

# 科技部專題研究計畫申請書

申請條碼：104WFA0250423

## 一、基本資料：



計畫類別(單選)		一般型研究計畫		
研究型別		個別型計畫		
計畫歸屬		人文司		
申請機構/系所(單位)		國立政治大學資訊科學系		
本計畫主持人姓名		陳恭	職稱	教授
		身分證號碼		M12010****
本計畫名稱	中文	跨平台社群媒體巨量資料蒐集與分析		
	英文	Cross-Platform Massive Data Collection for Social Media		
整合型總計畫名稱				
整合型總計畫主持人				身分證號碼
全程執行期限		自民國 104 年 09 月 01 日起至民國 105 年 02 月 29 日		
研究學門	學門代碼		學門名稱	
	H55		網路社群巨量資料研究	
研究性質		<input type="checkbox"/> 純基礎研究 <input checked="" type="checkbox"/> 導向性基礎研究 <input type="checkbox"/> 應用研究 <input type="checkbox"/> 技術發展		
<b>【請考量己身負荷，申請適量計畫】</b> 本年度申請主持科技部各類研究計畫(含預核案)共 <u>1</u> 件。(共同主持之計畫不予計入)				
本計畫是否為國際合作研究： <input checked="" type="checkbox"/> 否； <input type="checkbox"/> 是				
本計畫是否申請海洋研究船： <input checked="" type="checkbox"/> 否； <input type="checkbox"/> 是，請務必填寫表C014。				
1. 本計畫是否有進行下列實驗/研究：(勾選下列任一項，須附相關實驗/研究同意文件) <input type="checkbox"/> 人體試驗/人體檢體 <input type="checkbox"/> 人類胚胎/人類胚胎幹細胞 <input type="checkbox"/> 基因重組實驗 <input type="checkbox"/> 基因轉殖田間試驗 <input type="checkbox"/> 第二級以上感染性生物材料 <input type="checkbox"/> 動物實驗(須同時加附動物實驗倫理3R說明) 2. 本計畫是否為人文司人類研究(非生物醫學研究)計畫： <input type="checkbox"/> 是(請檢附已送人類研究倫理審查之證明文件)； <input checked="" type="checkbox"/> 否				
計畫連絡人		姓名： <u>陳恭</u> 電話：(公) <u>02-29387170</u> (宅/手機) <u>0920467590</u>		
通訊地址		台北市指南路二段64號 政治大學資訊科學系		
傳真號碼		02-22341494	E-MAIL	chenk@cs.nccu.edu.tw

計畫主持人簽章：\_\_\_\_\_

日期：\_\_\_\_\_

## 二、申請補助經費：

- (一) 請將本計畫申請書之第四項(表C004)、第五項(表C005)、第六項(表C015)、第七項(表C006)、第八項(表C007)、第九項(表C008)所列費用個別加總後，分別填入「研究人力費」、「耗材、物品、圖書及雜項費用」、「國外學者來臺費用」、「研究設備費」、「國外差旅費-執行國際合作與移地研究」及「國外差旅費-出席國際學術會議」欄內。
- (二) 管理費為申請機構配合執行本計畫所需之費用，其計算方式係依本部規定核給補助管理費之項目費用總和及各申請機構管理費補助比例計算後直接產生，計畫主持人不須填寫「管理費」欄。
- (三) 「貴重儀器中心使用額度」係將第十項(表C009)所列使用費用合計數填入。
- (四) 請依各年度申請博士後研究之名額填入下表，如於申請時一併提出「補助延攬博士後研究員額/人才進用申請書」(表CIF2101、CIF2102)，若計畫核定僅核定名額者，計畫主持人應於提出合適人選後，另依據本部訂頒之補助延攬客座科技人才作業要點規定向本部提出進用申請，經審查通過後，始得進用該名博士後研究。
- (五) 申請機構或其他單位(含產業界)提供之配合項目，請檢附相關證明文件。

金額單位：新台幣元

執行年次 補助項目		第一年 (104年9月 ~105年2月)	第二年	第三年	第四年	第五年
業 務 費		476,496				
研究人力費		410,496				
耗材、物品、圖書及雜項費用		66,000				
國外學者來臺費用		0				
研 究 設 備 費		244,390				
國 外 差 旅 費		0				
執行國際合作與移地研究		0				
出席國際學術會議		0				
管 理 費		108,133				
合 計		829,019				
貴重儀器中心使用額度		0				
博士後研究	國內、外區	共 0 名	共 _____ 名	共 _____ 名	共 _____ 名	共 _____ 名
	大陸地區	共 0 名	共 _____ 名	共 _____ 名	共 _____ 名	共 _____ 名
申請機構或其他單位(含產業界)提供之配合項目(無配合補助項目者免填)						
配合單位名稱		配合補助項目	配合補助金額	配合年次	證明文件	

### 三、主要研究人力：

(一) 請依照「主持人」、「共同主持人」、「協同研究人員」及「博士後研究」等類別之順序分別填寫。

類別	姓名	服務機構/系所	職稱	在本研究計畫內擔任之具體工作性質、項目及範圍	*每週平均投入工作時數比率(%)
主持人	陳恭	國立政治大學資訊科學系	教授	整體系統架構之規劃與專案控管	25%
共同主持人	陳百齡	國立政治大學新聞學系	副教授	資料蒐集範圍與分析議題	25%
共同主持人	鄭宇君	玄奘大學大眾傳播學系	助理教授	資料蒐集範圍與分析議題	25%

※ 註：每週平均投入工作時數比率係填寫每人每週平均投入本計畫工作時數佔其每週全部工作時間之比率，以百分比表示（例如：50%即表示該研究人員每週投入本計畫研究工作之時數佔其每週全部工時之百分五十）。

(二) 如申請博士後研究，請另填表CIF2101及CIF2102(若已有人選者，請務必填註人選姓名，並將其個人資料表(表C301～表C303)併同本計畫書送本部)。

#### 四、研究人力費：

- (一) 類別/級別欄請依專任助理(含碩士、學士、三專、五(二)專及高中職)、兼任助理(含博士生、碩士生、大專學生、講師及助教)及臨時工等填寫。
- (二) 專任助理每月工作酬金請參考本部補助專題研究計畫專任助理人員工作酬金參考表填寫；兼任助理每月工作酬金，不得超過本部補助專題研究計畫兼任助理人員工作酬金支給標準表之規定。
- (三) 申請專任助理者，除依工作月數填列工作酬金及至多1.5個月年終工作獎金外，須另填列投保勞保及健保之「雇主應負擔之勞、健保費」（於線上填列工作酬金時，系統會自動列入勞、健保費）。
- (四) 請分年列述。

第 1 年

金額單位：新台幣元

(一) 專任助理						
類別/級別	人數	姓 名	工 作 月 數	月支酬金 (含勞健保費)	小計	請述明：1. 最高學歷 2. 曾擔任專題研究計畫專任助理之經歷 3. 在本計畫內擔任之具體工作性質、項目及範圍
碩士級專任助理第四年	1	黃智彬	6.75	38,840	262,170	1. 碩士 2. "尋找媒體創用者的數位足跡：社交媒體研究之新方法探索－社交媒體重大事件之傳播模式比較分析：鉅量資料數位人文分析取徑"之專任助理 3. 協助建置社群媒體資料蒐集平台。
分擔勞健保費(碩士級專任助理第四年)	1	黃智彬	6	4,721	28,326	碩士級專任助理第四年 4721元 x 6月 x 1 名
合 計 (一)					290,496	
(二) 講師及助教級兼任助理						
類別/級別	人數 (1)	每人每月 酬金(2)	獎助月數 (3)	小計 (4)＝ (1)×(2)×(3)	在本研究計畫內擔任之具體工作性質、項目及範圍	
合 計 (二)				0		
(三) 博士班研究生、碩士班研究生及大專學生兼任助理						
級別/姓名	人數 (5)	每人每月 單元數(6)	獎助月數 (7)	小計 (8)＝ \$ 2000×(5)×(6)×(7)	在本研究計畫內擔任之具體工作性質、項目及範圍	
碩士班研究生研究助學金/王佑霖、李欣穎	2	4	6	96,000	協助蒐集社群媒體資料之程式撰寫、文獻檢閱、資料分析等。	
合 計 (三)				96,000		

(四) 臨時工		
臨時工資	小計	在本研究計畫內擔任之具體工作性質、項目及範圍並請敘明約用人數、按日/按時計酬數額
24,000	24,000	大學部學生協助社群媒體資料蒐集與分析，共計200小時，時薪120。
合 計 (四)	24,000	
總計 (五) =合計(一)+(二)+(三)+(四)	410,496	

五、耗材、物品、圖書及雜項費用：

- (一) 凡執行研究計畫所需之耗材、物品(非屬研究設備者)、圖書及雜項費用，均可填入本表內。
- (二) 說明欄請就該項目之規格、用途等相關資料詳細填寫，以利審查。
- (三) 若申請單位有配合款，請於備註欄註明。
- (四) 請分年列述。

第 1 年 金額單位：新台幣元

項 目 名 稱	說 明	單 位	數 量	單 價	金 額	備 註
電腦使用費	租用Amazon Web Service雲端伺服器及計算資源，計算與分析社交媒體大量資料	月	6	3,000	18,000	
消耗性器材	雷射印表機碳粉匣	個	1	6,000	6,000	
消耗性器材	隨身碟、滑鼠等電腦周邊耗材	批	6	2,000	12,000	
專家諮詢費	邀請社群媒體資料蒐集與分析專家提供諮詢建議	人次	6	2,000	12,000	
雜支	影印、郵寄、文具、紙張、會議餐費等雜支。	批	6	2,000	12,000	
圖書經費	採購社群媒體資料分析之相關圖書或研究報告	本	6	1,000	6,000	
合 計					66,000	

## 七、研究設備費：

- (一) 凡執行研究計畫所需單價在新台幣一萬元以上且使用年限在二年以上與研究計畫直接有關之各項設備屬之。各類研究設備金額請於金額欄內分別列出小計金額。
- (二) 購置設備單價在新台幣二十萬元以上者，須檢附估價單。
- (三) 若申請機構及其他機構有提供配合款，請務必註明提供配合款之機構及金額。
- (四) 儀器設備單價超過六十萬元(含)以上者，請詳述本項設備之規格與功能(諸如靈敏度、精確度…等)，其他重要特性與重要附件，以及申購本設備對計畫執行之必要性。本項設備若獲補助，主持人應負維護保養之責，並且在不妨礙個人研究計畫或研究群計畫之工作下，同意提供他人共同使用，以避免設備閒置。
- (五) 計畫主持人執行本項研究計畫，如欲申請本部補助單價新台幣一千萬元(含)以上之大型儀器，請填表C006-1。該項設備單獨核給一個規劃計畫。
- (六) 請分年列述。

第 1 年

金額單位：新台幣元

類別	設備名稱 (中文/英文)	說明	數量	單價	金額	經費來源	
						本部補助 經費需求	提供配合款之機 構名稱及金額
儀器及資 訊設備	2U超薄型伺服器 一組/Octo-core Intel Xeon E5系 列2.0 GH2二顆	蒐集與儲存大量社 群媒體資料之用途	1	244,390	244,390	244,390	
合 計					244,390	244,390	



# 商丞科技股份有限公司

Unifosa Corp.

