

# 操作系统支持矩阵国际化与自动翻译应用

## 基于 GPT 模型的自动国际化生成

测试小队

2024 年 7 月 3 日





- Github: wychlw,



# Support Matrix

- Mainstream operating systems × Mainstream RISC-V development boards
- Installation guide with test results for each combination
- [ruijsdk.org/supported](https://ruijsdk.org/supported)

Mainstream operating systems for RISC-V hardware adaptation (mainstream RISC-V development boards)

CPU	Product number	Arch Linux	Debian/RevyOS	Fedora	FreeBSD	Gentoo	openAnolis	OpenBSD	openCloudOS	openEuler	openKylin	openSUSE	Ubuntu	Tina-Linux	Android 13	Armbian	BuildRoot	OpenHarmony
SG2042	<a href="#">Pioneer Box</a>	N/A	Good	Good	N/A	N/A	N/A	N/A	WIP	Good	Good	N/A	N/A	N/A	N/A	N/A	N/A	WIP
CV1800B	<a href="#">Milk-V Duo (64M)</a>	Basic	Basic	CFH	N/A	N/A	N/A	N/A	N/A	Basic	N/A	N/A	N/A	N/A	N/A	N/A	Basic	N/A
TH1520	<a href="#">LicheePi 4A</a>	Good	Good	Good	N/A	N/A	N/A	N/A	N/A	Good	Good	N/A	WIP	N/A	N/A	Good	N/A	WIP
JH7100	<a href="#">VisionFive</a>	N/A	N/A	Good	N/A	N/A	N/A	Basic	N/A	Good	Good	Basic	Basic	N/A	N/A	Basic	Basic	N/A
JH7110	<a href="#">VisionFive 2</a>	Basic	Good	N/A	WIP	Basic	N/A	Basic	N/A	Good	Good	Basic	Basic	N/A	WIP	Good	Basic	WIP
K230	<a href="#">CamMV K230</a>	N/A	Basic	Basic	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Basic	N/A	N/A	N/A	N/A	N/A

# Support Matrix

- Check compatibility
- All guides in one place
- Provide files and scripts

将源码与工具链解压后放到工作区中。

由于官方示例不带有 Makefile 或任何构建脚本，而是采用官方 IDE 进行构建，若要直接使用工具链，请下载这个改好的 [Makefile](#)，并放到源码文件夹下 `EVB/EXAM/FreeRTOS/FreeRTOS` 下。

将刚才获取的 Makefile 中的工具链路径进行修改，即替换 `TOOL_CHAIN_PATH` 和 `OPENOCD_PATH` 中的内容。

接下来进行 `make prepare` 来复制一些必要的代码。

- Report and fix error

## 更新/引导 boot

若 u-boot 无法引导而是进入了命令行，需要更新 u-boot（详见：（pr 31）[\[starfive-tech/u-boot#31\]](https://github.com/starfive-tech/u-boot#31)

官方文档：<https://starfivetech.com/uploads/VisionFive%20Single%20Board%20Computer%20Quick%20Start%20Guide.pdf>

u-boot 下载：[https://github.com/starfive-tech/Fedora\\_on\\_StarFive/releases](https://github.com/starfive-tech/Fedora_on_StarFive/releases)

或者，手动输入以下命令以引导系统：

```
ext4load mmc 0:1 0x84000000 /boot/Image
ext4load mmc 0:1 0x88000000 /boot/dtb/starfive/jh7100-starfive-visionfive-v1.dtb
setenv bootargs "root=/dev/mmcblk0p1 console=ttyS0,115200n8 console=tty0 earlycon=sbi rootflags=data=writeback stmmaceth=chain_
booti 0x84000000 - 0x88000000
```



# Internationalization And Problems

- 50 development boards now with approximately 3-4 operating system each  
actually: 219 reports in total
- Huge amount of file to be translated: adding a new language is a huge workload without considering the human mistakes
- Have some kind of similarity between the reports, but also lots of variations
- With technical terms and context



# Ideas of Using LLMs

- Good quality <-> Traditional doc translation service
- Fast and cheap <-> Professional translator

- Good quality <-> Traditional doc translation service
- Fast and cheap <-> Professional translator

But need to pay attention...

- Hallucinations
- Secretly change minor details
- Manual review is still needed

# Auto Translation With OpenAI GPT

- Proj: [github.com/wychlw/report-i18n](https://github.com/wychlw/report-i18n)
- Support 3.5 and 4o ( Balance between quality and price )
- – Specialized for test report translation, may need to be adjusted for other purposes

- Automaticlly translate docs into all target languages
- `\%s/README.md/README_(lang).md/` , with link changed inside
- Will extract everyting in codeblock and not translate them
- Will not translate (mostly) URLs

- Tweaked model parameters included
- Config:
  - env.py : API\_KEY and API\_URL  
OpenAI official and FREE API provided by GPT\_API\_free
  - conf.py : Config lang and (part of) promote

- `max_length` : Max length for each iteration, generally token size if logogram else  $\{3, 4\} * \text{tokensize}$
- `target_lang/tips_translated_by_chatgpt` : Language to translate to and the tips
- `model`: 4o if logogram to ideogram (or vise versa), else 3.5
- `template`: !important  
eg: 测试成功 -> Test Passed | Test Successful | Test Succeed | Successful Test...  
To regulate general structure of the translation.



Marker is used to change default behavior (will be removed automatically):

- `r"\[translate\]"` : Force re-translate
- `r"\[update\]"` : Force update translation
- `r"\[skip_lang\[(.*)\]\]"` : Skip translate language
- Any other good ideas?

A template for CI, automatically translate and push the repo.  
(Notice the potential API leak)



Let's now try translating a report to French?  
(Hopefully the network is good enough)

# Thanks for your listening!