

河南台 AIGC 工具项目合同

合同编号：

需方（采购人全称）：河南广播电视台

供方（中标人全称）：腾讯云计算(北京)有限责任公司

供方持河南广电招标有限公司签发的中标通知书，根据招标文件、投标文件等约定（项目名称：河南广播电视台大象AI技术支撑平台项目，采购编号：豫财招标采购-2025-940，按照《政府采购法》、《政府采购法实施条例》、《民法典》等有关法律、法规，经供需双方协商一致，达成以下合同条款：

一、本合同名称：

二、本合同总价为人民币 1806652.80 元（大写 壹佰捌拾万陆仟陆佰伍拾贰元捌角），包括货物(服务、软件)报价和标准附件、备品备件、专用工具、运输、装卸、保险、检测、验收合格、售后服务、培训、维护所需的各种费用及必要的保险费用和各项税金等一切费用。合同不含税金额合计 1704389.43 元，增值税金额合计为 102263.37 元。（以上价款以人民币进行结算）

三、合同清单及履行地点要求：

履行地点：采购人指定地点。

货物及伴随服务的内容：

序号	品牌型号	规格	数量	单价	总价
1	算力调度平台	套	1	400000	400000
2	AIGC 工具-AIGC 创作	套	1	500000	500000
3	AIGC 工具-智能门户	项	1	120000	120000
4	AIGC 工具-智能体平台	套	1	95132	95132
5	AIGC 工具-服务内容	项	1	31000	31000
6	AIGC 工具-运营管理	项	1	100000	100000
7	AIGC 工具-知识库引擎	套	1	8000	8000
8	AIGC 工具-AIGC 引擎	套	1	19000.80	19000.80
9	AIGC 工具-大模型能力	套	1	147720	147720
10	AIGC 工具-基础 AI 能力	套	1	385800	385800

其中，公有云产品超出部分按照协议单价据实结算。

四、质量要求及供方对质量负责条件和期限（采购需求中的内容）

1、供方提供的系统必须是自主研发的符合采购要求的规格型号和技术指标。在产品质保保修期（3年）内非因需方的人为原因而出现质量问题的，由供方负责包修、包换或者包退，并承担调换或退货的实际费用。

2、供方对供货范围内的产品质量负责。供方需免费提供技术咨询服务，并对整个技术方案和系统质量负责，如出现因技术方案或系统不合格，影响验收，供方担负赔偿金。

3、交付时间

合同签订后的60日内系统交付至招标人指定现场、完成系统的安装、调试工作。

五、货款支付

1、供方须提交以下文件：合同；需方出具有项目负责人签字的验收报告；增值税专用发票。

2、付款方式：合同签订后10日内后，供方提供增值税专用发票、需方60日内通过银行转账支付合同总金额的30%；核心产品的软件全部到货并检验合格后，需方向供方出具《设备到货签收单》，10日内供方提供增值税专用发票，需方60日内通过银行转账支付合同总金额的40%；合同内软件安装调试完成并验收合格后，10日内供方提供增值税专用发票、需方60日内通过银行转账支付合同总金额的20%；需方出具项目负责人签字的验收报告后，供方提供合同总金额5%的银行履约保函[为期1年，到期后自动结束保函（不含利息）]，供方提供增值税专用发票、需方60日内通过银行转账支付合同总金额的10%。

3、合同签订后，因财政支付原因导致的付款延迟，不视为需方责任。

4、供方银行账户信息如下：

户名：腾讯云计算(北京)有限责任公司

账号：招商银行北京上地支行

开户行：110907316810601

汇款备注：2025-河南台-AIGC工具项目

六、违约责任

1、供方逾期履行合同的，自逾期之日起，向需方每日日历日偿付合同总价款千分之五的违约金；供方逾期20日历日不能交货的，需方有权终止合同。到期未能按时完工、无法验收或供货的，造成需方无法支付货款，由此造成的经济损失，由供方承担。由于到期未能完工造成的一切不良影响和法律责任，由供方承担。

系统到货后，所购设备指标、功能不满足合同规定的有关要求，需方有权拒绝签收，并

按供应商违约处理，采购人有权终止合同，并追索因供方造成工期延误而给需方产生的经济损失和货款，发生合同解除时，双方应于合同解除时核算供方已完成工作量，需方按供方实际完成工作量支付相应报酬”

如果测试验收失败次数超过两次或测试验收时间超过合同规定的有关期限，需方有权拒绝验收，并按供应商违约处理，采购人有权终止合同。

2、供方在项目验收合格之日起保修期内违反投标文件中有关承诺保证的，造成需方损失的，需方有权主张供方赔偿。

3、供方必须按照招标文件要求、投标文件响应标准和合同要求供货，擅自更改超越合同范围对需方造成损失的，由供方负责。

4、如发现供方违反合同的有关规定，需方有权追究供方违约责任，并有权终止合同。

5、因不可抗力导致本合同全部或部分不能履行时，双方各自承担其因此而造成的损失、损害。

七、调试和验收

供方免费负责送货及调试安装。验收主体为需方，需方可自主验收或委托的第三方进行验收，验收报告由需方出具，出具项目负责人签字的验收报告。

八、争议的解决

因服务质量和售后服务等发生的争议，应首先通过协商解决，如协商不成，依法向需方所在地人民法院提请诉讼。

九、其他

1、本合同经需方、供方法定代表人或其委托人签字并盖章后生效。

2、本合同一式陆份，需方、供方各执叁份。

3、合同未尽事宜，双方应依照谈判文件、谈判投标文件签订补充合同。

4、需方指定 张谦 为该项目负责人，供方指定 刘毅 为该项目负责人，共同推进合同项目的实施。

5、项目的招标文件和投标文件均是本合同的组成部分。

6、公有云服务的价格，以附件合同所载产品参考价格的折扣为准；若后期提供服务一方对价格进行调整，则以调整后价格与原折扣价格中较低者为准，公有云产品超出部分按照协议单价据实结算。

需方：河南广播电视台

地址：郑州市金水区纬一路2号

法定代表人或被授权人（签字）：

电话：0371-65605995

开户银行：中信银行郑州花园路支行

账号：8111101010836936966

供方：腾讯云计算（北京）有限责任公司

地址：北京市海淀区西北旺东路10号院西区9

号楼4层101

法定代表人或被授权人（签字）：

电话：010-82173652

开户银行：招商银行北京上地支行

账号：110907316810601

签约时间：2025年9月12日

附件一：

系统详细功能表

1. 算力调度平台

1.1 支持 N 卡、天数智芯、沐曦、华为昇腾等适配功能。

1.2 平台架构设计及核心功能。基于 AFS 平台 LXC 容器显示。支持容器化隔离任务环境，并分别对不同人员创建的不同任务进行隔离和管理。支持用户使用裸金属方式作为边缘侧节点，使资源的配置符合任务的计算需求，以保证训练任务能够顺利完成。

