第十組-反洗錢設計期末專題

- ●謝東景
- ●謝孟均
- ●吳承翰

目 录 CONTENTS

 01
 工作進度大綱

 分析流程
 02

 03
 討論結果

 成果展示
 04



工作進度大綱

工作完成情况 Job completion

1.以關鍵字搜尋,爬下相關新聞資料	0
2.用NER格式標記完成後上傳至MySQL	0
3.下載各組的標記資料,以ML方式訓練至NER	O
4. 擷取新聞中的人名, 若不具完整人名, 則捨棄該文章	O
5. 以Tree分析犯罪關鍵字與人名的距離	O
6.建立出犯罪文章及犯罪者人名的資料	0
7.寫網頁串接MySQL	O
8.Final Demo	0



分析流程 & 討論結里

Google API進行新聞爬蟲

- 1. 以作業2的爬蟲程式進行新聞資料擷取,找出特定行為之潛在犯罪人,
 - 風險評估報告之關鍵字
 - 程式碼連結
- 2. 擷取結果如下,連結

```
news, keywords, title, link, content, content on
appledaily,人口販運,台灣捲入跨國人口販運風暴 | 蘋果新聞網 | 蘋果日報,http://www.appledaily.com.tw/appledaily/article/headline/201
appledaily,人口販運,内政部:今年1到9月查緝人口販運被害人安置印尼人最多 | 蘋果新聞網 ..., http://www.appledaily.com.tw/realtimenews/
appledaily,人口販運,打擊人口販運台成立防制人口販運國際工作坊 | 蘋果日報,http://www.appledaily.com.tw/realtimenews/article/new/20
appledaily,人口販運,赴美留學成奴隸!洛城破獲大宗人口販賣 | 蘋果新聞網 | 蘋果日報https://,https://www.appledaily.com.tw/realtimenews
appledaily,人口販運,美人口販運報告台列最佳等級、中國最差 | 蘋果新聞網 | 蘋果日報, http://www.appledaily.com.tw/realtimenews/article/
appledaily,人口販運,非典型人口販運怎麼查|蘋果新開網|蘋果日報,http://www.appledaily.com.tw/realtimenews/article/new/20170223/
appledaily,人口販運,打擊人口販運法務部研議被害人遠距視訊| 蘋果日報,http://www.appledaily.com.tw/realtimenews/article/politics/
appledaily,人口販運,AIT也說讚! 台灣打擊人口販運有世界級標準| 蘋果日報https://,https://www.appledaily.com.tw/realtimenews/arti
appledaily,人口販運,美人口販運報告台連續5年列最佳| 蘋果日報,http://www.appledaily.com.tw/realtimenews/article/new/20140620/42
appledaily,人口販運,美人口販運防制報告台灣連8年第一級評等 | 蘋果新聞網 | 蘋果日報, http://www.appledaily.com.tw/realtimenews/article
appledaily,人口販運,防制人口販運台灣不及格| 蘋果日報,http://www.appledaily.com.tw/appledaily/article/headline/20061117/30378
appledaily,人口販運,防制人口販運我成效列最優| 蘋果日報,http://www.appledaily.com.tw/appledaily/article/headline/20100615/325
appledaily,人口販運,史上最大宗17越童販運來台 | 蘋果新聞網 | 蘋果日報,http://www.appledaily.com.tw/appledaily/article/headline/20
appledaily,人口販運,美防制人口販運有功獎項首度有台灣人獲獎| 蘋果日報,http://www.appledaily.com.tw/realtimenews/article/new/2017
appledaily,人口販運,涉虐外籍漁工、漁獲回報不實福甡11號調查結果明出爐|蘋果新聞網 ...,http://www.appledaily.com.tw/realtimenews/a
```

洗錢防制法

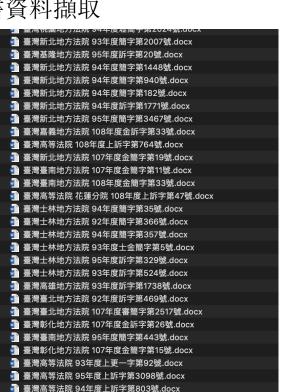
Selenium下載判決書

1. 以作業2的爬蟲程式進行判決書資料擷取

搜尋 (1)



