RONG · 融合模型

春雪初融
万流归一

一个模型身份，承接全部调用。

产品白皮书 · 2026

版本 1.0 · 2026 年 8 月

01 / 为什么需要统一模型入口

模型能力越来越丰富，但应用团队面对的接入成本也同步增长：不同供应方的认证、请求格式、模型命名、可用性和计量方式互不相同。每增加一条模型链路，都会把复杂度带入业务代码、运维流程和成本管理。

融模型的目标不是替代模型本身，而是在应用与模型之间提供一个清晰、可管理的统一入口。开发者只需维护一种兼容的调用方式，并通过一个账户完成令牌、日志与用量管理。

> “统一入口让业务关注能力选择，而不是重复建设接入层。”

接入碎片化	认证、端点和请求语义不同，切换模型需要反复改造。
凭证难治理	密钥分散在不同环境，轮换、撤销与权限边界难以统一。
用量难解释	调用记录、失败状态与消耗分布在多个控制台。
连续性风险	上游波动直接暴露给业务，定位链路长、恢复成本高。

02 / 产品定位与核心体验

融模型面向需要稳定使用大模型能力的开发者和团队，提供 OpenAI 兼容的调用入口。产品围绕“一个模型身份”组织体验：从注册到创建令牌，再到查看日志与用量，关键操作都在同一界面完成。

核心能力

统一调用	以兼容接口承接应用请求，降低已有工具与代码迁移成本。
令牌管理	独立创建、识别和撤销访问令牌，减少长期共享密钥。
用量可查	按请求查看时间、模型、状态与计量信息，帮助理解消耗。
账户安全	集中管理会话与账户设置，为异常处理提供明确入口。

典型路径

注册账户 → 创建访问令牌 → 使用兼容客户端发起调用 → 在日志中确认结果 →
在用量页面追踪消耗。整个路径强调最少概念和清晰反馈，让首次接入与日常治理使用同一套心智模型。

03 / 架构原则

融模型以边界清晰、凭证最小暴露和可追踪为基本原则。浏览器负责账户与管理体验，服务端负责敏感凭证处理和上游交互；应用调用通过专用令牌完成，避免把账户会话当作长期 API 凭证。

兼容优先	减少专有调用约定，为现有 SDK 与工具保留迁移空间。
服务端边界	敏感配置与上游凭证保留在可信服务端，不进入浏览器包。
最小权限	令牌可独立撤销；账户、会话与调用凭证采用不同生命周期。
可观测	请求状态、时间和用量形成可查询记录，支持定位与核对。
渐进演进	模型与上游可以变化，面向应用的核心接口保持尽可能稳定。

请求链路

应用携带令牌发起兼容请求；平台完成身份校验、路由与必要的协议适配；模型服务返回结果；平台向应用传递响应，并记录用于安全、故障定位和用量展示的必要元数据。实际能力和可用模型以线上状态为准。

04 / 安全、隐私与责任边界

模型调用天然涉及输入内容、访问凭证和第三方处理。平台通过访问控制、传输保护、日志审计和最小权限降低风险，同时明确提醒使用者：不要在提示词中提交非必要的敏感信息，也不要把概率性模型输出直接作为高风险决策的唯一依据。

安全实践

凭证	为不同应用使用独立令牌；避免写入公开仓库；发现泄露立即撤销。
数据	只提交完成任务所必需的信息；对敏感内容先脱敏或匿名化。
输出	对医疗、法律、金融和安全相关内容进行人工与专业复核。
依赖	模型能力可能由第三方提供，处理范围随所选模型和链路变化。

平台提供的是模型能力入口与管理工具，不承诺模型输出必然真实、完整或适合特定目的。用户仍需对输入的合法性、调用行为以及输出的具体使用负责。详细规则请以网站最新用户协议与隐私政策为准。

05 / 面向未来

模型生态仍在快速变化。融模型将围绕兼容性、稳定性、可管理性与透明度持续演进：在不增加应用侧负担的前提下扩展可用能力，提供更清晰的运行状态与用量反馈，并逐步完善团队级治理能力。

演进方向

更广的能力接入	在统一调用心智下承接不同类型的模型能力。
更清晰的成本理解	让调用、状态与用量之间的关系更容易核对。
更完整的团队治理	围绕成员、权限、配额和审计建立可控边界。
更可靠的连续性	完善状态反馈、异常定位与上游变化应对机制。

开始使用

访问

rong-ai.com，注册账户并创建访问令牌。最新功能、可用模型、服务状态和规则以网站实时展示为准。

融 · 一个模型身份，承接全部调用。

本文档用于介绍产品理念与能力边界，不构成服务等级、法律或商业承诺。