TEL:(02)29148001 FAX:(02)29147975

## 報價單

業務代表 : 柯志偉 (0963058730)

TEL:(02)2914-8001 Ext:373

客戶名稱	政治大學-資科系	報價日期	104.06.18
聯絡人	楊昇財 先生	報價單號	CTC DellR730 1040618
聯絡電話	02-2939-3091-62276	交易幣別	新台幣
FAX/E-Mail		付款條件	月結30天
報價期限	一個月	營業稅(5%)	

項次	貨品名稱	貨品規格	數量	單價	總價
		<b>台銀採購系統標 第四組 第30項次</b>			
		<b>Dell EdgeR730</b>			
1		<b>2U超薄型伺服器(Octo-Core Intel Xeon E5系列2.0GHz二顆)(無作業系統)</b>	1	143,390	143,390
		Intel Xeon E5-2630 v3 2.4GHz *2	1		
		8GB RDIMM, 2133MT/s, 雙陣列,*1			
		300GB 10K RPM SAS 6Gbps 2.5in Hot-plug Hard Drive*4			
		Dual, Hot-plug, Redundant Power Supply (1+1), 750W			
	<b>增購項</b>				
2		8GB RDIMM, 2133MT/s, 雙陣列, x8 Data Width	1	9,000	9,000
3		300GB 10K SAS HDD 升級為1.2T SAS 10K HDD	2	12,000	24,000
4		1.2TB 10K RPM SAS 6Gbps 2.5 吋熱插拔硬碟	4	17,000	68,000
		(以下空白)			

送貨地點:

附註:

總計(含稅): 244,390



1. 本報價單經客戶確認簽名並蓋章回傳後, 視同正式訂單。
2. 客戶若未按時給付任一筆貨款, 全部未到期貨款及票據自遲延之日起均視同已到期, 客戶應立即全部清償。
3. 客戶尚未付清貨款或票據未獲兌現前, 本公司仍保有已交付貨品之所有權, 客戶同意本公司有權以任何方式取回貨品, 絕無異議並願提供必要之協助。
4. 雙方合意以台灣台北地方法院為第一審管轄法院。

客戶確認簽名並蓋章

F-SA0301-1-B



## 十二、研究計畫中英文摘要：請就本計畫要點作一概述，並依本計畫性質自訂關鍵詞。

### (一) 計畫中文摘要。(五百字以內)

#### 跨平台社群媒體巨量資料蒐集與分析

##### 摘要

針對人文及社會科學領域之研究需求，本計畫結合資訊科學專家與傳播學者的領域知識，設計一個友善的跨平台社群媒體資料蒐集與分析系統，以便人文社會學者進行巨量資料(**big data**)取徑研究。此系統具有以下幾個目的：一、研究者可針對研究議題屬性為在地性或跨國性，選擇台灣最多人使用的 Facebook 或跨國訊息流動快速的 Twitter，或是二種平台並行，做為資料蒐集來源；二、針對特定議題，本系統可有效地幫助研究者收集社群媒體上公開的原始資料(**raw data**)，供研究者下載進行深度分析，以維持資料蒐集與分析流程的透明性；三、提供友善的資料查詢界面與基礎統計功能，幫助研究者快速檢視所蒐集資料狀況，以便於第一時間調整資料蒐集條件。我們希望透過此一系統的建置，可以減少目前巨量資料研究中，人文社會學者與資料專家的資料分析落差。

關鍵字：社群媒體、巨量資料、跨平台資料蒐集系統、臉書、推特

**十二、研究計畫中英文摘要：**請就本計畫要點作一概述，並依本計畫性質自訂關鍵詞。

(二) 計畫英文摘要。(五百字以內)

### **Cross-Platform Massive Data Collection for Social Media**

#### **Abstract**

This project aims to design a cross-platform social media data collector for the academia. The system will be designed based on the following objectives: (1) this system shall help the researchers to collect data from Facebook and Twitter for the agenda locally and internationally; (2) this system shall help the researchers to acquire the raw data and to keep the data processing as transparent as possible; (3) The system shall provide a user-friendly interface for viewing the basic statistics of the collected datasets, so the researchers are able to scan the data profiles and to make better query decisions.

The tentative output of this project will include: (1) a prototype of the data collector system, (2) an analysis of the operation and management mechanisms required for a practical data collector system.

**Keywords:** Social media, Big data, Cross-platform data collection, Facebook, Twitter

### 十三、研究計畫內容（以中文或英文撰寫）：

（一）近五年之研究計畫內容與主要研究成果說明。（連續性計畫應同時檢附上年度研究進度報告）

#### 一、水火計畫概述與研究團隊組成源起

水火計畫團隊成立自 2009 年底，研究範疇從網路媒體對災難訊息的傳佈，擴展至全球社交媒體華人社群所關注的重大事件，試圖描繪快速變遷中的社群媒體角色和議題如何轉變。由於網路上鉅量資訊（big data）的蒐集與分析，需要資訊科技的協助，發展出一套蒐集與處理資訊作業流程。因此，本團隊從成立之初，便是由傳播學者與資科專家組成的跨領域團隊，主要參與成員為：陳百齡（政治大學新聞系副教授，同時為本團隊召集人）、鄭宇君（玄奘大學大眾傳播學系助理教授）、陳恭（政治大學資訊科學系教授）、李蔡彥（政治大學資訊科學系教授）。

本團隊的成立源起於政治大學與澳洲昆士蘭科技大學（Queensland University of Technology, Australia，以下簡稱 QUT）合作研究計畫之先導研究計畫。雙方的研究團隊決定以 2009 年台澳兩國發生的重大天然災難事件為先導個案，分析災難時期之公眾傳播過程與動態變化，自 2010 年 1 月起展開為期三年的國際合作研究計畫。由於雙方分別研究澳洲火災與台灣水災，因此以水火計畫（Flood and Fire）做為此一國際合作團隊的名稱，後來並發展為國科會專題研究計畫（2011/08-2013/07）。

#### 二、階段研究進展

本研究計畫的先前知識積累，從 2009 年迄今已經有六年時間，其間轉折大致可以分為三個階段。

##### 1. 第一階段（2009-2011）：災難情境下的資訊流動

本團隊在此階段以「災難情境下浮現的媒體」為題，探討莫拉克風災個案中的災情網站文本分析，觀察風災當下資訊流通現象。該計畫研究成果可歸納為三部份：（1）莫拉克風災期間地方災防單位的通報資料，發現災難資訊流動有「瞬間鉅量」現象，也就是災難資訊流量驟然增加和擴散。（2）在風災這種高度不確定性情境下，人們使用不同類型的社交媒體和自發性災情通報平台替代既有傳媒做為溝通管道；（3）群眾在災難初期的大量訊息傳佈，主要內容類型為：告知、求助、轉述、討論、抱怨、協調、發佈情境資訊、募集物資人力等，資訊流動從仰賴大眾傳媒等垂直媒體轉往水平式媒體。

本階段所使用的資料主要由災情通報平台的管理者提供，加上對專家的深度訪談，本階段的發展重點在訓練研究人員處理大量資料的分析能力，包括：（1）透過實作，建立跨領域團隊的共識和默契；（2）建立資料處理標準作業程序（SOP），從資料篩選、清理、分析到視覺化處理；（2）建立人工內容分析的輔助平台；（4）建立災難及社交媒體研究資料的知識管理平台，由團隊成員定期閱讀文獻並製作摘要，並將資料上傳網站與公眾分享。

##### 2. 第二階段（2011-2013）：重大公眾事件情境下的社群媒體使用

本團隊開始聚焦於資料蒐集，同時蒐集對象也轉向重大公眾事件個案。在這個階段，我們針對社群媒體 Twitter 設計資料蒐集機制，Twitter 是國際間公眾事件討論的主要平台，也是兩岸三地華人社群共同使用的社群媒體，可以從中發覺兩岸三地的訊息交流情況。這段期間最主要的公眾事件是 2012 年臺灣總統大選，我們蒐集選舉期間兩岸三地的 Twitter 社群貼文資料進行資料分析。研究發現：（1）從時間序列分析發現，討論高峰在選舉日前後；（3）語系分析可做為區分語言次群體的指標；（2）Twitter

作為全球社交媒體平台，但不同語言社群受在地脈絡影響，有其意見領袖和關注議題；(4) 傳統新聞媒體透過用戶引述，鑲嵌在社群媒體的使用之中。

在本階段，研究團隊起初是自行撰寫程式透過 **Twitter API** 蒐集資料，後來改用開源軟體，在此過程中學習到許多鉅量資料處理的經驗，包括：(1) 社群媒體資料蒐集是動態過程，資料蒐集必須因應持續變動的情勢，調整資料儲存和供應的方式，例如，社群媒體的開發者介面 **API** 經常改版，研究者必須隨之調整程式。(2) 社群媒體的資料清理、分析和視覺化，為前後相關連的工作，特別是資料清理約佔 80% 工作內容，這部分必須仰賴跨領域合作，由傳播和資訊科學背景研究人員根據經驗積累、共同判讀如何進行資料清洗工作。

### 3. 在第三階段 (2013-迄今)：社群媒體傳播模式的比較

2013 年間，政治大學傳播學院和資訊科學系擴大合作研究規模共組團隊，促成雙方申請科技部數位人文整合型計畫〈尋找媒體創用者的數位足跡：社交媒體研究之新方法探索〉，並和其它研究團隊成員在政大研究總中心共同成立「數位足跡探索實驗室」。從鉅量資料數位人文分析取徑出發，發展資料分析工具應用於人文社會領域的可能模式。

