卡丁兔

https://www.kadingtu.cn

十项 AI 能力开放平台

十项AI能力开放平台

把十项 AI 能力
装进一行 API 调用

普通注册用户 · 完全免费 · 没有任何限制

开发者与中小企业 · 按量付费 / 包月不限次

十项能力 · 一行调用

包月不限次 · 放心调试

阶段性送一程

商务合作方案 | 面向 个人开发者 · 软件开发工作室 · 中小软件企业

CONTENTS

目录

从平台定位到商业模式、十项能力、技术底座与合作路线

01

平台与定位

- 一句话说清卡丁兔
- 中小企业的真实困境
- 三类用户 · 三种服务
- 十项能力全景图

02

十项能力详解

- 文档转换 · 2 项
- 智能感知 · 3 项
- 智能生成 · 5 项
- 输入输出与参数选项

03

商业模式

- 免费策略 · 一分钱不收
- 按量付费 · 用多少算多少
- 包月订阅 · 当月不限次
- 怎样选：按量还是包月

04

技术与生态

- 开发者接入 5 步
- 管理后台与风控体系
- 技术架构与开源自主
- 阶段性送一程 · 合作路线

平台与定位

一句话说清卡丁兔

一个入口，十项能力，三类用户，两种付费方式

卡丁兔把文档转换、智能感知、智能生成三类共 **十项 AI 能力** 收敛成统一的 HTTP 接口：

注册用户打开网页就能免费用；开发者拿到一个 API Key 就能接进自己的产品。

十

十项能力

文档 2 项 · 感知 3 项 · 生成 5 项，覆盖研发与内容生产主链路

一

统一调用

同一套鉴权、同一套返回结构、同一套错误码，接一次即可复用

免

普通用户免费

网页直接使用，完全免费、无次数限制、无任何附加条件

省

开发者很省

按量付费 + 包月不限次，最低 29 元/月，预算可控不失控

普通注册用户	网页端直接使用	**完全免费 · 没有任何限制**
个人开发者 / 工作室 / 中小企业	API 直接调用，嵌入自己的应用或服务	**按量付费 或 包月订阅**

平台与定位

中小团队的真实困境：不是不想用 AI，是用不起、不敢用

在把产品跑通之前，AI 能力的成本与试错代价往往先压垮了预算

01 大厂 AI 接口按量计费，调试期就是烧钱期

验证一个想法要反复调参数、换提示词，几百次调用下来账单已经很难看；
当调用还是免费的，团队才敢大胆试。

02 价格不透明，预算难以预估

按量计费没有天花板，一次上线活动的峰值可能吃掉全月预算；
中小企业最怕的不是单价高，而是花多少不由自己说了算。

03 能力分散，集成成本高于使用成本

文档、语音、图像、视频往往来自不同厂商，鉴权、返回结构、错误码各成一套；
团队要花大量时间写胶水代码而不是做业务。

04 技术选型被锁死，后期切换代价高

一旦深度绑定某家接口，后续换供应商要重写调用层；
团队需要一个过渡期，先跑通业务再决定长期方案。

结论 这正是卡丁兔要解决的问题：把 AI 能力的试错成本压到最低，让中小团队先活下来、先跑通、先拿到收益。

平台与定位

三类用户 · 三种服务方式

同一份能力，按使用方式切分为免费网页通道与付费 API 通道，互不干扰

免费通道

普通注册用户

网页端直接使用

¥0

完全免费

- ✓ 打开网页即可使用，无需申请 Key
- ✓ 没有任何次数限制与功能限制
- ✓ 十项能力全部开放，结果可下载
- ✓ 感知类报告支持 PDF / Word 导出

