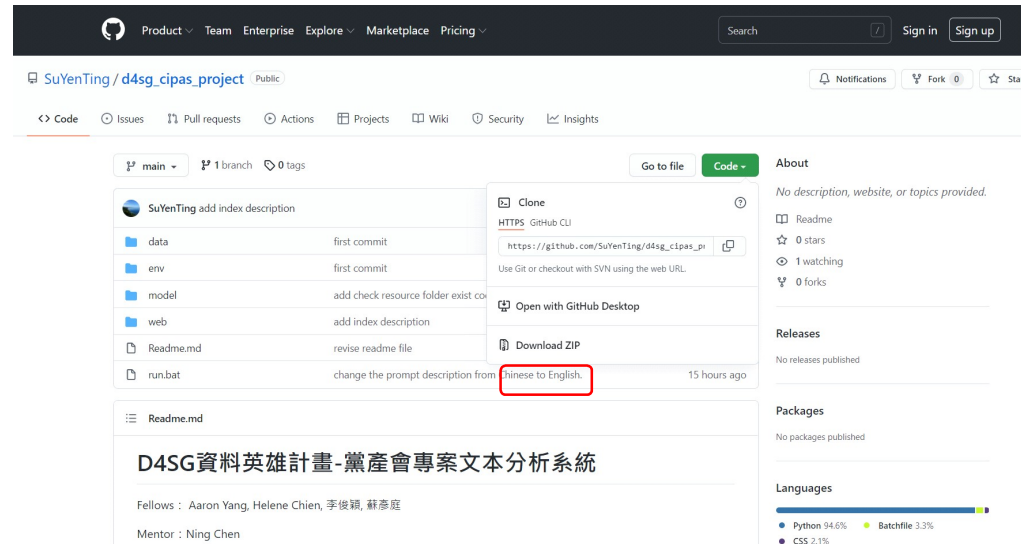
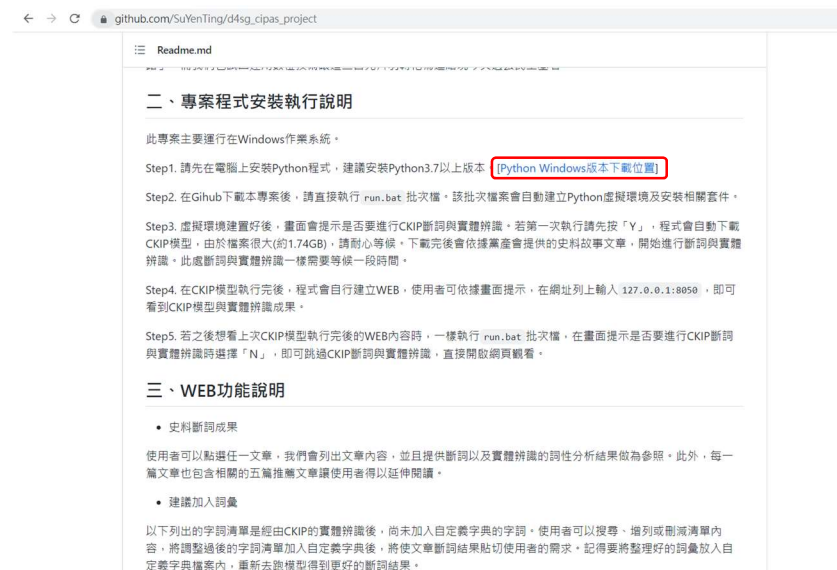


