

NDC-104-035-003（委託研析報告）

**政府巨量資料分析與政策端應用
效能提升之研析**

**國家發展委員會編印
中華民國 105 年 3 月**

NDC-104-035-003（委託研析報告）

政府巨量資料分析與政策端應用 效能提升之研析

受委託單位：電子治理研究中心

研究主持人：陳敦源

研究主持人：蕭乃沂、廖洲棚、陳恭

研究助理：陳揚中、林威志、吳昱明、蔡宇祥

國家發展委員會編印

中華民國 105 年 3 月

目次

目次	I
表次	III
圖次	VII
提要	IX
Abstract.....	XIII
第一章	緒論
第一節	計畫背景與動機
第二節	研析目的
第二章	文獻檢閱
第一節	企業與政府運用巨量資料分析技術之經驗
第二節	政府運用巨量資料分析有效應用之關鍵：組織議題管理的導入
第三節	群眾外包與政府治理
第四節	政府服務網站分析之技術與應用
第三章	研究設計
第一節	政府跨部門巨量資料分析
第二節	網路輿情分析平台與群眾外包之實驗設計
第三節	政府服務網之網站分析技術與應用
第四章	跨部門巨量資料分析之研析成果
第一節	資料治理與應用
第二節	巨量資料分析的特質及預期效益
第三節	政府機關導入巨量資料的搭配資源與因素
第四節	跨政府機關導入巨量資料的相關議題
第五章	網路輿情分析平台與群眾外包之實驗成果
第一節	群眾外包輿情分析之實驗流程以及平台架設過程

政府巨量資料分析與政策端應用效能提升之研析

第二節	網路輿情平台之活動推廣與運作概況	89
第三節	群眾與人工之輿情立場判定結果比較分析	95
第四節	群眾輿情判定之機構效應與從眾效應分析	107
第五節	輿情平台之使用者瀏覽行為分析	111
第六章	政府服務網之網站分析技術應用研析成果	123
第一節	政府為民服務網站深度訪談分析	123
第二節	中央政府機關網站服務調查分析	145
第三節	應用 GA 於政府服務網站能提出哪些問題解方	152
第七章	政策建議與結論	175
第一節	政府機關巨量資料分析應用的未來	175
第二節	政府機關巨量資料分析應用的實務建議	181
第三節	政府機關巨量資料分析的後續研析建議	186
參考文獻		189
附錄		199
附錄一、	政府為民服務網站深度訪談結果摘要	199
附錄二、	輿情平台群眾分析結果表	219
附錄三、	跨部門巨量資料分析之訪談結果摘要	235
附錄四、	PyBossa 平台架設之系統文件	243
附錄五、	中央政府各機關網站服務現況調查規劃	257
附錄六、	政府電子治理內涵與品質之前瞻發展焦點座談相關資料	271

表次

表 1：公私部門巨量資料計畫的異同比較.....	10
表 2：商業智慧與分析應用---從巨量資料到巨大影響（節錄）	11
表 3：研究設計摘要表	30
表 4：第一次訪談重點及訪談題綱對照表.....	39
表 5：第二次討論重點及討論議題對照表.....	40
表 6：第三、四次訪談重點及訪談議題對照表	41
表 7：政府機關跨部門巨量資料應用者背景	42
表 8：專家學者焦點座談重點及訪談題綱對照表	43
表 9：專家學者焦點座談與會者背景	44
表 10：網路輿情分析平台實驗研析時程規劃表	79
表 11：不同之常見網路付費廣告比較	90
表 12：輿情平台 FB 廣告成效.....	92
表 13：輿情平台 FB 廣告受眾年齡與成效分析	93
表 14：輿情平台 FB 廣告受眾年齡/性別之成效分析	94
表 15：觀察值處理摘要-群眾判斷結果 * 助理判斷結果.....	97
表 16：群眾判斷結果 * 助理判斷結果交叉表（原始）	98
表 17：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（原始） .	99
表 18：群眾判斷結果 * 助理判斷結果交叉表（合併無法判斷資料）	101
表 19：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（合併無法 判斷資料）	102

表 20：群眾判斷結果 * 助理判斷結果交叉表（只看判斷為支持與反對立場）	104
表 21：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（只看判斷為支持與反對立場）	105
表 22：輿情平台實驗設計之 4 種專案介面	107
表 23：四種專案介面的群眾判定概況	109
表 24：群眾之組內和全體判定結果一致性	110
表 25：公私部門服務網站經營者深度訪談重點及訪談題綱對照表	123
表 26：公私部門服務網站經營者背景	125
表 27：愛料理（iCook）食譜社群網經營者訪談重點整理	128
表 28：財政部電子申報繳稅服務網經營者訪談重點整理	132
表 29：內政部不動產實價登錄網經營者訪談重點整理	134
表 30：法務部全國法規資料庫經營者訪談重點整理	137
表 31：國發會 e 管家 Plus 網站經營者訪談重點整理	141
表 32：公私部門服務網站的經營落差分析	142
表 33：中央政府各機關網站描述性統計	148
表 34：Pearson 相關係數分析結果	150
表 35：網站經營績效相關指標與使用者需求調查之獨立樣本 T 檢結果	152
表 36：AARRR 分析架構與 GA 功能關聯一覽表	154
表 37：中央政府各機關網站經營者的網路流量分析資訊需求	157
表 38：中央政府二級機關網站經營者的網路流量分析資訊需求...	158
表 39：中央政府三級機關網站經營者的網路流量分析資訊需求...	159

表次

表 40：中央政府機關網站使用 GA 服務需求表	170
表 41：政府機關巨量資料推動進程	182

政府巨量資料分析與政策端應用效能提升之研析

圖次

圖 1：巨量資料分析 vs. 政策應用需求落差圖	15
圖 2：全方位的落差分析概念圖	33
圖 3：我國人事政策與公務人力資源管理組織體系	36
圖 4：資料治理的組成元素	59
圖 5：群眾外包之網路輿情分析雛形系統	74
圖 6：實驗平台之系統架構圖	81
圖 7：群眾判讀輿情資料的可能順序	83
圖 8：本次實驗之主要流程圖	85
圖 9：輿情立場判定之任務執行介面	86
圖 10：提供判讀結果統計的專案執行介面	88
圖 11：群眾判斷結果 * 助理判斷之長條圖	100
圖 12：群眾判斷結果 * 助理判斷結果長條圖（合併無法判斷資料）	103
圖 13：群眾判斷結果 * 助理判斷結果長條圖（只看判斷為支持與反對立場）	106
圖 14：使用者進入平台網站之管道	112
圖 15：使用者之性別數量與比例分析	113
圖 16：使用者所處網域之地區分布	114
圖 17：使用者停留時間以及跳出率	115
圖 18：使用者連線之網路瀏覽器分析	116
圖 19：使用者利用網站服務之紀錄	117

圖 20：不重複之新網站訪客比率	118
圖 21：網站使用者操作網站之過程記錄.....	120
圖 22：愛料理網站首頁	127
圖 23：財政部電子申報繳稅服務網首頁.....	131
圖 24：內政部不動產實價登錄網首頁	134
圖 25：法務部全國法規資料庫首頁	137
圖 26：國發會 e 管家 Plus 首頁.....	140
圖 27：訪客到站來源範例	162
圖 28：訪客性別範例	163
圖 29：訪客居住地範例	163
圖 30：訪客停留時間範例	164
圖 31：訪客使用瀏覽器類型分析範例	164
圖 32：最熱門網頁分析範例.....	165
圖 33：網站回應速度分析範例.....	166
圖 34：訪客再訪率分析範例.....	167
圖 35：訪客類型區隔分析範例.....	168
圖 36：訪客到訪流程分析範例.....	169

提要

一、計畫緣起與目的

巨量資料不論在公私部門的管理領域，掀起一陣旋風，但是，公部門的應用，不論從內部與外部的管理，都還有創新的瓶頸，其中最關鍵的議題，就在於離開表面上炫麗的個案，資料分析與公共治理專業的對話與深度合作，縮小「分析出來沒有價值」以及「有需求但作不出來」的雙重落差，本計畫選擇三個方向進行前述消弭「落差」的相關研析：第一、網路輿情分析平台的群眾外包應用、第二、跨部門巨量資料庫整合與應用—考試院與人事行政總處、第三、政府服務網站記錄品質(log)的標竿分析等，這三項研析工作基本上都是應用落差分析(gap analysis)來彰顯政府在巨量資料時代，可以努力走向以數據分析作各種內外政策管理決策的基礎之一。

二、研析方法及流程

本計畫的三個研析方向，將分別以文獻回顧、實驗應用、深度訪談與標竿學習等方法達成目標；基本上，三個研析方向各有不同的分工與流程；首先，跨部門巨量資料庫整合的部分，研析上將先經過深度訪談，接著進入資料庫整合分析，找出政府應用巨量資料精神整合跨部門資料的落差所在，接著輔以人力資源管理相關應用作為經驗探索，以期能形塑巨量資料時代公部門人力資源管理的新模式；再者，輿情分析的群眾外包，本計畫將以去年研析自由經貿區所掌握的資料，找出政府應用與資料分析的落差，建構一個可以操作的群眾外包網路平台，進行相關實驗，該實驗的結果可以用來評估政府網路輿情分析中之立場分析應用群眾外包的可行性與限制；最後，政府服務網站分析之品質和技術的落差分析，將以計畫團隊與國發會共同選擇之大型政府服務網站的管理者訪談開始，並且選定一個企業服務網站作為標竿，找出兩者之間的落差。此外，並利用發放網路問卷的方是針對我

國中央三級機關的網站管理人員進行問卷調查。最後根據上述方法所得之結果，試著對政府服務網頁記錄分析提出實務應用建議。

三、計畫成果

透過本計畫進行的「跨部門巨量資料分析」、「群眾外包應用在網路輿情分析」、以及「政府服務網站的網站log分析」等三個議題研析發現，本計畫可以歸納整理出兩個彼此獨立卻也相互關連的結論，包括：（一）政府巨量資料應用落差與（二）技術導入的分析結果。在政府巨量資料應用落差上，配合理論與此次三個議題的研析發現，分別針對決策應用的落差、資料治理的落差、以及議題導向的落差進行討論。而在技術導入的分析結果上，承接上述發現之決策應用、資料治理、以及議題導向等落差，本計畫再進一步提出對政府應用巨量資料導入跨域、群眾外包、以及網站log分析上的研析結果發現。

四、政策建議

對於實務政策建議，本計畫也針對三個議題的研析發現提出建議。（一）就跨部門巨量資料分析上：試著規劃政府機關推動巨量資料推動建議與進程，以應用分析的範圍配合業務領域，提出4種巨量資料推動於政府機關的可能類型。（二）群眾外包應用於網路輿情分析：以此次研析操作輿情分析平台的經驗從起初的平台建立、過程中的運作管理以及對引入群眾外包應用於輿情分析上提出建議。（三）政府服務網站的網站log分析：本計畫分析並歸納政府網站效能提升的決策支援資料，並發現這個部份因為公私部門的差異關係，過去在政府部門服務品質提升上並沒有受到應有的重視。因此，此部分主要針對網站服務效能評估、網站服務的品質提升、以及網站營運的策略分析三面項提出建議。

除上述外，本計畫認為後續政府機關巨量資料分析之研析，應視其應用案例主動尋求與相關學術領域的研究團隊合作，透過行動學習與行動研究（action learning and action research）途徑累積實作經驗而

提出建議。而就績效管理，建議在初步導入階段主軸應以：鼓勵試辦與不論是成功或挫折經驗的分享與推廣。最後，目前我國政府機關推動巨量資料分析應用的模式，仍有加強與外部技術與學術團隊合作交流的空間，或考量透過開放政府資料（Open Government Data, OGD）徵求群眾智慧、過程中以log分析確認政府與民眾的连接，並且最後應用回饋刺激政府內部的資料治理流程不斷成熟。

後續研究的建議包括：第一，政府資料治理的跨議題與跨部門機制原型（prototype），讓政府的資料分享架構可以實現「無界限」（borderless）的可能；第二，網路輿情分析的最後一哩路－「輿情論述萃取機制」（Internet Arguments Extraction Mechanism, ITEAM）的研究，與政府現有的政策行銷與公關部門的工作內容接軌。第三，「全觀式外部輿情資訊系統」（Holistic External Opinion System, HEROS）架構之研究，整合政府外部輿情資訊的多元管道，成為善於回應民眾意見的治理實體。

關鍵詞：巨量資料分析、網路輿情分析、跨部門、政府服務網站、網站記錄（log）分析、實驗設計

Abstract

In the era of big data, government needs to introduce the technique of data science as well as is obligated to detect and grasp the gap existed between the expectations of public governance and the reality big data technique can offer to solve problems of public governance. In this research, we utilize gap analysis in three chosen areas of applying big data technique to improve government responsiveness and performance. They are (1) an experimental content analysis with analyzing internet opinion by utilizing crowd-sourcing technique; (2) an application of boundary-spanning database integration in fulfilling public sector human resource management needs; and (3) a government webpage log analysis and benchmarking research to improve government web-based services. Based on the three petite-case researches, authors intent to explore and suggest possible solutions to the “gaps” of applying big data in helping government to better governance.

Key Words: Big Data Analysis, Internet Opinion Analysis; Boundary-spanning Management, Government Service Website, Website Log Analysis, Experimental Design

政府巨量資料分析與政策端應用效能提升之研析

第一章 緒論

第一節 計畫背景與動機

不論是流行還是工具的成熟使然，巨量資料（big data）分析已經成為公私部門組織的創新改革的利器之一；我國政府於2015年公佈「*ide@ Taiwan 2020*（創意臺灣）政策白皮書」之後，政府組織如何善用兼具Volume（數據資料的大量）、Velocity（資料分析的時效）、Variety（資料格式的多樣）、與Veracity（資料內容的真實）所謂4V特質的巨量資料分析能力（Desouza & Jacob, 2014; Dumbill, 2012; Zikopoulos et al., 2013），在政策e化參與，以及數位政府服務等面向，政府各機關已經積極展開各種應用的工作，包括網路輿情分析、行政內部資料的巨量資料分析、以及公共服務「物聯網」的規劃和應用等，然而，由於這種技術應用的效能，需要創造出具備跨專業領域技術應用的雙向回饋機制，才能真正落實巨量資料在政府機關的應用實效，本計畫以公共政策的「規劃、執行與評估」的三階段論出發，從政策端的「需求」出發，先清楚定義巨量資料應用的政策目標，再從這目標中尋找結合巨量資料分析與政策端應用效能提升的連接點所在，進一步開發所需要的分析技術組合，如此才能真正落實巨量資料在公部門的應用目的。

舉例而言，目前巨量資料分析網路輿情的技術，與傳統政府民意搜集的機制比較起來，基本上可以補強政府對於「婉君」資料的缺空，目前分析方法重點，主要是建構社群網站的言論資料庫，依照政府各機關的「關鍵字」需求，用資料探勘技術萃取出相關網路言論資料，進行長期「討論量」、「正負情緒（電腦自動分析）」、「言論來源」、以及「關鍵字暴風圈」的趨勢分析。然而，這些資料對於政府運作的公共政策端，需要政策立場的分析、政策論證內容的掌握、以及如何與其他輿情蒐集管道整合建議的需求，仍然有一段距離。當然，內部

的行政巨量資料分析（含物聯網的內部資料），也有相同的提升應用效能的問題，包括政府網頁log的上線民眾行為分析、行政資料庫民眾使用內涵分析、以及物連網民眾使用或觸動「感測器」（sensor）的分析等，要如何與政府政策端的目的進行相關連結。