付费通道 · 主力客群

个人开发者 / 软件开发工作室

API 直接调用

按量

最低 29 元/月起

- ✓ 一个 API Key 接入全部十项能力
- ✓ 按量付费，用多少算多少
- ✓ 包月订阅，当月不限次数
- ✓ 月费便宜且预算完全可控

付费通道 · 成长客群

软件创业者 / 初创中小软件企业

API 直接调用（生产级）

包月

按能力独立订阅

- ✓ 产品验证期：按量付费控制成本
- ✓ 推广期：包月不限次扛住流量
- ✓ 调试期：不限次放心改参数
- ✓ 业务有收益后平滑切换大厂接口

十项能力

十项能力全景图

统一端点 /v2 · 统一鉴权 Authorization: Bearer yck-xxx · 统一返回结构

文档转换 Convert · 2 项

文档转 Markdown
同步

doc2md 21 种格式一键转

Markdown 转文档
同步

md2doc 输出 PDF / Word

智能感知 Sense · 3 项

音频感知
异步

audiosense 转写 + 说话人 + 情绪

图像视觉
同步

imagevision OCR + 画面描述

视频感知
异步

videosense 场景 + 关键帧 + 事件

智能生成 Generate · 5 项

文本转语音
同步/流式

text2audio 5 种音色可选

文本转图像
同步

text2image 7 档尺寸 1~4 张

图生图
同步

image2image 6 类转换 + 局部修改

视频生视频
异步

video2video 风格化转换

脚本转视频
异步

script2video 脚本一键成片

说明 普通注册用户：十项能力在网页端全部免费、无次数限制。开发者：同一批能力通过 /v2 接口调用，按量付费或包月订阅。

文档转换：把客户手上的文档与 Markdown 双向打通

输入端覆盖 21 种文档格式，输出端直出 Markdown / PDF / Word，是知识库与报告类产品的地基

文档转 Markdown Convert · d2m

把 PDF、Word、Excel、PPT、HTML、EPUB 等 21 种格式统一转成 Markdown 。

- 输入：PDF / DOCX / DOC / XLSX / XLS / PPTX / PPT / HTML / HTM / TXT / MD / CSV / JSON / XML / EPUB / ZIP / IPYNB / MSG 等 21 种
- 输出：Markdown 全文、文档标题、使用的转换器名称
- 亮点：支持直取模式，纯 Markdown 输入零转换耗时；自动识别扫描件并走 OCR 通道
- 接口：POST /v2/doc2md/convert

典型应用场景

- 知识库：批量把客户资料转成 Markdown 喂给大模型
- 技术文档迁移：把历史 Word 手册转成站点可直接发布的 Markdown
- 内容归档：把散落的 PDF 报告统一成可全文检索的文本

按量 **¥0.01 / 次 + ¥0.0001 / 字符 + ¥0.0001 / KB**
包月 **¥29 / 月 · 当月不限次**

Markdown 转文档 Convert · m2d

把 Markdown 渲染成交付级 PDF 或 Word 文档，自动套用规范样式。

- 输入：Markdown 文本或 .md / .markdown 文件
- 输出格式：PDF（默认）或 Word（docx）
- 亮点：内置默认样式模板，支持 Doc 主通道 + Office 兜底，产物保留 30 天
- 接口：POST /v2/md2doc/convert

典型应用场景

- AI 报告交付：把模型输出的 Markdown 渲染成正式 PDF 报告
- 合同与方案：Markdown 源文件一键生成 Word 版本供客户批注
- 批量出稿：内容团队用 Markdown 写作，自动产出 PDF 成品

按量 **¥0.03 / 次 + ¥0.01 / KB**
包月 **¥29 / 月 · 当月不限次**

十项能力详解 · 智能感知

音频感知：把声音变成可检索、可分析的结构化文本

一句话输入一段音频，同时得到逐字稿、说话人区分、情绪与场景标签

会议记录、访谈整理、客服质检、课程归档类产品的核心能力：不只是转写成文字，还要知道「谁在说、说了什么情绪、处于什么场景」。

POST `/v2/audiosense/transcribe` 调用模式：**异步任务（返回 task_id）**

参 用户可配置项

说话人数量	0 ~ 10 可调；填 0 表示自动识别说话人数量
输入方式	本地上传（MP3 / WAV / M4A / AAC / FLAC / OGG / WMA）或文件 URL
大文件支持	超过 15MB 自动分片上传，单文件上限 500MB
输出粒度	整篇全文 + 按时间分段的明细（时间 / 说话人 / 内容）
结果导出	转写全文可下载 .txt，识别报告与感知报告可导出 PDF / Word

输入

音频文件（三通道任选其一）：
audio_url 音频地址 / audio_base64 + audio_name 直传 / audio_file_path 分片上传后引用。

输出

transcript.full_text 全文与 segments 分段；diarization 说话人分离结果；analysis 情绪与场景标签；audio events 音频事件（含置信度）；duration_seconds 音频时长。

典型应用场景

- 会议纪要自动生成：录完自动出带发言人的逐字稿
- 访谈与用户调研：批量转写并统计发言人时长占比
- 客服质检：识别通话情绪，筛出高风险对话
- 课程与播客：生成字幕文本，用于检索与二次创作

能力编码 audiosense · 计费按音频秒数，¥0.01 / 秒；包月 ¥99 当月不限次

十项能力详解 · 智能感知

图像视觉：一次调用，同时拿到文字与画面理解

OCR 文字识别 + 画面描述 + 场景与氛围标签 + 主色调提取，四类结果一次返回

面向需要「看懂图片」的产品：证照与票据识别、素材库自动打标、内容审核辅助、电商商品图理解。

POST `/v2/imagevision/analyze`　调用模式：**同步请求（也支持 SSE 协议）**

参 用户可配置项

输入格式	JPG / JPEG / PNG / WebP / BMP / GIF
输入方式	本地上传或文件 URL
可配置参数	无必填参数，上传即可识别（参数由平台统一调优）
识别能力	OCR 文字、画面描述、场景标签、氛围标签、主色调
结果导出	识别报告与感知报告均支持导出 PDF / Word

输入

图片文件（三通道任选其一）：
image_url / image_base64 直传（≤15MB）/
image_file_path 分片上传后引用。

输出

full_text 与 Text 列表为 OCR 文字；description 画面描述；analysis 内含场景标签、氛围标签与 DominantColors 主色调；另有结构化 OCR 明细。

典型应用场景

- 证照与票据：自动提取文字字段，减少人工录入
- 素材库打标：按场景与氛围自动归类图片
- 电商运营：理解商品图内容，生成卖点描述
- 无障碍与合规：为图片生成替代文本与内容摘要

