Perseworks — API 文档

perseworks.hk/zh/docs · 更新于 2026-10-26 · API 地址 https://api.perseworks.hk/v1

入门

简介

Perseworks API 用一个 key、一个预付余额调用 Seedance 视频、Seedream 图像和豆包 Seed 对话模型。接口遵循 REST：JSON 请求体、JSON 响应、标准 HTTP 状态码、Bearer Token 鉴权。

对话接口**兼容 OpenAI**——已在用 OpenAI SDK 的，只需改 `base_url` 和 API key。视频、图像接口沿用同一套约定。

所有请求必须走 HTTPS；明文 HTTP 或不带 key 的请求会失败。

BASE URL

```
https://api.perseworks.hk/v1
```

OPENAI SDK

```
from openai import OpenAI

client = OpenAI(
    api_key="sk-...",
    base_url="https://api.perseworks.hk/v1",  # 唯一需要改的地方
)
```

首次调用前

"key 用不了"的问题，绝大多数出在下面五步之一。请按顺序完成：

1. **登录** app.perseworks.hk（邮箱验证码）。
2. **充值点数。**新账号余额为 **0**。每次请求会*先预扣*点数再执行，余额为 0 时即使 key 正确，所有调用也都会返回 `402 insufficient_points`。在 Billing 页购买点数包，或兑换兑换码。
3. **创建 key**：API Keys 页。完整的 `sk-…` **只显示一次**，请当场复制。若无需要，*允许模型*和 *IP 白名单*保持留空。
4. **使用正确的 Base URL：** `https://api.perseworks.hk/v1`。请求必须发到 **`api.perseworks.hk`**——不是 `perseworks.hk`（官网）或 `app.perseworks.hk`（控制台）。发错地址的请求根本到不了 API，因此不会收到我们的任何报错，控制台里也不会有用量记录。
5. **运行右侧冒烟测试。**第一条验证网络连通，第二条验证 key 与余额，花费不到 1 点。

仍然失败？把完整的 HTTP 状态码和 JSON 错误体，连同 key 的前 11 位（`sk-` + 8 位，切勿发送完整 key）发到 support@perseworks.hk。

Request

```
# 1 · 网络连通—无需 key
curl https://api.perseworks.hk/v1/models

# 2 · 验证 key 与余额—最便宜的一次调用
curl https://api.perseworks.hk/v1/chat/completions \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -d '{"model":"seed-2.0-mini","max_tokens":16,
       "messages":[{"role":"user","content":"ping"}]}'
```

响应 — 余额为 0

```
HTTP/1.1 402 Payment Required

{
  "error": {
    "code": "insufficient_points",
    "message": "Not enough points for this request",
    "required": 1060
  }
}
```

鉴权

每个请求都需在 `Authorization` 头中携带 API key：

`Authorization: Bearer sk-…`

key 拥有你整个余额的消费能力。只在服务端使用——不要放进浏览器、移动端 App 或公开仓库。如果泄露，在 API Keys 页吊销并新建一个，吊销立即生效。

每个 key 可选设置限制：

- **允许模型** —— 调用其他模型返回 `403 model_not_allowed`。
- **IP 白名单** —— 匹配的是发起调用那台机器的公网出口 IP（不是内网地址）。其他来源返回 `403 ip_not_allowed`。

`GET /v1/models` 是公开接口，无需 key。

Request

```
curl https://api.perseworks.hk/v1/chat/completions \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -d '{"model":"seed-2.0-mini","messages":
[{"role":"user","content":"ping"}]}'
```

响应 — KEY 无效

```
HTTP/1.1 401 Unauthorized

{
  "error": {
    "code": "invalid_api_key",
    "message": "Unknown or revoked API key"
  }
}
```

错误

Perseworks 使用标准 HTTP 状态码：`2xx` 成功；`4xx` 请求本身有问题（修正后再试）；`5xx` 服务端或上游问题（可退避重试）。所有错误的 JSON 结构一致。