第二節 研析目的

依據上述背景與核心問題，本計畫以下列三點作為研析目的，同時也據以區分研析活動的階段：

1. 巨量資料分析與政策應用需求的「落差」（gap）分析：找出已經導入巨量資料分析的政府單位，從應用效能的角度，尋找「落差」存在的可能因素，可能包括技術可能性、資料取得、專業跨域、政策目標不明確、人工編碼需求未導入等問題。

2. 導入政府巨量資料應用政策效能提升方案：分別針對政府機關內部的服務系統資料及其外部的網路輿情，以特定政府服務系統與公共議題為範圍，與主管機關合作導入「巨量資料應用政策效能提升方案」，過程中計畫團隊將關注下面三個議題：（1）資料處理專業與公共政策專業之間，針對巨量資料的技術應用的溝通落差在哪裡？（2）在資料技術發展的過程中，如何將「政策論述」納入分析核心，「質量並重」、「人機整合」分析框架的重點為何？（3）巨量資料分析的結果，要如何與其他政府內部既存的政策資訊蒐集機制相互結合？

3. 巨量資料應用政策效能提升的策略建議：以上述的各國發展趨勢與我國政府機關導入經驗為基礎，針對巨量資料分析應用於我國政府的決策分析與政策規劃提出建議，其範圍包括：（1）政府資料收集結構如何離開行政方便主導，走向巨量資料時代的行政系統設計與應用，也就是如何擴大可蒐集資料的應用範圍，這部分可能來自於物聯網（Internet Of Things, IOT）的應用；（2）在政府政策應用端不變的前提下，巨量分析政策產出是否能夠更以政策而非技術導向為核心，

也就是巨量資料分析產出，可以協助擴大政策應用與資料分析交疊的範圍；（3）在巨量資料分析技術不變與可用資料有限的前提下，政府政策需求端要如何依照巨量資料時代釐清其目標與論述，以致於可以更有效地應用巨量資料的技術，完成政策目標。

政府巨量資料分析與政策端應用效能提升之研析

第二章 文獻檢閱

二十世紀末1980年開始的網際網路普化以及之後儲存設備的快速發展，人類生活的各個層面因著各式組織依照行政目的的蒐集，以及二十一世紀初期社群網站興起之後的Web 2.0網路環境下，個人自願上傳與記錄的結果，各式各樣來源的資料，包括結構性或非結構性、文字或非文字、程式或普通語言的資料大量累積，加上某些個案資料分析的新聞性所吸引。比方說，Google用其搜尋引擎資料分析和預測流行病（請參Lazer and others, 2014），人類社會進入一個所謂「巨量資料」（Big Data）時代，以資訊科技與管理學門引領的各種資料分析技術的發展，包括運用分析技術分析所具有大量（Volume）、快速（Velocity）、多元（Variety）與真實（Veracity）的資料，越來越受到公、私業界及學界的重視；然而，由於資料特性難以用傳統分析方法進行分析，必須用進階的技術和演算法來解讀、儲存、分析與管理，當然，不論是企業還是政府組織，都不能不注意這樣的發展。接下來，我們先來看巨量資料的技術在企業與政府部門的運用現況，最後再針對其計畫內容與分析應用上的異同，進行比較。

第一節 企業與政府運用巨量資料分析技術之經驗

巨量資料分析牽涉到資料形式與分析應用等兩個方面，而企業與政府在巨量資料導入工作上的比較，將有助於在政府部門內的研究或實務者，選擇不同資料與分析應用計畫時的重要參考。以下先就企業與政府的應用面進行討論，接著再比較企業與政府在資料形式與分析應用上的異同。

一、企業方面

首先企業界對於運用新科技增加績效及獲利一向有著強大的動機，在商業界如Google、Amazon與Facebook等業者，已經利用巨量資料分析開發出新的服務模式和創新的經營型態，在行銷上也都整合巨量資料分析產出高附加價值的經營模式¹。最早使用巨量資料分析的是幾位Google工程師在科學期刊（*Natural*）中發表關於流感的研究，他們從Google將近30億筆搜尋關鍵字中，篩選出常用的五千萬筆搜尋關鍵字，再與美國疾病管理局（Centers for Disease Control, CDC）在2003到2008年的實際流感感染案例進行比對，發現經過數學演算法（Algorithm）進行的相關性（correlation）分析與美國疾病管理局官方資料相當吻合（林俊宏譯，2013）。

前述經過數學演算法與大量資料的比對分析，資訊價值產生了突破性的發展，巨量資料分析技術提供了企業部門發展與生存的新利基。由Google的經驗可知，單單只有儲存資料量龐大是沒有任何用處的，因為資料並不會自己進行分析，任何一位企業高階領導者都想知道他們究竟能從大量的資訊中得到什麼價值，而新技術的出現雖然比以前蒐集資料容易了許多，許多企業仍持續地尋找更好的方式去從他們的資料中取得價值，以利於在市場上與其他企業競爭（LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011）。另外，在巨量資料的運用上如果結合了地理資訊系統（Geographic Information System, GIS）後，又大大的增加了GIS應用的潛力，位置訊息也是巨量資料分析一個非常重要的分析標的，加上行動裝置的普及，除了「人」的位置資訊以外，「物體」的位置也可以利用這類的裝置進行大量的蒐集，這些訊息也成為重要的分析資料，用戶的地理位置變成非常寶貴的資料（林俊宏譯，2013）。

¹ 資料來源：

http://www.digitimes.com.tw/tw/dt/n/shwnws.asp?cnlid=13&packageid=8326&id=0000371025_K8R1HK2V1NWQAC67YFO4N&cat=。檢閱日期：2014年4月26日。

最後，舉例而言，Xiang, Schwartz, Gerdes, Jr 與 Uysal (2015) 運用巨量資料中的文字分析 (text analytics) 的方法，分析 Expedia.com 這個世界上最大的旅遊網上顧客回應的資料，先用單字擷取的方法找出80個最常被使用的單字，再用因素分析出包括服務態度在內的六個顧客關心的消費面向，最後再用顧客的滿意度作為依變數，這六個面向作為自變數，進行簡單回歸分析找出影響旅館顧客滿意度最重要的面向為何，這樣的資料分析當然會對管理階層在服務變革的方向上造成重大影響，進而影響到商業市場的競爭模式。

二、政府方面：

隨著巨量資料的分析運用在企業界蓬勃發展，政府機關也逐漸開始運用巨量資料以追求良善治理 (good governance)；首先，白宮在2012年公告了一份行政機關運用數位資料的「概念性架構圖」，要求在進行資料的使用時必須遵守四項要點：(1) 以資訊為中心 (Information centric movement)、(2) 以顧客為中心 (Customer-centric)、(3) 共享平台 (Shared platforms) 與 (4) 安全與隱私 (Security and privacy)。學者Shindelar (2014) 分別以美國住房與城市發展部 (Department of Housing and Urban Development, HUD)、聯邦緊急事務管理署 (Federal Emergency Management Agency, FEMA) 與國家海洋與大氣管理局 (National Oceanic and Atmospheric Administration, NOAA) 三個機關為例，說明在使用巨量資料分析時政府機關的共同需求和個別要件。因為巨量資料對於機關的顧客而言是非常有用的資訊，管理者必須在符合白宮要求的條件下，提供最佳的資訊內容，來達到更精進的公共服務內容。另外，美國有國家科學基金會 (National Science Foundation, NSF)、國家衛生研究所 (National Institutes of Health, NIH)、國防部 (Department of Defense, DoD)、國防部先進研究計畫機構 (Defense Advanced Research Projects Agency, DARPA)、能源部 (Department of Energy, DoE)、美國地質調查局 (United States

Geological Survey, USGS)等六個聯邦機構²共同參與白宮所發起的「巨量資料研究與發展倡議」(Big Data Research and Development Initiative)計畫,其目的是提升巨量資料分析工具與技術以增進美國國防安全、教育、生物醫學等國家能力(王豐勝、黃彥文,2013)。澳洲政府亦於2013年3月成立「資料分析卓越中心」(Data Analytics Centre of Excellence),進行跨單位資訊、技能的共享,以發展具全局性的政府政策(劉龍龍、葉乃嘉、何志宏、余孝先,2013)。

再者,台灣政府也正在研擬未來對於巨量資料的應用,於2012年起正式將「政府公開性資訊(open data)開放」³來達到強化民眾監督與公私資訊流通,以發展新的加值服務與商業模式,最後於2015年公佈「ide@ Taiwan 2020(創意臺灣)政策白皮書」,更具體推動相關網路政策。不論在防災救護、公共衛生、交通分流資訊、水資源運用與長期照護等等公共政策都可以將GIS結合巨量資料分析,以精進公共服務過程(林蔚君、詹雅慧,2013)。另外,詹大千與金傳春(2013)也針對GIS系統應用在流行病學領域上進行討論,認為行動載具的普及與社群媒體的發展,對於流行病的防治與監測,可以利用GIS與巨量資料的應用來得到較過去權威式公共衛生治理更佳的防治效果;而台灣電子治理研究中心的團隊,近年也努力將巨量資料分析導入政府的第一線服務價值提升、政府網頁的運作資料分析、以及網路輿情分析等面向(廖洲棚、陳敦源、蕭乃沂、廖興中,2013;蕭乃沂、陳敦源、廖洲棚,2014)。

事實上,已有若干國家的政府開始重視前述巨量資料分析的潛力,並逐步將其應用在下面三個方面:首先,應用巨量資料分析在公共的決策。例如:美國衛生人力部(Department of Health and Human

² 資料來源:

http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf。瀏覽日期:2014年5月21日。

³ 資料來源:101年1月18日行政院科技會報辦公室舉辦「我國公開資料加值(Open Data)推動策略會議」,網址:

<http://www.ndc.gov.tw/m1.aspx?sNo=0027913>。檢閱日期:2014年4月26日。

Services) 利用巨量資料分析過去行政訴訟的資料，藉以改善其決策品質；美國國家衛生研究院 (National Institutes of Health) 則是透過建立病患巨量資料系統，協助研究人員儲存與分析病患的健康數據；美國能源部 (Department of Energy) 則是建立巨量資料的平台，收集國際間相關的大氣數據，藉以提升對全球大氣過程與氣候模型的瞭解；英國 Age UK 慈善機構，運用巨量資料分析發掘郵寄資料和捐贈金額本益比最高的潛在捐贈者，除節省郵寄資料的費用開支外，並為其設計適當的宣傳活動，以提高募款成功機率。

此外，巨量資料也被大量應用在了解不同政策領域的民意結構 (public opinion)。包括從 Twitter 上的意見，了解到美國國內傳統支持共和黨的州內，選民在 Twitter 上討論氣候變遷的議題時，比起民主黨掌握的州以及其他西方國家的討論，傳統支持共和黨州內的討論對科學證據是較為懷疑的，這可能可以解釋美國為何在全球各國努力降低二氧化碳排放的進程中，美國總是慢半拍的原因 (Jang and Hart, 2015)；另外，美國的太空政策須要大量的政府資源挹注，因此，民眾是否對太空探險抱持熱情成為太空政策能否存續的重要基礎，學者 Cobb (2015) 應用 Google Trends 與 Twitter 上民眾發言是否包含「太空探險」(space exploration) 一詞進行分析，可以看出這個用詞的區域差異以及下降趨勢，與民調的趨勢比起來十分類似。

最後，巨量資料分析也被應用在交通政策的層面分析，兩位土木工程學者 Shi 與 Abdel-Aty (2015)，應用物聯網 (Internet of Things, IOT) 的概念在美國一條城市高速公路上裝設大量的交通資訊蒐集裝置，再將回收的巨量資料進行各種統計分析，最後找出交通尖峰時間降低追撞車禍的宣導最佳地點與時間，有效增加城市高速公路在尖峰時間的安全性。

三、 企業與政府比較：

三位巨量資料研究者Kim, Trimi 與 Chung (2014) 認為，巨量資料的應用在公私部門的運作是有其根本的差異，這差異可以從各種巨量資料的計畫之目標、願景、決策、決策者、組織、資源、以及活動本質等七個面向看出，如表 1所示。

表 1：公私部門巨量資料計畫的異同比較

比較特徵	企業部門	政府部門
目標	利害關係人的獲利	國內的平安與永續發展
願景	競爭態勢的發展與顧客滿意	基本需求的保障（平等、自由與公義），促進一般福祉以及經濟發展
決策	短期的自利決策，主要在極大化收益與縮小成本	長期的決策同時滿足個人與公共利益
決策者	有限的成員參與	廣大的多元民眾都可參與
組織架構	層級節制的組織	治理的組織
資源	收益	政府稅收
活動本質	競爭與接觸	合作與制衡

資料來源：Kim and others (2014: 80)。

另外，Chen、Chiang 與 Storey (2012) 從商業經營與應用的角度出發，針對下列五種研究領域整理出不同的資料應用方法與技術：（1）電子商務與市場行銷（e-commerce and market intelligence）；（2）電子化政府與政治2.0（e-government and politics 2.0）；（3）科學與科技（science and technology）；（4）智慧型醫療與安寧照護（smart health and well-being）；（5）公共安全（security and public safety）。本計

畫從商業界與政府角度出發，在面臨各種不同的資料類型時，能夠以何種資料分析方法進行分析，如表 2 所示。

表 2：商業智慧與分析應用---從巨量資料到巨大影響（節錄）

	應用範圍	資料類型	資料特性	分析	影響
電子商務與市場行銷	1. 推薦系統 2. 社群媒體觀測與分析 3. 群眾分析系統 4. 社群媒體虛擬小遊戲	1. 關鍵字搜尋與登入日誌 2. 顧客交易紀錄 3. 顧客自創內容	結構化的網頁資料、顧客自創內容資料、豐富的網際網路資料與非結構化的顧客意見資訊	1. 關聯規則探勘 2. 資料庫集群分析 3. 異例偵測系統 4. 圖像探勘 5. 社群網絡分析 6. 文本與網頁分析 7. 語意與影響分析	1. 長尾行銷 2. 個人化與針對性行銷 3. 更高的顧客與銷售滿意度
電子化政府與政治 2.0	1. 無所不在的政府服務。 2. 平等接觸與公共服務 3. 公民接觸與參與 4. 政治競選與電子投票	1. 政府資訊與服務 2. 法令規章 3. 公民意見反饋及評論	分散的資料來源以及老舊資訊系統、豐富的文字內容、非結構化的且非正式的公民對話	1. 資訊整合 2. 內容與文字分析 3. 政府資訊語意服務與知識本體分析 4. 社群媒體監控與分析 5. 社群網絡分析 6. 語意及影響分析	1. 轉型政府 2. 授權公民 3. 改善透明、參與及平等

資料來源：修改自 Chen, Chiang and Storey (2012: 9)。

接著，三位來自歐洲與美洲的公共行政學者 Criado, Sandoval-Almazan 以及 Ramon Gil-Garcia (2014)，經過整理近年興起的政府社群媒體研究，發現這些研究雖然處在一個初步發展的階段，但是可以分為工具使用、應用目標、以及研究主題等三個面向，另外，他們也提出下面七個未來研究的大方向：(1) 社群媒體的創新與制度/組織變革；(2) 測量社群媒體在政府中的影響；(3) 新的研究與分析技術開發；(4) 社群媒體的需求面分析；(5) 公民接觸、參與、和合產 (co-production)；(6) 社群媒體、開放政府與巨量資料：智慧政府；(7) 社群媒體的限制與挑戰，本計畫主要落在 (1)、(2)、(5) 與 (6) 等四項方向中，主要是調和各種政府科技的應用與傳統政府運作領域，建構更具知識應用背景、且能預先介入的智慧型政府 (anticipatory and smart government) (Milakovich, 2013)。