能力编码　imagevision · 计费按调用次数，¥0.02 / 次；包月 ¥49 当月不限次

十项能力详解 · 智能感知

视频感知：把整段视频拆成场景、关键帧与事件时间线

语音转写 + 场景分割 + 关键帧抽取 + 事件识别 + 画面描述，输出可检索的视频档案

面向视频内容平台、课程平台与安防/巡检类产品：需要在不看完视频的前提下知道视频里发生了什么。

POST `/v2/videosense/analyze`　调用模式：**异步任务（返回 task_id）**

参 用户可配置项

输入格式	MP4 / MOV / AVI / MKV / WebM
输入方式	本地上传或文件 URL
可配置参数	无必填参数，上传即可分析
分析粒度	按场景切分，逐场景输出画面描述与检测物体
结果导出	视频识别报告与感知报告支持导出 PDF / Word

输入

视频文件（三通道任选其一）：
video_url / video_base64 直传 / video_file_path 分片上传后引用（单文件上限 500MB）。

输出

概要信息：视频时长、场景数、关键帧数、事件数；
识别结果：语音转写全文、场景时间轴（编号 / 时间段 / 持续时长）、事件时间线（时间 / 类型 / 描述）；
感知结果：逐场景画面描述与标签、整体氛围归纳。

典型应用场景

- 长视频建档：为视频库生成可全文检索的文字索引
- 课程与培训：自动生成章节化的场景与知识点时间轴
- 短视频运营：批量提取场景标签，用于推荐与分类
- 内容审核辅助：结合事件时间线快速定位风险片段

能力编码　videosense · 计费按视频秒数，¥0.02 / 秒；包月 ¥199 当月不限次

十项能力详解 · 智能生成

文本生成：语音、图像、改绘三项基础生成能力

输入端都是「一段文字 + 少量选项」，输出端直接是可播放的音频与可下载的图片

文本转语音 Generate · t2a

同步 / SSE 流式

- 文本上限：5,000 字
- 音色 5 种：默认女声、成年男声、成年女声、儿童女声、男童声
- 本地推理音色无需联网，首次使用某音色时加载模型
- 输出 WAV 音频，支持内联播放与下载；也支持 SSE 流式返回音频块

配音、有声书、课程音频、无障碍朗读类产品的首选能力。

按量 **¥0.01 / 100 字符**
包月 **¥39 / 月 · 当月不限次**

文本转图像 Generate · t2i

同步 / 流式进度

- 提示词上限：1,000 字；负向提示词上限：500 字
- 尺寸 7 档：512×512（默认）/ 768×768 / 1024×1024 / 768×512 / 1024×768 / 512×768 / 768×1024
- 生成张数：1 ~ 4 张批量生成
- 输出 PNG 图片，返回实际宽高与随机种子；异常情况返回非致命警告

营销配图、文章插图、创意草稿、批量素材生产的入口能力。

按量 **¥0.05 / 张**
包月 **¥79 / 月 · 当月不限次**

图生图 Generate · i2i

同步 / 流式进度

- 画风倾向：通用画风（默认）/ 动漫风
- 转换类型 6 类：风格转换、场景转换、氛围调整、色彩调整、内容重绘、自定义
- 局部修改：支持局部重绘（一期暂不支持涂抹选区）
- 重绘强度 0.1 ~ 1 可调（默认 0.5，调画风建议 0.75 ~ 0.85）；描述上限 1,000 字

把已有素材快速改成想要的样子，是投放素材 A/B 测试的效率利器。

按量 **¥0.05 / 次**
包月 **¥79 / 月 · 当月不限次**

十项能力详解 · 智能生成

视频生成：从已有素材改风格，或从一段脚本直接成片

两项视频生成能力均为异步任务模式，提交后返回 task_id，可轮询进度、可取消、可回调通知

视频生视频 Generate · v2v

把已有视频做风格化转换，保留原始动作节奏与时长。

- 输入格式：MP4 / MOV / AVI / MKV / WebM
- 输入限制：分辨率长边 ≤ 1920px；时长 ≤ 5 分钟（CPU 模式 ≤ 30 秒）
- 可选参数：最长 1,000 字的转换描述
- 输出：MP4 视频，返回时长；短片段通常几分钟，长视频可能 10 分钟以上
- 接口：POST /v2/video2video/transform

按量 **¥0.03 / 秒**
包月 **¥299 / 月 · 当月不限次**

脚本生视频 Generate · s2v

输入一段带镜头标记的脚本，自动生成带旁白与字幕的分镜视频。

- 脚本上限：5,000 字
- 脚本约定：只有「旁白：」标记的内容会朗读；
「开场 / 中段 / 镜头N / 中景」等标签作为画面描述；
「字幕：」用于指定画面字幕
- 输出：MP4 成片，返回时长、宽高与文件大小
- 容错：个别场景生成失败时成片仍会输出，并返回失败场景数
- 接口：POST /v2/script2video/generate

按量 **¥0.50 / 次**
包月 **¥499 / 月 · 当月不限次**

异 异步任务怎么用

机制 提交请求即返回 task_id 与查询 / 流式 / 取消三个地址；可轮询 /v2/tasks/{id} 或订阅 /stream 获取进度；任务完成后可配置回调地址，平台以 HMAC-SHA256 签名推送结果。

