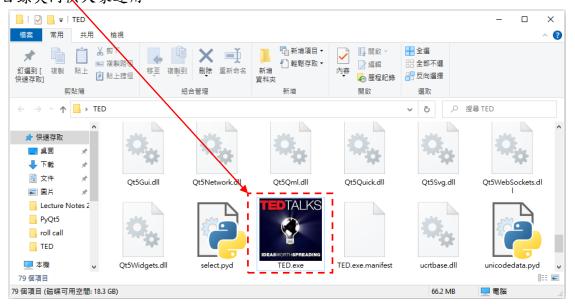
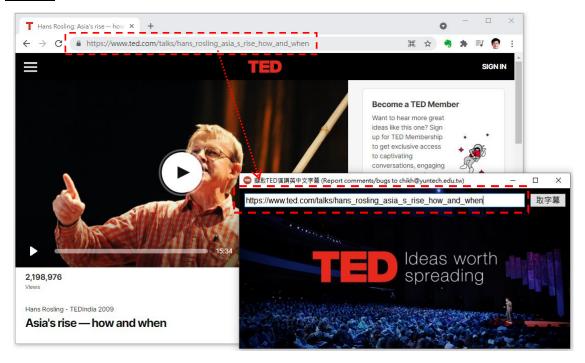
取得 TED 演講字幕檔、英中文並陳的方法 (紀光輝, 2024/5/28)

TED (Technology Entertainment and Design)網站是紀老師非常推薦的優質平台,這個平台把全世界特定領域表現傑出的人士邀來作演講,大部分的講演僅十多分鐘,符合一般人注意力集中的時間長度,內容精采豐富,同學可觀摩這些優秀心智的思維以及創意表現的方式,這些演講本身也是學習英語文的良好媒介。假設同學欲下載 TED 網站某場演講的影音檔與字幕作為研習的素材,可透過底下步驟完成。

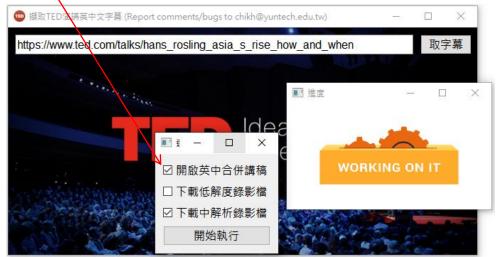
1. 準備:先<u>下載並安裝 VLC 媒體撥放器</u>,再連線至 <u>https://bit.ly/3DIMYUt</u>,下載並解壓縮至選定的目錄夾(目錄夾位置可任選)。解壓縮後的目錄夾內含多份檔案,只需以滑鼠點擊啟動其中的 TED.exe 執行檔。該執行檔由紀老師以 Python+PyQt5 套件撰寫編譯而成,原始碼 TED.py亦在目錄夾內供大家運用。

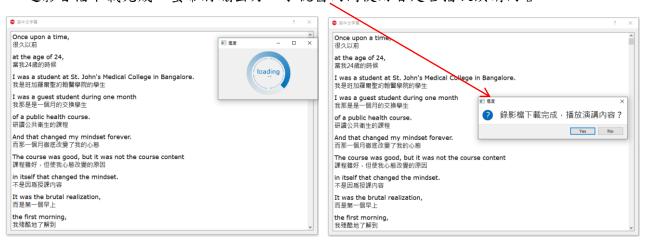


2. 取得演講超連結:以網頁瀏覽器查看演講所在網址,複製該網址並貼入TED.exe的空白欄中,再按取字幕鈕,程式將自動連線至TED伺服器抓取所需資料。



3. 螢幕將蹦出二個小視窗,一為顯示進度的動畫視窗,另一則供同學勾選所需功能(此例勾選「開 啟英中合併講稿」及「下載中解析度錄影檔」),再按開始執行鈕,隨後將依勾選的項目進行 處理。中高解析度錄影檔經常超過120M,受制於網路速度,偶需等候數分鐘方下載完成。







倘同學選擇播放演講,結果畫面將如上所示,原文講詞與中譯文字被配置在同一視窗,演講播放器則配置另一個視窗。播放演講的視窗內建「滑動條」(slider)、「重播」、「暫停」、「加速」、「減速」等功能。拖曳滑動條可令影片從某時刻播放,隨選操控不受限制。比較特別的是,同學閱讀講稿過程中,若對某句的英文詞彙不熟悉,想聽聽講者如何發音,可直接以滑鼠連續點擊二次(double click)該詞彙或以滑鼠選擇單行內某一區段,程式將自動進入「跟隨模式」,亦即將跳轉影音檔到該詞彙所在位置並播放該片段。按重播鈕可重聽該詞的發音,若聽不清楚,可先按減速鍵再重播,將有助於同學記憶該詞的英文唸法,加深印象、達到學習效果。拉動滑動條即可脫離跟隨模式,回復到一般播放的狀態。演講播放當中按鍵盤右方向鍵(→)可令講演往前跳轉約10秒鐘再繼續撥放;按左方向鍵(←)將倒退10秒鐘再恢復播放。

程式產出的英中文字幕合併檔與下載的影音檔,存放於電腦的「下載」(Downloads)目錄夾之中,隨時可複習取用。字幕合併檔為純文字的.txt 檔案,可以「記事本」或任何文書軟體開啟。

後記

- TED 或因網路安全考量,每隔一二年改版一次,每次改頭換面經常調整素材位於內部伺服器的存取方式。字幕檔則從之前的.srt 格式改為現今 json 格式,我摸索許久才理出大概。這套程式係依據現行 TED 維運架構而寫,未來 TED 若再改版或許便不再完全適用。
- 本程式自動判斷 TED 網頁描述可下載的演講影音檔案解析度有哪些選擇,自動整理成一份選單讓使用者勾選。TED 內部可能有作流量管控,多數演講提供低中高三種解析度,但部分熱門演講卻只提供低或中解析度的檔案可下載。
- 本程式適用於 TED 平台,部分較新的演講可能尚未有中文翻譯字幕,這個情形請瞭解。另提醒大家,英國國家廣播公司(BBC)亦有許多優質的影音出版品,甚至可直接在 YouTube 發現大量 BBC 富知識性的紀錄片,取得這類字幕需以其他方式進行,但尚未摸索出簡易的方法。
- 因可用的時間有限,在此只實作一個版本,程式功能尚有優化的空間。可延伸的功能包括取得其他語系(譬如日語、德語)的講詞作為研習素材,或容許使用者以滑鼠揀擇超過單行的文句再重播該段落的演講內容等。
- 以上說明希對大家有幫助。