

知识引擎原子能力 购买指南



腾讯云

【 版权声明 】

©2013–2025 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

购买指南

计费概述

购买方式

欠费说明

续费说明

退费说明

购买指南

计费概述

最近更新时间：2025-02-27 09:05:52

知识引擎原子能力提供灵活可配的 API 接口，根据用户所使用的能力按量进行计费，根据接口功能不同支持预付费和后付费。开通大模型知识引擎产品后即可获得一定量的免费额度。如需购买，请前往 [购买页](#) 购买。

开通方式

知识引擎原子能力为大模型知识引擎子产品，可在 [控制台](#) 开通权限后进行使用。该产品的开通使用需要先通过 [腾讯云企业实名认证](#) 或者 [腾讯云个人实名认证](#)。通过实名认证后，首次在 [大模型知识引擎产品介绍页](#) 单击 [产品体验](#) 或在控制台单击 [开通](#)，即可开通知识引擎原子能力使用权限。

免费额度

通过实名认证的主账号首次登录开通大模型知识引擎产品时，获得一定量的免费体验额度，详情如下表所示。

⚠ 注意：

免费额度仅针对2025年1月4日及以后开通产品的账号生效。

资源类别	免费额度
原子能力-多轮改写	开通大模型知识引擎服务即获赠累计100万 tokens 原子能力 token 体验资源，有效期2个月，以资源包的形式发放到您的腾讯云账号中，优先扣除。
原子能力-Rerank	
原子能力-文档拆分	
原子能力-Embedding	
原子能力-文档解析	开通大模型知识引擎服务即获赠累计500页文档解析体验资源，有效期2个月，以资源包的形式发放到您的腾讯云账号中，优先扣除。
原子能力-RAG 套件-知识库检索	开通大模型知识引擎服务即获赠累计500次 知识库检索调用额度，有效期2个月；以资源包的形式发放到您的腾讯云账号中，优先扣除。
原子能力-RAG 套件-知识库容量	开通大模型知识引擎服务即获赠累计300万字符数的免费知识库容量，有效期6个月。 到期后，知识库中的知识将被处理成“超量失效”状态，“超量失效”状态的知识在检索中不生效，且“超量失效”状态的知识30天后将被自动删除，如需要继续使用知识库，请在到期前及时购买。

⚠ 注意:

- 在账单结算时，系统将按照免费资源包 > 预付费资源包的顺序进行结算，即免费资源包是优先扣除的。
- 若您欠费或因违禁原因停服后，将不能继续享受免费额度，只有服务重新开启后才可继续享受免费额度。
- 调用量的扣费顺序为：免费额度 > 资源包 > 后付费，即先消耗账号中的免费额度，免费额度耗尽后再消耗资源包，资源包用量耗尽后支持后付费。

计费模式

接口名称	基础计费（必选）	增值计费（可选）
对话（支持 DeepSeek 模型）	后付费	无
Embedding	资源包/后付费	无
多轮改写	资源包/后付费	无
Rerank	资源包/后付费	无
文档解析	资源包/后付费	文档解析 QPS 叠加包
文档拆分	解析：资源包/后付费 拆分：后付费	文档解析 QPS 叠加包
RAG 原子能力套件	知识库字符数：包年包月 检索：资源包/后付费	无

预付费资源包

Embedding

Embedding 原子能力可通过 API 进行调用，单击 [Embedding](#) 查看 API 文档。预付费资源包自购买日起一年内有效，一年内若资源包未使用完，则过期作废。

资源包大小	价格
5000万 token	¥100
1亿 token	¥200
5亿 token	¥900
10亿 token	¥1,800

多轮改写

多轮改写原子能力可通过 API 进行调用，单击 [多轮改写](#) 查看 API 文档。预付费资源包自购买日起一年内有效，一年内若资源包未使用完，则过期作废。

资源包大小	价格
1000万 token	¥60
5000万 token	¥300
1亿 token	¥600
5亿 token	¥2,800
10亿 token	¥5,600

文档解析

文档解析预付费资源包

文档解析原子能力可通过 API 进行调用，单击 [文档解析](#) 查看 API 文档。预付费资源包自购买日起一年内有效，一年内若资源包未使用完，则过期作废。

资源包规格	价格
500页	¥100
1000页	¥180
1万页	¥1,700
5万页	¥5,600
10万页	¥9,800
50万页	¥40,000
100万页	¥75,000
500万页	¥290,000