本團隊在這階段的工作為：(1) 利用 **Twitter API** 蒐集和儲存在地重大公眾事件的 **Tweet** 資料（例如，2014 年的 318 反服貿事件、香港雨傘運動、復興航空澎湖與南港墜機事件等），以做為後續資料分析之儲備；(2) 測試不同社群媒體（如：**Facebook**, **PTT**）的資料蒐集機制，並發展出相對應的資料分析模式，以及(3) 強化與國際其他社交媒體研究者交流經驗，初步達成未來交換社群媒體（微博 **Weibo**）的蒐集資料。

歷經此三階段，我們的研究經驗歸納可為以下三點：(1) 社群媒體為一傘狀概念 (**Umbrella concept**)，個別社群媒體平台的設計理念和資料結構不盡相同，因此資料蒐集的條件與模式亦有很大差異；(2) 社群媒體平台 (**Twitter**, **Facebook**) 以資料為其營運模式核心，儘管它們開放 **API** 供外界下載公開資料，但也設定資料蒐集條件之限制（例如：**Facebook** 限制每次撈取留言的上限為 25 則、**Twitter** 使用關鍵字查詢所得資料上限為 1%），因此學術界蒐集社群媒體資料也只能符合此遊戲規則進行，除非直接向 **Twitter**, **Facebook** 購買才能取得完整資料，但購買資料價格高昂，非學術單位所能負荷；(3) 學術界對社群媒體資料的需求，不同於企業或政府。企業或政府對於社群輿論的需要，追求快速比精準重要；但學術界更重視資料透明度與研究的信效度，因此希望獲取完整的原始資料，以使研究人員能夠掌握與調整資料的分析方式，以確保研究品質的信度與效度。

### 4. 小結：團隊知識積累和傳承

本研究團隊自 2009 年成立迄今所整理的文獻（包括所有過程中團隊產出的文摘、論文和簡報）均收藏於網路平台，以備學界同儕取用參考。目前研究團隊共架設三個網站，皆以「水火計畫」為名，茲分述如下：

- (1) 計畫入口網站：包含計畫緣起、重要參考書目，以及部分研究成果。（網址：<https://sites.google.com/a/newliteracies.co.cc/floodfire/>）
- (2) 資料分析網站：主要用於線上內容編碼分析，以及資料視覺化處理之用。（網址：<http://station.ninthday.info/designING/WFProject/>）

- (3) 研究產出網站：團隊收錄的相關文獻、論述摘要、工作論文和電腦簡報等。（網址：<https://sites.google.com/a/newliteracies.co.cc/floodfiredata/>）

上述第一個網站開放供外界查詢，後二者網站主要提供團隊協作平台使用，需用帳號和密碼登入。登入帳號：[floodfire2013@gmail.com](mailto:floodfire2013@gmail.com)，密碼：2013floodfire

另一方面，本團隊也強調研究經驗傳承，主要表現在「課程開設」和「助理訓練」。

- (1) 開設研究所課程：本團隊同時致力將研究轉化為教學資源。政大傳院將在 2013 年春季起開設「社群媒體概論」、「社群媒體資料分析」跨領域課程，此一課程由傳播學院和資科系合力開設，已經進入第三年，由多位團隊成員教師聯合授課，修課學生來自新聞、廣告（傳播學院）、資訊科學（理學院）、以及資訊管理（商學院）等領域。教學方法強調「做中學」，教學資源主要來自本團隊的研究成果累積。
- (2) 助理訓練：研究助理透過內部和外部工作坊學習新知識。由於新技術不斷崛起，本團隊鼓勵並補助研究助理參與校內外各種相關工作坊（例如，2015 年政大 DSP 冬令營等）學習新知，並要求助理將所修習或研發的技術或研究心得透過內部工作坊傳遞給同儕（例如，Gephi 軟體工作坊）。

### 三、跨國合作研究（2009-2015）：

本計畫團隊與澳大利亞昆士蘭科技大學創意產業學院暨卓越創意產業與創新中心的 Axel Bruns 教授所帶領的團隊從 2010 年迄今已累積多年的國際合作基礎。Bruns 教授是全球社交媒體研究的領導人物之一，他是最早提出 Prodisage 概念的學者，近年致力於政治傳播、Twitter 研究，結合傳播研究與資訊技術發展社交媒體鉅量資料的新分析方法。

Bruns 帶領的 QUT 研究團隊具有社交媒體鉅量資料的豐富研究經驗，雙方過去的國際合作交流包括二次國際合作研究工作坊（2010 年澳洲布里斯本；2011 在台北）、兩次在國際會議合組論壇（2010 年在香港，2014 年在澳洲布里斯本 ICA 亞洲年會）、共同發表一篇期刊論文（2012），並獲邀參與社交媒體與政治之書籍篇章撰寫（預計 2015 年下半年出版）。

前述合作源起於 2010 年開始的台澳雙邊合作計畫，目前雙方在社交媒體研究領域都有初步研究成果與專長：澳洲團隊擅長結合資訊技術進行鉅量資料的分析與比較，我方團隊則偏重結合社會文化脈絡因素來詮釋不同社群在社交媒體上的使用行為，特別是中文 Twitter 資料分析，由於中文與英文語系的特性不同，造就英文與中文 Twitter 社群有不同的社交媒體使用習慣。雖然我們受限於經費和技術，目前無法從事即時資料分析，但本團隊目前跨出的這一步，應屬國內傳播學界前沿研究。我們希望未來能發展出跨國社交媒體研究的比較分析架構，進一步在同樣框架下進行各國社交媒體使用者的跨國比較研究。

本研究團隊 2014 年的合作重心鎖定在東亞地區，開始尋找華文學術圈的其他研究團隊合作，我們在 2014 年主要進行的跨國合作研究，包括：一、澳洲雪梨大學聶依文（Prof. Joyce Nip）教授、香港大學傅景華（Prof. King-Wa Fu）合作，香港大學的微博視野計畫（<http://weiboscope.jmsc.hku.hk/>），具有對於新浪微博資料蒐集與分析的專業，本團隊與香港大學的合作，研究焦點放在兩岸三地的重要事件，包括台灣的太陽花學運、香港的雨傘運動等，比較中、港、台的網路公眾在 Twitter、Facebook、

Weibo 上對於同一事件的討論，並將進行跨社交媒體平台、跨語言社群的比較分析，此一研究計畫已獲得蔣經國基金會二年期的研究計畫補助（計畫執行期間：2015.07-2017.06）；二、本計畫澳門易研網的張榮顯博士進行教學合作，讓個別研究生或教師透過易研網（<http://www.esocialhub.org/#/home>）進行 Twitter, Weibo, Youtube 等小規模的資料收集與視覺化分析。

#### 四、小結：

我們體認到即時性社交媒體對於災難事件、社會運動及政治傳播都有重大影響，這是當代學界的重要議題，值得進行長期追蹤探索，本團隊五年多以來持續探討社群媒體上的災難傳播、政治傳播等事件，已經累積一定資料蒐集和呈現的知識基礎，我們希望將此經驗分享給其他研究團隊，建立一個可長久有效利用的社交媒體資料收集平台。

#### 五、團隊研究成果：

##### • 期刊論文

1. 陳百齡、鄭宇君（2014）。〈從流通到聚合：重大災難期間浮現的資訊頻道〉。《新聞學研究》。121：89-125。
2. 鄭宇君、陳百齡（2014）。〈探索 2012 台灣總統大選之社交媒體浮現社群：鉅量資料分析取徑〉。《新聞學研究》，120: 121-146。
3. Cheng, Y.-C. & Chen, P.-L. (2014). Global social media, local context: A case study of Chinese-language tweets about the 2012 presidential election in Taiwan. *Aslib Journal of Information Management*, 66 (3). 342-356. (SCI & SSCI)
4. 鄭宇君（2014）。〈災難傳播中的群體力量：社交媒體促成新型態的公民參與〉，《傳播與社會學刊》，27：179-205。
5. 鄭宇君（2014）。〈向運算轉：新媒體研究與資科技術結合的契機與挑戰〉，《傳播研究與實踐》，7：45-61
6. Bruns, A & Liang, Y. E. (梁煜銜) (2012). Tools and methods for capturing twitter data during natural disasters. *First Monday*, 17(4)
7. 鄭宇君、陳百齡（2012）。〈溝通不確定性：探索社交媒體在災難事件中的角色〉，《中華傳播學刊》，21：119-153。

##### • 專書論文篇章

1. Cheng, Y.-C. & Chen, P.-L. (in press). Interactions between different language communities on Twitter during the 2012 presidential election in Taiwan. In Axel Bruns, Gunn Enli, Eli Skogerbo, Anders Olof Larsson, & Christian Christensen (Eds.), *The Routledge Companion to Social Media and Politics*. (預計 2015 出版)
2. 鄭宇君、陳百齡（2013）。〈超越在地脈絡的全球社交媒體：以 2012 年台灣總統大選的中文 Twitter 討論社群為例〉，吳玫（編）《全球化下的社交媒體與數字網絡：2012 中國網絡傳播學年會澳門國際會議論文選集》，頁 161-179。澳門，澳門大學出版中心。ISBN:978-99965-1-0519