1.3 管理运维模块。

系统支持对集群资源进行实时监控，支持单节点资源实时监控。

支持 CPU、内存、磁盘、GPU 卡、GPU 显存的资源使用率、使用量监控。

支持多维度资源统计，支持按时间段统计的功能

1.4 模型训练。支持用户根据训练完成的模型提交审核，审核通过后进入模型超市，其他用户可以在实例创建页面中选择需要的模型，快速部署。

1.5 模型推理。提供图形化的一键部署功能，用户可选择训练完成的模型或模型超市中的模型，通过简单的配置，快速将其部署为在线推理服务。

1.6 支持查看所有创建的实例，能够查看实例的状态和运行时间等基本信息。

1.7 运维管理节点无数量限制。

1.8 支持多租户和计费功能。支持多租户管理和计费功能。

1.9 应用层调度管理、优化以及异常处理。独家自研算力调度优化算法，支持任务层面的调度优化，基于内存模型、应用调度历史、应用队列等数据，实时进行任务分配，可以显著降低应用频繁加载和卸载模型时间 50% 以上，大幅提高算力的利用效率。支持常见的私有化部署应用平台，如 ComfyUI 的显存和内存的自动释放功能。

2. AIGC 创作

2.1 文案创作。支持选择不同的语言模型进行文稿生成、支持知识库、支持通过预设提示词搭建各种细分场景的智能体，支持模型一键切换。

2.2 智能体平台-AI 应用无代码封装引擎，支持插件化组件开发和接入。预设丰富的通用组件和 AI 场景专用组件，支持完全零门槛、无代码简单拖拽搭建 AI 应用，可以满足所有的业务场景 AI 应用搭建需求。支持组件插件开发，支持使用 Vue、React 等开发自己的组件，并且接入智能体平台使用。（需要按照相应插件规范开发）

2.3 智能体平台-Agent 创建、管理和 AI Chat 问答系统。

支持自建原生 Agent 应用。原生支持 Agent 搭建，支持为智能体预设提示词、提供知识库和模型选择能力，支持变量设置。

支持 Dify 等平台的 Agent 接入。支持 Dify 平台智能体和工作流的接入，作为一个智能体来进行使用，Dify Chat 应用、workflow 应用和知识库应用。

完全自研 AI Chat 问答系统，支持普通文本、Markdown、代码、数学公式等各种内容的混合渲染和流式渲染。

文本渲染。可实现标题、列表、引用、表格等结构化排版，提升内容可读性与规范性。

代码和脚本显示。支持 Python、GO、JAVA、C、C++、C#、Rust、Bash、Shell、JSON、Yaml、Markdown、SQL、JavaScript、HTML、CSS 等代码、脚本等渲染和语法高亮。

支持图文、视频等媒体内容链接的自动解析和渲染。

支持多模态文件上传、分析、识别和生成，支持图像遮罩编辑功能。

支持多模态的上下文。

支持 MCP 工具调用。

2.4 智能体平台-知识库引擎。

接入 GraphRAG 知识库引擎。相比较传统对知识库内容进行分块处理，GraphRAG 在分块的同时同时保留了分块内容在整个文档的上下文关系，在检测的质量上具有巨大

的优势。

传统 RAG

将用户问题作为查询。

从向量数据库中检索出最相关的几个文档（DOC 1、DOC 2、DOC 3）。

把这些文档直接输入到模型中，模型生成回答。

特点：

检索是点对点的，直接依赖向量相似度。

文档之间没有额外的结构化关系。

检索结果依赖单次查询的精确匹配，可能漏掉隐含相关的信息。

GraphRAG：

文档在预处理阶段被解析成知识图谱（图上的节点和连线）。节点可以是实体、概念、事件等。连线表示它们之间的语义或逻辑关系。

用户提问时，不仅直接匹配文本，还会利用图谱结构找到关联节点和间接信息。

通过图谱推理，将多条分散信息综合后，再交给模型生成答案。

特点

检索是网络化的，通过节点间的关系可扩展检索范围。

能发现跨文档的隐性联系。

对需要推理、多跳关系、上下文关联的场景更有优势。

总结：GraphRAG 更加适合对于上下文关联更加复杂的场景，例如公文写作和基于历史经验的萃取和模仿等。

平台同样支持 Dify 知识库的接入。

2.5 智能体平台-Agent Flow 智能体和 AI 应用混合编排引擎。支持将前面创建的 Agent 和 AI 应用再次进行复杂任务编排，例如可实现自动的分镜脚本策划，到图像生成再到视频的全流程内容创作自动化。

2.6 智能体平台-计费体系。支持平台、应用级别的计费体系，支持按照平台或者应用进行计费。支持免费、按照 token 计费、按照时长计费和固定计费等多种计费模式。

2.7 AI 门户。

架构设计及选型：基于 Vue 平台的 SSR 服务端渲染解决方案和微前端主应用架构平台。微前端架构设计以满足后续持续的接入需求。作为总台的 AI 门户，考量未来可能接入的不同平台的项目、不同的技术栈以及可能对外需求。同时，实现项目间的隔离，子项目例如智能体平台的升级更新可以同步更新，在 AI 快速发展的目前阶段非常重要。相比较传统 iframe 方案，基于 Webcomponents 的微前端方案在路由同步、应用加载优化、应用间通信和子应用保活等多方面具有无可替代的巨大优势。SSR 服务端渲染可以在未来可能面向 C 端提供服务时提供更快首页加载速度和 SEO 优化。

平台与 AI 门户的通信。通过微前端提供的原生通信方式进行用户鉴权和任务状态通信，并提供 webhooks 后端通信等多种通信方式。

模块化设计。授权管理、应用管理和业务系统（如知识库等）均采用模块化设计，均通过接口设计进行二次封装，便于后续的扩展或者更换具体的 Provider。

平台运营管理体系设计。立足于总台的 AI 平台的设计定位，平台运营管理系统设计采用多租户、可自定义权限、支持计费模式的设定与整合（规范各第三方 AI 应用平台的 AI 任务进度及结果通信规范）

开发实现计划。具体开发实现计划见项目进度控制部分的项目计划表。

2.8 运营管理平台。

应用管理。支持使用 ComfyUI 工作流封装 AI 应用，支持应用商店，可以下载最新的各种第三方面应用模版，并支持基于模版进行 AI 应用的编辑和二次封装。支持应用增删改查。

智能体管理。支持新增智能体封装，支持预设提示词、预设模型和知识库，支

持变量，并通过变量渲染智能体应用界面，打造更加灵活的智能体应用。

算力服务器管理。

支持动态添加服务器，无节点数量限制。

显示实时的服务器队列状态。需详细统计系统日周月运行情况，包括平台用户账号注册、登录、用户资源使用情况。

支持并发设置，可根据不同的应用设置并发控制(ComfyUI 类型应用默认为 1)。

支持设置将服务器设置为某个组织专属。

支持异常管理，包括清理队列等。

支持手动更新服务器模型（当模型列表更新时）。

AI 能力管理。AI 能力平台将用户所需的各第三方 AI 能力整合为统一的服务和接口。并通过同一个的接口或者应用模版进行调用。

支持目前主流的大语言模型、文生图、图生图、视频生成、3D 模型生成、音频生成等能力集成。兼容 Google、OpenAI、BFL、火山引擎、即梦、阿里云百炼、Liblib、RunningHub、硅基流动、Dify、Trip3D、Ollama 等平台，并支持持续更新和迭代升级。