商业模式

普通注册用户：完全免费，没有任何限制

免费不是试用，而是长期承诺 —— 网页端十项能力全开放，不限次数、无附加条件

¥0 网页端十项能力全部免费使用

不限次数 · 不限功能 · 不需要 API Key · 不需要充值 · 不需要信用卡

不

不限次数

没有每日上限，没有“额度用完”这回事，想用多少次都行

全

功能全开放

十项能力没有阉割版，参数选项与开发者接口保持一致

零

零门槛

注册登录即可使用，不需要申请 Key，不需要绑卡充值

同

结果同等

免费用户与付费客户使用同一套引擎，结果质量一致

页面就是这么呈现的

全部用“次数 + 成功/失败”的通俗口径

- **“所有能力免费敞开用 —— 不必焦虑”** — 看板横幅
- **“看板里的每一个能力都可以直接使用”** — 副文案
- **“费用：免费，不限次数”** — 使用帮助
- **“今日未使用 • 点此体验”** — 能力卡片

技术 免费通道按登录会话直接调用，不占用也不消耗 API Key 与余额；付费通道独立计量，两条线互不干扰。

商业模式

按量付费：用多少算多少，最低 0.01 元起

价格全部公开在能力列表与接入指南中，没有隐藏费用，没有最低消费

本次消费 = 主维度单价 × 计量单位数 + 第二 / 第三维度（可选）

支持 按次 / 按秒 / 按字符 / 按 KB / 按张 五种计量方式，引擎失败不计费

能力	编码	按量计费单价	包月订阅(不限调用次数)	计量方式
文档转 Markdown	doc2md	¥0.01 / 次 + ¥0.0001 / 字符 + ¥0.0001 / KB	¥29 / 月	三维度
Markdown 转文档	md2doc	¥0.03 / 次 + ¥0.01 / KB	¥29 / 月	按次 + KB
音频感知	audiosense	¥0.01 / 秒	¥99 / 月	按秒
图像视觉	imagevision	¥0.02 / 次	¥49 / 月	按次
视频感知	videosense	¥0.02 / 秒	¥199 / 月	按秒
文本转语音	text2audio	¥0.01 / 100 字符	¥39 / 月	按字符
文本转图像	text2image	¥0.05 / 张	¥79 / 月	按张
图生图	image2image	¥0.05 / 次	¥79 / 月	按次
视频生视频	video2video	¥0.03 / 秒	¥299 / 月	按秒
脚本转视频	script2video	¥0.50 / 次	¥499 / 月	按次

数据来源：平台能力主数据配置 yac_capability_info 与包月套餐表 yac_subscription_plan，与网页端展示完全一致。

商业模式

包月订阅：当月不限次数，放心调试放心用

一次付费、整月敞开用；订阅期内该能力调用不再扣余额，价格便宜且完全可控

当月不限次数

订阅生效期内，该能力的调用次数不做任何限制。反复改参数、反复试提示词都不再产生费用。

不扣余额

订阅优先于按量计费：命中订阅的调用消费直接记 0，账户余额不会被动用。

预算封顶

一个月一个固定数字，最贵的能力也只要 ¥499/月。花多少事先就知道，不存在超支。

算 包月 vs 按量：调试一个能力的真实成本

场景（以文本转图像为例）	按量付费	包月订阅 ¥79/月
调试 100 张（试提示词与参数）	100 × ¥0.05 = ¥5.00	¥0（含在月费内）
调试 500 张（调风格与构图）	500 × ¥0.05 = ¥25.00	¥0（含在月费内）
调试 2,000 张（压力测试）	2,000 × ¥0.05 = ¥100.00	¥0（含在月费内）
产品上线后月调用 10,000 张	10,000 × ¥0.05 = ¥500.00	¥79（封顶，省 84%）

