

源 1.0 大模型开源开放 FAQ

本文对源 1.0 使用过程中，常见的问题及解释进行了梳理，请各位开发者参考使用，如不能解决请邮件联系：air_service@inspur.com 或大模型小助理微信号：zhoupika。

1. Q：源 1.0 API 是如何收费的？

A：API 服务一直是免费申请的，目前没有收费计划。

2. Q：源 1.0 服务的响应时间有时候不尽如人意，怎么解决？

A：我们正在协调更多的服务器，提供更稳定的 API 服务，同时支持更多客户，响应时间也会得到优化。

3. Q：源 1.0 支持本地化部署（私有化部署）吗？

A：支持本地化部署（私有化部署），如果自有配套硬件环境的话，可以通过与浪潮签署针对源 1.0 模型套件（推理服务）的 POC（Proof of Concept）合作协议，免费试用 3 个月。试用满意后，可联系购买源 1.0 模型套件。

4. Q：如何进行模型在自有数据上的微调？微调模型是需要收费的是吗？

A：目前还不支持在线调用 API 进行微调。您可以向源大模型官方提供数据，经过查验数据质量可用的话，由源大模型官方来进行微调训练得到微调后的模型，微调模型会更加适合不同垂直领域场景下的特定业务。微调的模型同样支持在线 API 服务和本地化部署。微调模型属于定制化服务，具体收费以双方商定为准，详询源大模型小助理。。

5. Q：Github 示例代码中 add_example 有没有容量限制，似乎加到一定程度之后就会报错。

A：输入模型的总 tokens 长度（example 包含在内的 prompt 长度）不可以超过 2048tokens。

6. Q：代码提示“验签失败”。

```
RuntimeWarning: {'errCode': 'IBASE_TOKEN_ERROR', 'errMessage': 'token验签失败', 'flag': False}
```

A：

1) 首先，查看自己的使用权限，源 1.0 提供时限 3 个月，总次数 3 万次的免费权限。

2) 使用权限正常的话，请查看自己是否正确的输入了用户名和手机号。

3) 以上都排除的话, 请查看使用 API 服务的设备时间是否是正确的北京时间。

如以上方法不能解决, 请详询大模型小助理。

7. Q: 使用 API 代码显示无法找到 yuan_api

```
-----  
Traceback (most recent call last):  
  File "dialog.py", line 7, in <module>  
    from yuan_api.inspurai import Yuan, set_yuan_account, Example  
ModuleNotFoundError: No module named 'yuan_api'
```

A: 引用的库层次不对, 可在 yuan_api 同级目录, 使用相对路径调用目标代码, 即可导入 yuan_api 库。

8. Q: 安装 hashlib 库的时候, 报错

```
D:\Yuan-1.0\yuan_api>pip install requests hashlib json  
Collecting requests  
  Using cached requests-2.28.1-py3-none-any.whl (62 kB)  
Collecting hashlib  
  Using cached hashlib-20081119.zip (42 kB)  
  Preparing metadata (setup.py) ... error  
error: subprocess-exited-with-error  
  
× python setup.py egg_info did not run successfully.  
  exit code: 1  
  [7 lines of output]  
    Traceback (most recent call last):  
      File "<string>", line 2, in <module>  
      File "<pip-setuptools-caller>", line 34, in <module>  
      File "C:\Users\Administrator\AppData\Local\Temp\pip-install-uz07_g_h\hashlib_48e7d660be3442458cac24a4dc38eb18\setup.py", line 68  
        print "unknown OS, please update setup.py"  
  
    SyntaxError: Missing parentheses in call to 'print'. Did you mean print("unknown OS, please update setup.py")?  
    [end of output]  
  
note: This error originates from a subprocess, and is likely not a problem with pip.  
error: metadata-generation-failed  
  
× Encountered error while generating package metadata.  
  See above for output.
```

A: 在 pip list 中查看是否已经安装了对应库, 如果当前已经安装了对应库, 就跳过该库的安装。后期可以正常使用源 1.0 API 服务。