支持多账号轮转和限流管理。自动根据账号设置进行限流、超过请求的任务自动排队，如果存在多账号，进行自动轮转。

支持针对不同平台进行不同的计费设置。支持免费、按照应用计费和按照 token 计费等多种计费方式，支持设置 token 计费的比率（不同平台的 token 计算方式不一样，通过设置 token 计费的比率统一换算到平台的计量单位）

云服务能力管理。云服务能力提供短信、邮箱、翻译等通用支持辅助能力，支持阿里云、腾讯云两大主流云服务器平台，用于登录验证、异常告警、提示词自动翻译等场景，无感集成到 AI 应用中。短信验证功能

邮箱验证功能

机器翻译功能接入

支持仙宫云、AutoDL 两大商用算力平台的弹性部署功能。

支持扩展其他通用云服务能力。

存储管理。兼容主流云存储厂商，如阿里云、腾讯云、七牛云、华为云等，兼容私有化的 S3 存储，支持多存储自动重试（一个存储节点异常自动使用另外可用的存储节点自动重试）

组织构架管理。支持自定义组织架构、支持设置组织管理员、支持同一组织内共享组织余额。

用户管理。支持常规的用户增删改查、重置密码和禁用等操作。

内容审核，支持前置和后置审核。支持前置的提示词过滤，支持调用第三方内容审核对生成资源的内容审查，包括涉黄、涉政、涉暴、涉恐等各种不适宜展示的内容进行实时审核和屏蔽

资源和素材管理。支持上传公共的资源和素材，在 AI 应用场景进行直接使用。

支付和套餐管理。接入支付系统，支持自定义套餐和在线支付。

备份与还原。系统原生支持自动备份（自动保留最近 7 天的备份数据）、支持本地备份和远程备份。

其他系统设置。提供丰富的系统设置，包括主题颜色、界面自定义、公众号对接、小程序对接、功能开关等系统配置项目。

2.9 开发及扩展能力-插件系统、OpenAPI 和 MCP 服务。

支持通过插件的方式扩展 AI 无代码封装组件和界面组件，并通过远程组件的形式接入。支持使用 Vue3、React 等前端开发语言，基于 provide 的数据和开发规范，进行组件开发和注册，根据自身的需求扩展无代码封装引擎的组件。

提供整合后的同步、异步的兼容 OpenAI 标准规范的 AI 接口服务。平台的所有 AI 能力和 AI 应用均支持通过开发 API 进行调用。支持 Chat Competition 文生生成支持流式和非流式响应接口，支持 ImageGeneration 图像生成，ImageEdits 图生图及图

像编辑, VideoGeneration 视频生成。针对耗时较长的任务场景提供异步的接口调用和 Websocket 实时进度消息推送。

支持标准 SSE 和 StreamableHTTP MCP 服务。MCP 协议(Model Context Protocol) 是 Anthropic 推出的一个开放标准, 旨在为 AI 模型(尤其是大语言模型) 提供一种标准化的方式与外部工具、数据源(知识库等)或插件进行交互。它的目标是统一接口, 让不同的模型和服务在调用外部资源时无需关心具体实现细节。

核心特性包括:

标准化接口

MCP 定义了客户端(Client)与服务器(Server)之间的通信协议, 包括数据格式、请求类型和响应结构。目前主流的通信为 Stdio、SSE 和 StreamableHTTP 三种, 其中 Stdio 本地通信方式, 不适用与企业场景。对于 SSE、StreamableHTTP 两种远程的 MCP 调用方式, 智能体平台全部提供符合规范的接口支持。

采用 JSON-RPC 风格通信, 方便跨语言和跨平台实现。

模型上下文增强

模型可以动态请求外部数据或工具结果, 并将其纳入对话上下文, 避免依赖训练数据的静态信息。

支持“按需调用”, 减少无关数据传输, 提高性能和安全性。

可扩展性

可以为模型接入数据库、文件系统、API、搜索引擎等自定义资源。智能体平台的所有封装的应用, 均可作为 MCP 工具进行调用, 在支持 MCP 的工具平台, 例如 Dify 等, 工具例如 CherryStudio 等, 经过简单的授权配置, 即可调用智能体平台所有封装并且公开的应用。

不限于特定模型, 任何遵循 MCP 标准的模型和工具都能互通。

安全与权限控制

支持权限声明和用户确认机制，避免模型随意访问敏感数据。智能体平台支持 query、Header 等多种方式的 token 授权验证，保证通过 MCP 调用应用的接口权限和安全。

分布式和微服务系统架构设计，具有很强的扩展能力和弹性伸缩能力。基于大多数 AI 应用比较耗时的特性，普遍的工作流运行时间都在 30 秒到 5 分钟，甚至更少，传统 HTTP 请求已经无法满足长任务的请求结果返回。智能体平台，全部采用 Websocket 消息+HTTP 结合的方式来同步任务提交和结果取回。系统从应用层、AI 能力层、网关层(如 Websocket 网关)采用分层设计，支持基于实际的使用负载情况进行弹性伸缩。系统之间完全解耦，采用 MQ 消息网关进行通信。系统单节点可稳定支撑 1000 QPS 的并发访问，在集群部署下可轻松扩展至 10000 QPS 以上，满足高并发业务需求。开启 websocket 情况下，可以大幅度优化请求延迟，(P95 可以达到可以达到 200ms 以内(不开启 websocket 的情况，P95>30000ms，本地测试条件下数据))

2.10 定制开发展示页面、资源计费、组织管理和内部系统对接等内容，定制开发和我台业务相关的智能体数量 ≥ 10 个。软件定制化交付部分，全部源码交付，后续可以根据我台业务发展需要自主选择进行自由扩展。

2.11 AI 能力集成及模型能力管理。集成目前主流的 AI 能力平台，包括 OpenAI、Gemini、阿里云百炼、火山引擎(豆包)、即梦、可灵、硅基流动、BFL、Trip3D、LiblibAI、RunningHub 等，以及本地部署系列，如 ComfyUI、Ollama 等平台的 AI 能力整合，AI 能力范围涵盖文本、图像、音视频、3D 全模态范畴。提供 dify 平台的商业版本账号。

平台、模型限流和账号轮转功能。支持平台和模型维护的限流，支持并发数、RPM(每分钟请求数)、TPM(每分钟 token 限制)等多种 AI 模型和平台的并发限制，并依据设置自动控制智能体平台的请求数量。超出限制的请求将放入队列等待前序任务执行完成或通过自动寻找下一个模型相同可用平台，自动进行账号的轮转。节省下来额外购买并发请求数的费用。