注意 包月仅免除调用费，API Key 的每分钟限速、并发数与日配额仍然生效 —— 这是为了保护平台资源，也保证每位客户的服务质量稳定。

商业模式

怎么选：按量还是包月

产品阶段不同，最优解不同。卡丁兔两种方式并存，随时可以切换，互不锁定

按量付费

用多少算多少

- 调用量小、频次低
- 处于可行性验证与预研阶段
- 只想先试一两项能力
- 调用量波动大、难以预估

先充 ¥50 起，边用边看

推荐给：刚接触、还在评估的团队

包月订阅

当月不限次数

- 进入密集调试与联调阶段
- 产品已上线、调用量稳定且偏大
- 需要反复压测与回归验证
- 希望预算绝对可控、不超支

按能力单独订阅，¥29 起

推荐给：正在冲刺的团队（主力推荐）

规则 包月优先规则：同时存在按量计费与有效订阅时，系统优先走订阅（调用费记 0）；订阅到期自动回落到按量付费，业务不中断。

充值单次范围 ¥10 ~ ¥10,000，订单 30 分钟内有效，支持微信支付与支付宝扫码。

价值主张

阶段性送一程：我们不做最好的，我们做最合适的

十项能力的结果不是最优的 —— 但对预算有限的中小团队，它恰好覆盖最需要的那一段路

卡丁兔的十项能力，适合频繁调试、资金预算有限的中小企业与个人开发者。

覆盖产品验证阶段、产品推广阶段、开发调试阶段。待产品有一定收益之后，我们建议您切换成大厂的相关能力收费接口 —— 大厂可以提供更优质的产品结果。

阶段一

产品验证

0 ~ 3 个月

- 证明想法可行，先跑通最小闭环
- 调用量小、频次低、方向还在调整
- 按量付费，用多少算多少
- 最低 ¥50 充值即可开始

按量付费

→

阶段二

开发调试 + 产品推广

3 ~ 12 个月

- 密集改参数、反复压测、上线推广
- 调用量快速上升且难以预估
- 包月订阅，当月不限次数
- 预算封顶 ¥29 ~ ¥499 / 月

包月订阅（主力）

→

阶段三

业务有收益后

12 个月以后

- 产品已验证，用户与收入稳定
- 对结果质量提出更高要求
- 平滑切换大厂付费接口
- 调用层结构不变，改端点即可

切换大厂接口

我们不追求把客户绑一辈子 —— 卡丁兔只负责在最难的那一段路上陪您把产品跑起来。

技术与生态

开发者接入：5 步，半小时内跑通第一行调用

从注册到第一次成功响应，全流程在网页控制台内完成，不需要联系销售、不需要签合同

01 创建 API Key

在控制台「API Key 中心」创建，形如 yck-xxxx；
仅创建时完整可见一次，之后只显示脱敏前缀。

02 充值或开通包月

充值单次 ¥10 ~ ¥10,000，支持微信 / 支付宝扫码；
或在「订阅列表」按能力单独开通包月。

03 确认能力授权

调用 GET /v2/capabilities 查看该 Key 已授权的
能力清单与计费方式，避免 403 无权限。

04 发起第一次请求

请求头 Authorization: Bearer yck-xxx，
POST 对应能力端点，返回统一 JSON 结构。

05 查用量与账单

/v2/billing/usage 看用量，/v2/billing/balance
看余额；控制台的请求日志可逐条回看。

```
# 调用「文本转图像」
curl -X POST https://api.dwen.cn/v2/text2image/generate \
  -H "Authorization: Bearer yck-xxxxxxxx" \
  -H "Content-Type: application/json" \
  -d '{"prompt":"一只戴头盔的兔子在赛道上","size":"1024x1024"}'
# 统一返回：{code, message, data, request_id, timestamp}
```

接入指南里都有

- ✓ 23 个端点的完整清单与参数说明
- ✓ 5 种语言示例：cURL / Python / Java / JavaScript / C#
- ✓ 同步、SSE 流式、异步任务三种调用模式
- ✓ 11 条错误码含义与排查建议
- ✓ 十项能力请求参数与响应结构速查表

技术与生态

控制台：开发者自己就能完成全部操作

不用等客服、不用发邮件 —— 创建、调试、充值、查账、看日志全部自助

1 仪表盘

余额 / 今日请求量 / 成功率 / 平均延迟 / 今日消费 五张核心指标卡；
无 Key 时给出三步引导，有 Key 时直接列出最近请求记录（含计费方式与状态）。

2 能力测试台

按「文档转换 / 智能感知 / 智能生成」分组选能力，自动渲染该能力的参数表单；
支持 JSON 与 Markdown 双视图、流式逐行追加、图片逐帧预览、音频内联播放。