失败的请求不扣费：为其预扣的点数会全部退回。

状态码	错误码	含义
`400`	`invalid_request`	请求体不是合法 JSON，或缺少必填字段 / 取值越界。
`400`	`unknown_model`	模型 id 不存在，或与接口不匹配（例如把视频模型发到 `/chat/completions`）。
`400`	`unsupported_resolution`	该模型不支持此分辨率。
`400`	`invalid_duration`	`durationSec` 超过模型上限。
`400`	`too_many_references` `too_many_reference_videos` `too_many_reference_audios` `too_many_images`	参考素材数量超过模型上限。
`400`	`input_rejected`	上游拒绝了某个输入（例如参考图含真人）。需更换输入，重试无效。不扣费。
`401`	`invalid_api_key`	API key 缺失、格式错误、不存在或已吊销。
`402`	`insufficient_points`	余额不足以预扣本次请求。`required` 为所需点数。充值后重试。
`403`	`model_not_allowed`	该 key 限定了其他模型。

状态码	错误码	含义
`403`	`ip_not_allowed`	该 key 设置了 IP 白名单，请求来源 IP 不在其中。
`404`	`not_found`	任务不存在，或属于其他账号。
`413`	`payload_too_large`	请求体超限（chat 8 MB · video 100 KB）。
`422`	`content_rejected`	提示词未通过内容审核。不扣费。
`429`	`rate_limited`	该 key 请求过于频繁，等待 `Retry-After` 秒后再试。
`500`	`internal_error`	服务端错误，预扣点数会退回。
`502`	`upstream_error`	上游模型服务失败或超时，预扣点数退回，可重试。

属性

`error.code` `string`

机器可读的错误码，见下表。请以此判断分支，不要依赖 message。

`error.message` `string`

人类可读说明，可能调整，请勿解析。

`error.required` `number`

仅 `insufficient_points` 时返回：预扣本次请求所需点数。

错误对象

```
{
  "error": {
    "code": "model_not_allowed",
    "message": "API key not scoped for model \"seedance-2.0\""
  }
}
```

点数与计费

用量从预付点数余额中扣除——**13,800 点 = 1 美元**，所有模型共用一个余额。

响应头 — 成功

```
HTTP/1.1 200 OK
X-Credits-Charged: 1485
X-RateLimit-Limit: 120
X-RateLimit-Remaining: 119
```

1. **预扣。**执行前按本次请求可能的最大费用预扣（chat：输入 + `max_tokens`；视频：`durationSec` × 每秒单价）。余额不足返回 `402`。
2. **结算。**成功后按实际用量扣费——真实 token 数、真实视频时长——多余预扣部分释放。
3. **退还。**任何失败都会全额退回预扣。

成功响应会在 `X-Credits-Charged` 响应头和 `usage.credits` 中报告扣费（单位：点）。各模型实时单价以 `GET /v1/models` 为准，亦可参考价格页。

提示：chat 的 `max_tokens` 设得很大只会抬高预扣，不会多扣费——但余额较低时可能导致 `402`。按实际需要设置即可。

响应 — 余额不足

```
HTTP/1.1 402 Payment Required

{
  "error": {
    "code": "insufficient_points",
    "message": "Not enough points for this request",
    "required": 1060
  }
}
```

限流

限流按 API key 计算，各接口族独立：`/v1/chat`、`/v1/images`、`/v1/video` 各 **每分钟 120 次**（轮询计入 video 额度——每 5–10 秒轮询一次即可，不要死循环）。

每个响应都带 `X-RateLimit-Limit`、`X-RateLimit-Remaining`、`X-RateLimit-Reset`（秒）。超限返回 `429 rate_limited` 并附 `Retry-After` 头——等待相应秒数后重试。

需要更高吞吐？请联系 support@perseworks.hk。

响应 — 被限流

```
HTTP/1.1 429 Too Many Requests
Retry-After: 12
X-RateLimit-Limit: 120
X-RateLimit-Remaining: 0

{ "error": { "code": "rate_limited", "message": "too many requests" } }
```

幂等请求

视频任务成本高，因此 `POST /v1/video/generations` 支持幂等：带上 `Idempotency-Key` 头，用同一个 key 重试会返回*原任务*，不会再新建（并扣费）一个。