黨產會專案文本分析系統安裝說明

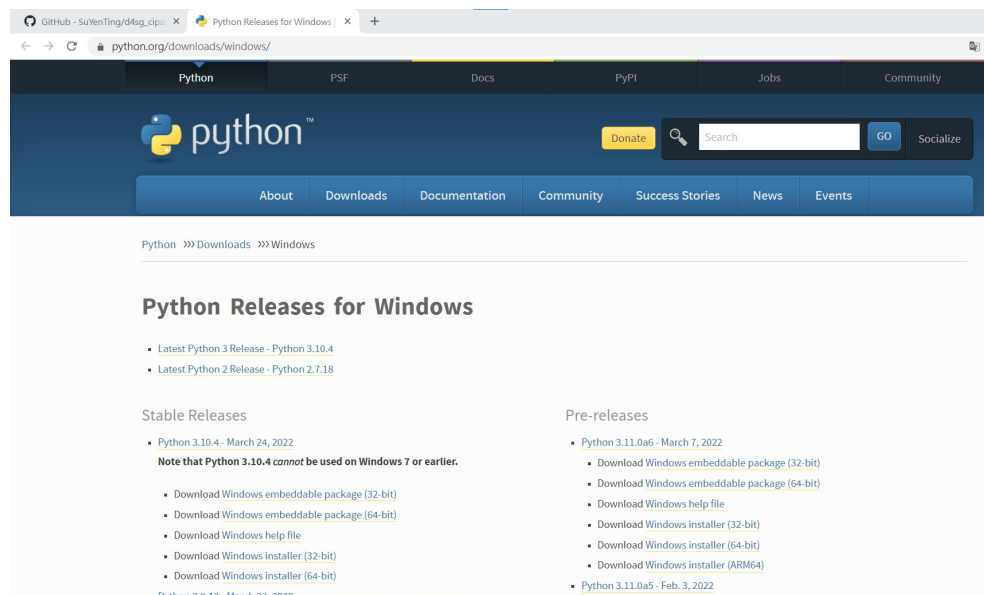
Step1. 連上：https://github.com/SuYenTing/d4sg_cipas_project，點擊 Code(綠色)，選擇 Download ZIP，下載主程式。



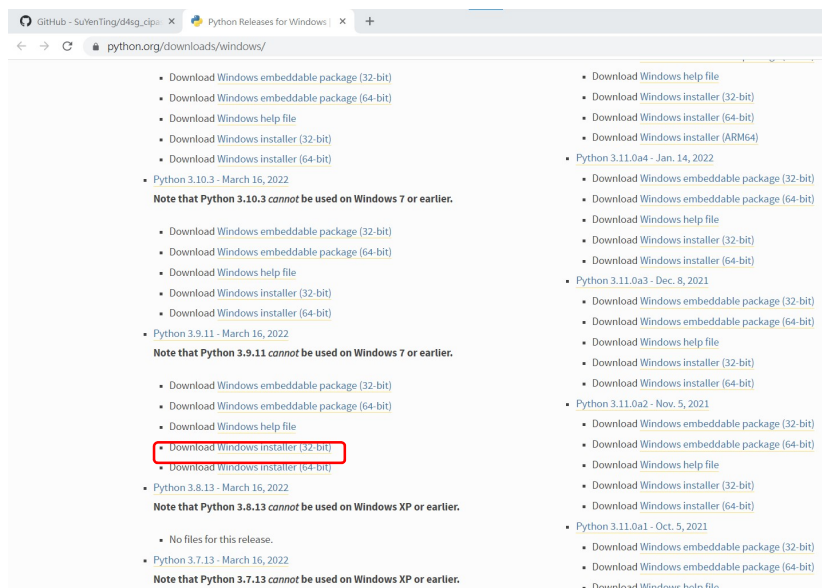
Step2. 網頁下拉，找到並點擊 Python Windows 版本下載位置。



Step3. 進入 Python 程式語言頁面，網頁下拉。



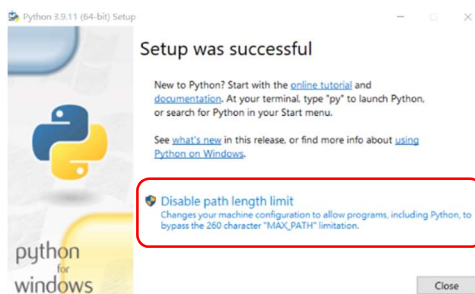
Step4. 找到 Python 3.9 的段落點擊下載 Windows installer (64-bit) 。