3 API Key 中心

创建 / 停用 / 启用 Key，查看脱敏前缀、授权能力、日配额与今日调用量；
异步任务的回调密钥也可在此重置。

4 充值中心

快捷金额 50 / 100 / 200 / 500 / 1000 / 2000 元或自定义；
扫码支付并自动轮询到账；充值记录、发票、退款三个页签齐备。

5 订阅列表

查看每项能力的包月套餐、生效与到期时间、用量进度；
支持按 1 ~ 12 个月购买，一键取消，临期自动提醒。

6 日志与任务

请求日志含端点、状态码、耗时、消费金额与错误信息；
任务列表含进度、当前步骤、预计剩余时间，可强制取消或重试。

设计 控制台与网页使用看板并存：普通用户只看到「次数 + 成功/失败」，开发者看到「余额 + 费用 + 延迟 + 端点」，互不干扰。

技术与生态

管理后台：13 个模块，把平台当成产品来运营

能力上下线、计费规则、客户配额、充值订单、风控规则与人工复审，全部可视化操作

能力管理

能力标识 / 端点 / 项目 / 调用模式 / 排序 / 上下线

计费规则

7 种计费方式 / 单价 / 二三维度 / 包月月费 / 每日免费配额

API Key 管理

查看全部 Key、调整日配额、每分钟限速、并发上限、启停

客户财务

客户档案、余额、日配额；余额调整带备注与流水留痕

客户配额

客户级兜底配额：Key 配额为 0 时回退使用这里的配置

充值订单

充值订单 / 支付单 / 退款审核 / 发票开具 四个页签

点数流水

充值 / 消费 / 退款 / 调整 / 订阅扣费 五类流水全记录

能力监控

今日请求、成功率、平均延迟、消费、活跃用户、异常数；可强制下线

任务管理

全平台异步任务进度、回调通知状态；强制取消与失败重试

风控规则

按能力或全局配置规则，优先级 P0~P3，动作与严格级别可调

敏感词库

7 大类别分级管理，支持批量导入与在线热重载

风险日志

总日志 / 拦截 / 审核 / 替换 / 标记 五类统计与明细

人工复审

待审核队列、今日已处理、总审核数；通过 / 驳回并支持批量

扩展 最后一张卡片的位置留给扩展 —— 平台以插件方式加载业务模块，新增能力或新增管理页都不需要改动主程序。

管理后台模块清单与源码一一对应，任何配置项都可追溯、可审计、可回滚。

技术与生态

内容风控：客户敢用、平台敢开的底气

输入与输出双向检测，文本 / 图片 / 音视频三引擎，命中后按规则自动处置

用户请求 ▸ 输入检测 ▸ 能力引擎 ▸

引 三大风险识别引擎

文本风险引擎

AC 自动机敏感词匹配 + PII 隐私信息识别（身份证 / 手机号 / 银行卡 / 邮箱 / 护照），命中自动掩码替换；支持繁体与全角归一化。

图片风险引擎

NSFW 分类模型，可配置拦截阈值与复核阈值；高风险直接拦截，中风险转人工复审。

音视频风险引擎

按固定间隔抽帧 + 音频转写回调，视频与语音内容一并纳入检测范围。

命中后的五种处置动作

Pass 通过	无风险命中，正常返回结果
Replace 替换	PII 等敏感信息掩码替换后继续
Flag 标记	记录风险但放行，供事后分析
Review 复审	结果转人工审核队列，暂不返回
Block 拦截	直接拒绝请求，输入拦截不调引擎

机制 规则优先级 P0 系统硬编码 > P1 能力级规则 > P2 全局规则 > P3 词库默认动作；敏感词与规则均支持在线热重载，无需重启服务。

技术与生态

技术架构：插件化底座 + 全本地引擎

YookingBiz* 插件化业务路由，AI 引擎本地部署，数据不出私有环境

接入层

- 网页使用看板（普通用户 · 免费）
- 开发者控制台（计量 · 计费 · 调试）
- 第三方应用（/v2 OpenAPI）

网关层

- 统一鉴权 Bearer yck-xxx
- 三级限流：每分钟 / 每日 / 并发
- 统一响应 {code, message, data, request_id, timestamp}

业务层

- 插件化业务路由（能力Hub · 能力Use · 任务 · 支付 · 会员）
- 计费结算 · 配额校验 · 风控检测 · 任务调度

引擎层

文档引擎 · 语音引擎 · 图像引擎 · 视频引擎 · 编排引擎

1 行 接入成本

一套鉴权与返回结构复用十项能力

3 种 调用模式

同步 / SSE 流式 / 异步任务按需选择

500MB 文件上限

URL、base64 直传或自动分片上传

5 级 限流兜底

分钟限速、日配额、并发数、余额预检、风控

MySQL（月分表）· Redis（缓存与限流）· 本地文件存储 · 结构化日志

技术与生态

开源自主：成本可控，长期可用

十项能力全部基于开源模型与自研编排层构建，没有昂贵的第三方 API 依赖

能力	核心引擎与技术	部署形态
文档转 Markdown	开源代码 + Office引擎预转换，支持 21 种输入格式	本地进程 / CLI
Markdown 转文档	开源代码 + Office引擎，输出 PDF / Word	本地进程 / CLI
音频感知	SenseVoice / Paraformer 语音识别 + Pyannote 说话人分离	本地 ONNX 推理
图像视觉	PP-OCRv5 文字识别 + BLIP 画面描述 + CLIP 语义理解	本地 ONNX 推理
视频感知	FFmpeg 抽帧 + TransNetV2 场景分割 + BLIP / CLIP 分析	本地 ONNX 推理
文本转语音	sherpa-onnx 系列语音合成模型，5 种音色	本地 ONNX 推理
文本转图像	ComfyUI + SDXL Turbo 工作流	GPU 推理服务
图生图	ComfyUI 图生图工作流，支持 6 类转换与局部重绘	GPU 推理服务
视频生视频	ComfyUI 视频转换工作流 + FFmpeg 编解码	GPU 推理服务
脚本转视频	自研编排层：复用文生图 + 文本转语音 + FFmpeg 合成	编排 + GPU

优势 全本地引擎带来三个好处：单位成本极低所以能提供 29 元起的包月、数据不出私有环境所以合规风险小、不依赖第三方接口所以不会因为上游涨价而被动调整价格。

平台概览

平台已经建成的能力规模

以下数字全部来自已上线的源码与数据库配置，不是规划

10

项 AI 能力

文档 2 · 感知 3 · 生成 5

23

个对外 API 端点

10 能力 + 3 任务 + 8 计费 + 2 系统

13

个管理后台模块

能力 · 计费 · 客户 · 订单 · 风控全覆盖

7

种计费方式

免费 / 按次 / 按秒 / 按字符 / 按KB / 按张 / 订阅

细 已实现的机制细节

调用与配额

- 鉴权：Authorization: Bearer yck-xxx，也兼容不带 Bearer 前缀
- 三种调用模式：同步请求、SSE 流式（data: [DONE] 结束）、异步任务（task_id 轮询）
- 异步任务回调：HMAC-SHA256 签名，失败自动重试 3 次指数退避
- 文件通道：URL 直取 / base64 直传（≤15MB）/ 自动分片上传（≤500MB）
- 配额体系：API Key 级与客户级双档，0 表示不限制，Key 为 0 时回退客户级