3. 智能门户

构建统一的 AIGC 智能门户，为用户提供一个集智能体应用、智能工具、AIGC 作品展示为一体的平台。本次建设的 AI 智能门户平台除了支持本次建设内容外，且具备扩展性，支持后续其他应用和能力集成，支持二次开发。

3.1 应用门户首页

智能门户首页提供功能导航、智能对话入口、推荐智能体、智能媒体工具、智能写作工具、智能办公工具等，方便用户浏览与使用智能体业务。其中 AI 工具应用入口为可修改的快速应用入口位，可跳转到指定智能体或 AI 工具；智能工具箱需集中展示已上架的智能体应用，包括对话型智能体和工作流型智能体；智能写作工具，要求提供各类写作风格的智能应用，满足各种不同创作需求。智能办公工具，要求提供各类办公公文写作智能应用，满足不同办公公文写作需求。

3.2 AI 智能体对话页面

提供 AI 智能体对话模块，通过智能体对话页面实现标题生成、稿件撰写等功能。

提供智能体入口，可查询历史对话并设置快捷指令。

3.3 智能体推荐及管理

提供智能体推荐功能，通过上下架操作展示推荐位的智能体，上下架功能权限可控制，集中接入、展示并管理面向媒体生产全链路的精选 AI 智能体，推荐前沿智能体工具。提供智能体分类、标准接入与管理功能，在智能体管理中实现智能体的灵活上下架管理，赋能编辑、策划、运营人员快速调用稿件生产、视频剪辑、热点创作、多模态审核等垂直工具智能体，实现智能生产力无缝集成。

3.4 智能媒体工具

提供智能媒体工具功能，支持工具展示和链接跳转。依赖本项目采购的 AI 大模型和 AI 基础能力，提供支持封装的工具包括：文案创作、基于 comfyui 封装的无代

码前端应用及 API 调用、AI 绘图工具、AI 视频工具、AI 音频工具、AI 基础识别理解、AI 数字人等。

3.5 智能写作工具

提供智能写作的工具快捷入口，功能为模版化写作应用场景。提供的写作类型，包括：

(1) 新闻类智能写作：新闻评论写作、新闻报告生产、新闻通讯写作、新闻消息写作、新闻缩写、新闻扩写、新闻改写；

(2) 新媒稿件创作：分镜编写、公众号文章写作、新媒体旁白文案、新媒体文案、个人风格仿写；

3.6 智能办公工具

提供智能办公工具快捷入口，通过智能办公工具可以快速完成日常办公需求，提供的智能办公工具包括：通报、请示、决议、公告、意见、通知、纪要、汇报、调研报告等智能编写。

3.7 文案创作

本次项目提供文案创作应用，具体功能如下：

- 支持选择不同语言模型进行文稿生成。（根据部署情况视模型选择）
- 支持对创作内容进行专业设置（例：新闻报道、新闻评论、新闻专访等）；
- 支持对创作内容进行网络风格设置（例：小红书、微博等）；
- 提供文本智能增强能力（改、扩、缩、润、摘、翻等）。

3.8 ComfyUI

本次项目所提供的 ComfyUI 接入能力，为相关创作与开发工作搭建起一座高效且智能的桥梁，极大拓展了创作的可能性与灵活性。支持基于 comfyui 封装的无代码前端应用及 API 调用功能。支持使用拖拽的方式将 comfyui 工作流封装为前端应用或者可以调用的 API，支持绝大部分 comfyui 组件、支持多算力的负载均衡及工作流级调度

优化、支持多租户灵活的计费及支付、支持 MCP 协议、支持对接 dify/coze 智能体平台、支持直接接入 liblibAI 及 Runninghub 等云平台的工作流。

3.9AIGC 工具

本项目配置不少如下量级的 AI 生成能力。提供 AI 绘图工具包括文生图、图生图、AI 扩图等功能；提供 AI 视频工具包括文生视频、图生视频等功能；提供 AI 音频工具包括文生配音、音乐、音效等功能；集成即梦、可灵、海螺等大模型产品；接入台内本地部署的大模型。要求图像生成及处理 ≥ 25 万张；1080p 视频生成（视频加音频） ≥ 74 小时；720p 视频生成（视频加音频） ≥ 2.5 小时；图像、视频数字人 1080p 生成 $\geq$ 约 18.5 小时；音乐生成 ≥ 10.7 万首；48G 显存模型 workflows 调用时长 ≥ 1429 小时。

依赖以上 AI 能力，可封装 AI 绘图、AI 视频生成、AI 作曲、数字人播报等工具能力。

本项目提供 1 年热点事件、内容的调用服务，支持 10QPS 并发。依赖此能力，可完成热点分析类智能体的封装。

3.10 新闻写作智能体，包括：新闻评论写作智能体、新闻评论写作智能体、新闻报告生产智能体、新闻通讯写作智能体、新闻消息写作智能体、新闻缩写专家智能体、新闻扩写专家智能体、新闻改写专家智能体。

3.11 媒体应用智能体，包括：分镜编写智能体、公众号文章写作智能体、新媒体旁白文案智能体、新媒体文字文案智能体、个人风格仿写智能体。

3.12 办公应用智能体，包括：通报智能体、请示智能体、决议智能体、公告智能体、意见智能体、通知智能体、纪要智能体、工作汇报智能体、调研报告智能体。

4. 智能体平台

腾讯云智能体开发平台（Tencent Cloud Agent Development Platform）是基于大模型的智能体构建平台。产品提供 LLM+RAG、Workflow、Multi-agent 等多种智能体开发框架，助力企业能够结合专属数据，更高效地搭建稳定、安全、符合业务需求的

智能体，助力企业改进沟通交互体验，增强服务能力，提升经营效率。

应用接待体验上，实现问答覆盖率及精准率大幅提升，“预训练”步骤积累了足够海量的行业知识，对接行业大模型开箱即用，能够更加准确理解行业垂直知识；使用客户领域知识，回答跟客户有关的业务问题，将不准确的答案概率减少到最低，全面提升客户端体验。

知识运营上，实现知识生产与更新效率跨越式提升。冷启动阶段即可对接行业大模型开箱即用，并进一步通过快速导入企业级知识，包括多种格式的知识文档、帮助中心网页，进行高效、智能的知识库自动化扩充，构建企业级大模型机器人，替代传统的问答梳理与对话树设计等繁重人力工作，大幅降低知识生成与更新成本。

同时，产品也设计了 Agent 模式以积极探索大模型应用的边界和潜力。在 Agent 模式下，我们把应用对话的主动权更多地交由给模型，发挥模型的主动性，适用于更灵活、开放的场景。

基于大模型的应用开发平台，旨在通过 LLM+RAG 、Workflow、 Multi-agent 等多种应用框架以及原子能力（API），促进高效的大模型应用构建。产品着重于支持客户重塑和升级其企业级知识服务系统。

