	<u> 編長2</u> 期运推事件	<u>祖頃</u> 王志管理	费用统计 服务资源 ECS 1 服务实例名称 创建时间 标签	si-9 2024年10月10日 10:09:18 ⊘	

4. 注册账号或登录已有账号,即可使用LibreChat。



输入邮箱密码后,可进入下面页面。

③ 创建新对话	OpenAl × chatgpt-4o-latest × II	9
Q 搜索消息		p
日間		Ø
		Д
		⇒
	\$	
	我今天能帮你做什么?	
	1(
	 ● 废送消息给 GPT-4o ● 	
	LibritChat v0.7.5-rc2 - Every AI for Everyone.	

快速入门LibreChat

多AI模型

首次进入LibreChat,需要设置API Key才可使用对应模型。如下图所示,点击左上角下拉框,按图示操作输入API Key后,即可完成密钥设置。

ď	OpenAl > gpt-40 > E	\odot	
	ගු OpenAl හි	0	
	☆ Assistants		
	G Google		
	Bing		
	D Plugins	V	
	A Anthropic	设置 API 密钥: OpenAl	×
		您的密钥将永远有效	
		Expires in 12 hours ~	
	I see a second	OpenAl API Key	
		输入值: OpenAl API Key	
		撤销	取消提交

LibreChat 支持多种 AI 提供商,包括 OpenAI、Azure、Anthropic、Google 等。可通过左上角切换服务商与模型。





点击图示位置,可轻松上传文件。

[● 发送消息给 GPT-40 LibreChat v0.7.5-rc2 - Every AI for Everyone.	
效果:		
	8 整理后文档自然段.txt Document 这个文档包含什么内容 ())	
	GPT-4 这个文档主要内容是一个直播间的推销台词,具体推销的产品是月饼。主播通过详细的描述、 互动和促销活动,向观众推荐月饼。以下是文档中涉及的一些主要内容和要点:	
	1. 产品介绍:	
	・主要推销的是传统老五仁月饼。	

快捷导出

简单快捷的聊天记录导出和分享给别人,可导出为屏幕截图、markdown、文本、json。

~	chatgpt-4o-latest v	
	u 你好	☆ 导出
	<່ 10	∝。共享
	🌀 GPT-4o	

语音输入

使用语音转文本和文本转语音功能,实现免提聊天。自动发送和播放音频支持 OpenAI、Azure OpenAI 和 Elevenlabs。

❶ 发送消息给 GPT-4o		Ŷ	Ŷ
	LibreChat v0.7.5-rc2 - Every Al for Everyone.		

其他问题

1. 如何进入ECS

点击服务实例详情,进入ECS:

☰ (-) 阿里云 🗠 🗉	作台	Q 搜索	费用	ICP 备案 企业 支持 工单 🖸	: 🏆 🕜 🛟
计算巢	计算巢 / 服务实例 / 服务实例详情				
欢迎页	← libre-m8ca (si-dc0a9547080d4f9ba3 _{远程连接 ☉}	×			
服务目录 我的服务	概览 资源 事件 监控 远维管理 通过会话管理远程过	E接			
服务实例	透波会活管理可以免密 通过会活管理可以免密 了解详情。	码,兔公网,无需开放入方向端口,对ECS实例进行远程控制。			
服务审核服务分销	全部操作 ∨ 请输入资源 ID 的关键词提索 Q ECS实例 SL 免回登录				
服务部署物	会话管理已开启(全地	藏)	5式 到期时间 1	续费方式	操作
运维公告 学习中心	□ + i-bp1fi6okv0d4dme5nu7z 云服务器 iZbp1fi6okv0d4dme5nu7zZ 云服务器	取消	₫	-	远程连接
设置	□ + d-bp1eobnbp9p547lxrrgi 云服务器 工量	■ 1000 PP	-		
	□ 王 Sg-bp1eatn3tkrmx8eds94t 云服务器 安全组 libre-m8ca		-		
应用管理 私网互联				#	转有3条, 每页显示