计费与安全

- 计费结算：订阅优先 → 免费配额 → 按量计费，引擎失败不计费
- 数据留存：请求与响应明细按月分表存储，参数可按能力单独调整
- 风控链路：输入检测 → 能力引擎 → 输出检测，五级处置动作
- 人工复审：高风险结果进入审核队列，内容仅存哈希不存明文
- 日志体系：请求日志、风险日志、回调日志、余额流水四线并行可追溯

价值场景

六类典型场景：客户拿这十项能力做什么

每一项能力都对应一个真实的产品功能，组合起来就是完整的产品能力矩阵

知识库与文档中台

doc2md + md2doc

把客户手里的 PDF、Word、PPT、Excel 批量转成 Markdown 喂给大模型，再把模型输出的 Markdown 渲染成规范的 PDF 或 Word 交付文档。

会议与访谈记录

audiosense + text2audio

录完的会议音频自动转写成带说话人区分的逐字稿，提取情绪与场景标签；再把纪要文字转成语音便于通勤时收听。

内容审核与素材理解

imagevision + videosense

短视频与图文平台用图片感知做 OCR 与画面理解，用视频感知抽取场景、关键帧与事件时间线，构建可检索的素材库。

电商与营销物料生产

text2image + image2image

按商品卖点批量生成主图与场景图，再用图生图把同一张图快速改成不同风格或局部替换元素，支撑 A/B 投放测试。

短视频与课程成片

script2video + video2video + text2audio

写好脚本一键生成带旁白与字幕的分镜视频；已有素材用视频转视频统一风格，旁白用文本转语音批量重录。

AI 应用与 SaaS 集成

全部十项能力

把十项能力当成自己产品的 AI 底座：一套鉴权、一套返回结构、一套错误码，用最少工程量补齐产品的 AI 能力。

价值场景

按输入与输出组织：客户只需要关心「我手上有什么」

十项能力按输入端形态归为三个入口，按输出端形态归为三类产物

出 输入端：我手上有什么

文本

prompt / 脚本 / Markdown / 逐字稿

text2image text2audio script2video md2doc

文档

PDF / Word / Excel / PPT / HTML / EPUB 等 21 种

doc2md

图像

JPG / PNG / WebP / BMP / GIF（≤15MB）

imagevision image2image

音频、视频

MP3 / WAV / M4A / FLAC / OGG；
MP4 / MOV / AVI / MKV / WebM

audiosense videosense video2video

输出端：我需要什么产物

结构化文本

Markdown 全文、带说话人的逐字稿、OCR 文字、
画面描述、场景与事件时间线、情绪与氛围标签

媒体文件

PNG 图片、WAV 音频、MP4 视频、
PDF / Word 文档，均提供可下载地址

感知报告

图片 / 音频 / 视频识别报告与感知报告，
支持一键导出 PDF 或 Word

价值主张

关于结果质量：我们把话说在前面

十项能力的结果不是最优的 —— 这是我们主动告知客户的第一件事

对比维度	卡丁兔 · 十项能力	大厂收费接口
结果质量	满足产品验证与推广期需求，不是行业最优	调优更充分，质量稳定领先
计费方式	按量低至 ¥0.01 起；包月 ¥29 ~ ¥499 当月不限次	以按量为主，单价显著更高
调试成本	包月订阅下反复调试不产生额外费用	每次调试都计入账单
预算可控性	包月封顶，花多少事先就确定	用多少算多少，峰值可能超预算
数据合规	全本地引擎，数据不出私有环境	需评估数据出域与合规要求
适用阶段	产品验证期 · 推广期 · 开发调试期	产品成熟期 · 追求极致质量

我们的定位很清楚：在客户预算最紧、最需要频繁试错的那一段路上，提供一个便宜、够用、不限次的 AI 能力底座。

待客户产品有了收益、对结果质量提出更高要求时，我们主动建议切换到大厂的相关能力收费接口 —— 只负责阶段性送一程。

因为调用层结构统一（同一套鉴权、路径风格与返回字段），客户从卡丁兔切换到大厂接口时，改动集中在端点与字段映射，业务代码基本不动。

这正是「阶段性送一程」的工程含义：我们提供一个低成本的过渡层，而不是一个长期绑定。

价值主张

五类客户，各自得到什么

同一套平台能力，按客户的预算与阶段提供不同价值

客户类型	当前处境	在卡丁兔得到什么	成本
普通注册用户	想体验 AI 能力，但不想付费、不想申请 Key	网页端十项能力免费用，不限次数，结果可下载	¥0
个人开发者	接私活或做副业，预算有限	一个 Key 调用十项能力，按量付费可控，包月便宜	¥29 ~ ¥499 / 月
软件开发工作室	同时服务多个甲方，项目周期短	按项目开通包月，交付后取消，不产生长期成本	¥29 ~ ¥499 / 月
软件创业者	产品还没验证，最怕把钱烧在调试上	包月当月不限次，反复改参数不额外花钱	按能力单独订阅
初创中小软件企业	已有产品，需要补齐 AI 能力	统一鉴权与返回结构，用最少工程量补齐能力矩阵	按量 + 包月组合