全部 法規 裁判 函釋 決議 法律問題座談 訴願決定書 起訴書 大法官解釋



以NER標記人名

- 1. 先透過人工標記,以Rule based的方式訓練NER模型
- 2. 將爬下來的文章用訓練好的模型進行人名標記

	news	keywords	title	link	content	content_cn	PER
0	appledaily	人口販運	台灣捲入跨國人□販運風 暴 蘋果新聞網 蘋果日 報	http://www.appledaily.com.tw/appledaily/articl	《報導者》獨家授權記者 /《報導者》蔣宜婷、李 雪莉 台北-柬埔寨報導台 灣已捲入遠洋漁業的人口 販	《报导者》独家授权记者 /《报导者》蒋宜婷、李 雪莉 台北-柬埔寨报导台 湾已卷入远洋渔业的人口 贩	蒋宜婷,李雪莉
1	appledaily	人口販運	內政部:今年1到9月查緝 人口販運被害人安置印尼 人最多 蘋果新聞網	http://www.appledaily.com.tw/realtimenews/arti	內政部公佈統計數據指 出,今年1到9月司法及警 察機關查緝人口販運案件 共116件,75%為性剝削	内政部公布统计数据指 出,今年1到9月司法及警 察机关查缉人口贩运案件 共116件,75%为性剥削	器文船
2	appledaily	人口販運	打擊人口販運台成立防制 人口販運國際工作坊 錆 果日報	http://www.appledaily.com.tw/realtimenews/arti	內政部「2018防制人口販 運國際工作坊」今(25) 日盛大揭幕,邀請副總統 陳建仁、美國在台協會	内政部「2018防制人口贩 运国际工作坊」今(25) 日盛大揭幕,邀请副总统 陈建仁、美国在台协会	陈建仁,谷立言,陈 建仁,谷立言,罗秉 成,徐国 勇,Jean,Patrick Taran,吴玉琴
3	appledaily	人口販運	赴美留學成奴隸!洛城破 獲大宗人口販賣 蘋果新 聞網 蘋果日報https://	https://www.appledaily.com.tw/realtimenews/art	(新增配音影片)洛縣人口販賣執行組周二宣佈, 協破一宗全國性人口販賣 案,3日的掃蕩行動中拘 捕5	(新增配音影片)洛县人 口贩卖执行组周二宣布, 捣破一宗全国性人口贩卖 案,3日的扫荡行动中拘捕 5	伍国庆,张紫茵,洛 杉矶

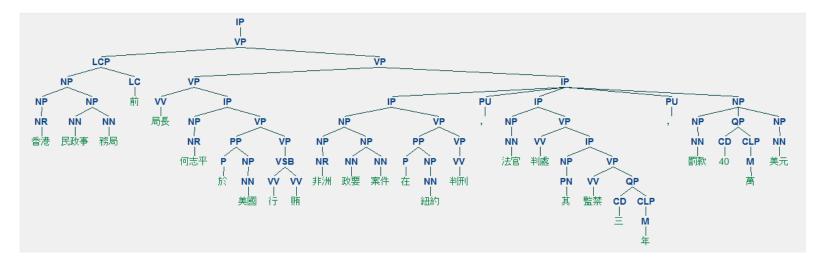
找出犯罪嫌疑人及其資訊

- 1. 從文章中出現過的所有人名找出犯罪嫌疑人
 - 找出人名前後的字詞,與自行定義的犯罪關鍵詞清單進行比對。
 - 一 將用此方法找到的人名認定為犯罪嫌疑程度高
 - 計算人名與犯罪罪名之間的最短距離。
 - 一 將距離罪名最近的人名認定為犯罪嫌疑程度低

	name	crime	crime_risk	suspect_probability	key_sentence	news	news_title	other_suspect	occupation	age
6107	賴清 徳	證券 犯罪	非常高	低	行政院今(2)日會同洗錢防制辦公室、 法務部及金管會,公布首部國家洗錢及資 恐風險評估評估報告,	chinatimes	《金融》台洗錢風險評估,8類型犯罪具非常高度威脅- 財經- 時報資訊	NaN	院長	NaN
6108	邱太	證券 犯罪	非常高	低	邱太三說,目前台灣深受洗錢威脅程度, 也有「低、中、高、非常高」4個等級, 其中毒品販運、詐欺、	chinatimes	洗錢防制OBU、國銀最 大漏洞- 中時電子報	NaN	法務部長	NaN
6130	邱垂 貞	貪汙 賄賂	非常高	高	檢調在2002年底接獲檢舉,指本案有多位 立委收錢護航修法,但直到特值組於2007 年4月成立後	appledaily	中藥商涉行賄立委案審9 年最高院又發回 蘋果新 聞網 蘋果日報	赿	前立委	64

找出犯罪嫌疑人及其資訊

- 2. 找出犯罪嫌疑人的個人資訊
 - 以NLTK文法樹結構找出與人名相連的形容詞子句,與定義的身分或職業 清單進行比對,取得身分資訊
 - 從人名前後查找與年齡資訊有關的敘述





成果展示

Demo 呈現

• Github: https://github.com/DennyHsieh/fintech-assignments/tree/master/final_demo

Setup Setup on Python 3.6. SQLite Create the database manually. python setup_db.py Then, Run Server python demo.py

Will run a server listening http://localhost:5000

Demo 呈現-首頁



首頁

反洗錢防治資料庫

請輸入想查詢的人名

Demo 呈現-搜尋結果

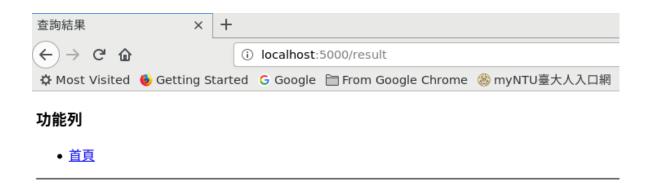


反洗錢防治資料庫

查詢結果

ID 姓名	年齢	職稱	關鍵字	犯罪風險等級	疑似犯罪內文	新聞標題	新聞連結	其他可能嫌 疑者	嫌疑可能性
59 張樵寬	34		人口販運	低	南檢今天偵結,主嫌張樵寬、任志偉及居中介紹印尼女賣 淫的王嘉琪、莊南希、胡嘉益、林漢錡、洪有仁、蔡佩諭 等8人,被依觸犯《人口販運防制法》、《圖利強制使人為 性交猥褻罪》等罪起訴。	女子賣淫檢方	<u>link</u>		高

Demo 呈現-搜尋結果(無犯罪紀錄或無新聞資料)



反洗錢防治資料庫

找不到人名: '我是誰'

Thanks