最後，在運用巨量資料分析時，行政部門所面臨的限制將會較企業來得多。廖洲棚等 (2013) 的研究指出行政部門在運用巨量資料之前，必須先具備四項能力與條件：(1) 技術能力、(2) 跨業務協調能力、(3) 規畫與分析能力與 (4) 相關法規掌握能力。行政部門在推行巨量分析技術之前，除了基本的科技基本設備與技術能力必須先行具備外，更須思考相關資料蒐集的手段是否牴觸現行法律的規定。為此，廖洲棚等人 (2013) 認為行政部門在推行巨量資料可參照國外 (尤其是美國) 的經驗，鑑往知來可成為臺灣內部推行時的標竿，並且加強行政人員對於巨量資料的認知，強化行政人員新科技重要性的體認，再者增加與研究機構與企業界的合作，將成功推行巨量資料分析技術之經驗進行交流，減低行政人員的不安與疑慮。

第二節 政府運用巨量資料分析有效應用之關鍵：組織議題管理的導入

藉由巨量資料分析，可幫助政府部門發掘潛藏在資料中的重要資訊，協助管理者改善服務遞送流程，降低遞送成本，以提高服務效能

和滿足服務對象的需要（Australian Government Information Office, 2013）。Yiu（2012）認為巨量資料分析有以下五種可應用於公部門的面向，包括：分享（sharing）、學習（learning）、個人化（personalizing）、解決（solving）以及成長創新（innovating for growth）等五項。首先，在分享方面，由於政府機關組織的數量與性質十分龐雜，各自所擁有的運作知識或資訊相當豐富。透過巨量資料的平台，可以有效地協助政府各機關分享彼此的資訊，提升其解決問題的能力並縮減反應民眾需求的時間。其次，在學習的面向上，管理階層可以透過巨量資料的分析，發掘單位在管理過程中未發現的可能問題或重要關鍵，進而針對可能的方向提出修正與調整。在個人化方面，則強調當服務提供者透過巨量資料分析發現特定民眾的特質，以便針對其個人的特質與需求進行客製化的微調。在問題解決方面，巨量資料一方面可以針對過去大量的資料進行分析並發掘隱藏其中的可能樣態或趨勢，進而提出相對應的解決方案。此外，針對這些過往的資料進行分析，可協助政府對於未來的趨勢進行預測，並研擬可能的因應對策。最後，在成長創新方面，利用巨量資料分析可以發掘以往政府運作浪費或低效率的部分，俾透過資源重新分配的方式，達到整體績效最大化的程度。

當然，巨量資料分析方法有下列三點獨特之處，並因此改變我們理解及組織社會的方式：第一，大改變是能夠取得、分析的資料量大為增加，使用所有資料分析可以讓我們清楚看見最細緻的地方；第二，大改變是不會堅持一切都要做到精準，因為巨量資料分析雖減少了抽樣造成的誤差，卻必須對於測量上的誤差給予些許妥協，而放棄百分之百的精確；第三，大改變是放下長久以來對於因果關係的堅持，而專注於發現事物的相關性，只找尋事情「正是如此」的答案，而不一定要瞭解某件事「為何如此」（林俊宏譯，2013：22-24）；然而，這個「正面與樂觀」的應用圖像，並不能改變政府真正存在的目的，就是解決公共問題，如果這新興的資料分析方法，無法從根本解決公共問題，長期來說，可能只是另一個改革風潮而已。

因此，從政府電子治理的角度來看，研究者與實務工作者有義務要挖掘這個巨量風潮中許多潛藏且需要被解決的問題，本計畫團隊認

為，其中最重要的就是對於巨量資料可以用「資料相關性」跳過理論意圖的問題，這個議題導因於Wired Magazine的主編Chris Anderson在2008年的一篇短文中曾經這樣說：「巨大的資料量讓我們可以這樣說：『相關性就足夠了...』，我們可以將這些資料扔進這世界從來沒有看過巨大的電腦連線組合中，讓統計運算自己找到沒有其他科學研究者見過的模式」⁴。

然而，也因這樣的過度樂觀且科學主義傾向的思維邏輯，使得政治學者針對近年許多資料分析師應用網路意見的情緒分析來預測選舉結果，提出循證的質疑（Clark & Golder, 2015; Gayo-Avello, 2011），社群網站的民意分析，不論在代表性、因果關係、甚至討論範圍都有方法論上的偏誤存在，因此，這些網路輿情資料與傳統民意調查比較起來，並無法真正預測或推論真正選舉的結果，更重要的，政府解決公共問題的作為，必須基於「對症下藥」的因果關係，而不只是表面的「退燒藥」作用而已；因此，公共行政學者Desouza 與 Jacob（2014: 16）認為：「即便巨量資料分析的確提供有價值的內容給政府每日的運作參考，但是公共管理者與決策者必須要能確認，他們不能因為這些日常運作資料分析，就忽略掉去處理更廣泛且複雜的公共問題。」

從圖 1當中，政府部門應用巨量資料分析要如何真正解決廣泛而複雜的公共問題關鍵，應該從不同的領域進行深入的思考，本計畫聚焦在政府不論在管理或是政策上的「回應性」（responsiveness; 陳敦源，2009）價值，從政府存在解決公共問題的角度出發，以服務輸送過程以及網路政策討論等兩種「公眾接觸」（civic engagement; 陳敦源等，2007）的互動類型所留下來的數位資料為基礎，進行巨量資料的應用研究，這樣的研究主題，包含理論與實務的雙重任務，在理論上，巨量資料的導入，對政府內部的組織與管理，會產生甚麼樣的效果；另外，在實務上，這些可能的效果對於政府治理需求的滿足，是

⁴ 原文如下：“Petabytes allow us to say: ‘Correlation is enough’ ... We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot.”

否存在應該注意的落差？這些落差要如何整合學術研究與實務操作，來共同消彌之。

因此，本計畫主要的理論應用範圍在於下面三點，包括「議題管理」、「數位內容分析」與「巨量治理」等，這三點議題的討論，將有助於政府在巨量資料的風潮當中，不致於迷失在過度樂觀的科技改造論，又能夠有足夠的理論深度，與實務結合，進行有效的政府行政改革：

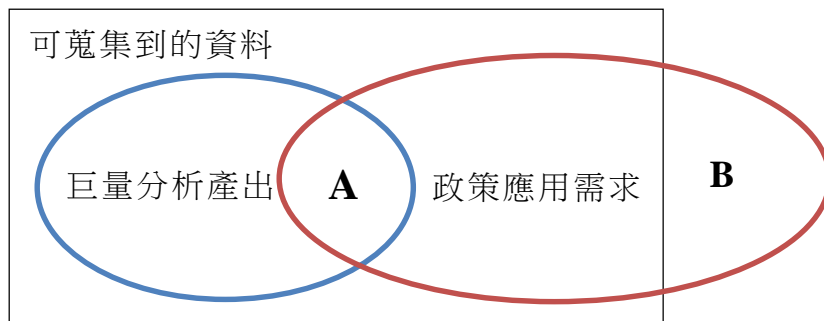


圖 1：巨量資料分析 vs. 政策應用需求落差圖

資料來源：本計畫自行繪製。

一、 議題管理：

以「政策需求的目標」而非「巨量資料的分析」為應用核心：政府部門在導入巨量資料分析以前，應先釐清公共問題的內涵，並且據此引導巨量資料的分析應用的方向，基本上，巨量資料分析的工具性價值必須受到公共政策問題解決的本質性價值的領導，本計畫當中，將會以公部門「議題管理」（issue management; 吳宜蓁，1998）的政策需求為主軸，議題管理是一種組織對外關係的管理系統，最早由學者Ansoff（1980）年提出，主要功能是對與組織生存發展有關的議題，進行「提早偵測」（early identification）與「快速回應」（fast response）的管理活動，其中外部的趨勢與議題的「發掘-描繪-測量」功能

(surveillance-monitoring-measurement)，學者Perrott (1996)認為，利害關係人(stakeholders; Roloff, 2007)的指認與分析，也是議題管理的一個重要環節，因為議題自己不會引起注意，而是新聞中的利害關係人，議題管理的作為也是組織變革的關鍵之一(Perrott, 2011)；當然，1990年間，由Heath 與 Cusino (1990)提出當代社會各類組織落實議題管理可以達到的功能有四：議題傳播、議題偵測、實踐社會責任、以及策略規劃。另外，國內學者陳恆鈞(2011)指出，議題管理可分為四步驟流程，包括議題界定、議題分析與衝擊性評估、擬定策略行動、與績效評估。不論是議題的發掘與測量、議題分析與衝擊性評估等工作，基本上都是巨量資料分析可以引進的地方，但是重要的問題是，巨量資料分析的應用目的，必須依據組織議題管理的需求而定，而巨量資料分析發展的方向，亦必須以公部門組織議題管理的需求為核心。

二、數位內容分析：

在既有的政策應用需求之上，巨量資料發展的目的，應該是設法擴大圖 1當中「巨量分析資料產生」與「政策應用需求」圖形的交集A。如何彌平分析產出與應用需求之間的「落差」，是政府部門應用巨量資料分析的運作目標，因此，本計畫將從議題管理的角度，導入目前傳播學門為因應新聞媒體網路化的趨勢，所發展出結合「電腦輔助」與「人工編碼」的議題分析工具，進行相關的分析(Lewis and others, 2013; Mahrt and Scharkow, 2013; McMillan, 2000; Zamith and Lewis, 2015)。學者Sjovaag 與 Stavelin (2012)認為，傳統的媒體內容分析，必須要重新設計以適應新聞網路化的趨勢，其中最要的差異在於傳統媒體有空間與時間的限制，但是網路的世界並沒有這些限制，因此，傳統的內容分析必須面對巨量資料處理的效率問題，這部分必須引進「電腦協助資料處理」(computer-assisted data analysis)的方式，但是，他們認為，巨量資料除了應用電腦進行字彙或資料結構分析所獲得快速分析大量資料的效率之外，網路資料背後的社會文化環境的意

義，往往人工的編碼還是很難被完全取代的。另一方面，學者Manvich（2012: 466）認為，機器與人工內容分析的平衡發展，主要是為了建構一個「混同」（hybrid）的數位媒體分析工具，其目的是要讓巨量資料的分析者今後「不用在資料量與資料分析深度兩者之間作出抉擇」，這其中包括如何用電腦進行網路資料中的各類公共議題的「立場」與「論述」的分析，這個部份目前是許多巨量資料分析師努力的重點，比方說，一個受到Microsoft與國際發展銀行（International Development Bank, IDB）經費補助的稱為「決策2.0」（DECIDE 2.0）的計畫，就意圖應用相關最新的分析技術，建構一個可以迅速處理民眾包括網路在內意見的「情資處理系統」（intelligent processing framework; Estevez and others, 2012）；因此，如何應用電腦輔助與人工編碼的任務，分不同階段進行資料統整的分析，也將會是巨量資料分析能否為組織所用的關鍵，本計畫也將針對外部輿情分析的部分，進行混同分析模式的測試，包括找出網路意見的立場和論述的內容，如此才能進一步縮小圖一中的A區域的落差。

三、巨量資料治理：

對一個組織而言，巨量資料分析的最重要價值有二，其一是分析使用（analytical use），巨量資料分析可以揭露資料隱藏的洞見（insight），如顧客間的同儕影響、消費者的交易習慣以及社會及空間資料的關係，這些洞見在過去往往因為資料分析的成本太高而被忽視；其二是開發新產品（enabling new products），巨量資料分析可以即時的處理與分析資料，以發現新的需求而刺激服務的創新（Dumbill, 2012）。因此，本計畫規劃透過對研析標的機關資訊服務系統的交易紀錄分析，除協助該機關瞭解使用者的服務偏好外，更進一步嘗試發掘使用者類型與服務使用習慣間的關聯性，以協助該機關設計更符合特定使用者偏好的服務內容。然而，這個分析的過程並不是組織管理的精隨所在，而是這個分析的方法必須放置在一個更為複雜的「服務導向的決策支援體系」（the Service-Oriented Decision Support System;

SODSS, Demirkan and Delen, 2013），這個系統當中絕對不只於一個巨量分析的計算架構而已；當然，從組織管理的角度出發，巨量資料分析與管理的問題應該要有更好的融合，根據 *Academy of Management Journal* 在 2014 年期刊中的編輯議題文章中，特別強調下面幾個部份：包括對巨量資料的目標進行組織的定位，其中包含巨量資料的來源管理、以及資料分享、隱私權與倫理的議題管理等，前面這個部份，可以看作是圖一當中 B 區域的管理問題，有些組織有興趣的政策管理議題，目前組織並沒有相關的資料分析，因此，如何從組織而非分析師的高度，進行資料收集的變革，包括可能增加數位「感應器」(sensors) 的設計，以便蒐集資料，這個部份可以是物聯網(internet of things, IOT) 的發展範圍；接著，在巨量資料分析的部分，該論文也強調巨量資料分析仍然必須面對組織所在意的因果關係，因此，資料分析單純展現相關性往往無法滿足組織應用的目標，因此，巨量資料的治理的功能，主要還是應該轉換到成為政府績效管理的新工具，近年美國地方政府建構「績效數據公開系統」(PerformanceStat; Behn, 2014) 的風潮，或許是一個值得觀察的發展；最後，該文章中對於相關的研究應該著重在哪些面向，研究者重視的因果關係與管理者所重視的管理有效性仍然是巨量資料分析治理的重點 (George and others, 2014)。

第三節 群眾外包與政府治理

一、 群眾外包的概念與機制

群眾外包 (Crowdsourcing)，在此概念的創發者之一的 Jeff Howe (2006a) 的定義下，其代表的是藉由公開徵求的形式，將過往應由組織或公司員工進行的工作，外包給一群不特定的網絡群眾；而群眾外包的關鍵前提便是，以公開徵求的形式尋求存在於廣大網絡之中的潛在工作者完成其工作。認為，群眾外包的主要概念就是透過簡單、可用與容易使用的介面及平臺，創造了有別於傳統將工作外包給廠商的方式，而將工作透過社群分享的機制，外包給社群組織的成員，並於

時間內交付完成工作的內容（王聖銘，2015：34）但一般而言，群眾外包並非真正的引入群眾的勞動力，而著重在意見、方案的徵求。而在群眾意見的引入（idea input）上，大致可分為兩種形式：一為基於需求的訊息（need-based information）或消費者的意見（voice of the customer）；另一種則為尋求解決方案的資訊（solution-based information），及不僅只是詢問群眾或使用者的問題，而更希望可透過他們得到問題的解決方案（Poetz and Schreier, 2012: 247）。

雖然，群眾外包在其概念的創始者Jeff Howe與Mark Robinson（2006b）的原始概念下，將其視為一種創新的網路商業模式，以相當於公開徵求建議的形式，藉此掌握那些散落在群眾網絡之中的各種創新性解決方案。但如今其概念上的應用以不僅侷限於商業模式，不同於在網絡之中找尋潛在的員工；群眾外包的概念也被應用在徵求志願者的協助的非營利（non-profit）模式。在Brabham（2008: 75-76）認知下，群眾外包便不僅只能是一種網路商業模式，而將其更廣泛地定義為一種可跨疆界、跨領域的問題解決模式（a model for problem solving）；並認為群眾外包除了商業模式的應用外，也能在非營利的應用層面上，對於世界的健康、社會、環境正義的問題解決具備深遠的影響。