请使用随机值（如 UUID v4），每个逻辑任务生成一次，重试时复用——例如网络超时、无法判断第一次提交是否成功时。

Request

```
curl https://api.perseworks.hk/v1/video/generations \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -H "Idempotency-Key: 5f1d6c1e-9b0a-4c36-8d1e-2a7f3c9b4e10" \
  -d '{"model":"seedance-2.0-fast","prompt":"一只猫在海浪上冲浪，电影感","resolution":"720p","durationSec":5}'
```

模型

Model 对象

描述一个可调用的模型及其实时价格。根据 `category` 不同，只会出现其中一种价格字段。

属性

`id` `string`

调用时传入的 `model` 值。

`name` `string`

显示名称。

`category` `enum`

`video` · `image` · `chat`

`status` `string`

可调用模型为 `active`。

`billing_unit` `enum`

`per_second` · `per_image` · `per_1m_token`

`pricing` `array`

仅视频。每个支持的分辨率一项。

▼ pricing ▸

`resolution` `string`

`720p` · `1080p` · `4k`

`points_per_sec` `number`

每输出秒点数。

`usd_per_sec` `number`

对应美元价格。

MODEL 对象


```
{
  "id": "seedance-2.0",
  "name": "Seedance 2.0",
  "category": "video",
  "status": "active",
  "billing_unit": "per_second",
  "pricing": [
    { "resolution": "720p",  "points_per_sec": 2760,
"usd_per_sec": 0.2  },
    { "resolution": "1080p", "points_per_sec": 6210,
"usd_per_sec": 0.45 },
    { "resolution": "4k",    "points_per_sec": 13800,
"usd_per_sec": 1    }
  ]
}
```

image_pricing object

仅图像。

▼ image_pricing ▸

points_per_image number

每张生成图点数。

usd_per_image number

对应美元价格。

extra_input_image_points number

第一张之后每张参考图的附加点数。

token_pricing object

仅对话。

▼ token_pricing ▸

points_per_1m_input number

每百万输入 token 点数。

points_per_1m_output number

每百万输出 token 点数。

points_per_1m_cached_input number

每百万缓存命中输入 token 点数。

context_window number

最大上下文 token 数。

列出模型

`GET` `/v1/models`

返回所有可调用的模型及实时价格。公开接口，无需 API key。当前模型：

模型 ID	类型	说明
`seedance-2.5`	video	旗舰。720p · 4–30 秒 · 文本 / 图片 / 视频 / 音频参考 · 编辑与延长。
`seedance-2.0`	video	720p · 1080p · 4k · 最长 15 秒 · 原生音频。

模型 ID	类型	说明
`seedance-2.0-fast`	video	720p · 出片更快。
`seedance-2.0-mini`	video	720p · 成本最低。
`seedream-5.0`	image	Seedream 5.0 Pro · 文生图 + 图像编辑 · 按张统一计价。
`seed-2.1-pro`	chat	最强对话模型 · 支持识图 · 262K 上下文。
`seed-2.0-mini`	chat	快速、低成本 · 262K 上下文。

返回

`list` 对象，其 `data` 为 Model 对象数组。

```
GET /v1/models
```

```
curl https://api.perseworks.hk/v1/models
```

响应


```
{
  "object": "list",
  "data": [
    { "id": "seedance-2.5", "category": "video", ... },
    { "id": "seedance-2.0", "category": "video", ... },
    { "id": "seedream-5.0", "category": "image",
      "billing_unit": "per_image",
      "image_pricing": { "points_per_image": 1485, "usd_per_image": 0.1076,
                         "extra_input_image_points": 49 } },
    { "id": "seed-2.1-pro", "category": "chat",
      "billing_unit": "per_1m_token",
      "token_pricing": { "points_per_1m_input": 15220,
                         "points_per_1m_output": 76103,
                         "points_per_1m_cached_input": 3044,
                         "context_window": 262144, ... } },
    ...
  ]
}
```