4.1 应用配置

应用配置功能是智能体开发平台的核心能力之一，旨在帮助用户快速构建、配置和管理智能体应用。应用配置的功能包括以下：

1) 基础设置

- 支持开场白与欢迎语：支持设置用户首次进入应用时的引导话术，提升交互体验。填写欢迎语后，应用的用户端界面与用户交流的开场白将显示为欢迎语。可以使用 AI 一键优化生成欢迎语。
- 图标与名称：自定义应用头像和名称，用于用户端展示。
- 支持三种应用模式：

- 标准模式：平台内置最佳实践流程，综合知识库、插件和工作流能力，达到更稳定和精确的知识问答效果。适用于企业知识服务、产品咨询等严肃问答场景。
- 单工作流模式：使用指定的工作流来响应用户所有对话。适用于有唯一且明确的业务流程的企业服务场景。
- Agent 模式：启用多 Agent 协同与全局意图识别能力，基于大模型完全自主规划任务，或者根据用户自然语言指令来规划任务。回复更加灵活和广泛，适用于生活咨询、百科知识等非严肃的日常场景。

2) 模型配置

● 双模型驱动机制

思考模型：用于意图识别、任务规划（如 DeepSeek-R1），支持复杂逻辑推理。

生成模型：用于答案生成（如 DeepSeek-V3），优化回复的自然性与准确性。用户可接入第三方模型作为生成模型，只要模型支持 OpenAI 接口协议即可快速接入，用户可在 TI 平台上部署第三方模型，然后在此处进行接入。

● 参数调优

上下文记忆轮数：控制对话历史长度（影响 Token 消耗与多轮连贯性）。

温度与 Top-P：调整生成内容的随机性（高温适合创意生成，低温适合确定性任务）。

上下文改写：自动识别指代省略词，提升问答连贯性。

3) 知识库设定

1. 多源知识接入控制

a. 文档库、问答库、数据库启用开关：在应用设置页面对是否接入文档库、问答库、数据库进行设定。

2. 智能检索增强

a. 混合检索策略：结合关键词与向量检索，平衡语义与字符匹配精度。用户还可以自定义重排序模型、召回增强策略。

b. Excel 专用解析：通过 Text2SQL 技术，用自然语言查询表格数据并秒级返回结果。

4) Agent 与插件生态

1. 多 Agent 协作

a. 零代码转交配置：设定 Agent 间任务转交规则，形成“专家协同体系”（如客服转接技术顾问）。

b. 全局意图 Agent：支持意图达成优先级设定，支持节点回退（如用户修改地点时自动返回上一步），避免流程僵化。

2. 开放插件接入

a. 兼容 MCP 协议，内置腾讯地图、医疗插件等，支持企业自定义插件快速扩展。

5) 意图与权限控制

1. 意图优先级配置

a. 按业务需求设定响应路径（如优先走文档库而非联网搜索）。

b. 对易混淆意图提供多候选选项，引导用户选择。

2. 多层权限体系支持平台级、应用级、知识库级权限隔离，按组织架构、角色或黑白名单精细管控。

6) 输出设置

输出方式配置：

1. 流式输出 (Streaming)

a. 定义：答案逐字生成并实时返回给用户，用户可边等待边阅读（类似打字机效果）。

b. 适用场景：交互性强、回答较长的场景（如客服对话、内容生成），减少用户等待焦虑。

2. 非流式输出（Non-Streaming）

a. 定义：模型完整生成答案后一次性返回，用户需等待全部内容生成完毕才能看到结果腾讯云

b. 适用场景：回答简短、需保证内容完整性的场景（如数据查询、代码生成）。

未知问题处理策略：

当用户提问超出知识库、工作流或插件能力范围时，提供两种应对方案：

1. 智能回复所有问题即使问题超出预设知识范围，仍由大模型自主生成答案腾讯云

a. 优点：灵活性高，避免频繁回复“无法回答”。

b. 风险：可能产生不准确或无关内容，需配合后续人工修正（如通过修正功能录入问答库）。

2. 预设回复兜底

a. 对知识范围外的问题统一返回预设内容（例如：“抱歉，我暂无法解答该问题，请咨询人工客服”）

b. 适用场景：需严格管控回答范围的场景（如法律、医疗咨询），避免误导用户。

7) 测试与调试

1. 实时验证工具

a. 对话测试窗：支持文本、图片（JPG/PNG）、文档上传问答，即时验证效果。

b. 运行过程监控：展示链路耗时、检索片段、模型中间结果，便于定位问题。

2. 反馈优化闭环

a. 修正功能：人工校正错误答案后直接录入问答库，下次优先匹配修正结果。

b. 重新生成：多次调整参数优化回复质量。

4.2 应用评测

应用评测模块通过“标准化数据集+多维度自动化评估+可解释性优化闭环”，为企业提供智能体性能的量化诊断工具。其设计核心在于以业务场景为导向，通过维护测试集模板，结合大模型自动执行测试答案生成，并支持人工标注修正，形成“测试-定位-优化”的完整迭代链条，确保智能体上线前满足企业级质量标准。

1) 样本集管理

1. 样本模版：提供上传模版进行样本集合的创建。
2. 创建样本集：支持对测试样本进行细颗粒度设置，支持按照模版格式设置与上传每个样本的 query、角色 prompt、用户标签。
3. 样本集管理：支持查看样本集的样本量、导入时间，支持增加、删除样本集

2) 应用评测

1. 新建评测任务：支持批量导入样本集（excel）验证 C 侧问答效果。
2. 查看评测任务：查看评测进度、标注进度，支持查看测试结果的参考来源，支持向运营人员反馈错误结果。

3) 评测结果查看与标注

1. 批量验证、标注：支持对批量测试的结果进行标注。标注完成后可查看当前样本集的准确率。
2. 测试结果查看：支持查看测试结果的准确率以及参考来源，支持向运营人员反馈错误结果。
3. 标注导出：支持批量导出标注好的样本集（excel）。

4.3 应用发布

应用发布模块通过“预检-部署-监控”一体化流程，实现智能体从开发环境到生

产环境的无缝迁移。其设计核心在于“企业级安全可控”：支持多环境配置及全链路权限管控，确保发布过程零中断、服务变更可追溯。通过实时监控看板，满足高严谨性场景的发布需求。

1) 待发布

该板块聚焦发布前的合规性与变更校验，界面中将显示应用的所有待发布内容，包括文档、问答、拒答问题和配置。用户确认无误后点击发布即可令待发布内容在正式环境生效。

1. 待发布内容预览：查看上次发布后修改的配置项，包括知识库变更、配置变更等。

2. 发布上线：将处于待发布的文档、问答、拒答问题进行发布，发布后支持调用 API 和 web 端问答。

2) 发布历史

该板块提供全生命周期版本追踪与回滚能力：按时间轴展示历次发布记录，支持对比任意版本差异（如节点配置变更、知识库更新内容）。所有操作留痕审计，满足企业合规要求。

1. 发布历史查看：支持查看历史版本更新时间、发布状态。

2. 发布详情查看：支持查看发布成功的文档、发布成功的配置和发布成功的工作流

3) 调用信息

成功发布之后将获得体验分享链接并支持创建 API 调用密钥，您可通过分享该链接给用户或相关测试运营人员，使他们能够在 web 体验页立即体验您的大模型应用 demo；或通过分享二维码在手机上进行体验；同时也可以直接通过应用 API 调用接口，以 API 形式进行调用。