此外，尋求群眾的智慧（Wisdom of crowds），是群眾外包模式得以成功的關鍵，以及運作此模式的關鍵理論。如同「三個臭皮匠勝過一個諸葛亮的概念」，Surowiecki（2004: xiii）發現，在適當的環境下，群眾的合作將能展現非同尋常的智慧，並往往能更勝於群體中最聰明的兩個人。但群眾的智慧，並非尋求群眾意見的平均，甚至認為平均將可能帶來平庸；其認為群眾外包所追求的群眾智慧，應是藉由多元群眾意見的聚合，在不斷精益求精後，藉由眾人的智慧得到最佳的解決方案（Brabham, 2008: 79-80; Surowiecki, 2004）。因此，群眾外包並不僅僅是Web2.0世代的流行詞，而實際上應視其為一種策略形式，以思考如何吸引有興趣、且具動機的群眾有辦法提供我們不論是在質或量上皆更勝於傳統商業模式下的創新方案（Brabham, 2008: 79）。然而也因對於群眾外包有如此之期待，不論為何種形式的群眾外包應用，

將廣大的一般群眾和少數專業在結果產出進行比較（Poetz and Schreier, 2012 ; Borromeo and Toyama, 2015），一直以來皆是探討群眾外包是否具備意義的關鍵問題。

二、 群眾外包與民主治理：供給與需求關係的轉換

群眾外包的機制運用，對於政府治理而言，除了具有上述藉由聚合外部群眾的多元意見，藉由網路平台而進行群眾的腦力激盪，而可能從中獲取更佳的問題解決方案，幫助政府解決不論是社會政策或公部門管理等面臨的各種問題。這樣的觀念與提倡公民參與（citizen participation）以提昇政府民主治理的正當性上，具有相互呼應的可能性。舉例來說，美國的聯邦總務署（General Services Administration , GSA）為了從私部門引入開放創新方案，以解決政府管理的問題而於2010年9月10日推動的「Challenge. gov」（Desouza, 2013）；以及英國中央政府藉由群眾外包的觀念所推動的「Red Tape Challenge」（Lodge, Wegrich, 2014），皆是著名的政府藉由群眾外包機制以尋求更好的政府治理案例。

此外，雖然群眾外包在起初的觀念上，是強調藉由群眾的智慧以得到更好的方案，主要是希望能從群眾之中得到更好的創新概念或問題解決方案，著重在資訊的引入層面。但隨著此概念被廣泛的應用與創新，除了訊息的獲得外，某些同時也將群眾視為具備勞動性或工具性質的群眾外包機制也逐漸的發展。以台灣為例，在2009年8月的莫拉克風災期間，台灣災區內外的民眾便主動的設置各種災害資訊交流平台，補足了政府單位在第一時間資訊應變與處理能力的缺口，而快速地彙整各地的災情資訊（王聖銘，2015：35）。因此，群眾外包的途徑下，政府與民眾之間的「供給—需求」關係，將具備流動性，政府將不再完全扮演供給者的角色，民眾也不全然是接受服務的顧客，相反的，群眾的智慧與行動力，在群眾外包的運作機制下，政府反而可能會是那不斷向群眾提出需求的角色。

然而，群眾外包機制運用於政府治理之上，特別是在強調其民主合法性與公民參與的層面上，如同視電子化政府為提升公民參與的途徑一般，目前必須面臨的關鍵問題之一便在於數位落差(digital divide)的現況與影響。不論是電子化政府或是群眾外包應用於政府治理，其在新媒體技術對於民主化以及公民培力上的所提供的前景，往往被過於高估了(Brabham, 2008: 85)。數位落差對於電子化政府或群眾外包而言，可能產生的最大問題在於僅讓特定的群體較有機會與能力參與其體系。而對於群眾外包而言，數位落差所帶來最大的傷害，便在於基於尋求群眾的智慧以達到更好的問題解決的理論上，我們無法確認我們所公開徵求的群眾，其意見是否具備足夠的多樣性(Brabham, 2008: 85-86)，而群眾來源以及其意見的多樣性不足，將對於群眾外包的成效具有嚴重影響。而我們也必須留意我們藉由群眾外包所產生的那個更好的意見或方案，其過程是否具備足夠的民主性，而非僅由特定的群體意見所產生(Brabham, 2008: 86)。

另外，在英國「Red Tape Challenge」的案例中，Lodge與Wegrich (2014: 18-19)認為群眾外包於此的應用上，仍未擺脫傳統協商(consultation)以及價值權衡(value trade-offs)的問題，也並非造就了一個近乎無成本之提供公共評論與打擊繁文縟節的機制，特別是在該機制運作上，政治的觀點遠強過於理想上的競爭邏輯時。相反的，若對於公民在政府諮詢平台的參與上能給予合適的機會與環境，公民們將會有助於民主的發展以及政府部門的活化(Desouza, 2013: 889)。

因此，若欲將群眾外包視為引入提升政府治理效能，或是獲取更佳的政策方案的途徑時，如何減少數位落差對於群眾參與屬性的偏誤以及參與機會不平等的影響、提高群眾的關注與鼓勵其積極的參與投入、以方案競爭為先而降低政治邏輯的影響、而意見產生與聚合的過程也具備民主性，將是未來必須持續探究的問題。而對於群眾外包的應用上，我們也必須不斷強調其本質，即尋求前在廣大網絡群眾之中的智慧，以及不斷追求更佳的問題解決方案。而政府與民眾之間的供給與需求關係的轉換，也或許將對於未來的民主治理帶來新的樣貌。

第四節 政府服務網站分析之技術與應用

一、政府服務網站於電子化政府治理之供給與需求

我國從民國70年代建立大型行政資訊系統以來即陸續推動電子化政府服務迄今已逾30年，在邁入第四階段電子化政府計畫（101年至105年）的此刻，不論中央政府或地方政府機關幾乎都已設有自己專屬的網站。政府設立網站的目的有二，其一是提供完整的公共資訊來滿足人民知的權利；其二則是透過網站提供人民服務，提高民眾和政府互動的便利性，因此提供民眾需要的「資訊」和「服務」就是政府網站核心的營運目標（Gill, 2012）。

然而，基於政府組織與政府服務的特性，民眾參與使用政府服務網站不論在動機、目的、或需求上，皆與利用民間商業網站有所差異，並不全然僅是主動地、有選擇性地「尋求政府的服務」，而在部分情況下使用政府服務網站，是在不得不、或僅此一家不可有分號的情況下被動地只能使用特定政府網站。因此，相較於商業性資訊科技的採用，對於電子化政府的相關應用必須留意（1）政府角色的平衡、（2）媒介的政治性、（3）使用者包含民主社會中多元的民眾，而可能並非特定的顧客，等三項特殊本質必須予以考量（羅晉，2008：190）。雖然公共電子服務的規劃通常被認為主要是替那些「宅在家」的民眾服務，在電子化政府的脈絡下，其實是必須為了我們所有人（all of us）而發展，而並非像電子商務服務般僅針對特定的消費族群（Axelsson, Melin, and Lindgren, 2010）。因此在平等的可近用性上，將是電子化政府與電子商務服務在發展上的一大區別。

進一步從公共參與的角度論述，羅晉（2008：185-186）認為，民眾參與和使用政府網站的目的，即為民眾可能藉此獲得的參與效用之認知；其並歸納國內外學者對於民眾使用政府網站的目的，將其參與目的概分為（1）參與成本的降低；（2）政府公開資訊的近用（access to information）；（3）公共諮商（public consultation）；（4）決策制定參與（decision making）等四大類型。因此，民眾在利用政府網站時，

除了尋求特定的服務、公開資訊的獲取外，也可藉由參與、使用政府網站的過程，或透過特定的網站功能進行線上的公共參與。

但在電子化政府的政策與策略指引上，以及各國政府在制定電子化政府議程的行動計劃中，著重公民的觀點，而將其作為顧客導向的一部分此種基於公民的立場，雖然是一再被強調的，但這些公民的需求卻往往在電子化政府的實務當中並沒有被回應，而往往在電子化政府的規劃與實施過程中被忽略（Axelsson, Melin, and Lindgren, 2010）。這樣的發現揭示了政府網站作為一公共性媒介的特殊性，所反映的不僅只是民眾不同的特質、科技風險意識、參與需求以及實際媒介使用行為間的互動關係，而深入理解民眾對於政府網站的認知與採用行為，更是作為未來我國政府改善與精進電子化政府的重要依據（羅晉，2008：201）。如果政府能使公民們對於電子化政府的議題感興趣，並發展容易組織和具吸引力的方式去貢獻政府的服務，那麼對於更多元的使用者觀點之需求、要求與挑戰，也將可更完整的被聚集（Axelsson, Melin, and Lindgren, 2010）。

二、網站分析技術作為一種研究方法

從網站搜尋研究的觀點來看，電腦網站的日誌分析（Transaction Log Analysis, TLA）是一種為了瞭解網路使用者的行為，而從大量的使用者進行非侵入式（non-intrusive）的資料蒐集方法（Davis, 2004: 327）。其被定義為藉由網站日誌資料的收集，而進一步調查關於在搜尋事件的過程中，在網站使用者、網站搜尋引擎、或者與網站內容之間互動的相關研究問題；而其目標在於對使用者、內容、以及系統之間的互動，或基於任何研究問題而對於這些結構元素之間的互動能有更清楚的理解，進而得以強化系統設計、改善搜尋服務、或者更確實地識別使用者的使用行為（Jansen, 2006: 409）。不同於使用者滿意度調查或者網站內容評比，網站日誌的分析能夠捕捉到即時的（real-time）數據資料與反饋（feedback），而讓我們能直接觀察與感受到我們所做的任何創新或改變所伴隨的影響（Nicholas, Huntington,

Lievesley, and Withet, 1999: 265)。此種即時且基於使用者於網站系統之行動的資訊蒐集，是網站日誌分析與傳統資料蒐集方法上最大的優勢與差異。

Jansen (2006:409-410)認為，從方法論的角度看待網站搜尋日誌分析 (web search transaction log analysis)，其主要包含了從資料蒐集 (collection) 開始、到資料準備 (preparation)、以及最後的資料分析 (analysis) 此三大階段；並認為該研究途徑因強調源自資料所進行的比較與分析，而其結果係根據對於真實世界 (real word) 的觀察所產生系統性的理論發現，因此在方法論上具備紮根理論 (grounded theory) 的特質。此種研究方法的優勢在於，其提供了一種便利、成本低廉、且可即時地蒐集大量且複雜之使用者與系統互動的相關數據，此外，這種數據蒐集的行為是自動化且具隱蔽性 (unobtrusiveness) 的，因此所得關於使用者與系統間互動的資訊是並未被數據蒐集的行為而影響 (Jansen, 2006: 424)。

然而，當網站日誌資料的蒐集方式隨著人們對於資訊科技的理解以及個人隱私和數位監控的意識程度越高，這種自動化的資料蒐集方式還是否能具隱蔽性，或者甚至可能被視為監控個人的行為時，這樣的優勢是否還存在就變值得探究。雖然目前並非每個人都具有一定的科技知識與能力能躲避網站系統的記錄，但相關的技術門檻確實有逐漸下降的趨勢，例如利用虛擬私有網路 (Virtual Private Network, VPN) 避免實體IP位址被紀錄或者避開防火牆的限制 (俗稱之翻牆)，對於時下年輕人來說，或許早已不是新鮮事。相關的各式應用軟體如今在網路上也是唾手可得。此外，Jansen (2006: 424-425)，也指出網站日誌分析在研究方法上具有蒐集資料的類型受限於系統設置、電腦IP與使用者的不一致、僅記錄使用者於網站系統上的行為，而無法得知其背後的原因與動機、紀錄可能因系統規劃與使用者的使用方式而不完整等限制。

不過Groves (2007)認為，網站日誌檔案對於網站管理員或研究者來說，是關於系統的行為，而並非是使用者的行為，而網站日誌分

析僅僅是一種基於推測而建立的格式中蒐集這些數據，並旨在提供洞察使用者行為的方法。Groves（2007）也提醒我們，對於相關技術知識的理解不足以及缺乏對其方法上限制的認識，將影響我們所蒐集之資訊的可用性以及分析應用該資訊的效用性。但不可否認，就目前而言網站日誌分析確實對於使用者於網站系統的行動資訊的蒐集，不論是在蒐集資料的數量、即時性、便利性、經濟性、複雜性、與有效性上仍具有其獨特的優勢。相較於傳統的主觀意見調查、使用者意見調查，透過網站日誌的分析技術，我們可能精確、便利、且即時地蒐集極大量的使用者於網站系統上的各種行為紀錄，但為了充分理解民眾對於電子化政府在「供給—需求」之間的關係與原因，多元研究方法的配合仍是有其必要的。因此，提升對網站日誌分析此種資訊技術研究方法的理解與應用，對於電子化政府的相關研究上將必然有其助益。

三、消除「落差」以提升電子化政府治理的能力

延續本章前兩節與上述的討論，政府運用巨量資料分析而欲形成更好的治理，必須兼顧技術上巨量資料分析能力的提升，以及在策略上以議題管理的高度積極、主動、且具政策目標性的處理問題。這兩層面的提升分別代表了在技術上如何更有效地對可得之各種巨量資料進行分析而得到有意義的分析產出；而所謂有意義的分析產出，則也同時視乎是否能減少在分析產出對於實際治理需求之間的落差，即能否在策略上將議題管理的策略性思維與巨量資料分析做有效結合。

雖然在這30年來努力下，台灣公共行政學界在電子化政府的領域下，並不乏對於網站服務品質的評量研究。但相對於私部門研究長久以來對於人們在各種新興科技的認知與採用行為之關係的重視，在公共行政相關領域的研究對於憑藉著ICTs與網路技術所構築的電子化政府，對其「供給—需求」之關係與現況的理解仍有所不足（羅晉，2008：201）。

此外，過往對於政府網站內容與服務的評估多是以網站整體為分析範圍，主要的研究途徑大多是採使用者意見調查與網站內容的評比（林巧敏，2011：85-86）。對比其他應用資訊科技較為熟稔的學術領域，以圖書資訊學研究領域來說，僅是針對網路上資訊搜尋行為的研究主題，便涵蓋了檢索詞彙研究、檢索策略研究、人機互動或網站使用性研究、網路搜尋歷程分析、使用情境研究等五大類型（林珊如，2002：38）。相較之下，公共行政領域在電子化政府的研究上，雖然也有應用記錄檔分析法（log analysis），即針對網站的使用者行為進行分析之研究。但可能囿於學術能力培養過程中在相關技術上的認識不足，以及關注的研究取向差異，因此在公共行政研究領域的研究成果明顯較少，而通常需藉由跨領域的學術合作，或者由熟悉相關研究技術的其他學術領域研究者進行。例如陳祥與林明童（2002）的《我國「電子化政府整合型入口網站」使用者行為分析》，蕭元哲、葉上葆（2003），《電子化政府之使用分析—檔案分析法》。然而類似的研究技術與途徑並未快速且廣泛地影響台灣公共行政對於電子化政府的研究取向，特別是在政府網站的行者行為分析上，類似的跨領域合作或是以資訊技術為研究途徑的應用，至今仍不頻繁，而由公共行政學者獨立進行的相關研究更為少見。

