

# 大模型知识引擎 常见问题



腾讯云

## 【 版权声明 】

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

## 【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

## 【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

## 【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

# 文档目录

## 常见问题

产品常见问题

技术常见问题

# 常见问题

## 产品常见问题

最近更新时间：2024-08-23 14:58:21

### 大模型知识引擎是什么？

大模型知识引擎（LLM Knowledge Engine），是面向企业客户及合作伙伴的，基于大语言模型的知识应用构建平台，结合企业专属数据，提供知识问答等应用范式及原子能力接口服务，更快更高效地完成大模型应用的构建，推动大语言模型在云通信与企业服务场景的应用落地。各个行业的客户可以在行业大模型基础上，快速构建基于自然语言对话的自助问询服务，利用运营工具低成本提升服务质量。主要面向企业的运营维护人员，通过配置管理平台，编辑和维护客户级知识库、业务接口的有效性，通过机器学习的方式，提高运营人员训练机器人和维护知识库的效率。

### 大模型知识引擎知识库上限是多少？

大模型知识引擎开通试用即发放有效期为6个月的300万字符免费资源包，您可通过 [购买扩容包](#) 进行扩容。

### 大模型知识引擎如何收费？

知识引擎支持预付费资源包、知识库字符包、共享并发等计费模式。更多信息，请参见 [计费概述](#)。

### 大模型知识引擎支持导入的格式有哪些？

目前支持 pdf、docx、doc、pptx、ppt、xlsx、xls、png、jpg、jpeg、markdown、txt、csv 格式的文档导入，并支持网页导入；支持导入含有图片的文档。

### 答案可以支持哪些格式？

文字、图片、表格、超链接。

### 如何获取大模型知识引擎生成式人工智能服务备案信息？

#### 算法备案

在网信办的 [互联网信息服务算法备案系统](#) 中选择搜索备案编号即可查看：

1. 基于“腾讯混元助手大模型”提供文本生成服务：网信算备440305295988701230071号；
2. 基于“腾讯云行业大模型算法”提供文本生成服务：网信算备440305295988701230089号。

# 互联网信息服务算法备案系统

登录 / 注册

# 互联网信息服务算法备案系统

搜索

共有 0 条备案记录

主体名称	算法类型	算法名称	公示内容	备案编号(唯一)	发放日期	当前备案状态
						

## 大模型备案

1. 基于“行业大模型”提供文本生成服务，大模型备案号：[Beijing-TengXunYunHangYeDaMoXing-202404280005](#)；
2. 基于“腾讯混元助手大模型”提供文本生成服务，大模型备案号：[Guangdong-TencentHunyuans-20230901](#)。

## 如何对接用户标签实现一问多答场景？

可通过知识标签进行配置，知识标签功能用于实现一问多答场景，即针对不同身份属性的用户提问，可在同一个应用的知识库中检索跟该属性相关的文档内容并通过大模型生成对应的答案。具体操作步骤请参见 [知识标签](#)。

## 如何进一步了解产品细节？

如果您想进一步了解该产品更多的细节，您可以通过 [售前咨询](#) 进行提问，将有专属客户经理在线为您解答，为您提供专业的上云方案。

同时，您也可以直接致电销售人员，咨询热线：4009100100转1（工作日9:00 - 18:00）。

# 技术常见问题

最近更新时间：2024-08-09 09:17:41

## 1. 使用密钥调用 API 报错：用户未登录

常见原因是子账号操作 API，主账号未给子账号授权。

授权方式请参考：

- [签名方法](#)
- [子账号访问密钥管理](#)

## 2. 腾讯云上只能看到 SecretId，看不到 SecretKey

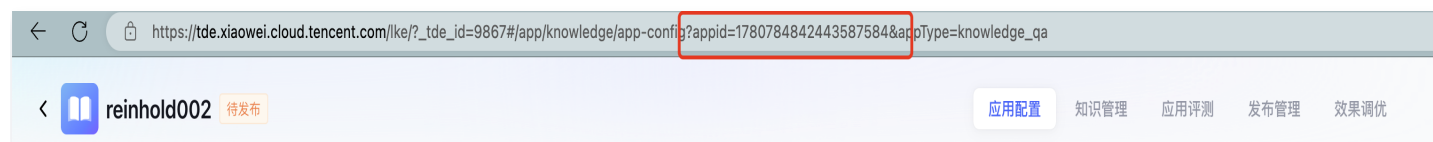
- 腾讯云的规则是只有新建的时候才能看到。
- 可以新建一个，然后自行记录保存。
- 原有的一个密钥禁用（如果不使用的话），然后删除；删除之后就可以重新创建。