1. 体验链接：发布后支持通过 API 形式体验应用效果，发布后支持通过链接（需

要内网，需要登陆账号密码）、二维码形式体验应用效果（需要手机在内网并登陆账号密码）。

2. 体验链接-立即体验：点击“立即体验”，进入B端体验链接。

3. 体验链接-重新生成：点击“重新生成”，会重新生成新的体验链接（旧的体验链接会失效）。

4. 体验链接-分享链接：分享链接到剪切板，可粘贴复制给其他用户。

5. 体验链接-分享二维码：发布后支持通过链接、二维码形式体验应用效果；二维码支持扫描使用。

6. API 管理：发布后支持通过API形式体验应用效果，以及使用appkey进行接口对接。

4.4 应用运营

应用运营模块通过“数据监控-问题溯源-反馈优化”闭环体系，为企业提供智能体上线后的全生命周期管理能力。其设计核心在于“以业务效果为导向”，通过实时埋点采集用户满意度数据，结合OP平台定位效果问题瓶颈，并联动人工标注与策略干预（如拒答规则配置、答案修正），驱动智能体持续迭代优化，最终实现服务稳定性与用户满意度的双重提升。

1) 数据分析

该模块聚焦量化评估：支持按时间、渠道维度统计消息量、互动用户数、点赞率等核心指标；通过可视化看板监控用户反馈，为运营人员不断优化知识库、问答对提供重要的数据参考。

1. 数据分析看板：支持查看自定义时间内、不同渠道调用方式下的各类数据（消息数量、互动用户数量、点赞数、点赞率）以及曲线图。

2) 对话记录

该模块实现全链路回溯与主动干预：完整存储用户对话上下文（含输入内容、输

出答案、意图、调用方式），并透传关键调试信息（如用户 id、sessionID、召回的知识片段、 workflow 节点路径、Agent 思考过程）；支持人工直接修正错误答案，修正后的内容将自动同步至知识库，确保同类问题下次命中优化结果，形成“纠错-学习”的自治闭环。

1. 查看对话记录：支持查看发布后应用的对话记录，支持修正。支持按时间导出对话记录。

2. 修改对话记录：支持修正发布后应用的对话记录，支持提交反馈。

3) 不满意回答

该模块负责构建负面反馈驱动优化机制：自动回收用户点踩的问题及其关联上下文（含调试日志），运营人员可标注失败原因（如“知识库缺失”“意图识别错误”），并选择修正答案、设为拒答问题或忽略无效反馈，系统自动归类高频问题生成优化任务。

1. 查看不满意问题：针对发布后的用户点踩问题进行回收，可在管理后台查看不满意问题机器上下文（包括调试信息），以及对问题进行修正、拒绝、忽略等处理。支持对点踩的消息记录进行查询细节、答案纠错、拒答、忽略。不满意问题列表支持搜索、筛选、批量忽略、批量导出。

2. 设置拒答：对于不满意的问题可设置拒答。添加拒答问题后，模型判断相似语义类问题将拒绝回答，并回复未知问题回复语，如“针对您这个问题，我暂时还无法进行回答，请换一个问题吧。”

3. 设置忽略：对于不满意的问题可设置忽略，可将问题从不满意问题列表移除。

4) 拒答问题

该模块强化风险管控与体验兜底，支持手动配置或从不满意回答转化拒答规则（如敏感问题、超业务范围查询），通过语义相似度匹配拦截高风险提问，并触发预设回复（如“该问题超出解答范围，建议转人工”）；结合权限隔离机制，确保金融、医

疗等场景的合规性要求。

1. 新建拒答问题：支持手动创建拒答问题。

2. 编辑拒答问题：拒答问题发布上线后，模型判断和拒答问题相似语义类的问题将拒绝回答，并提供兜底回复语。支持查看并编辑拒答问题（来自于不满意问题或者自行新建），需要发布之后才能生效。

5. 服务内容

5.1 要求支持无代码应用开发封装、任务算力调度，支持多模态混合输出编排，支持任务在线编排和设计，可以根据具体的场景和需求进行优化和调整。

5.2 需提供运营服务，包括基础运维服务、定制开发、功能升级等。要求定制开发展示页面、资源计费、组织管理和内部系统对接等内容，定制开发和我台业务相关的智能体数量 ≥ 10 个。软件定制化交付部分，全部源码交付，后续可以根据我台业务发展需要自主选择进行自由扩展。

5.3 要求支持整合开源生态和 ComfyUI、应用开发、任务调度、多模态混合输出、任务在线设计和运行编排能力并提供可视化的应用快速搭建的功能。

5.4 要求支持接入目前主流模型的 API 且参数可调，并能够支持无代码的方式封装应用界面，方便根据业务实际应用开发。

6. 运营管理

实现平台统一的用户管理、用户认证体系，针对本次项目提供完善的运营管理及计费结算功能，实现对算力资源、AI 模型服务的精细化运营，确保资源合理分配、成本可控。真正做到“全景可视，精算可控”。

运营管理模块专注于提供全栈、可视、精细的运营支撑能力，覆盖算力、模型、用户、成本四大维度，满足采购人对资源统一监管、配额动态分配、费用精准结算的综合需求。它内置多粒度仪表盘，可按日、周、月自动生成平台运营简报，直观呈现用户账号注册量、活跃登录曲线、GPU 使用率、模型调用峰值等关键指标；同时支持一键钻取，帮助运营人员快速定位异常波动。

该模块还具备智能配额引擎，能够依据部门、栏目、个人三级维度灵活设置算力上限与模型并发数，配合实时告警和弹性回收策略，既保障重点业务优先运行，又避免资源闲置浪费。计费结算中心则提供多维度账单：既可按 GPU 卡时、显存占用、推理次数进行分项计价，也支持打包套餐与阶梯优惠，满足不同业务线的成本核算需求。

此外，运营管理模块还提供权限分级管理，确保统计、配额、账单三类操作相互隔离，既方便协作又降低风险，真正实现资源看得见、控得住、算得清，为客户构建长期可持续发展的智算运营体系。

6.1 用户管理、权限管理、角色管理

在系统中对用户账户进行创建、维护和删除。包括管理用户的个人资料、登录凭证（如用户名和密码）、联系方式以及用户状态（如激活或禁用）。通过为用户分配角色来简化权限分配的方法定义和分配特定的访问权限给用户。

6.2 对接 SSO 单点登录及身份认证

对接 SSO 单点登录及身份认证，允许用户使用一套凭证在多个系统中实现无缝访问，提高了安全性和用户体验。

6.3 组织架构管理

在系统中创建和维护企业的部门、团队和职位层次结构，以便有效地管理和分配员工角色、权限和职责。

6.4 用户密码安全策略、登录安全策略、登录日志

安全管理功能，通过配置来确保用户密码的强度和复杂性，以防止密码被轻易破解；包括账户锁定策略、登录尝试的监控和限制；记录用户登录的详细信息，包括登录时间、登录地点、登录平台以及登录的 IP 地址。它们共同构成了信息系统安全的基础。