对话

创建对话补全

`POST` `/v1/chat/completions`

为一段对话生成模型回复。请求与响应遵循 OpenAI Chat Completions 格式，OpenAI SDK 可直接使用。

参数

`model` string 必填

`seed-2.1-pro` 或 `seed-2.0-mini`。

`messages` array 必填

对话历史，按时间正序。1–256 条。

▼ messages ▸

`role` enum 必填

`system` · `user` · `assistant`

`content` string | array 必填

文本（最多 128K 字符），或用于发送图片的内容块数组——见发送图片。

`stream` boolean

为 `true` 时以 SSE 流式返回，以 `data: [DONE]` 结束。默认 `false`。

`max_tokens` integer

最多生成的 token 数，1–8192，默认 `1024`。同时决定预扣点数。

`response_format` object

`{"type": "json_object"}` 强制输出合法 JSON；默认 `{"type": "text"}`。

`thinking` object

`{"type": "disabled"}` 关闭深度思考以降低延迟；也可传 `enabled` / `auto`。

返回

`POST /v1/chat/completions`

```
curl https://api.perseworks.hk/v1/chat/completions \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -d '{
    "model": "seed-2.1-pro",
    "messages": [{"role": "user", "content": "你好! "}],
    "stream": true
  }'
```

响应

```
{
  "id": "chatcmpl-x2Lq9ZtR0bWk4mNv8sYe",
  "object": "chat.completion",
  "created": 1785000000,
  "model": "seed-2.1-pro",
  "choices": [{
    "index": 0,
    "message": { "role": "assistant", "content": "你好! 有什么可以帮你? " },
    "finish_reason": "stop"
  }],
  "usage": {
    "prompt_tokens": 9,
    "completion_tokens": 8,
    "total_tokens": 17,
    "credits": 1
  }
}
```

响应 — STREAM: TRUE

```
data:
{"id":"chatcmpl-…","object":"chat.completion.chunk",
       "choices":[{"index":0,"delta":
{"content":"Hi"},"finish_reason":null}]}

data:
{"id":"chatcmpl-…","object":"chat.completion.chunk",
       "choices":[{"index":0,"delta":
{},"finish_reason":"stop"}]}

data: [DONE]
```

`chat.completion` 对象。`stream: true` 时为一串 `chat.completion.chunk` 事件。`usage.credits` 为本次扣除的点数。

发送图片

`POST` `/v1/chat/completions`

`seed-2.1-pro` 可以识图。把 `content` 从字符串改为内容块数组即可。

内容块

`{ type: "text" }` `object`

文本块：`{"type": "text", "text": "…"}`。

`{ type: "image_url" }` `object`

图片块。

▼ `{ type: "image_url" }` ▸

`image_url.url` `string` 必填

HTTPS 地址，或 base64 data URI（`…`）。

`image_url.detail` `enum`

`low` 每张按 512 输入 token 计费；`high`（默认）按 1,600 计费。

返回

限制：每个请求最多 32 张图，每条消息最多 64 个内容块，请求体总计 8 MB。

`POST` `/v1/chat/completions`

```
curl https://api.perseworks.hk/v1/chat/completions \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -d '{
    "model": "seed-2.1-pro",
    "messages": [{
      "role": "user",
      "content": [
        {"type": "text", "text": "这张图里有什么？"},
        {"type": "image_url", "image_url": {"url":
"https://example.com/cat.jpg", "detail": "low"}}
      ]
    }]
  }'
```

图像

生成图像

`POST` `/v1/images/generations`

根据提示词生成图像，可选参考图引导。该接口为**同步**调用，通常耗时 **70–90 秒**——请把 HTTP 客户端超时设为至少 120 秒。（如果你的服务端位于有 ~100 秒限制的代理之后，例如 Cloudflare，请放到后台任务中调用。）

`POST` `/v1/images/generations`

```
curl https://api.perseworks.hk/v1/images/generations \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  --max-time 150 \
  -d '{
    "model": "seedream-5.0",
    "prompt": "小熊猫咖啡师，电影级布光",
    "size": "2K"
  }'
```

参数

model `string` 必填

`seedream-5.0`

prompt `string` 必填

要生成的内容，最多 4,000 字符。

size `enum`

`1K` 或 `2K`，默认 `2K`，同价。

image `string | array`

用于编辑或多图参考的参考图：URL 或 base64，最多 10 张。

n `integer`

生成张数，1–4，默认 `1`，按张计费。

output_format `enum`

`png` · `jpeg`

seed `integer`

固定以获得可复现结果。

返回

返回对象的 `data` 为 `{ url }` 数组。需要长期保留的图片请自行下载存储。

响应

```
{
  "id": "img_4nT8vQ1sXw9cR2kL",
  "created": 1785000000,
  "model": "seedream-5.0",
  "data": [ { "url": "https://.../out.png" } ],
  "usage": { "images": 1, "credits": 1485 }
}
```