## 3. BotBizID，BotAppKey 是什么，从哪里获取？

BotBizID 是创建某个应用时的唯一标识。

BotAppKey 是分享或者通过 websocket，sse HTTP 调用的时候用来获取访问通行证的一个参数。

BotBizID：可以直接在当前页面的 URL 中获取【appid】。



BotAppKey: 应用列表，点击【调用】，就可以看到 appkey 了。

### ⚠ 注意：

应用发布之后，【调用】按钮才可以使用，否则按钮是置灰的状态



The image shows a screenshot of the Tencent Cloud console's application management interface. It features a table with the following columns: 应用 (Application), 发布环境/模型 (Release Environment/Model), 发布环境/服务状态 (Release Environment/Service Status), 修改时间 (Modification Time), 最后修改人 (Last Modified By), and 操作 (Actions). The '操作' column contains buttons for '查看' (View), '调用' (Call), and '删除' (Delete). The '调用' button for the first application, 'reinhold002', is highlighted with a red rectangular box.

应用	发布环境/模型	发布环境/服务状态	修改时间	最后修改人	操作
reinhold002	精调知识大模型高级版	运行中	2024-07-30 17:01:33	600000560583	查看 调用 删除
harry-0729	精调知识大模型高级版	运行中	2024-07-29 20:42:39	600000560583	查看 调用 删除
harry0723	精调知识大模型高级版	运行中	2024-07-29 14:39:20	600000560583	查看 调用 删除

## 调用信息



### 体验链接

[https://tde.xiaowei.cloud.tencent.com/webim\\_exp/?\\_tde\\_id=9867#/chat/FPKXGS](https://tde.xiaowei.cloud.tencent.com/webim_exp/?_tde_id=9867#/chat/FPKXGS)

立即体验

重新生成

分享链接

分享二维码

### API管理

#### 第一步：开通应用服务

应用根据你选择的模型收费，[点击查看费用详情](#)，按需购买。如您还有免费使用额度可以直接对接

#### 第二步：获取鉴权

使用appkey（见下文）及API密钥获取访问权限，密钥获取方式，[点击查看腾讯云控制台-访问管理-访问密钥](#)

#### 第三步：接口调用

[查看接口文档及示例代码](#)，接入到业务场景中使用



appkey

创建时间

k\*\*\*\*\*D

2024-04-18 10:26

复制

我知道了

## 4. 接口调用，报错：应用【机器人】不存在

- 调用一般的接口，传入的 BotBizId 不对，找不到对应的应用。
- 调用 GetWsToken，当 Type=5【API 访客】的时候，BotAppKey 必填。

## 5. sessionId 和 requestId 如何使用

- sessionId 用来标识一个会话，不同会话（不管是否是同一个应用）需要使用不同的 sessionId，同一个会话使用同一个 ID。sessionId 的生成规则是 uuid。

- requestID 用来标识一个请求，不同请求使用不同的 requestID。

↓ 40{"sid":"cpkk83tjgadm99r8eq0"}	32	14:45:35.554
↑ 42["send",{"payload":{"request_id":"FUfl5VtJhE-5353543786","session_id":"8dc87422-0308-4a38-8...}}	147	14:45:43.770
↓ 4295536["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":false,"content":"你是谁",...}}	672	14:45:43.908
↓ 4295537["token_stat",{"type":"token_stat","payload":{"elapsed":475,"free_count":5000000,"order_cou...}}	530	14:45:44.424
↓ 4295543["token_stat",{"type":"token_stat","payload":{"elapsed":1439,"free_count":5000000,"order_co...}}	531	14:45:45.385
↓ 4295546["token_stat",{"type":"token_stat","payload":{"elapsed":1774,"free_count":5000000,"order_co...}}	617	14:45:45.720
↓ 4295549["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":true,"content":"我是腾讯...}}	705	14:45:45.993
↓ 4295552["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":true,"content":"我是腾讯...}}	716	14:45:46.234
↓ 4295554["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":true,"content":"我是腾讯...}}	726	14:45:46.300
↓ 4295555["token_stat",{"type":"token_stat","payload":{"elapsed":2489,"free_count":5000000,"order_c...}}	612	14:45:46.435
↓ 4295557["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":true,"content":"我是腾讯...}}	728	14:45:46.529
↑ 42["send",{"payload":{"request_id":"rsceSDXvSq-1041443046","session_id":"8dc87422-0308-4a38-8...}}	148	14:45:49.530
↓ 4295571["reply",{"type":"reply","payload":{"can_feedback":false,"can_rating":false,"content":"你在干嘛...}}	673	14:45:49.644
↓ 4295572["token_stat",{"type":"token_stat","payload":{"elapsed":424,"free_count":5000000,"order_cou...}}	530	14:45:50.073
↓ 4295573["token_stat",{"type":"token_stat","payload":{"elapsed":1386,"free_count":5000000,"order_co...}}	531	14:45:51.043
↓ 4295574["token_stat",{"type":"token_stat","payload":{"elapsed":1697,"free_count":5000000,"order_co...}}	617	14:45:51.345
▼ 42["send", {, ...}]		