6.5 智算资源管理

需提供完善的运营管理及计费结算功能，实现对算力资源、AI 模型服务的精细化

运营，可对部门、栏目、个人设置智算配额，确保资源合理分配、成本可控。

6.6 运营统计分析

详细统计系统日周月运行情况，包括平台用户账号注册、登录、用户资源使用情况。

7. 知识库引擎

知识库引擎是智能体开发平台的另一个核心能力，旨在帮助企业高效构建、管理和应用多模态知识体系，实现基于大模型的智能问答与知识服务。

在本项目中，我方提供面向媒体行业的垂类知识库，通过自动识别文件类型、提取关键词，将散落的文档转化为结构化知识库。

知识库引擎的功能包括以下五大功能模块，通过五大模块的深度协同，腾讯云智能体平台将企业碎片化知识资产转化为可执行、可溯源、可演进的智能引擎，驱动业务场景的持续创新。在检索文件时，只需输入关键词，AI 助手就能从海量文件库中迅速定位到所需文件，支持多模态数据解析、动态更新及 RAG 增强生成，通过专题知识库构建和智能标签关联，实现高效知识检索与推理应用。

7.1 文档库

将企业海量非结构化文档（产品手册、合同、报告等）转化为可检索、可理解的结构化知识源。

1) 全格式解析

- a. 支持类型：PDF/Word/Excel/PPT/TXT/图片（PNG/JPG）等 20+ 格式；
- b. 文件限制：文档类 $\leq 200\text{MB}$ ，表格类 $\leq 20\text{MB}$ （支持 1 万行 $\times$ 100 列），图片 $\leq 50\text{MB}$ ；
- c. 智能解析：

- i. 通过 OCR+LLM 提取图文混排内容（如产品说明书中的图表）。

- ii. 对 Excel 表格启用 Text2SQL，用户用自然语言提问（如“Q3 销售额最高的产品”）即可秒级返回结果。

2) 版本控制与差异管理

a. 自动比对：高亮文档增删改内容，支持“保留/合并/删除”一键操作。

3) 检索增强配置

a. 支持混合检索策略

b. 支持参数调优：召回数量（1-10 条）、匹配度阈值（0-100%），控制信息广度和准确性。

4) 解析拆分干预

a. 切分策略动态配置

- 多策略支持：提供 5 种可选的切分算法，支持根据文档类型动态适配：
 - 固定尺寸切分：基础策略，按预设字符数/Token 数分块（如 512tokens），适合格式简单文档。
 - 语义感知切分：基于嵌入模型（如 BERT）计算段落相似度，在语义突变点自动切分，保障上下文连贯性（相似度阈值可设 0.1~0.5）
 - 结构保留切分：识别文档物理结构（标题、图表、表格），确保图文/表文关联内容不被分割。例如：将表格及其描述文本合并为独立块，避免数据错位。
 - 滑动窗口切分：设置重叠区（如前后 100 字符），减少边界信息丢失，适合长技术文档。
 - LLM 智能切分：调用大模型按语义单元（如“研究背景→方法→结论”）重组内容，适合论文、报告等结构化文本。
- 参数调优界面：可视化调整分块大小、重叠比例。

b. 复杂元素专项保护

- 表格跨页合并：自动识别跨页表格并合并为完整数据单元，保留表头与单元格关系。

- 公式与代码块锁定：通过语法规则（LaTeX/编程语言标记）保护数学公式、代码片段不被截断，确保逻辑完整性。
- 图文关联增强：识别图片与相邻说明文字，强制分配至同一文本块，避免“图不对文”。

c. 切片干预与修正

- 切分预览编辑器：上传文档后自动生成切分预览，支持手动修改、添加、删除文档切片。

7.2 问答库

将企业将高频业务问答（如客服 FAQ）标准化，提升回复效率与一致性。

1) 多源构建方式

- a. 手动录入：直接添加“问题-答案”对，支持图片、链接、视频形式的答案；
- b. 批量导入：通过 Excel 模板一次性导入千级问答。
- c. 自动生成：支持选择文档库中的文档生成问答对，模型根据向量数据库相关内容进行上下文片段理解，进行问答对生成。

2) 问答对管理与优化

- a. 问答对管理：支持在问答列表中查看问答来源、关联文档、修改时间、添加时间、字符数量、状态。问答状态包含未校验、待发布、发布中、已发布。操作项支持校验，已校验过的问答支持二次编辑。支持创建问答分类、批量移动问答、删除问答。
- b. 问答到期时间：问答知识的生效时间设置，可设置永久有效或自定义到期时间。若设定了自定义到期时间，在到期时间后，该问答知识将会自动失效。
- c. 问答校验：对文档生成的问答进行人工校验（采纳、不采纳），并支持修改生成的问答对；
- d. 冲突问答校验：针对文档中生成的相似度较高问题进行冲突问题提示，两两判断，用户可选择采纳其中一个、或者保留 2 个。

7.3 数据库

连接企业数据库，用自然语言替代 SQL 查询，降低数据使用门槛。进入数据库界面后支持【接入第三方数据库】，数据库类型支持 mysql、sqlserver。数据表的上限定每个库最多 100 个表。数据表中行列最大支持 10 万行 100 列。

7.4 共享知识库

实现知识跨团队/应用复用，兼顾权限隔离。支持文档库、问答库、和数据库，参考应用知识库中的功能细化，共享知识库支持被应用引用，一个应用可以引用多个共享知识库、一个共享知识库支持被多个应用引用。

创建共享知识库的时候可勾选自定义 embedding 模型。

7.5 标签管理

智能体开发平台通过标签体系实现知识智能分类与场景化路由。标签管理是维护标签名称和标签值的功能，标签名称和标签值可用于为文档/问答打标签。智能体开发平台支持建立不同维度的标签，支持批量导入标签。

支持新建、导入、编辑/删除知识标签。

8. AIGC 引擎

AIGC 引擎是低代码 workflow 编排平台，工作流是企业复杂业务的数字骨架，通过“可视化编排+规则化驱动+全链路可控”，将割裂的人工操作转化为自动化流水线。其本质是以流程为粒度的可编程生产力，在降低人为错误的同时释放创新效能，成为采购人数字化转型的核心中枢。

引擎支持拖拽式流程设计、多智能体协同及插件扩展，内置媒体处理专用节点，实现复杂任务的自动化编排与灵活控制。下文将描述该引擎 workflow 编排及管理能力。

8.1 工作流管理

工作流管理模块为用户提供全生命周期管控能力，支持工作流创建（手动录入/批量导入）、跨应用共享、版本回溯与三层权限体系（平台级→应用级→工作流级），

确保流程的可复用性与安全性。

8.2 workflow 画布

画布是零代码可视化编排的核心载体，通过拖拽节点与有向连线构建链式/并行/路由流程。其内置三大约束保障稳定性：

- 无环依赖校验：禁止循环引用（如 $A \rightarrow B \rightarrow C \rightarrow A$ 无效）；
- 并行分支限制：禁用信息收集类节点（避免逻辑冲突）；
- 强制结束节点：确保流程必有输出。