• 會議論文

1. Cheng, Y.-C. (2015.07). Analysis of news citations on Twitter during the 2012 Taiwanese presidential election. Paper will be presented at 2015 IAMCR Conference, Montreal, Canada. July 13-17, 2015.
2. 盧安邦、鄭宇君 (2015.07)。〈用方法說故事：議題框架分析方法綜探〉。中華傳播學會 2015 年會，高雄市，義守大學，2015 年 7 月 3-5 日。
3. 鄭宇君 (2015.06)。〈鉅量資料時代下的使用者研究〉。元智大學「大數據、新媒體、使用者」研討會。台北市，台大集思會議中心，2015 年 6 月 12 日。
4. 潘伯彥、陳恭 (2015.05)。〈臉書使用者行為資料蒐集系統之設計與實作〉。ICIM2015 第 26 屆國際資訊管理學術研討會，台北：大同大學，2015 年 5 月 23 日。
5. 鄭宇君、施旭峰 (2014.12)。〈探索 2012 台灣總統大選社交媒體之新聞來源引用〉。2014 第五屆數位典藏與數位人文國際研討會。台北市，中央研究院人文社會科學館國際會議廳，2014 年 12 月 1-2 日。
6. 吳君孝、李蔡彥、鄭宇君、陳百齡 (2014.12)。〈社群感測器：社群媒體分析工具之設計〉。2014 第五屆數位典藏與數位人文國際研討會。台北市，中央研究院人文社會科學館國際會議廳，2014 年 12 月 1-2 日。
7. 許矢勇、陳恭、潘宗佐 (2014.10)。〈資源感知之社群媒體資料搜集系統：以推特為例〉。「2014TANET 臺灣網際網路研討會」，高雄國際會議中心，2014 年 10 月 22-24 日。
8. Chen, P.-L. & Cheng, Y.-C. (2014.10). Contrasting patterns of social media users participation in different mode of disasters, in Panel: Digital transformations of crisis communication and disaster media coverage, 2014 ICA Brisbane, Oct. 1-3.
9. 施旭峰、李蔡彥、鄭宇君、陳百齡 (2014.05)。〈災難事件中社群媒體訊息之自動分類設計〉。ICIM2014 第 25 屆國際資訊管理學術研討會，台中：國立中興大學，2014 年 5 月 24 日。(榮獲大會最佳論文獎)
10. 陳百齡、鄭宇君 (2013.12)。〈災難期間網路媒體使用者報導文本的內容分析〉，2013 年中國網絡傳播學年會。上海，復旦大學，2013 年 12 月 6-8 日。(榮獲大會教師組最佳論文獎)
11. 鄭宇君、陳百齡 (2013.07)。〈探索 2012 台灣總統大選之社交媒體浮現社群：鉅量資料分析取徑〉。中華傳播學會 2013 年會，新北市，輔仁大學。
12. 袁涵郁、陳百齡 (2013.07) 〈由人肉搜索事件看網路集體行動：情節、角色與協力的尋人行爲〉，2013 年中華傳播學會年會論文研討會，台北：輔仁大學。
13. Cheng, Y.-C. \* & Chen, P.-L. (2013). Global social media, local context: A case study of Chinese-language tweets about the 2012 presidential election in Taiwan. Paper presented at 2013 IAMCR Conference in Ireland June 25-29, 2013.
14. 鄭宇君、陳百齡 (2012.12)。〈超越在地脈絡的全球社交媒體：以 2012 年台灣總統大選的中文 Twitter 討論社群為例〉，「2012 中國網絡傳播學年會澳門國際會議」論文。澳門，澳門大學，2012 年 12 月 6-8 日。
15. 鄭宇君 (2012.06)。〈探討社交媒體於災難期間的應用：以南瑪都颱風為例〉，「2012 文化創意與傳播科技研討會」論文。新竹，玄奘大學，2012 年 6 月 1 日。
16. 鄭宇君 (2012.01)。〈災難傳播中的群體力量：社交媒體促成新型態的公民參與〉，發表於「《傳播與社會學刊》五周年國際學術會議」，香港，香港中文大學新聞與傳播學院。

17. 陳百齡、鄭宇君 (2011.07)。〈災難情境下的新興媒體：莫拉克風災中的浮現頻道〉，「2011 中華傳播學會年會」論文。新竹，交通大學客家學院。
18. 鄭宇君、陳百齡 (2011.07)。〈在災難事件中透過社交媒體溝通不確定性：以莫拉克風災為例〉，「2011 中華傳播學會年會」論文。新竹，交通大學客家學院。(榮獲社會組優良論文獎)
19. 陳百齡、鄭宇君 (2010.07)。〈媒體創新：莫拉克風災中的浮現頻道〉，「2010 年創新與創造力會」論文。臺北：政治大學。

## • 研究計畫

1. Yee-man Joyce Nip (University of Sydney, Australia)、傅景華 (University of Hong Kong, HK)、陳百齡 (政治大學)、陳恭 (政治大學)、鄭宇君 (玄奘大學) (2015-2017)，《Online Public Opinion in Greater China: Similarities, Differences, and Mutual Influences》，蔣經國基金會委託二年期計畫。
2. 陳百齡 (2013-2016) (主持人)，《尋找媒體創用者的數位足跡：社交媒體研究之新方法探索－社交媒體重大事件之傳播模式比較分析：鉅量資料數位人文分析取徑》，數位人文整合型計畫之子計畫，科技部委託三年期計畫 (計畫編號：102-2420-H-004-048-MY3)
3. 鄭宇君 (2014-2016) (主持人)，《探索社交媒體之民意形成過程》，科技部委託二年期計畫 (計畫編號：103-2410-H-364-004-MY2)
4. 陳百齡 (2014) (主持人)，《跨社群平台訊息傳播與議題發掘之研究》，財團法人資訊工業策進會委託一年期計畫 (計畫編號：103-7001)。
5. 鄭宇君 (2012-2014) (主持人)，《在社交媒體時代追求真實性的新聞專業》，科技部委託二年期計畫 (計畫編號：101-2410-H-364-002-MY2)
6. 陳百齡 (2011-2013) (主持人)，《災難情境下浮現的媒體》，行政院國家科學委員會委託二年期計畫：(計畫編號：100-2410-H-004-156-MY2)。
7. Jean Burgess (Queensland University Technology, Australia)、陳百齡 (政治大學) (2010)《Flood and Fire: Understanding Public Communication in Times of Crisis》，ASSA International Science Linkages Bilateral Program (2010.03-2010.10)。

## • 碩博士論文

1. 鄭雅云 (2015)，《探討網絡化社運中社會網絡之形成：以苗栗大埔事件為例》，國立政治大學新聞研究所碩士論文，指導教授：鄭宇君、陳百齡。
2. 任喆鵬 (2015)，《憂鬱傾向的社群媒體書寫》，國立政治大學新聞研究所碩士論文，指導教授：陳百齡、陳恭。
3. 張瑋庭 (2015)，《基於貼文分享之臉書粉絲頁關聯性分析》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
4. 李燕宜 (2015)，《臉書粉絲頁超連結分析系統》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
5. 王力弘 (2015)，《社群媒體新詞偵測系統》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
6. 吳君孝 (2014)，《社群感測器：社群媒體分析工具之設計》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：李蔡彥。



7. 許矢勇（2014），《資源感知之社群媒體料搜集平台》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
8. 潘伯彥（2014），《臉書使用者行為資料蒐集系統》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
9. 周玉駿（2013），《實作推特社群媒體的資料蒐集與管理服務》，國立政治大學資訊科學研究所碩士在職專班論文，指導教授：陳恭。
10. 施旭峰（2013），《災難事件下新媒體資訊傳播方式分析與自動化分類設計 — 以八八風災為例》，國立政治大學資訊科學研究所碩士論文，指導教授：李蔡彥。
11. 黃淑琳（2013），《行禮如儀還是安慰劑效應？—探討 Facebook 互動儀式的運作機制》，國立政治大學新聞研究所碩士論文，指導教授：陳百齡。
12. 袁涵郁（2012）《由人肉搜索事件看網路集體行動：情節、角色與協力的尋人行為》，國立政治大學新聞研究所碩士論文，指導教授：陳百齡。

### 十三、研究計畫內容（以中文或英文撰寫）：

（二）研究計畫之背景及目的。請詳述本研究計畫之背景、目的、重要性及國內外有關本計畫之研究情況、重要參考文獻之評述等。本計畫如為整合型研究計畫之子計畫，請就以上各點分別述明與其他子計畫之相關性。

## 跨平台社群媒體巨量資料蒐集與分析

### 壹、前言

當代社群媒體的普及，連結了人們的線上與線下世界，跨越了國與國的疆界。無論是政治傳播、社會參與、廣告行銷等各個面向的應用，社群媒體都扮演了不可或缺的角色。然而，對於社群媒體資料的蒐集與分析，國內的人文及社會科學研究卻起步地相當晚，其中主要原因一方面受限於技術門檻，一般人文社會學者不易取得社群媒體資料進行分析，另一方面則是國內學界缺乏人文社會學者與資料科學者的跨領域合作經驗。

江奕瑄、林翠絹（2015）分析國內外使用大數據探討媒體使用的文獻，他們檢閱國內外的學術期刊資料庫發現，屬於傳播領域的大數據研究共有 80 篇，其中只有 40 篇屬於 SSCI 資料庫，最早一篇是 2009 年刊出，她們發現在這 40 篇論文中，採用數據資料進行實證研究的只有 12 篇，其餘皆為理論或概念分析，且這 40 篇論文當中，超過一半以上文章有共同作者。這足以顯現對於人文社會學者而言，要取得社群媒體資料進行實證分析具有一定的困難，因此需要透過跨領域或產學合作方式發展。

長期進行社群觀察的人文社會學家與挖掘大量社群資料的電腦科學家，他們所接近的是不同類型資料、問不同的問題、注意不同模式，最後獲得不同的 Insights，因此人文社會學者與資料專家之間具有資料分析落差（data analysis divide），雙方必須透過長期合作形成默契，並經由語言轉譯瞭解彼此思維，最終才能合作發現問題、用對工具、提出正確解題策略，降低人文學者與資料專家的資料分析落差（Manovich, 2011；鄭宇君，2014）。

從商業角度來看，目前國內雖然有幾家社群資料服務供應商，這些供應商提供社群媒體大量資料的分析服務，對企業而言，這些服務商可快速提供業主欲觀測主題的社群輿論變化與分析服務，但他們並不提供原始社群資料，也難以獲知他們所使用的社群分析演算法，服務商將資料收集與分析流程被視為商業機密，這使得社群資料服務供應商提供的分析服務，缺失學術研究的可驗證性，研究者也難以從社群資料供應商取得完整原始資料進行深入分析。總體而言，學術研究使用商業社群媒體服務具有幾個可能缺點：（1）未提供資料來源的透明程度；（2）缺乏指標品質的評估標準；（2）資料分析服務的價格高昂（Manovich, 2011；boyd & Crawford, 2012）。