## 6. 通过 websocket、HTTP sse 调用的时候如何重置会话

需要用户重新生成一个 sessionID 即可。

## 7. 实时文档解析后报错：msg:http client codec StatusCode: Not Found

这种情况是 cos 里面没有找到相关的文档导致的，请检查下上传文档是否传错了 cos 桶，着重检查下获取临时密钥的时候的 TypeKey，可参考以下文档：

- [实时文档解析](#)
- [离线文档上传](#)

## 8. 离线文档上传之后，怎么知道是否解析完成，是否可以发布

离线文档上传后，会有解析、审核、发布等状态，当前序步骤处理完成之后，文档的状态会处于待发布状态，这个时候可以调用发布接口执行文档发布。

可以通过轮询 ListDoc 接口，查看文档是否可以发布。

离线文档的状态可参考：[文档列表](#) 中的 Status.N 文档状态。



## 9. 获取用户历史消息，GetMsgRecord

接口逻辑：根据当前传入的会话的 session\_id 获取聊天，当消息过多的时候通过 Count 和 LastRecord 来做翻页。

例如：

session\_id: xxxxxxxx下有50条数据，count 设置为20，分页取消息时的传参：Count = 20 、 LastRecordId = 当前获取到的消息列表中的第一条 record\_id。

### 说明：

- 1、如果传入的 session\_id 不正确，会找不到相应的会话。
- 2、该接口只能获取某个 sessionID 下的聊天数据，不能获取当前应用下的所有聊天数据。
- 3、用户通过接口获取 API 访客数据 [ Type=5 ] 的时候，需要通过 ReplyMethod 来区分 IsFromSelf : 0表示用户请求，非0表示知识引擎的回复。

## 10. 如何在 websocket, HTTP sse 调用中带上知识标签

websocket 方式，在获取 token 的时候带上 VisitorLabels, 具体可参考：[获取ws token](#)

HTTP sse 方式，需要在入参的时候带上 VisitorLabels, 具体可参考：[对话端接口文档 \( HTTP SSE \)](#)

### 说明：

如果一个文档关联了多个属性，多个知识标签，API 调用传入了多个属性，多个标签，在做数据搜索的时候，标签的叠加逻辑如下：同属性下多个标签搜索为或的关系，不同属性之间搜索为且的关系。

## 11. 用户会话端不展示参考来源

知识引擎测试预览页面是默认展示参考来源的，即使文档未打开外部引用链接。

在用户端或者 api 接入的方式未展示参考来源，请从下面几个方向排查：

- 1、参考来源未打开，可以在导入文档的时候打开；文档导入后也可以选择编辑文档，修改外部引用链接的方式来定义应用链接。

## 文档设置 ×

适用范围 ?

全部  按条件设置

到期时间

永久有效  自定义到期时间

外部引用链接 ?

2、一段对话并不总是有参考来源，参考来源需要看问题的内容和知识库里面的文档，问答的相似程度来进行判断；可以对同样一个问题，分别在知识引擎的测试预览页面和用户页面进行问答，在打开了文档的外部引用链接之后，两个地方的回答应该是一样的。

## 12. 应用文档接口中给出的 Python 示例运行错误，无法进行问答

可能是客户未安装正确的软件版本，按照如下方法更新即可：

```
pip3 install -r requirements.txt
```

### ⚠ 注意：

执行 `pip3 install requests` 命令来安装依赖，不能解决上面的问题。

## 13. 通过 websocket、HTTP sse 进行图片问答时参数如何填充

1. 需要将图片上传到实时文档 cos 中, 请参考: [实时文档解析](#)
2. websocket 和 HTTP sse 中的 content : 如果发送图片, 在此传递 markdown 格式的图片链接, 例如, 其中图片链接需要可公有读。

#### content 示例:

```
{
  "request_id": "LHWObW2Sea-3173301679",
  "session_id": "f5652b8c-e88e-4abc-8629-7952b741a433",
  "file_infos": [],
  "content": "描述下下面这张图片"
}
```