支持可视化编排（拖拽式编排：从左侧节点库拖拽节点至画布，通过有向连线定义执行顺序（如参数提取→大模型处理→API 调用）；一键优化画布布局：支持一键优化 workflow 布局，使 workflow 展示的更清晰。c. 操作友好模式：支持鼠标和触控板用户，提供针对性的快捷键和交互体验。）和版本控制、workflow 回退等功能。

8.3 workflow 节点

节点是 workflow 中的基本构建单元，每个节点代表一个特定的操作或功能。节点通过互相连接形成一个完整的业务流程。根据功能和用途，腾讯云智能体开发平台 workflow 节点可以分为以下几类：

- 信息收集类节点：此类节点基于多轮对话、页面 UI 等方式收集信息，并将信息向后传递。例如参数收集节点，通过与用户进行多轮对话，收集必要的参数信息。
- 信息处理类节点：此类节点仅对信息进行加工处理，不再收集信息。例如大模型节点，支持通过编辑提示词，利用大模型能力处理信息。
- 基础节点：此类节点是 workflow 运行的基础，实现 workflow 结果输出、逻辑判断等功能。包含结束回复、条件判断等节点。

8.4 workflow 调试

workflow 调试模块为 workflow 开发提供全链路的可观测性，帮助用户实时追踪运行路

径、进行节点/变量级诊断和性能调优。

1) workflows 调试模式（整体调试：支持通过多轮对话对 workflows 进行完整测试，并展示每个节点的输入输出变量等运行结果；单点调试：支持单节点级别的调试，帮助用户测试某个节点的运行效果。）

2) workflows 调试工具（实时预览与校验：点击调试按钮后，系统自动检查配置错误（如未连接结束节点），并在画布实时显示执行路径；变量追踪与断点：观察节点输入/输出变量，定位数据流转异常点；问题反馈：支持在 workflows 调试过程中，反馈节点流转错误、节点执行错误等问题，便于运营人员排查问题。

3) workflows 调用方式配置（同步调用：采用流式调用 workflows，实时运行并返回结果，适用于对话等实时性要求高的场景，例如客服、员工咨询等；异步调用：支持创建工作流异步任务，适用于非对话的长时间任务处理场景，例如视频分析、社媒文章分析等。）

8.5 插件广场

插件广场是智能体开发平台生态体系的核心组件，致力于通过开源共创、即插即用、安全可控的插件生态，降低企业构建智能体的技术门槛，同时提升智能体的执行能力与场景适配性。插件是一个工具集，一个插件内可以包含一个或多个工具（API）。工具是一个 API，通过调用这个 API 完成某一项需求。

- 官方插件：腾讯云智能体开发平台官方提供的插件，您可以直接将其使用在应用中。使用这些插件，可以帮助您拓展应用能力边界。
- 第三方插件：收录在插件广场中，由非腾讯云智能体开发平台官方的第三方开发者（包括企业、开源社区、独立开发者等）开发的插件。
- 自定义插件：如果腾讯云智能体开发平台提供的插件不满足您的使用需求，您还可以创建自定义插件来集成需要使用的 API。将所需使用的工具/插件基于 OpenAPI3.0 协议注册到腾讯云智能体开发平台作为自定义的插件，

即可在应用配置时使用该插件。

9. AI 能力部分

提供以下公有云能力服务，按照实际调用量计费，耗完即止。

9.1 视频摘要能力：提供大模型视频摘要后台能力，支持基于 ASR 识别结果进行大模型摘要提取。支持离线分段和直播实时分段，自动提取段落摘要及关键词，帮助用户快速定位感兴趣的视频片段，提高使用效率。

9.2 视频理解能力：提供视频能力后台能力，提供的视频理解能力可基于 AI 技术和内容运营经验，对视频内容输出涵盖人物、场景、物体、事件的高精度、多维度的优质标签内容。通过对视频内容进行细粒度的结构化解析，应用于媒资系统管理、素材检索、内容运营等业务场景中。

9.3 高光集锦能力：提供高光集锦后台能力，通过将图像特征和文本特征映射到一个统一的特征空间，我们可以利用这些关键词精准地找到视频中的精彩场景，从而提取出最引人注目的片段。支持影视剧、足球、篮球等视频场景。

9.4 横转竖能力：提供智能横转竖后台能力，能够将视频裁剪成一定比例的适合移动设备播放的视频。横转竖提供批量生成短视频的能力，能够将现有的横屏视频资源转换成竖屏视频资源。

9.5 其他基础 AI 能力：包括文字信息提取、文字识别、图像标签、台标识别、语音信息提取、文字转语音、智能字幕、智能拆条、智能审核、智能擦除、智能标签、2D 小样本数字人 1 个（含 10 小时播报）。

附件二：合同产品参考价格折扣

序号	服务内容	单位	数量	价格	折扣
1	新闻评论写作智能体	tokens	1000 万	¥115	5 折
2	新闻报告生产智能体	tokens	1000 万	¥115	5 折
3	新闻通讯写作智能体	tokens	1000 万	¥115	5 折
4	新闻消息写作智能体	tokens	1000 万	¥115	5 折
5	新闻扩写专家智能体	tokens	1000 万	¥115	5 折
6	分镜编写智能体	tokens	1000 万	¥115	5 折
7	公众号文章写作智能体	tokens	1000 万	¥115	5 折
8	新媒体旁白文案智能体	tokens	1000 万	¥115	5 折
9	新媒体文字文案智能体	tokens	1000 万	¥115	5 折
10	个人风格仿写智能体	tokens	1000 万	¥115	5 折
11	通报智能体	tokens	1000 万	¥115	5 折
12	决议智能体	tokens	1000 万	¥115	5 折
13	请示智能体	tokens	1000 万	¥115	5 折
14	公告智能体	tokens	1000 万	¥115	5 折
15	意见智能体	tokens	1000 万	¥115	5 折
16	通知智能体	tokens	1000 万	¥115	5 折

17	纪要智能体	tokens	1000 万	¥115	5 折
18	工作汇报智能体	tokens	1000 万	¥115	5 折
19	调研报告智能体	tokens	1000 万	¥115	5 折
20	录音文件识别预付费包	小时	10000	10000	4 折
21	形象定制_2D 小样本	个	1	4000	7 折
22	播报视频生成服务小时包	个	1	2000	7 折
23	知识库字符扩容包	(字符) 年/套	1 亿_	16000	5 折
24	腾讯云智能体开发平台. 搜索服务包	千次	1 小时	22	5 折
25	腾讯云智能体开发平台 token 资源	Token	5 亿	5600	5.65 折