Bruns & Weller（2014）指出目前社群媒體研究所面臨的五大挑戰：（1）資料的代表性與效度，如何有效地收集貼文，以減少不相關貼文及垃圾資訊的比例；（2）跨平台研究，如何同時處理二個以上社群媒體平台資料；（3）比較研究，如何針對不同個案進行比較分析；（4）多重方法取徑，如何用不同取徑來處理相同的資料集，以獲得更全面的瞭解，（5）脈絡與意義：社群資料的蒐集與分析，如何反映使用者經驗與意義。上述 Bruns 等人提出的觀點，也正是臺灣在地學者在社群媒體資料分析上所面臨的主要問題。

因此，本團隊希望能結合資訊科學專家與傳播學者的領域知識，設計一個友善的跨平台社群媒體資料蒐集與分析系統供學術界使用，可以幫助更多人文社會學者參與社群資料的研究，我們希望這個

系統可以達成以下幾個目的：一、研究者可以針對研究議題屬性為在地性或跨國性，選擇台灣最多人使用的Facebook或跨國訊息流動快速的Twitter，或是二種平台並行蒐集資料；二、有效地幫助研究者針對特定議題蒐集大量的社群媒體原始資料（raw data），供研究者下載進行深度分析，維持資料蒐集與分析流程的透明性；三、提供基礎的資料查詢與統計功能，幫助研究者快速檢視所蒐集資料狀況，以便第一時間調整資料蒐集條件。最後一點是社群媒體資料研究有別於傳統大眾媒體研究之特殊性，大眾媒體由於頻道有限、內容有限，過期的報紙或電視節目內容仍可從機構的資料庫取得，但社群媒體資料的產生及變動十分快速，若未在第一時間蒐集資料，日後即難以取得。<sup>1</sup>

基於上述特點，本研究目的在建置一個跨平台社群媒體資料蒐集系統，針對 Facebook 與 Twitter 的公開資料進行蒐集，以便研究者針對研究主題收集大量的社群媒體原始資料，並可快速檢索找到所需分析資料，方便人文社會科學研究者進行後續分析，以減少目前鉅量資料研究中，人文社會學者與資料專家的資料分析落差。

## 貳、社群資料與社會科學研究之關聯

### 一、社群媒體的資料特性

社群媒體出現至今已十數年，從早期 BBS 到 Blog，再發展成 Twitter、Facebook。雖然社群媒體類型不一，但其仍有幾項共通特性。

第一個特性是使用者參與行為呈現冪律分布（Power law distribution）。Rhiengold（2012）認為人們在網路上的參與行為五花八門，若以參與門檻作為排序標準，則至少包括閱讀（read）、標籤（tag）、評論（comment）、訂閱（subscribe）、分享（share）、連結（network）、書寫（write）、定調（refactor）、協作（collaborate）、協調（moderate）以及領導（lead）等參與行為。更重要的是，Rhiengold 指出這些參與行為出現的頻率為冪律分布，也就是說，愈後面幾種門檻較高的行為參與人數愈少，但對群眾活動貢獻度最高，大多數人只是閱讀文章。這種使用行為的冪律分布在不同的社群媒體中都會出現。此特性具有對於網路社群的集體行為具有何種意義？帶來怎樣的影響？便成為值得關注的問題。

社群媒體的第二個特性，在於資訊流動快速而龐雜，因此資料佚失速度也相對變快。舉例而言，PTT 八卦版每日可能出現數百篇新貼文，而在平台能容納的貼文數量有限的狀況下，舊文章就會不斷地消失。此外，重大事件時也常在短時間產製出大量訊息，如 Burgess & Bruns（2012）搜集 2010 年澳洲大選選前一個月到選後 2 天（共 38 天）所有標示#ausvote 的相關 tweets，結果發現總數 41.5 萬則 tweets 中有 22% 都集中在選舉當天發出。這些短時間內大量的訊息流動仰賴適合的工具即時記錄。

第三個特性在於透過資訊工具（API、爬蟲程式等）蒐集的社群媒體資料，包含結構性與非結構性資料，如：文字、圖片、影像、互動等，這些資料被設計為以電腦可讀方式分類、呈現資料，若使用 API，其所提供的資料除了一般使用者介面可以看到的內容外，還包括大量後設資料（metadata），如貼文時間、使用者 ID、地理資訊等等。

社群媒體資料的這些特性，與傳統大眾媒體訊息集中在少數頻道（channels）不同，社群媒體扮演的是平台角色，它提供眾多個人或組織頻道在同一時間進行多對多的傳遞與溝通。因此，我們可以說社群媒體資料提供了新的可能性，使研究者在探問民意形成、訊息傳散速度、用戶互動及影響範圍時可取得新的資料來源，這些多樣性的社群資料，包含了 data 與 metadata，可幫助研究者從社群資料中挖掘更多線索，並提出新的問題。

---

<sup>1</sup> 目前透過 Twitter Rest API 僅能取得七天以內的資料，若要取得更久之前的歷史資料，則需向 Twitter 公司購買，所費不貲；Facebook API 所撈資料雖然不限歷史資料，但可能因為使用者調整隱私設定或關閉帳戶，而不再能取得資料，因此研究者最好在事件或議題開始之初即先蒐集社群媒體資料。

因而，我們希望建立社群資料蒐集平台與基礎分析工具，幫忙研究者打開自身的想像，探索更多先前不曾思索的問題。

## 二、社群媒體資料對應的傳播問題

當一個重大事件或議題發生時，許多公眾紛紛在社群媒體上表達自身意見，這些眾多意見形成了一股網路輿論，例如：2014 年反服貿事件促成了太陽花學運；2014 年台北市長選舉柯文哲因為網民的支持，創下素人參選成功的案例；以及 2015 年反黑箱課綱議題，促成了全國數十所高中生透過臉書串連。對於研究者或政府單位而言，關切的是這樣的運動如何進行動員及訊息擴散，以及在這些事件中究竟誰在發聲？誰又被影響了？

我們以傳播學研究的 SMCRE 模式來檢視社群媒體資料所能回答的問題，當上述這些議題發生時，透過社群資料的收集與分析，研究者大概能夠幫助回答下述問題：

表一：社群媒體資料對應的傳播研究問題

	傳播者(S) Sender	訊息(M) Message	管道(C) Channel	接收者(R) Receiver	效果(E) Effect
發文時間	使用者在甚麼時候說話？	甚麼時候會說甚麼話？	甚麼時候在甚麼平台說話？	甚麼時候、隔多久才會產生互動或回應？	甚麼時候說話會有甚麼不同效果？
發文內容	誰喜歡說甚麼？	說了甚麼？態度為何？運用甚麼訴求手段？	甚麼內容會在甚麼平台出現？	產生甚麼樣子的回應內容？	甚麼內容會有甚麼不同效果？
使用者檔案 (ID、國籍、性別、教育程度等)	甚麼類型的人會在社群媒體上說話？	甚麼類型的人會說甚麼話？	甚麼類型的人會在甚麼平台說話？	甚麼類型的人會進行甚麼類型的回應？	甚麼類型的人會產生甚麼樣的傳播效果？
來源平台/網站	使用者喜歡透過那些平台說話？	不同平台的使用者喜歡說些甚麼話？	不同平台的機制設計對使用者帶來的影響？	不同平台的使用者會有甚麼不同的回應模式？	不同平台的傳播效果？
點閱數/回應數	甚麼類型的人會獲得不同的點閱、回應數量？	甚麼類型的發文會獲得不同點閱/回應數？	不同平台獲得的點閱/回應數量與模式？	不同回應或分享所獲得的點閱/按讚/回應數量與模式。	點閱/回應的社會效果、對消費行為的影響。
引用連結	甚麼類型的人會引用甚麼不同類型的資訊？	引用內容可分為哪些類型？	不同平台的使用者會引用哪些不同平台、類型的內容？	回應或分享時會引用那些不同平台、類型的內容？	引用不同平台、類型的內容會產生甚麼樣的傳播效果？
標籤	誰在標籤？誰被標籤？哪些議題被標籤？	甚麼標籤會跟甚麼描述一起出現？	甚麼平台會出現甚麼標籤？	回應或分享時會產生甚麼樣的標籤行為？	標籤/被標籤的傳播效果？

轉推/分享	誰在轉推/分享？	轉推/分享會搭配甚麼樣的描述？會否改造、編輯、挑選原訊息內容？	不同平台的轉推/分享行為有何不同？跨平台的資訊分享模式？	閱聽人對轉推/分享內容的反應為何？	轉推/分享行為會產生甚麼樣的傳播效果與社會效果？
地理資訊	誰會在哪裡貼文、分享照片？	地理資訊的分享會搭配甚麼樣的內容？那些地點最常被分享？人們最常在哪些地點貼文？	不同平台在分享地理資訊的機制有何不同？有何影響？	閱聽人對人們分享地理資訊的反應為何？	分享地理資訊行為會產生甚麼樣的傳播效果與社會效果？

一般而言，我們可以將社群媒體資料分為數量資料、文字資料以及後設資料，數量資料包括回應數、瀏覽量、分享量、聲量等，而文字資料則為發文內容、評論內容、標籤內容以及使用者檔案等；後設資料則指文章中引用的超連結、發文地理資訊（常以座標表現）、使用者 IP、使用的載具等，這些資料常需要透過進一步資訊處理（如將座標化為地理位置、辨識引用連結網站的類型或區域等），方能彰顯資料意涵。

綜合使用上述社群資料，研究者能夠根據各自研究旨趣進行探索。在傳播效果典範下，勸服研究學者可分析意見領袖在社群媒體中的影響力，也可關注貼文時間、敘事策略、貼文類型（純文字、引用新聞、照片、影片）對傳播效果的影響；從行銷角度來看，可分析特定商品或是特定品牌在社群媒體中的討論熱度與態度，藉此觀察行銷活動之成敗，甚至根據社群反應進一步擬定策略；政治傳播可探勘特定候選人、政治人物或是政策議題在社群中的討論聲量與內容，藉此分析民眾對公共政策或政治人物的關心程度、面向、以及態度，研究者甚至可根據社群媒體中的討論內容發展各種傳播效果的預測指標與模型（鄭宇君、陳百齡，2014）。

文化研究者則可分析特定概念在不同場域中聲量的差異，甚至可深入探討特定概念在時序中意涵的轉變，如「文青」、「宅男」等詞語在台灣社會各領域經常出現，但意涵仍屬模糊，藉由社群媒體資料，研究者或可發現這些概念如何被人們所描述；「迷文化」牽涉在社群中逐漸形成的特殊互動模式，社群媒體資料能幫助研究者分析特定迷社群的互動模式與常用語彙，帶領研究者透過資料探索迷群的世界。