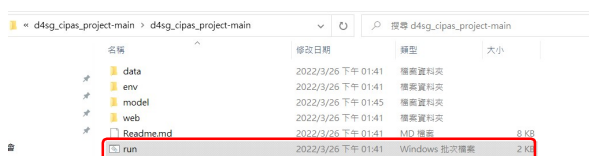
Step5. 下載後，點擊 Python 安裝，請勾選 Add Python 3.9 to PATH，再 Install Now。



Step6. 到達步驟顯示 Setup was successful 時，點擊 Disable path length limit，按是，再結束。



Step7. 解壓縮並點擊 d4sg_cipas_project-main 資料夾，決定好安裝的地方，例如：桌面（之後請不要移動資料夾），點擊 run 進行安裝，如果安裝過程中不幸中斷，請刪除這個資料夾，重新下載後再次安裝。



Step8. 安裝過程較為費時，除需要下載 CKIP 預訓練模型外，也還要對每篇文本進行斷詞與實體辨識，請耐心等待，出現以下畫面請輸入 Y，並繼續等待程式將所有文本辨識完畢。

```
C:\WINDOWS\system32\cmd.exe
Created wheel for gdown: filename=gdown-4.4.0-py3-none-any.whl size=14775 sha256=a45fba237fa24e8025c604eef807949556803
5a5488cae5e28083b2796e94d23
Stored in directory: c:\users\asusi7c\appdata\local\pip\cache\wheels\7d\37\b6\b2a79c75e898c0b8e46ff255102602d715
9a10d9af0d80641a
Successfully built gdown
Installing collected packages: tf-estimator-nightly, termcolor, tensorboard-plugin-wit, pytz, pyasn1, libclang, keras, f
latbuffers, dash-table, dash-html-components, dash-core-components, certifi, cached-property, Brotli, zipp, xlrd, wrapt,
wheel, Werkzeug, urllib3, typing-extensions, tensorflow-io-gcs-fsfilesystem, tensorboard-data-server, tenacity, soupsieve
six, rsa, PySocks, pyasn1-modules, protobuf, oauthlib, numpy, networkx, MarkupSafe, itsdangerous, idna, gast, filelock
et-xmlfile, colorama, ckiptagger, charset-normalizer, cachetools, tqdm, scipy, requests, python-dateutil, plotly, opti
einsum, openpyxl, Keras-Preprocessing, Jinja2, importlib-metadata, h5py, grpcio, google-pasta, google-auth, click, beaut
ifulsoup4, astunparse, absl-py, requests-oauthlib, pandas, Markdown, Flask, google-auth-oauthlib, gdown, Flask-Compress,
tensorboard, dash, tensorflow, dash-cytoscape, dash-bootstrap-components
Running setup.py install for termcolor ... done
Successfully installed Brotli-1.0.9 Flask-Compress-1.11 Jinja2-3.0.3 Keras-Preprocessing-1.1.2 Markdown-3.3.
6 MarkupSafe-2.1.1 PySocks-1.7.1 Werkzeug-2.0.3 absl-py-1.0.0 astunparse-1.6.3 beautifulsoup4-4.10.0 cached-property-1.5
.2 cachetools-5.0.0 certifi-2021.10.8 charset-normalizer-2.0.12 ckiptagger-0.2.1 click-8.0.4 colorama-0.4.4 dash-2.3.0 d
ash-bootstrap-components-1.0.3 dash-core-components-2.0.0 dash-cytoscape-0.3.0 dash-html-components-2.0.0 dash-table-5.0
.0 et-xmlfile-1.1.0 filelock-3.6.0 flatbuffers-2.0 gast-0.5.3 gdown-4.4.0 google-auth-2.6.2 google-auth-oauthlib-0.4.6 g
oogle-pasta-0.2.0 grpcio-1.44.0 h5py-3.6.0 idna-3.3 importlib-metadata-4.11.3 itsdangerous-2.1.1 keras-2.8.0 libclang-13
.0.0 networkx-2.6.3 numpy-1.21.5 oauthlib-3.2.0 openpyxl-3.0.9 opt-einsum-3.3.0 pandas-1.3.5 plotly-5.6.0 protobuf-3.19.
4 pyasn1-0.4.8 pyasn1-modules-0.2.8 python-dateutil-2.8.2 pytz-2022.1 requests-2.27.1 requests-oauthlib-1.3.1 rsa-4.8 sc
ipy-1.7.3 six-1.16.0 soupsieve-2.3.1 tenacity-8.0.1 tensorboard-2.8.0 tensorboard-data-server-0.6.1 tensorboard-plugin-w
it-1.8.1 tensorflow-2.8.0 tensorflow-io-gcs-fsfilesystem-0.24.0 termcolor-1.1.0 tf-estimator-nightly-2.8.0.dev2021122109 t
qdm-4.65.0 typing-extensions-4.1.1 urllib3-1.26.9 wheel-0.37.1 wrapt-1.14.0 xlrd-2.0.1 zipp-3.7.0
Python packages are all installed!

Python virtual environment setup has been completed.

Do you want to run CKIP word segmentation and name entity recognition? [Yes/No] [Y.N]?
```