视频

Video 任务对象

视频生成为异步。提交后返回 `queued` 状态的任务；轮询直至 `succeeded` 或 `failed`。

属性

id `string`

任务 id，前缀 `vid_`。

成功

```
{
  "id": "vid_8KqZ2mR4tY7wP1xN",
  "status": "succeeded",
  "video_url": "https://.../vid_8KqZ2mR4tY7wP1xN.mp4",
  "thumbnail_url":
"https://.../vid_8KqZ2mR4tY7wP1xN.jpg",
  "duration": 5
}
```

`status` enum

`queued` → `running` → `succeeded` | `failed`

`model` string

提交时返回。

`video_url` string

`succeeded` 时：MP4 地址。请及时下载。

`thumbnail_url` string

`succeeded` 时：封面图。

`duration` number

`succeeded` 时：实际时长（秒）——按此计费。

`error` object

`failed` 时：`{ code, message }`。预扣已退还。

失败

```
{
  "id": "vid_8KqZ2mR4tY7wP1xN",
  "status": "failed",
  "error": { "code": "upstream_failed", "message":
"Video generation failed" }
}
```

创建视频任务

`POST` `/v1/video/generations`

创建视频任务，立即返回 `202 Accepted`。此时按 `durationSec` 预扣点数，完成后按实际时长结算。支持 `Idempotency-Key` 头。

参数

`model` string 必填

`seedance-2.5` · `seedance-2.0` · `seedance-2.0-fast` · `seedance-2.0-mini`

`prompt` string 必填

最多 2,000 字符。

`resolution` enum 必填

`720p` · `1080p` · `4k`（小写）。`1080p` 与 `4k` 仅 `seedance-2.0` 支持；其余模型均为 `720p`。

`POST /v1/video/generations`

```
curl https://api.perseworks.hk/v1/video/generations \
  -H "Authorization: Bearer sk-..." \
  -H "Content-Type: application/json" \
  -d '{
    "model": "seedance-2.0",
    "prompt": "一只猫在海浪上冲浪，电影感",
    "resolution": "1080p",
    "durationSec": 5,
    "aspectRatio": "16:9"
  }'
```

响应

```
HTTP/1.1 202 Accepted

{
  "id": "vid_8KqZ2mR4tY7wP1xN",
  "status": "queued",
  "model": "seedance-2.0"
}
```

durationSec integer 必填

秒。`seedance-2.5`：4–30；其他：1–15。

---

aspectRatio enum

`16:9` · `4:3` · `1:1` · `3:4` · `9:16` · `21:9` · `adaptive`

---

image string

要驱动的图片 URL（图生视频）。

---

firstFrame string

用作首帧的图片 URL。

---

lastFrame string

用作尾帧的图片 URL。

---

references array

参考图 URL。`seedance-2.5`：最多 30；其他：最多 9。

---

referenceVideos array

参考视频 URL。`seedance-2.5`：最多 4；其他：最多 3。

---

referenceAudios array

参考音频 URL，最多 2 个。

---

seed integer

固定以获得可复现结果。

返回

`202`，返回 `queued` 状态的 Video 任务。素材 URL 必须能被上游公网访问。

查询视频任务

`GET /v1/video/generations/{id}`

返回你创建的任务的当前状态。每 5–10 秒轮询一次，直至 `status` 为 `succeeded` 或 `failed`。

```
GET /v1/video/generations/{id}
```

```
curl
https://api.perseworks.hk/v1/video/generations/vid_8Kq
Z2mR4tY7wP1xN \
  -H "Authorization: Bearer sk-..."
```

路径参数

id `string` 必填

创建时返回的 `vid_…` id。

返回

Video 任务。id 不存在或属于其他账号时返回 `404 not_found`。

响应


```
{
  "id": "vid_8KqZ2mR4tY7wP1xN",
  "status": "succeeded",
  "video_url": "https://.../vid_8KqZ2mR4tY7wP1xN.mp4",
  "thumbnail_url":
"https://.../vid_8KqZ2mR4tY7wP1xN.jpg",
  "duration": 5
}
```