文档解析 QPS 叠加包

购买预付费资源包或开通后付费后，默认提供 5 QPS，可按接口类型增购。

接口类型	价格
------	----

同步接口	¥10,000/QPS/月
异步接口	
实时接口	

知识库容量

定义：知识库容量计算用户在 rag 原子能力套件中上传的文档和问答的总字符数，删除的文档和知识库不占用知识库容量。

资源有效期：可选，即将过期如需继续使用请续费。已过期的知识库仅支持删除操作。

付费方式：包年包月。

知识库扩容包	包月	包年
1000万字符	¥160	¥1,800
1亿字符	¥1,400	¥16,000

⚠ 注意：

- 字符数不计算文档中的空格，文档中存在图片的将转存为图片链接，每张图大约换算为150个字符。
- 一篇3万字的硕士论文，大约相当于4万字符，50页文档。1000万字符约为250篇硕士论文，1亿字符约为2500篇硕士论文。

知识库扩容包到期说明：

- 到期后，知识库中的知识超量失效，将不再支持检索，需要删除知识库内容或购买扩容包使用。如不处理，知识库中超量的知识180天后将被自动删除。如需要继续使用知识库，请在到期前及时续费。

后付费

知识引擎原子能力支持开通后付费使用。

按量后付费根据实际调用接口消耗的资源量计费，每小时对腾讯云账户进行结算和扣费，您可在 [控制台](#) 中查看您的用量情况。每月1 - 3日系统会推送上个月账单。

接口名称	单价		
DeepSeek API 接口	DeepSeek-R1 模型	输入：0.004元/千token	输出（含思维链）：0.016元/千token
	DeepSeek-V3 模型	输入：0.002元/千token	输出：0.008元/千token
Embedding	0.002元/千tokens		

多轮改写	0.006元/千tokens
rerank	0.0001元/千tokens
文档拆分	拆分0.03元/千tokens
	解析0.20元/页 (pdf/docx/doc/ppt/pptx/jpg/png等格式)
文档解析	0.20元/页
RAG 套件-知识库检索	0.04元/次

说明：

- 文档解析按照文档调用成功的页数计费；不同文档按页计费规则：
 - doc、docx、ppt、pptx、pdf 按页计量。
 - jpeg、png等图片格式以一张图为一页。
 - xlsx、txt、md、csv 以1份为一页。
- 文档拆分接口调用时，根据文件类型的不同，将分别计算解析和拆分的费用，详情如下：
 - 输入文件为pdf/docx/doc/ppt/pptx文档格式和 jpg/png 等图片格式时，需要计算文档解析费用（按页计费），支持预付费和后付费。
 - 输入文件为txt/md/xlsx/xls格式时，仅需工程解析，不需要付费。
 - 两种方式输入的文件都需要计算拆分消耗的费用，按 token 消耗数量计费。

购买方式

最近更新时间：2025-02-26 14:35:22

预付费资源包

1. 在购买知识引擎原子能力之前，您需要先 [注册腾讯云账号](#) 并通过实名认证，详情请参见 [注册腾讯云](#)。
2. 进入知识引擎原子能力 [购买页](#)，选择您需要的服务类型、套餐包容量和数量，单击**立即购买**，完成支付。支付后即可使用对应 API 接口。

知识引擎原子能力 [返回产品详情](#)

购买须知

- 购买说明** 请根据您的需要购买对应的原子能力套餐包。详情参考[计费概述](#)
- 使用说明** 购买完成后资源为生效状态，通过[调用API](#)使用该产品
- 退费说明** 购买前请详细了解[退费说明](#) 腾讯云有权根据退费说明拒绝相应退款要求

选择配置

采购类型 多轮改写 Embedding 文档解析 Rerank RAG套件

可选套餐包

<p>1000万tokens 有效期自购买之日一年内</p>		<input type="text" value="0"/>
<p>5000万tokens 有效期自购买之日一年内</p>		<input type="text" value="0"/>
<p>1亿tokens 有效期自购买之日一年内</p>		<input type="text" value="0"/>
<p>5亿tokens 有效期自购买之日一年内</p>		<input type="text" value="0"/>
<p>10亿tokens 有效期自购买之日一年内</p>		<input type="text" value="0"/>

后付费

知识引擎原子能力支持开通后付费使用。购买预付费资源包后，如您有资源耗尽不停服需求，可开启后付费开关，即可在预付费资源包资源耗尽时，自动切换后付费使用，可在 [后付费设置](#) 打开计费开关后使用。