雖然近年在公共行政研究領域的學者在政府服務網站使用者行為的研究上，雖也持續有發展出以結合問卷調查與內容分析的實驗設計研究（潘競恆、陳敦源，2011），以及廖洲棚、陳敦源、蕭乃沂及廖興中（2013），在《運用巨量資料實踐良善治理：網路民意導入政府決策分析之可行性研究》的研究報告中也嘗試了應用網站日誌分析的技術。但在公共行政學界的電子化政府研究，以網站規劃內容的比較以及藉由問卷的方式，在數量上仍為大宗。若在研究技術與途徑的應用出現落差或傾斜，對於電子化政府的研究並非吾人樂見的情況。

網站日誌分析早已不是新鮮事，但若是對其應用沒有足夠的認識，無法針對問題需求而對系統設置必要的資料蒐集，或者應用適當的分析方式處理資料，並將結果與需求適當地連結，那麼也許它能夠為我

們帶來許多測量網站使用者行為的數據，但這些數據與測量卻可能是毫無意義的（Nicholas et al, 1999: 265）。因此，在電子化政府的研究領域上，除了公共行政學界研究者的持續投入外，引入更多跨領域的學術合作將是未來應該持續努力的方向。跨學科領域的合作能以更多元的研究觀點與途徑豐富相關研究的進展，而跨專業領域的學術或公私合作，也將能促使研究技術與觀點的相互學習與創新，進而同時減少跨領域之技術知識和研究與實務應用上的落差。

政府巨量資料分析與政策端應用效能提升之研析

第三章 研究設計

以第一章計畫目的與第二章文獻檢閱為基礎，以下再詳述本計畫與進行之計畫內包含的三個議題面向：

第一、巨量資料分析與政策應用需求的「落差」(gap)分析：找出已經導入巨量資料分析的政府單位，從應用效能的角度，尋找「落差」存在的可能因素，可能包括技術可能性、資料取得、專業跨域、政策目標不明確、人工編碼需求未導入等問題，這個部份除了文獻回顧的方法之外，將針對目前正在導入巨量資料分析的政府機關相關人員，進行訪談，以釐清相關問題，整合之後作為落差的分析基礎。

第二、導入政府巨量資料應用政策效能提升方案：分別針對政府機關內部的服務系統資料（包含物聯網）及其外部的網路輿情，以特定政府服務系統與公共議題為範圍，與主管機關合作導入「巨量資料應用政策效能提升方案」，過程中計畫團隊將關注下面三個議題：1.資料處理專業與公共政策專業之間，針對巨量資料的技術應用的溝通落差在哪裡？2.在資料技術發展的過程中，如何將「政策論述」納入分析核心，「質量並重」、「人機整合」分析框架的重點為何？3.巨量資料分析的結果，要如何與其他政府內部既存的政策資訊蒐集機制相互結合？這個部份的研析將以深度訪談與特定政府部門導入巨量分析過程中，進行人機整合的實作，主要希望從目前單純情緒分析進入「立場」與「論述」的分析。

第三、巨量資料應用政策效能提升的策略建議：以上述的各國發展趨勢與我國政府機關導入經驗為基礎，針對巨量資料分析應用於我國政府的決策分析與政策規劃提出建議，其範圍包括：（1）政府資料收集結構如何離開行政方便主導，走向巨量資料時代的行政管理系統設計與應用，也就是如何擴大可蒐集資料的應用範圍，這部分可能來自於物連網（internet of things, IOT）的應用；（2）在政府政策應用端不變的前提下，巨量分析政策產出是否能夠更以政策而非技術導向為

政府巨量資料分析與政策端應用效能提升之研析

核心，也就是巨量資料分析產出以政府部門的議題管理目標為主，可以協助擴大政策應用與資料分析交疊的範圍；（3）在巨量資料分析技術不變與可用資料有限的前提下，政府政策需求端要如何依照巨量資料時代釐清其目標與論述，以致於可以更有效地應用巨量資料的技術，完成政策目標。這個部份將根據文獻檢閱、深度訪談、以及舉辦三場焦點座談來建構一份政策效能提升的策略建議書（含標準作業程序 Standard Operation Procedure, SOP）。

根據以上三大計畫目的，本計畫並規劃三項研析子題，分別為一、政府跨部門巨量資料之應用效能提升；二、網路輿情分析平台與群眾外包之實驗設計；以及三、政府服務網之網站分析技術與應用效能提升等三項研析子題，以符合本計畫所規劃之三項計畫目的。三項研析子題皆以巨量資料分析的發展趨勢與應用落差議題出發，並以議題管理為核心理論主軸，以求各研析子題的不同的巨量分析個案皆能符合以議題管理為核心策略導向，而追求巨量分析技術應用於政府治理的效能提升。在研究方法上，除了文獻檢閱、深度訪談、問卷調查、以及焦點團體等途徑的應用外，在方案實作的導入上，本計畫將在輿情分析上，由計劃團隊自行規劃輿情分析平台，以群眾外包的機制進行實驗設計，以同時探究輿情分析在人機整合以及引入群眾外包的機制效能提升。最後，本計畫規劃資料蒐集與執行方法摘要如下表 3。

表 3：研究設計摘要表

計畫目的/階段	相關之研析議題	資料蒐集方法	資料來源
1. 巨量資料分析的發展趨勢與相關「應用落差」的議題： 整理歸納各國應用巨量資料分析於公部門或特定公共服務精進過程中，所遭遇的困難。	-政府跨部門巨量資料應用。 -輿情分析與群眾外包。 -政府服務網站分析。	(a)中英文文獻檢閱。	- 中英文期刊資料庫。 - 國內外政府出版品。 - 網際網路。

計畫目的/階段	相關之研析議題	資料蒐集方法	資料來源
2. 巨量資料分析的實際導入經驗： 分別針對政府機關內部的服務系統資料及其外部的網路輿情，以特定政府服務系統與公共議題為範圍，與其政府主管機關合作導入巨量資料分析政策效能提升方案 3. 巨量資料分析的治理策略建議： 以上述的各國發展趨勢與我國政府機關導入經驗為基礎，針對巨量資料分析應用於我國政府應用巨量資料分析提升政策管理效能的建議	-政府跨部門巨量資料應用 -輿情分析與群眾外包 -政府服務網站分析	(b)巨量資料分析政策效能提升方案導入實作	-政府內部特定服務系統 -政府跨部門巨量資料庫 -政府外部特定議題網路輿情 -自行設置輿情分析平台
	-政府跨部門巨量資料應用 -政府服務網站分析	(c)深度訪談	-導入機關人員（首長、輿情、業務、與資訊單位主管及人員）
	-政府跨部門巨量資料應用	(d)焦點團體座談	-其他中央及地方政府資訊主管及輿情分析人員
	-政府服務網站分析	(e)問卷調查	-具備巨量資料分析導入經驗的企業或非營利組織

資料來源：本計畫。

透過以上研究設計與其執行成果，本計畫首先可更深入瞭解目前各國政府運用巨量資料分析所面對的應用落差，輔以我國政府機關的實際導入經驗，不論是對於政府機關內部的服務系統或外部的網路輿情，計畫團隊預期可以提出經實際驗證且可行的標準作業程序（Standard Operating Procedure, SOP），協助我國政府機關善用與導入巨量資料分析，強化政策運用的效能。

最後，為能夠令本次計畫成果可對於實際政策應用具明確且實質的建議效用，真正使得政府巨量資料分析與政策端之應用效能得以提升。除了針對不同研析子題規畫各種方法進行研究外，在方法論上，

本計畫將更著重在落差分析（gap analysis）的概念，試著理解各項關鍵指標的現況基準，並分析從此現況到達現有標竿之間的落差。

關於進行落差分析，Balm（1996）更提出了我們應該擴大標竿的設定以及落差分析的範圍，從基準線到現有標竿之間，再從標竿到完全顧客滿意（total customer satisfaction）的程度之間，以提升我們對於追求卓越的洞察力，以及有助於我們考量對資源分配的優先順序，以最大化變革的能量與成果。另一方面，Silvestro（2005）則強調，應用落差分析此方法論，將有助於在期望與感知上進行關鍵性的差異指認。而從常見的落差分析工具之一，顧客滿意度調查的應用中，則須留意四項可能造成使用者經驗與知覺間落差的原因：1. 無法理解使用者的期望、2. 無法將管理者對使用者經驗的知覺轉換或落實到實際的服務規範上、3. 不遵守管理規範、以及4. 無法適當地管理顧客的期望（Parasuraman, Zeithaml, and Berry, 1985）。然而，進行使用者意見調查的落差分析，主要僅能幫助我們解決第一項的問題，後續仍需要一連串得以實際消除落差的管理作為。

此外，進行落差分析的目的也不僅只是從績效基準線-標竿-完全顧客滿意之間的線性進步，對於一個組織來說，應該基於各種關鍵測量項目而同時兼顧其落差，而進行全方位的落差分析。因此，落差分析的方法論，可同時應用在各種可評估的層面之上。舉例來說，上述對於落差分析的概念以及對於落差分析之範圍的擴展，Balm（1996）建議可以雷達圖的形式展現，例如圖2。

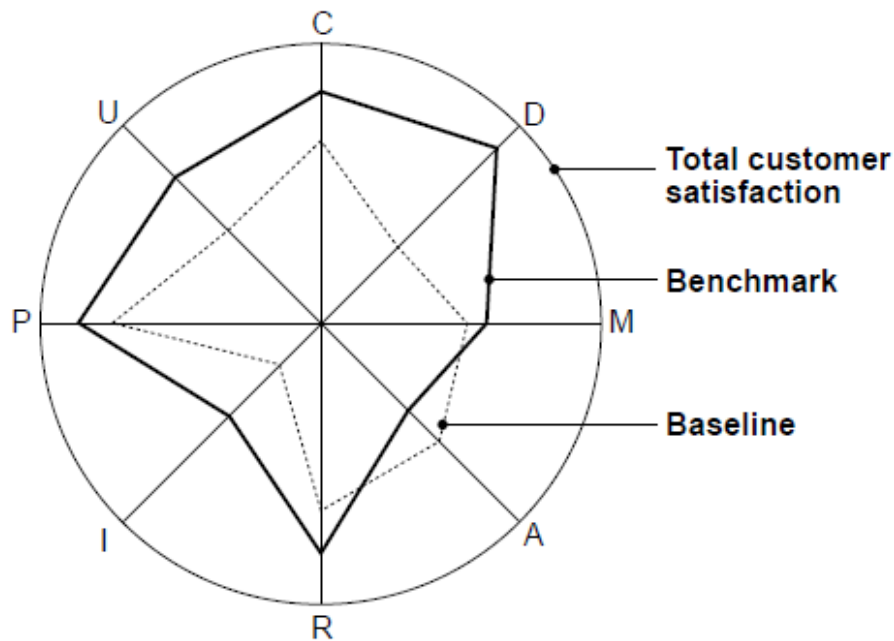


圖 2：全方位的落差分析概念圖

資料來源：Balm (1996：31)。

上圖中的虛線，表示現況的表現基準，中間的粗體黑線，則代表我們所設定的標竿，至於最外面的圓圈，即表示最佳理想的完全滿意程度。從上圖我們可以清楚理解到，進行落差分析與標竿學習的目的不僅只是為了單向和線性的趕上標竿，而應該是針對組織的各項關鍵績效指標同時的進行全面評估（包含標竿）。全面性的落差分析，不僅能夠讓我們瞭解跟學習標竿的差距，並可能提供我們思考為了有效提升效能，何處才是我們最應該先加強的。例如圖 2 之中來說，測量指標的D，便可能是在追求效能提升上時最應該優先改善的部分。因為，該項目不僅適合標竿之間落差最大之處，而標竿與完全滿意程度之間的落差也並不明顯。

然而，將落差分析應用在電子治理的研究領域，還必須要看待研究的內涵而定，比方說，學者Sushil（2000）曾經提出一個名為「SAP-LAP models」的落差分析模式，這個模式以六個程序性的概念：S = Situation（情境），A = Actors（行為人），P = Process（流程），L = Learning（學習），A = Action（行動），P = Performance（績效），討論可以進行方案落差分析的模型；之後也將之應用在電子化政府的方案分析（Suri and Sushil, 2012），可是由於本計畫主要都是探索性的意圖，主要針對在巨量資料時代多個電子治理可能面對的內外問題，因此，最多的落差分析只是落在「情境」的第一階段，因此，本計畫將聚焦在各項探索性主題的情境當中，可能出現「預期」與「實際」之間的問題。

因此，基於上述落差分析的方法論指引，本次計畫在政府巨量資料分析與政策端的應用效能提升上，將試著同時針對三大議題進行研析，首先，對政府應用巨量資料分析現況和技術進行落差分析，先試著了解政府在應用巨量資料的相關技術上的現況以及可能遭遇到那些問題，另外，相關分析技術上的應用效果而理想上的落差可能為何，並期望能在最後針對政府應用巨量資料分析上，對於各種不同的落差提出因應策略或建議。以下內容將詳述本次計畫中三項研析子題的研究設計與方法。

第一節 政府跨部門巨量資料分析

呼應前述的整體計畫目的，在此政府內部巨量資料分析方面，其具體研究內容包括：（1）挑選我國政府內部跨機關的個案，本團隊經考量後將以政府人事政策與公務人力資源管理為範圍（詳後文）；（2）與此個案的相關主管機關聯繫合作共同規劃執行巨量資料分析應用，以其雛形成果展現其決策支援價值；（3）也透過此共同合作的實際參與觀察，深入瞭解分析可能影響巨量資料分析效能的落差；（4）針對

以上三項研究成果，研擬提升我國政府跨機關內部巨量資料分析效能的策略。

經過計畫團隊內部討論，我們認為巨量資料於政府內部跨部門的應用上，有一特殊且重要的議題：人事管理。因此本團隊決定先行拜訪我國政府兩大人事主管機關—考試院銓敘部與行政院人事行政總處之資訊單位，瞭解巨量資料在於其機關的使用經驗、影響與相關應用，並進一步詢問未來的規劃與方向。首先，先針對我國人事主管機關之體系進行說明。

一、個案介紹

在五權憲法的架構之下，考試院考選部、銓敘部、公務人力保障暨訓練委員會（以下簡稱保訓會）、公務人員退休撫卹基金監理委員會（以下簡稱退撫基金監理委員會）、行政院人事行政總處、中央與地方政府各人事單位，組成我國政府的人事政策與人力資源管理體系（如圖 3）。由於當代資通訊技術的演進，由於人事管理相關作業與其資料量龐大、計算繁複、且異動頻繁，人事管理的自動化資訊處理，即成為我國政府業務資訊化的先驅者之一，並於1990年代至今的網際網路、行動化、雲端運算等新興技術與應用領域之一。

