准备与修炼 · 一页看懂

Key 与模型怎么配

调用身份、服务地址和模型，各司其职

四项配置的检查顺序

01 Base URL 找平台 → 02 Key 验证身份 → 03 model 选择模型 → 04 权限与额度放行

编号表示检查顺序；实际请求同时携带认证信息和模型字段。

同一把 Key

有权限时可以选择不同模型；备注名不选模型。

两条路线

APIGate Key 配 APIGate 地址；直连用厂商凭据。

最容易错

混用平台 Key 与地址，或把网页会员当 API 余额。

检查完成

收到测试回答，并核对对应模型的新用量记录。

记住这一点

换模型看 model，能否调用还要看权限。

APIGate.biz

把 APIGate 额度接到您的聊天工具

本课完成一条实际连接：在 Cherry Studio 输入一句测试文字，收到回答，再到 APIGate 控制台核对用量。您会分清账户、客户端和模型额度。安装与注册时间因设备而异，可以单独安排，不计入主轴课时。

01 先分清 Key、模型、地址和权限

打开 Kimi 或其他厂商的 Key 管理页时，您可能会疑惑：这里没有选择模型的地方，这把 Key 到底是给哪个模型用的?

Key 用来验证调用身份，模型通常在客户端或每次请求中另选。 一把 Key 可以调用哪些模型，还要看所属平台、组织或项目的权限。拿到 Key，不等于所有模型都已开通。

您看到的设置	它决定什么	该怎么填或检查
API Key（密钥）	以哪个调用身份访问服务	在对应平台创建，只填入指定密钥框
Base URL（服务地址）	把请求发到哪里	按该平台、该客户端的接入说明填写
Model ID（模型标识）	本次请求使用哪个模型	从当前模型目录复制精确名称
权限与限额	这个调用身份能用什么、能用多少	查看组织 / 项目授权、余额和限额；不能从 Key 名称判断

同一把 Key，怎样换模型

假设 Kimi 当前目录列出了 `kimi-k3` 和 `kimi-k2.7-code`，而您的 Key 对两者都有权限。两次请求可以使用同一把 Key、同一个服务地址，只改变 `model`：

```
{"model": "kimi-k3", "messages": [{"role": "user", "content": "请只回复：测试成功"}]}
```

```
{"model": "kimi-k2.7-code", "messages": [{"role": "user", "content": "请只回复：测试成功"}]}
```

这两个代码框只用于对照字段，不是可以直接粘贴到终端执行的命令，也不要求您现在调用。这一例子不代表您的账户已经开通这些模型。实际操作时，先核对当前目录、权限和价格；模型名称和可用性会变化。

如果使用聊天客户端，通常在供应商设置中保存地址和 Key，再在模型列表里选择模型。客户端会把您的选择写进请求中的 `model` 字段。**给 Key 改名不会切换模型**：把一把 Key 命名为“工作用”“备用”或“K3”，只是方便自己辨认。

通过 APIGate 与直接调用厂商，是两条配置路线

使用路线	在客户端填写的 Key	服务地址	模型名称从哪里找
通过 APIGate	您的 APIGate Key	本课 OpenAI 兼容入口：`https://api.apigate.biz/v1`	APIGate 当前模型目录
直接调用 Kimi	您在 Kimi 开放平台创建的 Key	Kimi 官方文档中适用于您账户与协议的地址	Kimi 当前模型目录

通过 APIGate 时，普通用户不需要再取得 APIGate 的上游厂商 Key。客户端把所选模型和请求发给 APIGate，网关再按模型及路由配置调用上游。不要把 Kimi Key 粘到 APIGate 的密钥框，也不要用 APIGate Key 请求 Kimi 官方地址。

费用要看实际使用的服务平台、模型、输入输出用量及其他适用计费项；Key 的备注名称不决定价格。APIGate 价格以站内当前价目为准，不等于厂商官网标价。Kimi 网页聊天会员、Kimi API 账户和 APIGate 模型余额也不能当作同一份额度。