重大事件如災難發生時，人們的不確定感增加，傳播交流的需求因而上升，各類傳播平台常在瞬間湧入大量資訊，這些資訊能幫助實務工作者看見人們的需求，並且根據群眾反應定位問題的急迫性。災難傳播研究便可利用社群媒體資料探索災難時期人們的資訊行為模式，甚至針對不同災難類型，設計即時觀測機制，讓社群討論進入實務決策過程中成為參考指標之一（陳百齡、鄭宇君，2014）。

社群媒體資料的另一優勢在於可讓社會關係可視化(visualization)，藉由追蹤、加好友、參與活動、共同關注特定人物或粉絲專頁、共同標籤、在特定貼文留言等等行為，不同使用者之間產生了各式各樣關係，這帶給社會網絡(social network)研究者大量可供分析的新資料。舉例而言，研究者可藉由前述社群媒體資料辨識出特定迷群的成員、關注特定議題、喜歡同類電影的使用者，並可找出這些人之間的社會網絡，甚至進一步分析成員間各項社會類屬之異同；加上了時間資料，研究者也可以社群成員網絡關係之變化。

### 三、社群資料收集與分析流程

本研究團隊結合了傳播與資料的跨領域學者，自 2009 年底開始進行社群資料的收集與分析，瞭解重大事件中社群媒體中的訊息傳散與公眾參與之情況。本團隊經過五年的研究探索，共發表了 7 篇期刊論文、19 篇會議論文、12 篇碩士論文，根據這些經驗，我們歸納出社群媒體資料蒐集平台需具備以下的基本功能，以幫助研究者進行社群媒體資料蒐集的基本流程：



圖一：社群資料蒐集的基本流程

1. **資料收集：**依照不同社群媒體屬性及資料結構方式，建立資料蒐集工具。我們初步規劃以台灣最多人使用的 Facebook、以及與國際訊息快速流通的 Twitter 做為社群媒體資料蒐集來源，透過這兩個平台官方所提供的程式開發者介面（API）收集特定議題的公開資料，未來可視需求擴充所蒐集的社群媒體來源。
2. **資料檢索：**依據個別社群媒體資料所具備的資料（data）與後設資料（metadata）格式，建立友善的查詢介面，方便非資訊背景之研究人員設定檢索條件與查詢資料。
3. **資料視覺化：**由於社群資料結構龐雜，研究人員難以一窺究竟，本研究透過視覺化工具呈現社群資料的基本概況，例如：時間趨勢變化、地理資訊分佈、按讚留言分享數量之時間變化等，並可匯出原始資料供研究者進行人工分析。

當研究人員透過此一平台蒐集社群媒體資料，並瞭解資料基本概況之外，研究者可匯出原始資料，根據所欲探詢之問題，進一步使用不同的工具進行後續分析。例如：結合統計軟體，可進行統計因果關係之探討；結合社會網絡分析工具，如 Pajek、UCINET、Siena、Pnet，以及 Gephi 等軟體進行資料處理，可瞭解特定議題中的社會網絡關係。

此外，透過社群媒體所蒐集的大量的文字資料，除了可使用傳統社會科學研究方法如：論述分析、內容分析、框架分析進行人工資料分析外，也可使用新興的文本分析工具進行，包括為關鍵詞分析 (Keyword Analysis)、詞頻分析 (Frequency Analysis)、詞語共現分析 (Co-occurrence Analysis) 以及情緒分析 (Sentiment Analysis) 等 (陳百齡、鄭宇君，2014；Stubbs, 2001；Liu, 2012)。其中關鍵詞分析試圖找出網路使用者針對某個議題較常使用的詞彙；詞頻分析則可看出特定議題相關討論中所有詞彙之出現頻率；詞語共現分析可找出研究者關心之詞彙與其他詞彙之間的關係；情緒分析則可幫助研究者看出人們談論特定議題時之態度為何。

因此，根據上述研究理念，我們認為設計一套跨平台之社群媒體資料蒐集與分析系統，可有效地幫助研究者解決收集大量社群資料的難題，並可透過友善的查詢介面，方便研究者獲得資料的基本概況，並匯出原始資料供研究者進行探索。

(三) 研究方法、進行步驟及執行進度。請分年列述：1.本計畫採用之研究方法與原因。2.預計可能遭遇之困難及解決途徑。3.重要儀器之配合使用情形。4.如為整合型研究計畫，請就以上各點分別說明與其他子計畫之相關性。5.如為須赴國外或大陸地區研究，請詳述其必要性以及預期成果等。

## 參、跨平台社群媒體資料蒐集系統之架構

本計畫目的在於設計一個跨平台社群媒體資料蒐集系統，供研究者收集社群媒體上之公開資料，本計畫選定兩種主要的社群媒體平台，做為資料蒐集之來源，一個是台灣地區使用人數最多的社交媒體平台 Facebook（臉書），目前約有 1300 萬用戶，另一個則是兩岸三地華人社群及國際社會經常使用的 Twitter（推特）。

相較於 Facebook 的訊息傳遞以朋友社交互動為主，Twitter 則是一個以訊息擴散為主的社交媒體，更有利訊息透過弱連結方式快速擴散，特別是在跨語言社群之間的流動（Cheng & Chen, in press）。特別像是台灣總統大選、亞洲投資銀行、日本 311 地震等重大事件，參與討論的線上社群不僅限於在地社群，還包括周邊國家社群，因此本計畫擬建置一個跨平台社群媒體資料收集系統，使用者可依議題屬性偏向在地性或跨國性，選擇收集 Facebook 或 Twitter 社群平台的資料，系統亦會提供所收集社群資料的基本統計及概況描述。

以下就臉書粉絲頁（Facebook fan page）與推特（Twitter）兩種具公開性質的社群媒體，根據 Facebook 及 Twitter 所提供的 API 服務，設計與開發一個社群媒體公開資料蒐集的服務系統。使用者可以透過本系統蒐集特定期間的指定粉絲頁的資料，或是包含特定關鍵字推文。以下為此系統主要的設計思維：

首先，依據本計畫之時程與資源限制，此系統之主要目的為提供議題導向之資料蒐集服務，而不是長時期監控用的資料蒐集服務。也因此本系統提供資料蒐集與基本的資料統計分析功能，並不提供使用者長期儲存資料與檢索之服務。但針對此部分，本系統將預留彈性架構，以利在未來進行擴充。

其次，不管是蒐集臉書或是推特的資料，都需要提供經由臉書或推特認證的使用者識別碼（access token），且每個 token 都有一定的資料蒐集限制，基於平等互惠原則，本系統將要求使用者提供個別的 token 來進行資料的蒐集。以 Twitter 為例，Twitter 在 2012 年 8 月為減少垃圾訊息與機器人使用者數量，將 API 使用更新為 1.1 版本，新版本對於第三方使用者增加許多限制性，對於資料的獲取數量也由無限制下修至每區段時間 180 次 API 請求（Requests）。

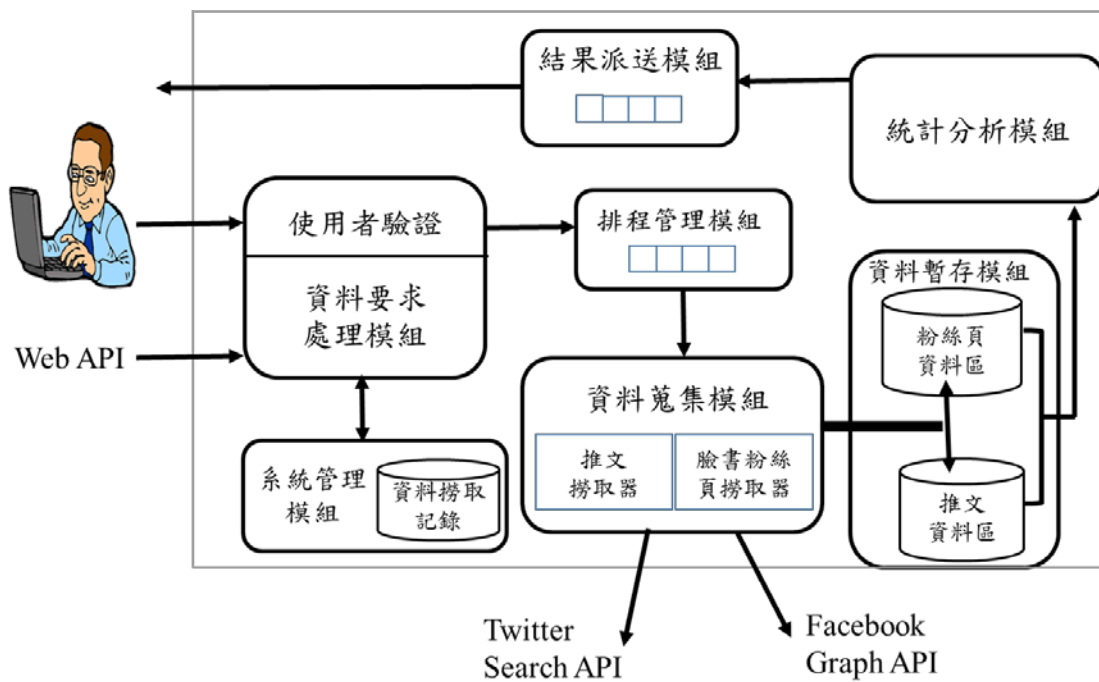
此外，為避免特定資料蒐集要求耗費過多系統資源與時間，影響其他使用者，本系統的服務以非同步、排程的方式進行，待服務完成後再通知使用者下載資料與瀏覽統計分析之結果，以充分運用系統的各項資源。最後，本系統將紀錄所有的資料蒐集要求，以利問題排除與確保資料蒐集人之可受問責性（accountability）。<sup>2</sup>

另外，除了 API 使用次數的限制性，有些關鍵用法在不同語系的使用也會造成英文語系外的開發者限制性。在 The Digital Methods Initiative（Borra & Rieder, 2014）所開發的開放原始碼工具 dmi-tcat 中，主要以關鍵字透過 Twitter Stream API 來追蹤與收集推文，而 Stream API 完全無法支援雙字元語系關鍵字的請求方式，以致於中文關鍵字收集推文時，必須另外再設計使用 Twitter Search API 的方法來收集。