省钱

包月封顶，最低 29 元/月；按量最低 0.01 元起

省心

统一鉴权与返回结构，接一次即可复用十项能力

省事

控制台自助完成创建、调试、充值、查账、看日志

价值主张

算一笔账：包月封顶让预算真正可控

以调用量最大的三类能力为例，包月订阅把不可预测的按量账单变成了固定支出

能力	按量单价	月调用量	按量月支出	包月月费	节省
文本转图像	¥0.05 / 张	10,000 张	¥500.00	¥79	84%
图生图	¥0.05 / 次	10,000 次	¥500.00	¥79	84%
视频生视频	¥0.03 / 秒	60,000 秒（约 16.7 小时）	¥1,800.00	¥299	83%
音频感知	¥0.01 / 秒	200,000 秒（约 55.6 小时）	¥2,000.00	¥99	95%
视频感知	¥0.02 / 秒	100,000 秒（约 27.8 小时）	¥2,000.00	¥199	90%
脚本转视频	¥0.50 / 次	2,000 次	¥1,000.00	¥499	50%

解读 调用量越大，包月的性价比越高；调用量很小的时候，按量付费反而更划算。这也是我们把两种方式并存、随时可切换的原因。

按量付费的风险

账单跟着调用量走，一次推广活动或压测就可能超出预期，预算上限不可控。

包月订阅的价值

月费固定，调用量再怎么涨也不额外收费，团队敢放开手脚调试与压测。

合作路线

怎么开始：当天就能用上

不需要商务谈判、不需要签署纸质合同、不需要预付年费，注册与充值全自助

第 0 天

注册账号

网页端直接免费用十项能力，先感受结果质量

第 1 天

创建 API Key

控制台一键创建，拿到 yck- 开头的密钥

第 1 天

充值或开包月

扫码支付即时到账，或被订阅，当天即可生产调用

第 1 周

接入与联调

按接入指南完成对接，测试台可逐项验证参数

第 1 个月

复盘与调整

根据用量看板决定继续按量还是转包月

诺 我们对客户的三条承诺

免费永久

普通注册用户的网页端使用权不会收费，也不会缩水，这是平台的长期承诺。

随时可退

包月订阅按能力单独生效，可随时取消；订阅到期自动回落按量，业务不中断。

不做绑定

当您需要更好的结果质量时，我们主动建议您切换大厂接口，并配合完成迁移。



卡丁兔

十项 AI 能力开放平台

普通注册用户	完全免费，没有任何限制
开发者与中小企业	按量付费或包月订阅，最低 ¥29 / 月
结果质量	不是最优但够用；有收益后建议切换大厂接口

把 AI 能力的试错成本压到最低，先让产品跑起来。

我们只负责阶段性送一程。

附录：网页端功能全清单

以下功能均已在前端页面与后台接口中实现，可作为功能核对清单使用

普通用户 · 能力使用

- 能力使用看板：今日已用 / 成功 / 失败 / 累计四张指标卡
- 三大分组能力卡片（文档转换 / 智能感知 / 智能生成）
- 最近使用记录表，点击行直达对应能力页
- 统一页面骨架：标题栏 + 我的记录 + 使用帮助按钮
- 输入区：本地上传 / 文件 URL 双通道
- 大文件自动分片上传，单片 4MB、并发 3、进度到 95% 后合并
- 配置区：每项能力独立的参数表单
- 执行区：进度动画 + 取消按钮（可中断后端任务）
- 结果区：文本复制下载 / 音频播放 / 图片预览 / 视频播放 / 产物卡片
- 我的记录抽屉：按能力与状态筛选、分页、回看详情
- 报告导出：识别报告与感知报告，支持 PDF 与 Word
- 状态持久化：离开页面再回来，输入与结果还在

开发者 · 控制台

- 仪表盘：余额、今日请求量、成功率、平均延迟、今日消费
- 无 Key 时的三步引导与「立即创建 Key」入口
- 能力列表：展示按量单价与包月月费，支持按类别筛选
- 能力测试台：按分组动态渲染参数表单
- 测试台支持 JSON / Markdown 双视图与流式逐行追加
- 测试台支持图片逐帧预览与音频内联播放
- API Key 中心：创建、停用、启用、查看脱敏前缀与配额
- 接入指南：5 步流程 + 23 个端点 + 5 种语言示例代码
- 接入指南：11 条错误码说明 + 计费规则表 + 单价表 + FAQ
- 充值中心：快捷金额与自定义、支付渠道、订单 / 发票 / 退款
- 订阅列表：套餐、生效期、用量进度、按 1~12 月购买、取消
- 请求日志、任务列表、用量看板、风险记录四个查询页

管理后台 · 运营

- 能力管理：端点、项目、调用模式、排序与上下线
- 计费规则：7 种计费方式、单价、二三维度、包月月费、每日免费配额
- API Key 管理：调整日配额、每分钟限速、并发上限、启停
- 客户财务：客户档案、余额调整（带备注与流水）
- 客户配额：客户级兜底配额，Key 配额为 0 时回退
- 充值订单：充值订单 / 支付单 / 退款审核 / 发票四个页签
- 点数流水：充值、消费、退款、调整、订阅扣费五类
- 能力监控：今日请求、成功率、延迟、消费、活跃用户、异常数
- 任务管理：进度、回调通知状态、强制取消、失败重试
- 风控规则：按能力或全局、优先级 P0~P3、动作与严格级别
- 敏感词库：7 大类别分级、批量导入、在线热重载
- 风险日志与人工复审：统计、明细、通过 / 驳回、批量处理