Step9. 出現以下畫面就表示程式已成功對文本進行斷詞與實體辨識，並且將分析結果整理好。接著打開瀏覽器，輸入 <http://127.0.0.1:8050/> 即可使用，使用網頁時，請不要關閉此程式。

```
C:\WINDOWS\system32\cmd.exe
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 288 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 289 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 290 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 291 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 292 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 293 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 294 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 295 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 296 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 297 / 298
目前程式正在執行CKIP斷詞與實體辨識 執行進度: 298 / 298
CKIP斷詞與實體辨識已執行完畢!
檔案已儲存: ./web/resource/20220326140625.pickle
Start building a WEB environment...

正在載入CKIP斷詞模型結果...
正在整理自定義字詞與內置辨識詞彙...
正在計算字詞相似度...
網站相關資料已載入完畢，請在網址列輸入 http://127.0.0.1:8050/ 即可使用
使用網站期間請勿關閉此視窗！
Dash is running on http://0.0.0.0:8050/

* Serving Flask app 'index' (lazy loading)
* Environment: production
* WARNING: This is a development server. Do not use it in a production deployment.
  Use a production WSGI server instead.
* Debug mode: off
* Running on all addresses.
* WARNING: This is a development server. Do not use it in a production deployment.
* Running on http://192.168.0.152:8050/ (Press CTRL+C to quit)
```

Step10. 之後若要重新開啟網頁，點擊 run.bat 即可。如果不需要再次對文本進行斷詞與實體辨識，請輸入 N，再輸入網址即可使用。如果有調整文本或自定義字典，請輸入 Y，讓模型能夠重新對文本進行斷詞與實體辨識。

```
C:\WINDOWS\system32\cmd.exe
Active code page: 65001
Check whether the current folder has a virtual environment installed.
A virtual environment is currently installed.
Do you want to run CKIP word segmentation and name entity recognition? [Yes/No] [Y,N]N
Skip CKIP segmentation and entity recognition.
Start building a WEB environment...
正在輸入CKIP斷詞模型結果...
正在整理自定義字典內實體辨識詞彙...
正在計算文章相似度...
請於指定資料已輸入完畢，請在網址列輸入 http://127.0.0.1:8050/ 即可使用
使用網路連線請勿關閉此視窗！
Dash is running on http://0.0.0.0:8050/
* Serving Flask app 'index' (lazy loading)
* Environment: production
* WARNING: This is a development server. Do not use it in a production deployment.
  Use a production WSGI server instead.
* Debug mode: off
* Running on all addresses.
* WARNING: This is a development server. Do not use it in a production deployment.
* Running on http://192.168.0.132:8050/ (Press CTRL+C to quit)
```