

語言支援：字體與文字轉語音

我的部落格現在支援九種語言：日語 (ja)、西班牙語 (es)、印地語 (hi)、中文 (zh)、英語 (en)、法語 (fr)、德語 (de)、阿拉伯語 (ar) 和繁體中文 (hant)。你可以在 <https://lzwjava.github.io> 找到這個網站。

在電腦環境中處理多種語言時，有幾個方面需要考慮。

字型處理

不同的語言需要特定的字型來正確顯示，尤其是在使用 LaTeX 生成 PDF 時。以下 Python 程式碼展示了如何根據作業系統和語言選擇合適的字型：

```
if platform.system() == "Darwin":
    if lang == "hi":
        CJK_FONT = "Kohinoor Devanagari"
    elif lang == "ar":
        CJK_FONT = "Geeza Pro"
    elif lang in ["en", "fr", "de", "es"]:
        CJK_FONT = "Helvetica"
    elif lang == "zh":
        CJK_FONT = "PingFang SC"
    elif lang == "hant":
        CJK_FONT = "PingFang TC"
    elif lang == "ja":
        CJK_FONT = "Hiragino Sans"
    else:
        CJK_FONT = "Arial Unicode MS"
else:
    if lang == "hi":
        CJK_FONT = "Noto Sans Devanagari"
    elif lang == "ar":
        CJK_FONT = "Noto Naskh Arabic"
    elif lang in ["en", "fr", "de", "es"]:
        CJK_FONT = "DejaVu Sans"
    elif lang == "zh":
        CJK_FONT = "Noto Sans CJK SC"
    elif lang == "hant":
        CJK_FONT = "Noto Sans CJK TC"
    elif lang == "ja":
        CJK_FONT = "Noto Sans CJK JP"
    else:
```

```

CJK_FONT = "Noto Sans"

command = [
    'pandoc',
    input_markdown_path,
    '-o', output_pdf_path,
    '-f', 'markdown',
    '--pdf-engine', 'xelatex',
    '-V', f'romanfont={CJK_FONT}',
    '-V', f'mainfont={CJK_FONT}',
    '-V', f'CJKmainfont={CJK_FONT}',
    '-V', f'CJKsansfont={CJK_FONT}',
    '-V', f'CJKmonofont={CJK_FONT}',
    '-V', f'geometry:{GEOMETRY}',
    '-V', 'classoption=16pt',
    '-V', 'CJKoptions=Scale=1.1',
    '-V', 'linestretch=1.5'
]

```

需要注意的是，這個解決方案並不完美。例如，程式碼塊註解中的印地語文字可能無法如預期般顯示。

文字轉語音

我使用 Google 文字轉語音來生成部落格文章的語音版本。以下程式碼片段展示了我如何為文字轉語音引擎選擇合適的語言代碼：

```

if filename.endswith('-zh.md'):
    language_code = "cmn-CN"
    voice_language_code = "cmn-CN"
else:
    language_code = "en-US"
    voice_language_code = "en-US"

text_to_speech(
    text=article_text,
    output_filename=output_filename,
    task=task,
    language_code=language_code,
    dry_run=dry_run,
    progress=progress
)

```

目前，語音生成僅針對中文和英文內容。要擴展到其他語言，必須配置相應的語言代碼。