有 Key 却仍然调用失败

- **401 / 认证失败**：先检查地址与 Key 是否属于同一平台，再检查复制是否完整、Key 是否停用、软件是否读到新凭据。
- **模型不存在 / 无权限**：检查 Model ID 的精确拼写，以及这个渠道、组织或项目是否提供并授权该模型。不能仅凭这句话断定模型已经下线。
- **额度不足 / 支出上限**：区分平台余额、账户或项目限额、单 Key 限额以及上游渠道问题。换一个 Key 名称不会增加额度。
- **聊天成功，读取文件或执行工具失败**：模型回答文字与客户端工具执行是两项能力，要分别验收。模型自称“我是某某模型”也不能证明请求走对了线路。

需要逐项处理时，打开接入验收与故障排查。不要把完整 Key 发给“问 APIGate”或放进作业截图。

两分钟检查：不用充值，也不用发送请求

先在纸上或本机笔记中回答，再展开核对您的理解：

1. 把 Key 的名字改成“K3”，是否会自动使用 K3?
2. 有一把 Kimi Key，是否就能调用它的全部模型?
3. 已通过 APIGate 配好客户端，改选另一个可用模型，是否一定要重新建 Key?

APIGate.biz

4. 出现“模型不存在或无权限”，先检查哪三项?

参考答案：① 不会，检查模型选择或 `model` 字段。② 不一定，受平台及组织 / 项目权限限制。③ 不一定；同一 Key 有权限且所选模型支持该接入协议时，可以继续使用。④ 精确 Model ID、服务地址 / 渠道、账户授权。

能说清这四题，再继续下面的连接练习。练习结束后，在控制台核对同一时间、对应模型的新用量记录，并保留不含密钥的结果说明。

本节调用机制参考 Kimi 官方 API 入门示例；官方旧示例中的历史模型名称不作为本课当前可用模型推荐。核对日期：2026-09-30。

02 先认识这次会打开的两个窗口

APIGate 网页负责账户、余额、模型价目和密钥管理；Cherry Studio 是安装在电脑上的聊天客户端，负责让您输入问题并查看回答。两者要靠 API Key 连接。它相当于发给软件使用的一把钥匙，完整内容只填进客户端的密钥框，不放进聊天正文。

充值增加 APIGate 的模型使用余额，不会变成豆包、Kimi、ChatGPT 或 Claude 的会员。不同模型的输入、输出价格不同，文件越长、反复生成越多，用量也会增加。先完成一次短测试，再做本章练习。

03 准备账户与客户端

1. 打开APIGate 首页，点击“免费注册”，按界面填写自己的邮箱和密码，完成邮箱验证。已经注册过就点登录。
2. 进入会员控制台，找到余额、充值和 API Key 管理。先阅读当前价目和充值金额，再由您自己选择是否充值。课程不会代替您提交付款。
3. 找到自己账户的可用 API Key。新建或显示完整 Key 时保存到密码管理器；若旧 Key 只显示前缀，那是为了保护完整内容，按控制台提供的创建入口生成一把练习用 Key，不要猜后半段。
4. 到Cherry Studio 官方下载页选自己的系统。Windows 通常选择对应架构的安装程序；Mac 在苹果菜单“关于本机”查看是 Apple 芯片还是 Intel，再选对应安装包。下载页若提供通用版本，按官方说明选择。

Windows 下载后，从浏览器下载列表打开安装程序，按安装向导完成；Mac 打开下载的磁盘映像，按照窗口提示把应用拖进“应用程序”，再从应用程序打开。若电脑由单位管理，按单位软件安装流程办理。遇到系统明确阻止运行时查看官方安装说明，不照陌生教程关闭系统保护。