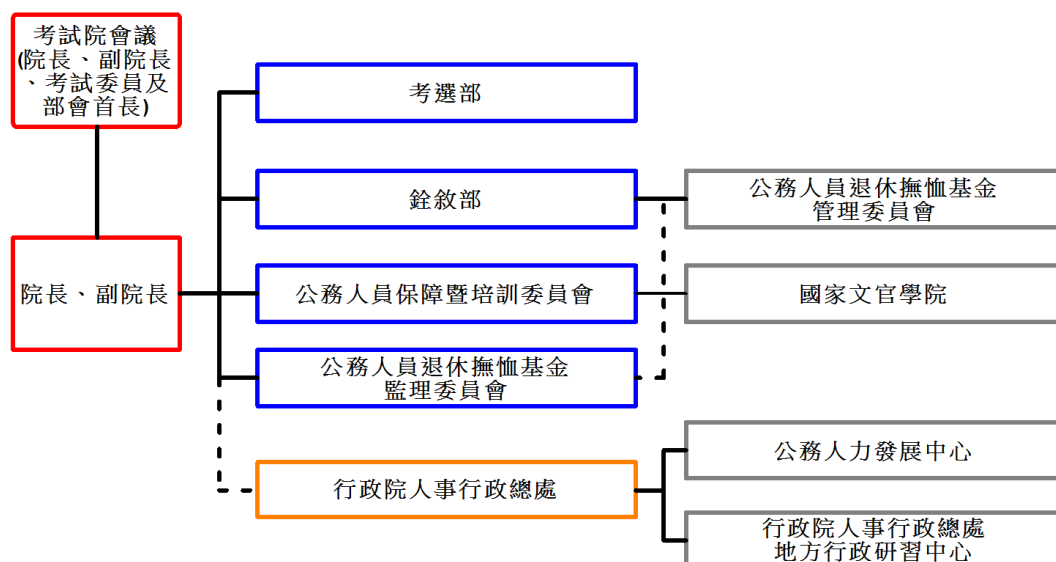


圖 3：我國人事政策與公務人力資源管理組織體系

資料來源：本計畫整理，修改自考試院⁵。資料時間：104年8月1日。

對於政府而言，公務人員是一項重要的資產，如何透過人事政策與行政機制妥善運用有限的公務人力資源，並借助其專業以提高政府績效，是各國執政當局的首當要務。為了蒐集全國公務人員之人事資料，以及辦理各項人事業務，我國政府目前所開發與建置的資料庫主要有二（莊素宜，2010；黃彥榮，2011；陳敦源等人，2014），其一是由行政院人事行政總處所管理維護之「全國公務人力資料庫」，另一則是由考試院銓敘部所建置的「銓敘業務資料庫」。前者約略從民國84年開始蒐集全國公務人員之人事資料，其來源系由各機關人事機構定期透過人事管理資訊系統，將人事資料彙整後送至資料庫。此資料庫之範圍包括了送銓敘審定公務人員、政務人員、民選首長、公立

⁵ 網址：<http://www.exam.gov.tw/cp.asp?xItem=13684&ctNode=559&mp=4>

學校教職員、公營事業人員、人事經費進用之約聘僱人員、雇員、技工工友等。

而銓敘業務資料庫大約從民國76年以後開始建置，因銓敘部承辦各項主管之銓敘業務，各機關之人事機構則透過銓敘業務網路作業系統，將各項資料彙送進資料庫，此為較為完整且正確之銓敘審定資料；資料庫的範圍係為送銓敘部審訂之公務人員，資料內涵包括了機關組織編制、銓審、考績、退休、撫恤、獎懲等資料，可以視為審視全國公務人員動態的基本背景。這些資料約有95%係由各機關單位的人事單位使用，故銓敘部後來便將此資料庫移轉由行政院人事行政總處負責維護與運用，銓敘部目前只保留了網路資料查詢的權力。

本團隊經審慎考量評估後決定以人事政策與公務人力資源管理作為巨量資料分析的應用範疇，主要是因為此橫跨中央與地方政府的人事政策與管理體系，符合本計畫先前所設定的跨政府機關的條件，其人事資料已具備高度電子化與自動化處理，且其資料量龐大、計算繁複、與異動頻繁符合一般巨量資料的特質；同時在有限的計畫資源（尤其是時間）限制下，相對於跨業務領域（例如整合戶政與社政）資料在政策方向上較有重點，可避免花費時間在摸索跨業務政策上的應用可能。

二、 研析規劃及預期成果

在以上針對跨政府機關內部巨量資料分析的研析內容與個案（即廣義人事政策與公務人力資源管理，Public Personnel Policy and Human Research Management，以下簡稱PPP-HRM）的基礎上，此部分研析設計也承襲本計畫整體研析設計的構想（表 3），其首要工作即是與我國人事主管機關共同合作選定若干特定案例（問題導向，problem-driven），或以資料探勘作為引導（資料導向，data-driven），以規劃與實作PPP-HRM的巨量資料分析雛形（prototyping），搭配參與觀察、個別訪談與焦點團體訪談我國人事主管機關的業務與資訊單位代表，

以整理分析相關經驗並研擬有助於政府跨機關內部巨量資料分析的策略，作為其他政府內部巨量資料應用案例的參考。

依據上述對於我國PPP-HRM相關主管機關的介紹，本計畫將首先訪談行政院人事行政總處，包括其人事資訊處與其他人事業務高階主管；同時在考試院方面，也將訪談人事相關法規主管機關之一的銓敘部，同樣包括其資訊室與考試院其他高階主管或考試委員。以上訪談企圖從當前的PPP-HRM業務重點中逐步聚焦決策範疇，同時為展示其跨機關應用的潛力，這些應用雛形希望不限於單一人事機關（考試院考選部、銓敘部、保訓會、退撫基金監理委員會、行政院人事行政總處等），而企圖具備跨人事機關的決策。

不論是透過共同規劃實作過程中的參與觀察、個別訪談與焦點團體訪談，為了能讓本計畫的經驗在未來可以作為其他政府機關運用其內部巨量資料分析的參考，本計畫將特別注重規劃實作的流程建立與經驗歸納，尤其聚焦於跨政府機關的資料管理（data management）與資料治理（data governance）的視角，包括資料品質、分析技術、行政流程、法規制度等面向。

三、 深度訪談設計

如前述規劃，本團隊認為人事管理的議題對於巨量資料於政府內部跨部門的應用上有其值得探究之處。因此本團隊針對我國政府兩大人 事主管機關—考試院銓敘部與行政院人事行政總處之資訊單位進行訪談，以試著瞭解巨量資料在於我國人事管理體系的使用經驗、影響與相關應用，並進一步詢問未來的規劃與方向。表 4為本次深度訪談之重點與訪談題綱：

表 4：第一次訪談重點及訪談題綱對照表

訪談構面	訪談分析重點	訪談題綱
跨部門巨量資料之應用	瞭解政府機關巨量資料使用經驗	1. 請問貴機關使用公務人力巨量資料之經驗為何？ 2. 使用後對於人事相關業務的決策目標與應用的影響或效益為何？ 3. 在目前的業務執行上，貴機關對於人力資料庫有無相關應用？
	瞭解政府機關巨量資料應用之未來規劃與方向	1. 對於貴機關而言，公務人力資料庫之應用，有無相關的未來規劃？或是大致上的方向為何？

資料來源：本計畫。

本團隊就前次訪談之結果進行討論後，發現公務人員退休年金是個重要的議題，且影響範圍廣大，社會上的討論也相當熱烈。此外，我國公務人員的退休資料相當龐大，其牽涉的主管機關橫跨行政與考試二院，並符合本計畫的研析目的之一：政府跨部門巨量資料分析，故選定此議題進行討論與分析。為此，本團隊二度前往人事資料之管理單位進行請益，並針對研析主題進行討論。表 5 為第二次討論之重點及討論議題：

表 5：第二次討論重點及討論議題對照表

討論構面	討論重點	討論議題
跨部門巨量資料之實際應用與相關議題之延伸	<ol style="list-style-type: none">1. 瞭解巨量資料於公務人員退休年金議題的使用經驗2. 巨量資料實際演練之設計3. 政府巨量資料相關議題	本團隊經內部討論後，決定以公務人員退休年金為範圍，希望能整合現有的相關資料，實際演練一次政府內部跨機關的巨量資料分析，並且記錄資料庫數據的應用與分析過程。除了量化分析成果之外，運用巨量資料的相關議題也會有所探討，例如資料治理（data governance）、隱私、開放等，期待也可以對政府後續的相關政策方案有些貢獻。想知道長官對於此議題有無指導與建議？

資料來源：本計畫。

經過先前兩個機關（考試院銓敘部、行政院人事行政總處），總共四次的訪談討論，以及計畫團隊所辦理的焦點座談後，我們發現不同機關或單位，面對政府巨量資料之應用，其看法與期待皆有所差異。因此，本計畫團隊分別前往業務單位與幕僚單位，進行請益與瞭解，希望能夠得到不同角度的觀點，找出跨部門或機關的巨量資料應用的矛盾所在。

表 6：第三、四次訪談重點及訪談議題對照表

訪談構面	訪談分析重點	訪談題綱
跨部門巨量資料應用之現況與未來展望	瞭解政府機關巨量資料應用之未來規劃與方向	<ol style="list-style-type: none"> 1. 從貴單位的角度出發，想瞭解貴單位，甚至是行政院是如何看待政府跨機關部門的巨量資料應用？針對這個議題，有何策略與想法？ 2. 在目前我國政府跨機關的巨量資料應用上，有無較為積極發展，或是較為特殊，值得分享之案例？ 3. 國網中心在我國政府跨機關資料應用上之角色？在實務運作上將面臨什麼障礙？有何解決之道？

資料來源：本計畫。

表 7：政府機關跨部門巨量資料應用者背景

政府機關名稱	受訪者職級	性別	訪談日期	代碼	訪談者
考試院銓敘部資訊室	主任	男	2015/8/06	A1	蕭乃沂老師
			2015/9/16		
行政院人事行政總處人事資訊處	處長	男	2015/8/07	A2	蕭乃沂老師
			2015/9/16		
衛生福利部中央健康保險署	組長	男	2016/1/12	A3	蕭乃沂老師
行政院科技會報	副研究員	男	2016/1/27	A4	蕭乃沂老師

資料來源：本計畫。

四、專家學者焦點座談

計畫團隊在經過先前兩次訪談後發現，政府機關跨部門資料的交換與應用上，皆面臨或大或小的困難與障礙。為釐清不同政府機關間應用巨量資料進行分析的整體現況，並比較跨部門或跨領域間應用上的差異，本計畫透過邀請巨量資料相關應用領域的學者專家，以及政府機關實務工作者進行焦點座談，目的欲瞭解政府部門實際應用巨量資料時所面臨到的障礙與困難，並分析目前有哪些業務具有巨量資料分析的特質，又進一步瞭解哪些業務適合使用巨量資料加以分析，並探討資料治理在公部門運用的可能性為何。因此，本計畫就政府巨量資料之分析與政府資料治理的應用這兩大構面進行討論，茲就就分析重點與題綱內容整理如下表 8所示：

表 8：專家學者焦點座談重點及訪談題綱對照表

訪談構面	訪談分析重點	座談題綱
巨量資料分析於政府機關之應用	瞭解政府機關中，有哪些業務的應用是符合巨量資料分析之基本特質？	1. 一般政府機關或特定於貴機關中有哪些業務領域的資料集（ datasets ）符合所謂巨量資料（或巨量資料， big data ）的3V基本特質？亦即 Volume （資料量龐大）、 Velocity （即時更新）、與 Variety （多元資料格式或來源）。
	何種業務類型可利用巨量資料的分析方法？	2. 您認為其中有哪些業務領域或資料集應該優先善用量化或文字的巨量資料分析（ big data analytics ）以提升公共服務績效或輔助公共議題決策？以其應用角度來看，其資料來源可能為跨機關的有哪些？
資料治理於政府機關之應用	瞭解資料治理在政府機關應用的障礙	3. 資料治理（ data governance ）被視為「確保組織中的多元資料足以提升組織整體績效的標準與過程，包括策略、管理、與技術等方面」，在政府機關運用巨量資料分析時，在資料治理上的可能挑戰為何？在跨政府機關資料的整合應用時，特別可能面臨哪些挑戰？

政府巨量資料分析與政策端應用效能提升之研析

	分析政府資料應用障礙之原因與解決之道	4. 針對以上特定政府機關內部或跨機關巨量資料應用的諸多挑戰，您認為有哪些原因？又有哪些可能的因應？
--	--------------------	----------------------------------------------------

資料來源：本計畫。

本次專家學者焦點座談之與談對象選擇採取立意抽樣的方法，目的在於瞭解政府業務應用巨量資料分析之現況，以及資料治理之可運用性。是以，計畫團隊經過專家之建議後，在學術界與產業界中，選擇了近年來在政府出版品與學術期刊中，在各領域有發表過巨量資料應用相關之專家學者；在政府機關部分，則是挑選其業務執行上大量使用巨量資料，且具有相關豐富經驗的業務與資訊單位主管作為與談對象，並在2015年11月20日於國立政治大學公共行政學系進行座談。茲就與會者背景及座談摘要分析代碼等資訊整理如下 表 9所示：

表 9：專家學者焦點座談與會者背景

機關名稱	代表領域	與會者職級	性別	代碼
衛生福利部中央健康保險署	政府機關	組長	男	B1
國立台北大學公共行政暨政策學系	學術界	教授	男	B2
台北市政府人事處	政府機關	主任	男	B3
行政院人事行政總處資訊處	政府機關	科長	男	B4
財團法人資訊工業策進會數據科技與應用研究所	產業界	組長	女	B5
		正工程師	男	B6

資料來源：本計畫。

第二節 網路輿情分析平台與群眾外包之實驗設計

一、 將群眾外包視為一種輿情分析技術

去年度TEG團隊曾透過網路輿情分析（國家發展委員會的「自由經濟示範區」，結案報告第四章）的實作經驗，提出了研擬提升網路輿情分析品質的人機互動模式的建議。具體而言，我們認為文字探勘之電腦演算法雖然在分析輿情中之正負情緒有相當的準確度，但是公共議題的界定與政策立場的指認，往往需要人類的主觀判斷才能竟其功，因此網路聲量或情緒的分析，雖然可以全然應用電腦的運算能力達到快速分析的效果，但是，意見立場的分析仍需借重人腦的協助。因此，本年度本計畫擬就如何開發出快速有效的人工評估技術，以結合電腦演算法之分析來提升網路輿情分析品質，進行深度探討。

傳統的人工評估作法，就跟本團隊去年的作法一樣，是以類似統計品管程序，隨機抽取網路輿情資料，以人工方式對這些樣本，進行以情緒意向與政策立場關聯等嚴謹檢測，來了解民眾透過網路媒體所發表意見的情緒意向與政策立場。但巨量資料時代標榜的是直接以母體操作而非僅僅以抽樣處理，意謂著我們應朝向全面性採用人工判讀的方式來分析網路輿情，問題是如何在有限的人力與時效性的限制下來執行人工判讀呢？既然這些輿情來自網路，秉持「網路的問題用網路的方式解決」（蕭乃沂、陳敦源、廖洲棚、楊立偉、呂俊宏，2014：217），本團隊規畫應用群眾外包（Crowdsourcing）（Howe, 2008）的作法來實驗人工判讀輿情的可行性，將大量的網路輿情資料的主觀判斷，經過精細分工讓龐大的鄉民一起參與並解決。

群眾外包泛指將工作切分，透過網路由不特定的多數人來處理的一種新興合作型態。經過過去十多年來的發展，群眾外包的模式已經十分多樣化，不僅有各種商業上的應用，也不乏用於公益性的志願工作與解決政府棘手難題的案例。從最近的文獻中，也看到應用群眾外包來進行評論性文字的情緒意向分析（Borromeo and Toyama, 2015），雖尚未見到有以群眾外包對網路輿情中的政策立場進行分析，我們相信這是值得嘗試的。緣此，本團隊將以上年度針對「自由經濟示範區」

主題所獲致的網路輿情資料與分析結果為基礎，探究以群眾外包方式來判讀輿情中之情緒意向與政策立場之可行性，並透過與上年度結果之比較，來深入剖析群眾外包在此課題上之優劣處，並進而研擬可提升網路輿情分析品質的人機互動模式。