本計畫所規劃之資料蒐集系統的架構如圖二所示。以下分別說明各模組之功能。

<sup>2</sup> 由於每位研究者所欲研究的議題與面向不一，所需使用的資料與後設資料欄位皆不相同，必須採取不同的資料清理與分析方式，因此本系統根據研究者設定的條件蒐集原始的公開資料（含基礎數量統計）後，交由研究者自行下載進行後續的資料處置，是故該批資料的使用者隱私保護及研究倫理等議題，需由研究者負起相關的研究倫理及個人資料保護之責任。





圖二：跨平台社群媒體資料蒐集系統架構

1. **使用者驗證模組**：使用本系統前，使用者必須經過認證，提供臉書或推特認證過的 access token 來進行資料的蒐集。因資源有限，本系統可依實際的系統狀況，設定使用者的用量限制。
2. **資料要求處理模組**：本系統以專案 (project) 方式，同時提供臉書粉絲頁與推特推文的蒐集服務。若是 Facebook 粉絲頁專案，使用者可以指定若干個粉絲頁，設定資料蒐集期間，由本系統協助蒐集這些粉絲頁的內容。至於 Twitter 推文專案，使用者則是給定若干個關鍵字，設定資料蒐集期間，本系統將協助蒐集並彙總這段期間內包含這些關鍵字的推文。此外，本系統也會提供 Web API，提供使用者以程式，仿照 REST (Fielding & Taylor, 2002) 方式提出資料蒐集服務。
3. **系統管理模組**：此模組負責設定一些系統管理參數，例如每位使用者同時間最多可建立的資料蒐集專案，以及資料蒐集期間的限制等。此外本模組也負責紀錄與管理所有的資料蒐集要求。
4. **排程管理模組**：本系統在接受使用者的資料蒐集要求後，會將要求送交排程管理模組的佇列 (queue) 中，由排程管理程式依序設定為資料蒐集工作，依系統排程周期定期執行之。意即本系統採用非同步、排程的方式來進行資料蒐集工作。
5. **資料蒐集模組**：此模組有兩個資料蒐集器：臉書粉絲頁蒐集器依臉書 Graph API 規格蒐集使用者指定期間的粉絲頁內容，包含貼文，留言以及分享數與按讚數等相關資料。<sup>3</sup>Twitter 推文蒐集器依 Twitter Search API 依關鍵字進行周期性的推文蒐集，<sup>4</sup>目前規劃以周為單位持續蒐集若干周，最大蒐集期限由系統參數統一規定。本模組將採多執行緒 (multithreads) 方式執行資料蒐集工作，並視系統資源狀況，動態調整執行緒的工作狀態。
6. **資料暫存模組**：本模組負責儲存與管理使用者所要蒐集的資料，蒐集過程中的資料暫存於此區，資料格式以 CSV 檔為主，未來可增加 JSON 格式。蒐集完成將暫存於系統中，等待使用者自行下載儲存。未來若要支援大量且長期的社群巨量資料蒐集與儲存，則可將此模組置換，改以分散式資料儲存系統，例如：HDFS 以及 HBase，來儲存管理。

<sup>3</sup> Facebook 資料蒐集的規定及限制，參見 Facebook Graph API: <https://developers.facebook.com/docs/graph-api>

<sup>4</sup> Twitter 資料蒐集的規定及限制，參見 Twitter REST API: <https://dev.twitter.com/rest/public>



7. **統計分析模組：**本模組針對使用者蒐集的資料提供基本之統計分析功能。以臉書粉絲頁為例，本模組提供貼文數、留言數、分享數與按讚數統計，並依時間軸，以視覺化工具呈現各類貼文( status, link, photo, video, event, note) 的數量變化，若是 Twitter 貼文，則可呈現推文數量、發文人數、發文地點等資料進行時間序列分析。此外，針對 URL link 分享，本模組將會進行短網址(tiny URL) 還原，並依網址 domain 進行統計。
8. **結果派送模組：**蒐集完成的資料集，在統計分析後，資料與結果會存入本模組之佇列，由本模組依序以電郵方式通知使用者來下載資料與瀏覽統計分析之結果。

## 肆、 資料保護與研究倫理議題

社群媒體具有鉅量資料提供學術研究者探究社會面貌的豐富資訊，透過數位資料來再現「現實」之樣貌 (Zwitter, 2014)。但鉅量資料之新資料型態也引起了重要的研究倫理議題，這些研究議題出現在整體研究過程，包括資料蒐集、資料分析、資料呈現、與資料保存之階段，它同時也牽涉許多技術議題。以下本提案分就知情權與隱私權之間的衝突，以及研究倫理涉及的議題分別加以敘述。

首先，「知情權」與「隱私權」是資料蒐集過程最重要而又相互衝突的兩個概念。社群媒體研究之資料大多來自網路上的公開資料，包括使用者在社群媒體上產生的個人內容，但多數使用者卻並不知道其資料被學術研究人員蒐集與分析。或者說，他們並非是知情並自願地提供資料的「研究對象」。而且，這不只是牽涉到個人使用者而已，也包括諸如「粉絲頁」等群體資料。對研究者來說，不管他們是以直接撈取或透過社群媒體之應用程式介面 (API) 來取得資料，都會遇到資料對象並不知情之狀況，也包括研究分析與出版後可能對資料對象有負面影響；例如，對某個人或群體之分析結果，可能暴露了其身份 (或群體中人們之社會網絡關係)，也可能道出了其原本不知情，卻可能有損個人或團體形象之分析結果 (鄭宇君，2014)。

因此，研究人員必須設法保護資料對象之隱私，包括以匿名方式呈現資料、隱藏其它可被辨識身份之欄位資訊，以及妥善保存所蒐集到的資料，以免外洩被用作其他用途。當然，以上只是一些初步的解決方法，仍未能徹底地解決「知情權」和「隱私權」之間的衝突，與可能對第三者造成傷害之難題；例如，對某重大社會事件進行社會網絡分析時，往往必須辨認出關鍵性使用者，但這辨認過程與分析 (儘管以匿名方式進行) 仍可能暴露出其身份或相關線索，影響其日常生活，甚至人身安全。對社群媒體研究而言，究竟如何定義、呈現，與使用「公開資料」這都是目前學術社群與法律專家尚在討論之議題，這些議題包括：不同屬性的資料 (例如，使用者的個人資料與登錄資訊) 是否有不同程度的隱私重要性，可否有不同程度的保密或呈現方式？不同的使用者帳號 (譬如說，政府部門、新聞機構、記者、公眾人物、知名部落客，與一般使用者) 之資料是否又有不同可被公開之程度？總之，社群媒體研究之倫理規範要如何落實還有待進一步釐清。

另外，研究倫理除了有關研究對象外，也關於研究者透過鉅量資料進行社群媒體研究之信度與效度問題，它同時也涉及技術議題。以目前學術機構的資料撈取技術來看，所取得的資料仍經常出現不完整的情形。這可能因為網路資料量過大難以全部擷取，也可能因為受限於社群媒體平台的應用程式介面所提供的資料量，對學術研究來說，這便容易造成取樣偏誤 (availability bias) (boyd & Crawford, 2012; Rivers & Lewis, 2014)。一方面，樣本的數量大小與代表性會容易受到挑戰，另一方面，如何對資料進行「清洗」，有效地去除雜訊干擾，才能獲得具有解釋力的資料分析結果。

就社群媒體研究來看，倫理議題也是技術問題，牽涉到資料蒐集者、資料分析者，與資料產製者之 (知情或不知情的) 互動。要如何兼顧知識之有效性、研究之正當性，與學術倫理之重要性，將是網路社群媒體鉅量資料研究所不能逃避的重要議題。

## 伍、 小結：

本計畫目的是將水火研究團隊過去五年在社群媒體鉅量資料分析的跨領域合作經驗，轉化為一套可供其他研究者使用的社群媒體資料蒐集平台與資料處理流程。這個平台的設計精神，一方面包括了本團隊經過多年嘗試所累積的默會知識，並提出鉅量資料分析的操作原則與處理流程；另一方面則是將原本分別由不同研究人員撰寫程式進行的資料蒐集與分析工作模組化後，將多個模組整合為一套跨平台社群媒體資料蒐集與分析系統，加上友善的人機界面設計，以提供不具資訊背景的人文社會學者可以自行設定條件蒐集、查詢與匯出社群媒體資料。

我們認為，在進行社群資料分析時，研究者不僅需要具有清晰的問題意識，以及熟稔資料分析的操作流程，同時必須兼顧到現實層面，在有限的人力及計算資源的條件下，達到最佳化的研究資源利用。在計畫設定的六個月期間內，我們預計建置一套跨平台的社群媒體資料蒐集與分析系統，除了供研究人員使用之外，我們同時以開放原始碼的方式釋出，供研究人員自行下載使用，並可透過社群力量共同維護，以因應 Facebook 與 Twitter 官方隨時改版更新 API 時所新增的條件限制。此外，我們也將提出未來系統延展性之規劃，針對該系統未來若擴充為更多研究人員使用時，所需的使用者層級之權限管理、資料蒐集與儲存之資源需求、研究人力及物力配置、以及資料保護及研究倫理等議題進行探討。

我們希望奠基在過去幾年對社群媒體資料的研究基礎上，架設一個可供研究者使用之社群媒體資料蒐集平台，以降低其他研究者進入此領域的摸索時間，讓更多研究者一起透過社群資料的分析來探索當代社會的各項重要議題。

（四）預期完成之工作項目、成果及績效。請分年列述：1.預期完成之工作項目。2.對於學術研究、國家發展及其他應用方面預期之貢獻。3.對於參與之工作人員，預期可獲之訓練。4.預期完成之研究成果及績效（如期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期績效）5.本計畫如為整合型研究計畫之子計畫，請就以上各點分別說明與其他子計畫之相關性。