第一次打开后先找到“设置”，再找“模型服务”。软件版本不同，设置可能位于左下角齿轮或菜单内；可对照官方模型服务设置。这一步不需要打开终端。

APIGate.biz

04 填好三个关键值

新增自定义供应商，名称可以写 `APIGate练习`，类型选 OpenAI 或 OpenAI 兼容。名称是您自己辨认用的；类型决定软件按什么方式发送请求。

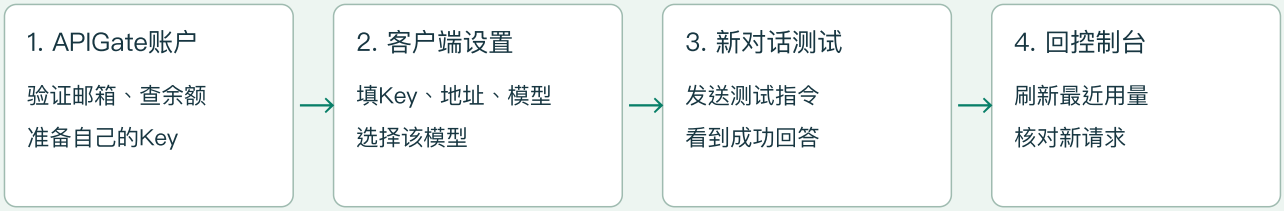
字段	填什么	不要填什么
API Key	您自己账户的完整 Key	示例星号、账号密码、别人的 Key
API 地址（Cherry Studio）	https://api.apigate.biz	APIGate 首页地址或完整聊天请求路径
Model ID / 模型 ID	从 APIGate 当前模型价目复制的精确 ID	自己翻译的中文名称或失效旧型号

> https://api.apigate.biz

打开模型接入与价目，选择一个当前可用的文字聊天模型，复制 Model ID。回 Cherry Studio 的该供应商下添加模型，粘贴 Model ID，并启用供应商。当前官方设置文档说明客户端会补常规请求路径，所以此处先填写根地址。其他客户端若要求以 `/v1` 结尾的 Base URL，应填 `https://api.apigate.biz/v1`。查看请求预览时，最终路径应为 `/v1/chat/completions`，不应出现 `/v1/v1`。不要同时改模型和地址，先修正一个字段再测试。

保存后回到对话界面，打开一个新对话。在模型选择处选刚加入 APIGate 的模型。只在设置里保存供应商，还不等于当前聊天已经使用它。

连接练习：每一步都有可检查的结果



原创连接关系示意；课程不展示或保存学员的完整API Key。

本图是连接关系示意。真实页面、字段名称和您可用的模型以当前账户为准。

05 发一条短测试

在聊天输入框粘贴以下内容，点击发送。Windows 粘贴用 Ctrl+V；Mac 用 Command+V。

> 这是一条连接测试，请只回复：APIGATE_OK

收到回答后，回 APIGate 会员控制台，点击刷新用量，确认出现相应模型的一笔新请求。回答和用量记录要一起看：如果回答来自客户端其他已配置供应商，APIGate 用量不会因此增加。模型偶尔多回复几个字并不自动表示连接失败，重点是请求成功且账户出现对应记录。

在本页作业区记下测试日期、客户端名称、Model ID、是否收到回答、是否看到新用量。只记 Key 名称或前缀，不写完整 Key。保存为 `连接检查.txt`。

06 再把任务材料放进去

回到第 1 章，复制包含材料的完整任务指令，在同一个客户端建立新的课程对话，把标题改成“01任务清单”。今后每章开一个新对话，避免上章材料混进下一章。模型切换前先记下当前结果，有些模型不支持图片、工具或特别长的文件。

费用按真实用量结算。您可以先在账户中设一个自己能接受的月度预算，再练习；额度不足时停止重试，查看余额和请求记录。课程不会承诺某个固定金额足够无限生成。给任务更多无关资料并不会自动让结果更好。

07 根据报错找到对应字段

看到什么	先检查哪里	下一步
401、未认证、无效密钥	API Key 是否完整、是否已撤销	重新复制有效 Key，注意前后空格
模型不存在、404	Model ID 和请求地址	从当前模型列表重复制，并检查重复路径
余额不足	账户余额和预算上限	查看用量记录，再决定是否充值或调整预算
429、请求太多	是否连续发送或多窗口一起测试	停止连点，稍后再试一条
超时或供应商不可用	网络、服务状态、该模型状态	保存错误文字，使用连接医生检查

打开AI连接医生可以先检查账户配置。它不等于在您电脑上完成了真实调用，所以修好后仍要重复短测试。提工单时提供客户端、模型 ID、时间、报错和 Key 前缀即可。

APIGate.biz

08 完成检查

- [ ] 我知道当前聊天选择的是 APIGate 供应商下的模型。
- [ ] 我收到短测试的回答，并在账户看到对应新用量。
- [ ] 我保存了不包含完整 Key 的连接检查文件。
- [ ] 我知道在哪里看余额和预算。

09 继续学习

开始第 1 章。希望让工具操作文件时，再看智能体工具上手。有关密钥和接口的解释见概念修炼课。

APIGate.biz