二、 輿情分析平台之建制與群眾外包實驗設計

雖然目前已經有許多成熟的群眾外包平台，例如：亞馬遜的 Amazon Mechanical Turk與CrowdFlower，但這些多平台是商業導向，以付費為主，也不易運用於中文文件的處理工作上。因此，本團隊將選取適當之開放源碼的群眾外包平台，根據我們的需求，自行客製化後建立一個輿情分析平台。為此，我們的重點將是訂定需求、建制平台以及實驗設計、執行與評估，而非輿情本身。這也是為什麼我們不另外進行其他議題的網路輿情搜集，而直接以上年度的資料與結果來作為我們實驗的基礎。我們的研析步驟與預期成果如下：

（一） **訂定需求與建制群眾外包平台**：擬定對群眾外包平台之基本需求，比較各開放源碼的群眾外包平台，選取能大致滿足我們需求的平台，並據以建立一個應用群眾外包進行輿情分析的雛形平台。

（二） **實驗規劃與設計**：就如何應用群眾外包於輿情分析選取實驗變項，並據以設計實驗步驟以及對平台雛形進行進一步的客製化與調整，以利實驗之執行。

（三） **實驗執行**：依據各實驗變項的不同組合情境，招募適當之群眾來進行各情境之實驗，對實驗結果進行整理，產出輿情分析結果。

（四） **實驗結果評估**：將整理後之各實驗結果，與上年度之結果進行比較與評估，歸納出使用群眾外包的優劣處，並據以設計可提升網路輿情分析品質的人機互動模式。

第三節 政府服務網之網站分析技術與應用

我們必須體認到，我們的網站已成為民眾需向政府問問題或需要和政府互動的第一站。我們的設計需要符合他們同時對於資訊和服務日漸增加的需求⁶（Gill, 2012:37）。

一、政府網站可用性概念與發展趨勢

隨著電子化政府政策的普及，研究者或許不容易指出民眾終其一生使用了多少個政府機關的網站，但要指出有哪些政府機關網站為其生活帶來便利性卻相對容易許多（Buie and Murray, 2012）。過去10多年來，主流的電子化政府網站營運績效評比方法，多藉由人工檢核途徑，評估政府機關網站於預設的網站營運績效指標表現成績，諸如美國布朗大學的電子化政府評比和日本早稻田大學的全球電子化政府評比（Obi and Iwasaki, 2015）。然而，在巨量資料時代，當網站使用者行為已愈來愈容易取得，有關政府網站營運績效的評估方式，是不是也應該有所調整？

政府設立網站的目的有二，其一是提供完整的公共資訊來滿足人民知的權利；其二則是透過網站提供人民服務，提高民眾和政府互動的便利性，因此提供民眾需要的「資訊」和「服務」就是政府網站核心的營運目標（Gill, 2012）。我國從民國70年代建立大型行政資訊系統以來即陸續推動電子化政府服務迄今已逾30年，在邁入第四階段電子化政府計畫（101年至105年）的此刻，不論中央政府或地方政府機關幾乎都已設有自己專屬的網站。然而，依據2014年國家發展委員會發布之我國個人及家戶數位機會調查顯示，我國12歲以上民眾雖已有高達78%的民眾有使用網路的經驗，但僅有46.2%網路族曾在最近1年

⁶原文為 “We must acknowledge that our web sites are becoming the place our citizens go to first when they have questions about or want to interact with government. Our designs need to match their growing need for both information and service.”

透過網路查詢政府公共資訊，也僅有32.4%曾在最近一年透過網路申請政府服務（國家發展委員會，2014）。本計畫進一步從這些數據換算後得知，我國12歲以上民眾在2013-2014年間，約僅有36%的民眾曾透過網路查詢政府公共資訊，25.3%的民眾曾透過網路申請政府服務。對比國外的發展情況來看，美國及澳洲分別在2010年進行之民眾上網行為調查，且調查結果顯示美澳兩國民眾各約有近80%的民眾曾瀏覽政府網站（轉引自Gill, 2012:22），我國民眾對於政府網站的利用度似乎較之為低，即使調查結果並未指出美澳兩國的民眾對於其政府網站有較滿意的使用經驗。由於政府網站提供的資訊及服務和民眾的生活息息相關，因此如何提高我國民眾對於政府網站的利用度，應是政府網站經營者在推動電子化政府服務過程最需關切的議題之一。

有關政府網站可用性的問題，學者蕭元哲及葉上葆早在2003年就曾運用日誌分析（log analysis）方法分析我國電子化政府入口網使用者在2002年6月及7月的使用行為。研究發現當時民眾對於電子化政府入口網的利用程度並不高，每日平均利用人數僅有3,481人，且近8成的使用者停留時間僅在一分鐘內且重複使用率也不高，故進一步呼籲政府應從顧客導向觀點強化網站內容以吸引民眾的停留與利用（蕭元哲、葉上葆，2003）。十年後，廖洲棚、陳敦源、蕭乃沂及廖興中（2013）等人亦同樣運用日誌分析方法（兩者差別在運用不同分析工具及網站標的）分析我國目前的電子化政府入口網---我的E政府網站（My eGov）的使用效益，其以2013年9月11日至10月11日的使用者瀏覽紀錄為例，舉出該月的每日訪客雖有18,929人次，但每日每位訪客平均閱覽頁面僅有1.64次，而最多瀏覽的頁面為網站的首頁⁷。這顯示我的E政府網站的瀏覽人次和十年前的調查結果相比，雖已成長5.44倍，但瀏覽行為

⁷值得注意的是，國發會提供的 iTaiwan 免費無線上網服務，在使用者成功連線後，會立即被連接至我的 E 政府網站。對比 2014 年國發會發布的數位機會調查結果顯示，有無線網路使用經驗的我國 12 歲以上民眾，約有 41.2%的使用者有使用政府機關提供的無線網路經驗，亦即有約 37.7%的網民（78%*42.2%）曾利用政府無線網路服務。由於 iTaiwan 網路服務遍及全國各地，因此可大膽猜測有一定比例的民眾可能是使用了 iTaiwan 服務而不自覺的連上我的 E 政府網站。在此情形下，網站的瀏覽人次統計紀錄可能不能反應真實的網站使用人次。

似未有明顯的差異，因為大多數民眾仍只停留很短暫的時間或瀏覽很少的頁面。換言之，十年來我國電子化政府的主要入口網內容雖已歷經多次改版，但網民對於此一網站的利用程度似乎並未有非常顯著的提昇。這個情況，也代表著過去十年來政府網站的經營者對於網站的可用性（usability）問題並未給予適當的重視。

Lo與Wang（2012）曾調查影響台灣民眾使用政府網站的因素，研究發現民眾對於政府的信任度、幫助民眾使用網路的條件、民眾感知網站有用、感知網站容易使用以及對於政府網站的態度等五個因素會影響民眾使用政府網站的意圖，而其中和網站本身可用性相關的因素就佔了三個。Coleman、Lieber、Mendelson與Kurpius（2008）等人透過與州政府預算相關之實驗性網站來研究公民參與行為後發現，當網站的設計能符合民眾對於內容、搜尋以及呈現方式的需要時，確實能增強民眾對於公眾接觸（civic engagement）的態度，其中民眾感知到網站的可用性（usability）是增強民眾公共接觸態度的關鍵。事實上，當有愈來愈多的民眾轉向網路來找尋所需的答案或處理面對的問題時，政府網站的設計者必須要網站的可利用性列為設計的第一要務（Gill, 2012:36）。

依據美國衛生及人群服務部（Department of Health and Human Services, HHS）設立的Usability.gov網站⁸的定義，所謂的可用性意指「使用者與產品或系統互動經驗的品質，包括網站、軟體、設備或應用程式。可用度包含了效能、效率以及使用者整體滿意度等概念。⁹」，

⁸ 美國衛生及人群服務部（Department of Health and Human Services, HHS）為協助提高美國聯邦政府機關網站的網站可用度，協助各政府機關建立以使用者需求為導向的機關網站，其掌管公共事務之助理秘書辦公室中的數位溝通處（Digital Communications Division）設立了Usability.gov網站，詳細說明提升網站可用度之使用者經驗分析、網站分析、人機互動設計等網站分析的方法、設計理念及原則。Usability.gov的內容十分豐富，也提供非常豐富的具體操作指引給關心政府網站可用度的實務工作者運用，例如這篇介紹網站分析的基本原理，對想提升網站使用績效的管理者而言，就具有高度的參考價值。

⁹ 原文為 Usability refers to the quality of a user's experience when interacting with products or systems, including websites, software, devices, or applications. Usability is about effectiveness, efficiency and the overall satisfaction of the user.

並可從符合直覺的設計(Intuitive design)、容易學習(Ease of learning)、使用效率(Efficiency of use)、被懷念程度(Memorability)、使用者使用發生錯誤頻率和錯誤嚴重程度(Error frequency and severity)以及主觀的滿意度(Subjective satisfaction)等六個面向衡量(U.S. Dept. of Health and Human Services, 2006)。陳玫真和黃營杉(2009 : 83)認為網站可用性意指「消費者對網站呈現的知覺」，要衡量網站的可用性可採用微軟公司(Microsoft)提出的網站可用性指導原則(Microsoft's Usability Guidelines, MUG)中的五個面向進行衡量，分別是：內容(content)、易用性(ease of use)、個人化(made for the medium)、情感訴求(emotion)以及推廣(promotion)等(轉引自陳玫真、黃營杉，2009)。Gill (2006)則認為一個具可用性的網站，除了具有讓使用者不需思考就能利用的條件外，還需考量具備有用(be useful)、可接近(be accessible)、具吸引力(be attractive)、具社交性(be social)和考慮移動性(think mobile)等六大因素。上述定義關注的焦點雖未盡相同，但可觀察出一個網站可用性的共同面向為使用者和網站互動的整體經驗品質，亦即從使用者經驗(user experience)來衡量網站品質，故衡量構面必須從涵蓋網站的效率、效能以及使用者主觀的使用經驗等多元面向。換言之，使用者的使用行為以及使用後的評價成為網站可用性衡量關切的焦點，而網站可用性衡量的結果更可作為網站後續改版設計的依據，亦即做為以使用者需求為導向來改善網站人機互動介面(human-computer interaction, HCI)的重要指引。

從前述的分析可知，一個具可用性的網站應從使用者需求角度進行設計和更新，研究者雖可從多元的研究方法來瞭解使用者需求，但忠實記錄造訪網站的使用者從何而來、點閱何種內容、停留多久以及完成哪些交易、查詢哪些資訊等使用行為資料之網站日誌檔案分析(log analysis)或網站流量分析(web analytics)是巨量資料分析專案中，最先被使用的一種分析方式(Eaton et al., 2012)。網站流量分析是指「分析者以取得網站造訪資料為前提，針對這些訪客行為資料進行分析與匯整，以便從中發掘出極具參考價值的網路行為規律性或網

路行為多變性（鄭江宇、張佳榮，2015：14）」。它主要是針對每一位網站使用者，其使用網站伺服器所留下的軌跡與連結Web伺服器後的一舉一動加以分析，協助管理者站在網站使用者的角度，對於使用者的行為與需求有更全面的認識與了解，並且可以作為未來網站規劃與改進的依據（廖洲棚等人，2013）。一般而言，網站經營者需蒐集和以下幾個面向有關的網站流量資料進行分析，分別是：網站經營目的（goals）、網站經營目標（objectives）、達成使用目標前需完成的行動（calls-to-action）、關鍵績效指標（key performance indicators, KPI）以及行為觸發標的（targets）等面向（U.S. Dept. of Health and Human Services, 2006），但由於網站流量分析是屬於一種情境依賴（context-dependent）的分析途徑，不同類型的網站對於網站流量分析有不同的應用需求，導致不同領域網站對於需蒐集的網站資料項目以及內容存有差異而需要依據個別需要進行資料的記錄及分析程序的設定。

二、 研究設計

在眾多的網站流量分析工具中，Google Analytics（GA）為近年來最受企業歡迎及使用的網站巨量資料分析工具（鄭江宇、張佳榮，2015）。例如，高度依賴網站和讀者互動的圖書館管理領域而言，研究指出GA提供非常多種對於網站使用者使用行為的重要視覺化資訊，包括訪客來源、訪客點閱頁面、訪客在每個頁面的停留時間、訪客探索網站的深度、訪客最後離去頁面以及離開後去了哪裡等資訊，能有效幫助網站經營者改善網站設計來滿足使用者需求（Fang, 2007）。GA為Google企業所推出的網站分析工具，最主要的目的在於讓一般大眾或企業可以針對自行架設的網站進行分析作業，以瞭解網路瀏覽者在網站中的行為以及瀏覽者來源，作為未來網站設計以及服務的參考（廖洲棚等人，2013）。GA可支援網站分析師進行分析工作，例如：追蹤許多標準的網站指標，如造訪數、個別訪客數、跳離率、放棄率以及追蹤業務成果等等；追蹤不同類型的行銷活動以及協助區隔不同客戶

的網站使用行為；提供豐富的報表方便閱讀分析結果（胡為君譯，2011），它具有幾乎免費、支援應用軟體介面、客制化流量分析報表、多查看功能以及強大的網站分析產品整合能力等優勢（鄭江宇、張佳榮，2015）。因此，本計畫為探索網站經營者對於網站使用者行為資訊的看法以及該資訊對於實際經營決策的幫助，除將研析聚焦在「政府網站使用者行為分析，對政府網站經營預期與實際管理作為的落差分析」外，計畫團隊亦進一步假設當以GA為分析工具，政府網站經營者需要進行哪些情境式的設定，才能符合其分析的需要。

綜合而言，本計畫的目的有四：

- 第一、瞭解政府網站經營者對於網站可用性的管理期望與實務的落差。
- 第二、透過標竿學習幫助政府網站經營者瞭解如何藉由使用者行為分析來協助網站經營決策。
- 第三、分析Google Analytics（GA）用於彌平前述落差的能力。
- 第四、產出一分可操作的服務需求檢核文件，協助政府辦理網站委外作業時可增加GA的設定。

為達成前述目的，本研析子題採用三種研究方法：

第一、深度訪談：針對重要政府服務網站及成功的商業網站經營者進行訪談，前者是協助計畫團隊對於政府網站管理現況的瞭解，後者則是提供網站經營者標竿學習的參考。依據國發會的建議，本計畫目前規劃的訪問對象為下列網站的主要經營者或管理者，包括屬政府部門的國發會e管家plus網站、法務部全國法規資料庫入口網、財政部電子申報繳稅服務網、內政部實價登錄網，至於民間的iCook愛料理食譜社群網則是本團隊選擇做為標竿學習對象。

此部分關於政府服務網站分析的研析子題，將採用深度訪談、文獻分析、以及焦點團體座談等研究途徑，聚焦在「政府網站使用者行為分析，對政府網站經營預期與實際管理作為的落差分析」以及假設

當以GA為分析工具，政府網站經營者需要進行哪些情境式的設定，才能符合其分析的需要。目前已針對相關文獻進行初步的整理，並已著手進行針對重要政府服務網站及成功之商業網站經營者的深度訪談。以下針對本計畫子題的訪談提綱設計、內容規劃、以及目前進度進行說明。

（一）訪談題綱設計架構

為瞭解網站經營者的經營情境，本計畫子題的深度訪談題綱分為兩個構面進行設計，分別是網站預期管理作為以及網站實際管理作為，其內容如下：

1、網站預期管理作為

- (1) 網站營運目標：網站設立背景、資金來源、經營目標、目前成就以及技術團隊和技術。
- (2) 網站績效要求：網站資源分配者對於網站營運的績效要求。
- (3) 網站服務對象：網站提供服務對象什麼服務，服務對象的特質（目標市場）。
- (4) 網站經營的問題：特別關注網站服務和使用者互動之間產生的問題，以及經營者對網站服務品質持續提昇的自我期許。