## 陸、 預期完成之工作項目、成果及績效

本計畫為半年期之先導型計畫，因此在六個月的期間內，我們希望完成以下工作項目，同時規劃這一先導性計畫的後續延伸方向：

1. 建置跨平台社群媒體資料蒐集與分析系統，可供研究人員蒐集、分析、匯出社群資料。
2. 將本計畫建置之社群平台資料蒐集系統，以開放原始碼（open source）方式釋出，供使用者自行下載使用或修改，透過開源社群的維護，未來可提供更多元的功能。
3. 規劃多專案執行之控管機制，包括系統維護、帳號權限管理、資源配置、隱私保護等議題，供科技部未來建置全國性社群資料蒐集平台之參考。
4. 規劃舉辦種子教師工作坊，針對有意從社群媒體資料研究者，本研究可分享如何蒐集與分析社群媒體資料之經驗。
5. 規劃建立社群鉅量資料分析成果分享平台，供不同團隊成員上傳並分享其研究成果，包含資料概況、資料視覺化呈現、學術論文等研究成果。
6. 強化系統的擴充性，未來可擴充到其他社群媒體（如：PTT 八卦版等）之資料蒐集。

## 參考文獻

1. 江奕瑄、林翠絹 (2015)。〈採用大數據探討媒體使用之學術期刊文獻分析〉，元智大學「大數據、新媒體、使用者」研討會論文，2015 年 6 月 12 日。
2. 陳百齡、鄭宇君 (2014)。〈從流通到聚合：重大災難期間浮現的資訊頻道〉，《新聞學研究》，121：89-125。
3. 鄭宇君、陳百齡 (2014)〈探索 2012 台灣總統大選之社交媒體浮現社群：鉅量資料分析取徑〉《新聞學研究》，120：121-165。
4. 鄭宇君 (2014)。〈向運算轉：新媒體研究與資訊技術結合的契機與挑戰〉，《傳播研究與實踐》，4 (1)：45-61。
5. Borra, E. & Rieder, B.(2014). Programmed method: developing a toolset for capturing and analyzing tweets, *Aslib Journal of Information Management*, 66 (3), 262-278.
6. boyd, d. & Crawford, K. (2012). Critical questions for Big Data, *Information, Communication & Society*, 15(5), 662-679
7. Bruns, A. & Weller, K.(2014). The pleasures and perils of studying Twitter, *Aslib Journal of Information Management*, 66 (3), 246-249.
8. Burgess, J., & Bruns, A. (2012). (Not) the Twitter election: The dynamics of the #ausvotes conversation in relation to the Australian media ecology. *Journalism Practice*, 6(3), 384-402.
9. Cheng, Y.-C. & Chen, P.-L. (in press). Interactions between different language communities on Twitter during the 2012 presidential election in Taiwan. In Axel Bruns, Gunn Enli, Eli Skogerbø, Anders Olof Larsson, & Christian Christensen (Eds.) , *The Routledge Companion to Social Media and Politics*. (預計 2015 出版)
10. Fielding, R. T. & Taylor, R. N. (2002, May), Principled design of the modern Web architecture, *ACM Transactions on Internet Technology (TOIT)*, Volume 2 Issue 2, May 2002.
11. Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167.
12. Manovich, L.(2011). Trending: the promises and the challenges of big social data, in M. K. Gold (Ed.). *Debates in the Digital Humanities*. Minneapolis, MN: The University of Minnesota Press. [Online] Available at: [http://www.manovich.net/DOCS/Manovich\\_trending\\_paper.pdf](http://www.manovich.net/DOCS/Manovich_trending_paper.pdf) (15 July 2011).
13. Rhiengold, H. (2012). *Net smart: How to thrive online*. London, UK: The MIT Press.
14. Rivers, C. M. & Lewis, B. L. (2014). Ethical research standards in a world of big data. *F1000Research*. Retrieved from <http://f1000research.com/articles/3-38/v1>
15. Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. New York, NY: Blackwell Publishers.
16. Zwitter, A. (2014). Big data ethics. *Big data & society*, 1(2). Retrieved from <http://bds.sagepub.com/content/1/2/2053951714559253.full.pdf+html>

#### 十四、近三年內執行之研究計畫

(請務必填寫近三年所有研究計畫，不限執行本部計畫)

計畫名稱 (本部補助者請註明編號)	計畫內擔任之工作	起迄年月	補助或委託機構	執行情形	經費總額
穿戴式互動展演創新應用與技術研究(1/3)(103-2218-E-004-001-)	共同主持人	2014/09/01~2015/07/31	科技部	執行中	3,763,000
SaaS環境下多租戶應用程式的SQL改寫機制：理論與實務(103-2221-E-004-005-)	主持人	2014/08/01~2015/07/31	科技部	執行中	603,000
探索社交媒體之民意形成過程(103-2410-H-364-004-MY2)	共同主持人	2014/08/01~2016/07/31	科技部	執行中	1,589,000
消費者性向偵測與發展趨勢預測技術及其創新應用研發(103-2221-E-006-257-MY3)	共同主持人	2014/08/01~2017/07/31	科技部	執行中	3,263,000
尋找媒體創用者的數位足跡：社交媒體研究之新方法探索—「流動的」手機與液態現代性：時空軟化與公私交錯(102-2420-H-004-047-MY3)	共同主持人	2013/08/01~2016/07/31	科技部	執行中	1,893,000
尋找媒體創用者的數位足跡：社交媒體研究之新方法探索—社交媒體重大事件之傳播模式比較分析：鉅量資料數位人文分析取徑(102-2420-H-004-048-MY3)	共同主持人	2013/08/01~2016/07/31	科技部	執行中	2,244,000
以依值型別與對話型別確保並行程式之正確性(100-2221-E-001-003-)	共同主持人	2011/08/01~2012/07/31	科技部	已結案	456,000
安全多方計算協定描述語言之設計與實作(100-2221-E-004-002-MY3)	主持人	2011/08/01~2014/09/30	科技部	已結案	1,702,000
合 計					15,513,000

# 科技部人文司近五年(98.8.1 至 103.7.31)專題計畫成果出版情形表

主持人：陳 恭

服務機關：國立政治大學資訊科學系

☐ 近五年未曾執行本部專題研究計畫，請打勾。

說明：1.102 年度計畫如仍在執行中請註明。

2.「會議論文集」其定義為會後編輯出版之論文集，此外之會議論文請勿列入。

3.出版情形將列為 104 年度專題計畫審查之重要參考資料，敬請詳實填寫。

(如擔任共同主持人、協同研究人員之計畫請勿列入)		計畫編號
計畫名稱	程式之模組性與擴充性:驗證方法與工具—子計畫五:剖面導向函數語言之模組化狀態處理與型態擴充研究	NSC97-2221-E-004-001-MY3 (三年期第二年, 98)
<input checked="" type="checkbox"/> 已出版 (請填寫論文題目、期刊名稱、卷號、頁數、年份或專書(含會議論文集)名稱、頁數、年份、出版社) Kung Chen, S.C. Weng, M. Wang, S.C. Khoo, and C.H. Chen (2010, Nov), "Type-Directed Weaving of Aspects for Polymorphically Typed Functional Languages", <i>Science of Computer Programming</i> , Vol. 75, Issue 11, pp.1048-1076. Nov. 2010. Elsevier. (SCIE)		
<input type="checkbox"/> 已投稿 (請填寫論文題目及期刊名稱或專書(含會議論文集)名稱) <input type="checkbox"/> 未出版		
計畫名稱	程式之模組性與擴充性:驗證方法與工具—子計畫五:剖面導向函數語言之模組化狀態處理與型態擴充研究	NSC97-2221-E-004-001-MY3 (三年期第三年, 99)
<input checked="" type="checkbox"/> 已出版 (請填寫論文題目、期刊名稱、卷號、頁數、年份或專書(含會議論文集)名稱、頁數、年份、出版社) Kung Chen, Shu-Chun Weng, Jia-Yin Lin, Meng Wang, Siau Cheng Khoo (2011, Jun). "Side-Effect Localization for Lazy, Purely Functional Languages via Aspects." <i>Higher-Order and Symbolic Computation</i> , Volume 24 Issue 1-2, page 151-189. Springer-Verlag. (EI)		
<input type="checkbox"/> 已投稿 (請填寫論文題目及期刊名稱或專書(含會議論文集)名稱) <input type="checkbox"/> 未出版		
計畫名稱	程式之模組性與擴充性:驗證方法與工具—子計畫五:剖面導向函數語言之模組化狀態處理與型態擴充研究	NSC97-2221-E-004-001-MY3 (三年期第三年, 99)
<input checked="" type="checkbox"/> 已出版 (請填寫論文題目、期刊名稱、卷號、頁數、年份或專書(含會議論文集)名稱、頁數、年份、出版社) Kung Chen, Y.C. Chang, D.W. Wang (2010, Mar). "Aspect-oriented design and implementation of adaptable access control for Electronic Medical Records." <i>International Journal of Medical Informatics</i> , Vol. 79, pp. 181-203. Elsevier. (SCI).		
<input type="checkbox"/> 已投稿 (請填寫論文題目及期刊名稱或專書(含會議論文集)名稱) <input type="checkbox"/> 未出版		
計畫名稱	安全多方計算協定描述語言之設計與實作	NSC100-2221-E-004-002-MY3
<input checked="" type="checkbox"/> 已出版 (請填寫論文題目、期刊名稱、卷號、頁數、年份或專書(含會議論文集)名稱、頁數、年份、出版社) Kung Chen, Tsan-Sheng Hsu, Churn-Jung Liao and Da-Wei Wang (2013, Jul). "A Scripting Language for Automating Secure Multiparty Computation." The 8 <sup>th</sup> Asia Joint Conference on Information Security (AsiaJCIS 2013), Seoul, Korea.		
<input type="checkbox"/> 已投稿 (請填寫論文題目及期刊名稱或專書(含會議論文集)名稱) <input type="checkbox"/> 未出版		

計畫名稱	SaaS 環境下多租戶應用程式的 SQL 改寫機制：理論與實務	MOST 103-2221-E-004-005
<input checked="" type="checkbox"/> 已出版（請填寫論文題目、期刊名稱、卷號、頁數、年份或專書(含會議論文集)名稱、頁數、年份、出版社） Chung-Feng Liao, <i>Kung Chen</i> , Jiu-Jye Chen. “A Service Framework for Multi-Tenant Enterprise Application in SaaS Environments.” International Conference on Software Engineering and Applications (ICSOFT-EA), Vienna, Austria, 2014. <input type="checkbox"/> 已投稿（請填寫論文題目及期刊名稱或專書(含會議論文集)名稱） <input type="checkbox"/> 未出版		