一登录实例			() × ()
服务实例: 📄 iZbp1fi6okv0d4dme5ni	u7zZ si-dc0a9547080d4f9ba35e		
实例: 🆂 iZbp1fi6okv0d4dme5nu	ı7zZ i-bp1fi6okv0d4dme5nu7z 华东	1(杭州)	
* 网络连接: 💿 会话管理-云助手	〇 8.149.246.240(公网)	〇 192.168.1.132(私网)	
		取消	确定

2. 如何添加转发地址

- 1. 进入ECS,参考问题一。
- 2. 输入命令

•		JSON
1 2 3	<pre>sudo su cd /home/admin/application/docker_compose sudo vim .env</pre>	

将会进入如下页面



3. 寻找需要修改的AI模型,本处以OpenAI为例,按下i键,左下角将会出现'--INSERT--'标志,通过 方向键移动光标,在第一个红框处(OPEN_REVERSE_PROXY)输入转发地址。

#=======# # OpenAI # #========#
<pre>OPENAI_API_KEY=user_provided # OPENAI_MODELS=gpt-4o,chatgpt-4o-latest,gpt-4o-mini,gpt-3.5-turbo-0125,gpt-3.5-turbo- bo-1106,gpt-3.5-turbo-instruct,gpt-3.5-turbo-instruct-0914,gpt-3.5-turbo-16k</pre>
DEBUG_OPENAI=false
<pre># TITLE_CONV0=false # OPENAI_TITLE_MODEL=gpt-3.5-turbo</pre>
OPENAI_SUMMARIZE=true # OPENAI_SUMMARY_MODEL=gpt-3.5-turbo
OPENAI_FORCE_PROMPT=true
OPENAI_REVERSE_PROXY=
OPENAI_ORGANIZATION=
#========# # Assistants API # #========#
INSERT

4. 输入完成后按 ESC 键, 左下角'---INSERT---'消失, 输入:wq(冒号别漏了), 保存退出。

3. 如何添加第三方API模型

1. 进入ECS,参考问题一。

2. 修改librechat.yaml文件:新增第三方模型,输入命令(librechat.yaml文件内容可参考链接)。

•		JSON
1	sudo su	
2	<pre>cd /home/admin/application/docker_compose</pre>	
-		

3 sudo vim librechat.yaml

进入librechat.yaml文件(即上面代码中的sudo vim librechat.yaml),此文件中默认已有groq、 Mistral、OpenRouster三个模型,如需添加,可按照其格式新增,如下新增了通义千问(Qwen)模型。 (URL参考链接,ApiKey参考链接)。修改方式为:

1. 按下i键, 左下角将会出现'---INSERT---'标志, 通过方向键移动光标, 输入红框处内容;

2. 输入完成后按 ESC 键, 左下角'---INSERT---'消失, 输入:wq(冒号别漏了), 保存退出。



3. **修改docker-compose文件**:配置读取第三方模型,输入命令:



此时将会新增红框处文件:



进入文件,输入: sudo vim docker-compose.override.yml,输入i进入修改模式,按照如下修改, 修改完成后按 ESC 键,然后输入:wq 退出。

修改前:



修改后:



4. 重启Docker: 输入

JSON

1 sudo docker compose up

•

5. 效果展示:进入LibreChat页面,第三方API模型已显示,选择Qwen即可对话。



了解更多

本文通过计算巢服务完成了LibreChat的快速部署,什么是计算巢服务?

计算巢是一个开放给企业应用服务商、IT集成服务商、交付服务商、管理服务提供商(以上都统称为服务商)和最终用户的服务管理PaaS平台。服务商能通过计算巢服务更好地在阿里云上部署服务、交付服务及管理服务。最终用户能通过计算巢管理在阿里云上订阅的各类服务商提供的服务。计算巢为服务商和用户提供了更高效、便捷、安全的服务使用体验。点击计算巢服务官方文档了解更多。

参考资料

计算巢服务官方文档

更多信息,请访问计算巢官网。