**Python Dcard Image Crawler**

**作者： BadBug**

**檔案說明：**

**分為兩部分**

* **exe執行檔資料夾**
* **.exe執行檔路徑:**

**10444207\_Python期末專案/exe執行檔/dist/CYCU\_final.exe**

* **Dcard\_Image\_Crawler.py 原始碼**
* **Python 原始碼，可透過python直譯執行**
* **！！！注意！！！**

**要執行此原始碼檔前請確定：**

1. **透過pip安裝requests套件**

* *pip install requests*

1. **透過pip安裝beautifulsoup4套件**

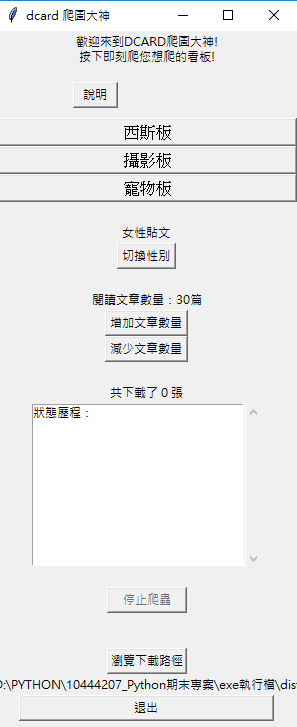
* *pip install beautifulsoup4*
* **DEMO\_output**

**>範例輸出( 以寵物版、女性貼文、30篇文章數為範例 )**

**pet 資料夾內含url\_information圖片來源文章網址資料**

**程式描述：**

**完整程式樣貌**

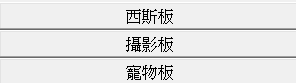


* **說明按鈕**



**按下後會開啟README.txt檔案給使用者閱讀說明**

* **下載鍵欄位**

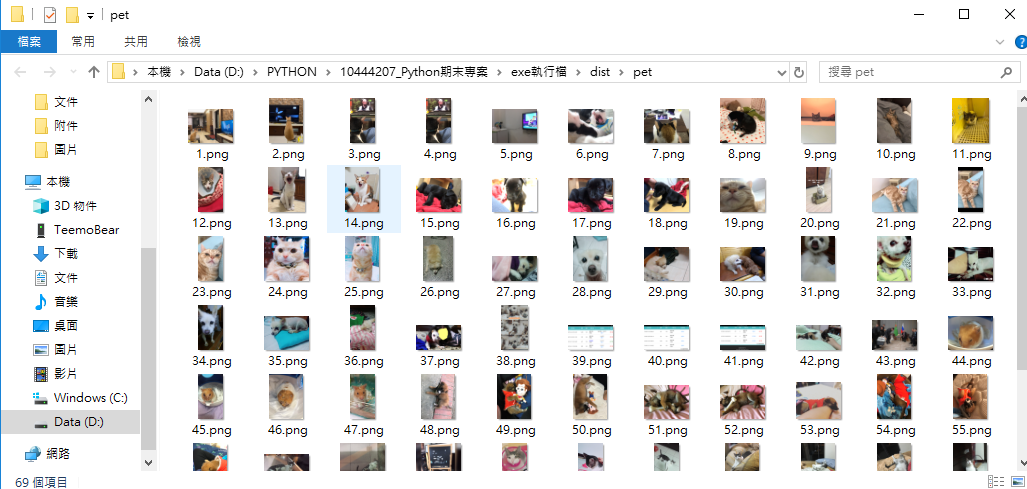


按此三個鈕其中之一程式立刻爬所選擇看板中文章的圖片。

若須自行選擇爬圖條件可以在下方眾多設定欄位做調整，調整完再回來按下載鈕即可。

**下載中會跳出>**

下載的圖片資料夾，直接讓使用者觀看



**下載完會跳出>**

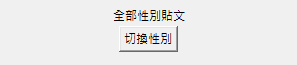
1. url\_information.txt 紀錄所下載的圖片文章來源，讓使用者可以查看



1. 工作完成的提醒按鈕。

*p.s. 若點選下載後所有設定及選擇路徑及下載按鈕會轉換成disabled狀態*

* **切換性別欄位**

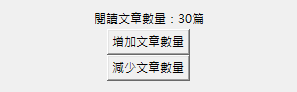


**按切換性別即可切換想爬圖的發文者性別，共有三種狀態：**

1. **2. 3.**



* **選擇爬圖文章數欄位**



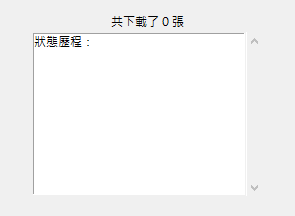
因應使用者需求，預設(最低)的爬文文章數量為30

可按”增加文章數量”來增加所要爬的文章數量(每次增加30篇)

可按”減少文章數量”來減少所要爬的文章數量(每次減少30篇)

若要減少到低於預設的30篇數量的話便會跳出警告視窗。

* **狀態歷程欄位**



* 共下載 x 張欄位 ：計算每按一次下載的圖片數量
* 狀態歷程：會記錄

1. 創建資料夾與否，若有則顯示Already have folder

若沒有則顯示 Making pic folder。

1. 使者變更的下載路徑
2. 現階段爬蟲的狀態，若持續下載中則狀態會以

動態… 來顯示



1. 錯誤狀況

* **停止爬蟲按鈕**



**預設關閉，一旦開始下載後便會轉換成可以點選的狀態**

**作用是若使用者覺得此次爬蟲過於漫長，可以直接停止該次下載。**

**並且記錄歷程會記錄爬蟲錯誤，提早停止。**

* **瀏覽下載路徑以及路徑顯示**



**預設下載路徑為使用者執行檔案時的路徑**

* **退出按鈕**



不論是下載中還是任何狀態按下即可強制退出。

**Functions()**

**類別Download裡面所擁有的函數：**

**\* DownloadProcess():顯示在歷程記錄中的…動畫(讓…像跑進度條一樣)。**

**\* RunProcess(): 同上。**

**類別MyWin裡所擁有的函數：**

**\* UnableAllButton(): 使下載中的設定，下載紐都不能作用(應該是 disalbed英文不好請見諒XD)。**

**\* EnableAllButton(): 非下載狀態時讓設定，下載按鈕可以作用，停止爬蟲按鈕不能作用。**

**\* CrawlChild(): 創建爬蟲的子執行緒，不然會造成主UI卡當。**

**\* ClickAction(): 開始爬蟲的函數，會傳入設定的參數給CrawlChild()。**

**\* ClickLeftChangeSex(): 對應切換性別按鈕。**

**\* ClickPlusReadNum(): 對應增加文章數量按鈕。**

**\* ClickMinusReadNum(): 對應減少文章數量按鈕。**

**\* InsertMsgBox(): 把歷程狀態的文字輸入進歷程狀態的紀錄框裡。**

**\* DeleteMsgBox(): 同上，只是是用來移除的。**

**\* StopAll(): 給退出按鈕用的，讓整個程式停止並跳出。**

**\* SelectPath(): 設定下載檔案路徑**

**\* StopCrawl(): 停止爬蟲**

**\* OpenReadME(): 對應說明按鈕，跳出說明文件**

**\* \_\_init\_\_(): 視窗初始化**

**\* MakeFolder():創建文件夾(下載資料夾)**

**\* Crawl(): 核心爬蟲程式**

**\* main(): 主程式**