☰ 计费管理

资源包
并发扩展
知识库容量扩展
后付费设置

未开启

1. 开启后，资源包耗尽时，以“后付费方式”按小时从账户余额中扣费（修改上限：每小时1次/每月5次）

2. 扣费顺序：体验资源包>付费资源包>后付费，可在[控制台-费用中心](#)查看账单

3. 同1个模型的“付费资源包”未购买/全部已耗尽/全部已到期后“开启后付费”，该模型无法使用后付费。需购买token资源包生效，其他模型的后付费可正常使用。

子产品名称	状态
大模型知识引擎_多轮改写_后付费	未启用
大模型知识引擎_Embedding_后付费	未启用
文档解析_后付费	未启用
知识引擎_Rerank_token后付费	未启用
腾讯云大模型知识引擎_知识引擎_RAG_知识库检索_后付费	未启用
腾讯云大模型知识引擎_知识引擎_文档拆分_token后付费	未启用

按量后付费根据实际调用接口消耗的资源量计费，每小时对腾讯云账户进行扣费和结算并推送相关账单。

⚠ 注意：

1. 开启后，资源包耗尽时，以“后付费方式”按小时从账户余额中扣费（修改上限：每小时1次/每月5次）
2. 扣费顺序：体验资源包 > 付费资源包 > 后付费，可在 [控制台-费用中心](#) 查看账单。

欠费说明

最近更新时间：2025-01-03 22:37:12

每月1 - 3日统计上个月接口调用总量进行结算并生成对应账单，系统自动扣费，如果结算扣费后，您账号中余额为负值，知识引擎原子能力会在24小时内停止服务，金额充正后，服务将恢复正常。停服后，即使您之前购买的预付费套餐包仍有余量，资源包也将暂停抵扣，当金额充正后，系统会自动恢复资源包抵扣功能。

当您的账号发生欠费时，系统将为您推送欠费提醒，请在收到欠费通知后，及时前往控制台 [充值中心](#) 进行充值，以免影响您的业务。

📌 说明：

- 若您对欠费有疑问，可在控制台 [费用账单](#) 页面查阅和核对您的消费明细。
- 若您对具体的扣费项有疑问，可以参见 [计费概述](#) 了解具体的计费项含义及计费规则。

续费说明

最近更新时间：2025-01-03 22:37:12

若您购买的产品资源耗尽或到达产品有效期，请 [点击购买](#) 或联系架构师，进行续费。选择采购的原子能力服务类型并确认资源包容量后，进行下单，完成支付后续费成功。

退费说明

最近更新时间：2025-01-03 22:37:12

知识引擎原子能力提供两类计费项：**按量计费**和**包年包月类计费项**。相关退费说明见下。

退费规则

- 正式采购7天内（含7天）支持退费：基于**资源包剩余比例**计算退款金额。
退款金额 = 新购折后价格 * （剩余量/总量）
- 超过7天不支持退费。

按量计费

计费项：**原子能力预付费资源包**。

退费规则：

- 正式采购7天内（含7天）支持退费：
 - 基于**资源包剩余比例**计算退款金额。
 - 退款金额 = 新购折后价格 * （剩余量/总量）
- 超过7天不支持退费。

包年包月

计费项：**包括文档解析并发包及知识库扩容包**。

退费规则：

- 正式采购7天内（含7天）支持退费：
 - 基于**使用时长折算**（以天为单位计算，不满24小时记为一天）退费金额。
 - 退款金额 = 支付金额 - （已使用时长 ÷ 总时长） × 订单原价 × 适用折扣
- 超过7天不支持退费。

退费流程

发起退费前，请先确定是否满足退费说明的条件，并确保基于大模型知识引擎服务产生的一切数据、资料已完成备份、迁移。若您未进行备份、迁移，在退费后产生的任何损失由您自行承担，与腾讯云无关。

1. 登录腾讯云官网，单击 [售后支持](#) 页面，反馈需要退款的产品并描述退款原因，腾讯云客服人员将会在线处理您提交的申请。
2. 您可在 [订单管理](#) 页面查看退货订单，订单状态为“已退费”，可在 [费用中心](#) 页面查看款项。若审核不通过，您可在工单里查看审核结果。

ⓘ 说明：

发起退款流程前，请您查看资源是否属于可退费范畴，若不符合，则无法退还。