2、網站實際管理作為

- (1) 目前做成網站服務改善決策的依據以及資料蒐集方式。
- (2) 目前分析網站使用者行為的方法。
- (3) 網站維運資源分配者對於目前網站經營所關切的重點。

（二）訪談題綱內容

為讓訪談分析結果有一個共同的比較基準，故規劃採用同一套訪談題綱進行，其內容初擬如下：

1、請您先簡單介紹您的網站，包括：設立背景、資金來源、經營目

標、目前成就以及技術團隊和技術？【1-(1)】

- 2、 請問您的網站的財務資助者對網站經營管理的績效要求為何？
【1-(2)】 目前最關切的績效目標為何？【2-(3)】
- 3、 請問您的網站主要的服務對象為何？【1-(3)】
- 4、 請問您目前如何進行網站的經營決策分析？（用什麼分析工具、
哪些是重要決策資訊、這些資訊代表什麼意義且影響哪些決策？）【2-(4)】
- 5、 請問您認為網站的使用者分析重不重要？為什麼？【1-(4)】 請問
您目前是透過何種方法來瞭解網站使用者行為？【2-(2)】
- 6、 請問您對於網站服務品質的未來期許為何？要達到您的期許，您
規劃從事哪些改善工作？使用者行為分析對於您從事這些改善工
作的幫助在哪裡？【1-(4)】

第二、 文獻分析：針對GA功能進行文獻探討，並依據訪談結果分析GA用於滿足政府網站經營者瞭解網站使用者行為需求的程度以及GA功能設定重點，目的在產出一分可操作的服務需求檢核文件，協助政府辦理委外作業。

第三、 問卷調查：本計畫子題規劃針對我國中央政府各機關（一～三級機關）的機關入口網，以及各機關瀏覽人次最多之主題網站（若無免填）為分析標的，以發放網路問卷的形式令中央各機關之網站管理人針對其負責之服務網站建置、運作、管理、與資料分析之概況進行調查，調查內容包含政府機關網站的服務功能、行銷方式以及推動使用者導向的需求。

問卷調查中有關政府機關網站回應使用者需求的需求，係參考AARRR分析架構設計，以結合GA設定需求，其內涵包括：

1. **Acquisition（獲得使用者）**：如何讓使用者從各種管道得知並連結到

網站？

2. **Activation（成為有效的使用者）**：如何讓已瀏覽網站的使用者使用網站提供的主要功能、服務或閱讀重要資訊，而不是隨意瀏覽即離開？
3. **Retention（留住使用者）**：如何讓使用者願意重複造訪網站？
4. **Referral（使用者推薦他人）**：如何讓使用者願意推薦他人使用網站？
5. **Revenue（使用者付費或執行預期行動）**：如何讓使用者增加網站的造訪率、停留時間或服務利用率？（考量政府網站特質，故調整原始定義，改參照免費提供內容，但藉廣告獲利的商業網站模式，設計衡量重點及調查問題）。

最後，在預期調查成果上，期望能第一、調查中央政府各機關官網及為民服務網站目前營運情況以及及關注使用者需求的現況；第二、瞭解中央政府各機關為民服務網站推動顧客導向服務的具體作法及未來需求；第三，瞭解中央政府各機關便民服務網站使用Google Analytics（GA）協助分析使用者行為的需求。

政府巨量資料分析與政策端應用效能提升之研析

第四章 跨部門巨量資料分析之研析成果

本計畫子題透過國內外相關文獻分析，並以個別深度訪談以及焦點團體訪談，瞭解我國當前政府機關推動巨量資料分析應用的現況，尤其是針對公共服務績效或輔助公共議題決策有貢獻的特定機關業務與跨機關議題，以及被視為「確保組織中的多元資料足以提升組織整體績效的標準與過程，包括策略、管理、與技術等方面」的資料治理（data governance）挑戰與因應。

計劃團隊原本預定以公務人力資料庫為實作範例以取得相關經驗，但在分別與考試院銓敘部與行政院人事總處多次洽商，得知在當前個資法的限制下，取得公務人員個資與退休相關原始資料所需程序與時間較預期為長。因此在經過審慎評估後，決定以領域更為多元的個別與焦點團體訪談，蒐集整理有關跨政府機關巨量資料庫整合的相關經驗與意見。

本章首先歸納整理資料治理的基礎學理，作為後續政府巨量資料分析應用的探討角度；接著以巨量資料分析的特質及預期效益、導入條件、與相關議題等三個相互關連的方面予以論述。

第一節 資料治理與應用

面臨數量龐大且結構複雜、類別多元的資料集合，一般傳統的資料處理方式早已無法因應，這些資料需要經過特別設計的處理與整合的模式來加以應用。為了妥善發揮這些資料背後蘊藏的價值，對於蒐集來的資料僅進行有系統的管理是不夠的，必須從資料的蒐集、管理與應用層面加以考量。就私部門的觀點來說，為了確保重要資料與其營運與發展的目標達到一致，企業會制訂一系列的政策與程序，此即所謂的「資料治理」（data governance）。

對一般企業而言，治理指的是管理與監督，從強調整體的公司治理（corporate governance）、企業內部的資訊科技治理（IT governance），

一直延伸至針對資料應用的資料治理，這些皆包含在治理的範疇當中（吳嘉川，2010）。所謂的資訊治理，指的是企業裡關鍵資訊活動的權力模式，意旨在闡明決策權與責任，以及組織對於其之期望行為；換句話說，治理係指決定由誰進行決策與負責，而資訊治理則是針對企業重要資訊，透過制度的設計與技術的發展，去釐清如何運作才能夠達到組織的最大利益，它是一套管理程序，也是一個周而復始的生命循環觀念，其需要組織成員共同參與，才能建構出完整的架構（framework），使資訊對組織帶來有效的利益（邱耀漢，2009）。因此，資訊治理可以說是藉由協調各使用者、流程及技術，使組織邁向成功之路。

然而，資訊是將組織蒐集到的資料進行整理與轉換後，產出有意義的資訊以利運用，所以資料治理與資訊治理其實擁有極為相似的概念，其最大的差別僅在於分析的層次—資料與資訊，其具有上下游的因果關係而已。是以，資料治理也是試圖建立一套程序與觀念，包含了組織分工、政策與評量標準等，以框架的形式呈現，讓資料治理成為組織順利運作的工具，使組織的發展有更多的可能性。資料治理這個詞結合了策略與執行的元素，從策略的角度來看，資料治理包含了蒐集、管理、轉換與應用等概念，其牽涉的不僅是組織內的資訊人員，也涵蓋了業務單位與幕僚單位，共同建構一個彼此能夠理解與使用的架構與規則，可以說是一種組織文化的轉變。另外在執行方面，這個治理架構也涉及到組織資料運用，包括影響管理流程與應用技術的發展，可以幫助組織更快且更明智地產生決策。

政府一方面是公共服務的提供者，另一方面也是執行公權力的管制者，其蒐集的資料相當多元且持續，也具備資料分析與再利用的潛力，若能妥善運用政府資料，將其轉換成有效的政策或服務，這將是巨量資料對政府帶來的最大效益之一。運用巨量資料實踐良善治理一直是政府的施政目標，但這樣的概念必須以良好的資料治理為基礎，相較於一般的企業，政府面臨的法規限制較為複雜，但其制定決策的影響範圍更為廣泛，資料治理的好壞與否將顯得更為關鍵。

一、 資料治理的組成元素

承接組織的使命、願景與策略目標，資料治理以相關技術為基礎，企圖以資料管理（data management）、業務流程管理（business process management）、標準遵循管理（compliance management）、角色管理（people management）等面向，確保組織中的多元資料足以提升組織整體績效的標準與過程（圖 4）。首先最常被提及的是資料品質（data quality）是指資料的正確（accuracy）、精確或顆粒粗細（precision）、即時更新（currency, updatedness）、與完整（completeness），另外在資料管理中仍須關注的還有其詮釋或元資料（metadata）、結構（architecture）、整合（integration）、同步（synchronization）、與必要的轉換（transformation），這些都是確保資料品質與可用性的必要元素。

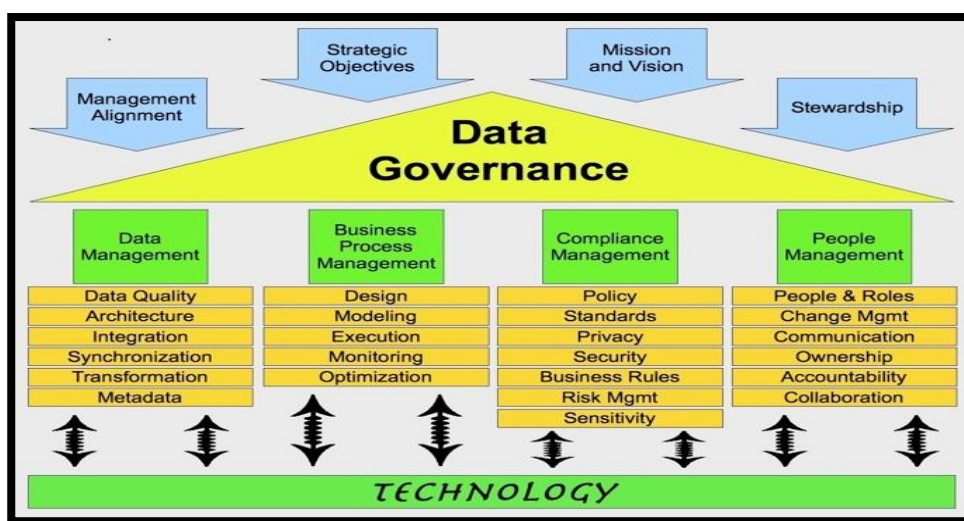


圖 4：資料治理的組成元素

資料來源：Loshin（2013）¹⁰。

與組織資料互為表裡的業務功能與流程管理，則包含其規劃

¹⁰ Loshin, D. (2013). *Data Governance for Big Data Analytics: Considerations for Data Policies and Processes*, URL: <http://data-informed.com/data-governance-for-big-data-analytics-considerations-for-data-policies-and-processes/>.

(design & modeling)、執行(execution)、監控(monitoring)、與最佳化(optimization)的回饋修正歷程，搭配各相關管理面向的標準遵循，包括整體政策(policy)、標準(standards)、隱私(privacy)、安全(security)、業務相關法規(business rules)、風險控管(risk management)、與敏感度(sensitivity)。最後則是與前三個資料治理面向有關的角色管理機制，包括人員與角色(people & roles)、變革管理(change management)、溝通(communication)、權責歸屬(ownership)、課責(accountability)、與協同合作機制(collaboration)。

二、 為何推動資料治理

近年來，私部門相當強調資料治理的重要性，原因不外乎可以配合企業的發展策略，整合資訊系統，使業務的流程達到最佳化，增加服務傳遞的效率，進而提升其利潤。資料治理對於組織最大的幫助，在於資料經過治理框架處理後，以及後續與其他資料的交叉分析，使得原本大量且呈現非結構化的數據價值提升，讓組織在運用資料上保有較佳的應用彈性與安全性，增加組織未來的競爭優勢，創造更大的利益。

受到企業化的風潮影響，政府越來越強調服務導向的概念，加上資訊科技的應用也落實在各層級的政府中，使得跨領域或跨層級以及不同使用型態的資料運用也更加常見，妥善利用這些數量龐大的資料，將使政府推出的政策更能符合民眾的需求。各國政府紛紛針對巨量資料的發展，提出因應的措施，例如澳洲政府於 2013 年 3 月發布的「巨量資料策略議題分析報告(Big Data Strategy-Issues Paper)」指出，政府運用巨量資料的機會在於利用預測性的分析改善政府運作的效率，不僅滿足人民的需求，也能以開放資料的形式與產業進行連結，以促進政策的推動(劉龍龍、葉乃嘉、何志宏、余孝先，2013)。由此可知，政府推動資料治理之目的，不僅在於使行政流程更為順暢，更重要的是將這些龐大的資料進行加值，讓政府與民間各界共享及公平利用。

三、 資料治理與政府巨量資料

巨量資料的特色在於其數量龐大，且未經過有意義的整理，若能經過有系統性蒐集、轉換及整合，並搭配分析處理技術，能夠使這些蘊含高度價值的資料，不會再被隱藏於資料金礦中，讓使用者與決策者有更多更好的方式來做出最佳的決策。我國推行已十餘年的電子化政府，自民國 101 年 1 月起邁入以「主動、分眾、全程服務」為核心的第四階段計畫，在電子資料流通要點與行動化服務發展作業原則之法源依據下，研考會（現國發會）針對資料的應用提出了四大推動策略，包括了資料提供、制定共同規範、建立共用平台、進行示範宣導等策略（宋餘俠、李國田，2012）。

由於政府提供的服務相當多元，其施政的資訊與統計資料十分豐富且分散，如何充分且正確的整合與運用，將影響政府決策品質的良窳。因應不同領域的多元資料，我國政府在不同議題上亦提出了類似資料治理的應用，例如有鑑於我國公務人事資料的業務需求漸趨多元，人事機關間資料交換應用的需求也與日俱增，銓敘部在其掌管的全國公務人力資料庫與網路業務處理系統的基礎下，進行「全國公務人力資料雲端服務平台」的設置及建置工作。另外，中央部會與地方政府常會因應不同領域業務的需求設計資訊系統，但由於需求與執行的目的不同，地方政府為了要發展其特色，資料對於決策支援的角色更加吃重，因此像高雄市政府為配合中央的國土計畫，以及整合其方案管理制度，建立了「高雄市政府社會經濟資料庫決策應用系統」，試圖以制度化及系統化的方式，整合其所屬各類原始及次級統計資料，並提升資料的品質與效率，使業務及決策主管人員，能夠即時掌握重要且經整合之市政資訊，提供制定決策的參考（方映均、林進富，2015；宋方捷，2015）。

第二節 巨量資料分析的特質及預期效益

首先，對於政府機關而言，巨量資料的價值不在於其資料量龐大的本身（何況有關特定領域的資料勢必限縮），而是在於其公共服務與政策上的決策支援（decision support）意義，因此與其將它視為另一個電子化政府甚至新公共管理的「流行運動」，更應該連結至以循證為基礎的公共治理（evidence-based public governance）訴求。

「並非需要同時符合這些 V 才屬於大數據，它只要有某一些類型的特性存在的時候」（受訪者 B5）

「資料量大小取決於一種相對的概念，如果要作主題式或未來有價值的分析一定是要切割領域」（受訪者 B1）

「所謂沒有分析上的價值，是指資料不符合需求，並非指資料是否正確」（受訪者 B5）

「以差勤系統的資料說服市長，大概撈了一整年的同仁出勤資料，約有上百萬筆，經過計算後向市長作簡報。出來的數據統計可以明顯得知，同仁的上班習慣有其固定性，在不同時間的上班情形也相當平均，並不會影響市民的洽公需求，所以建議不用改成三班制」（受訪者 B3）

受訪者也一致主張在巨量資料原始的 3V (volume, velocity, variety) 特質中，資料量（volume）雖然是巨量資料的命名基礎，但並非其最大的挑戰。仰賴日新月異的技術平台，除了在少數的即時應用情境之外，龐大資料量的蒐集與處理通常足以因應。如同本章前節資料治理的相關文獻，資料的更新與傳播速度（velocity）與多元格式與來源（variety），所導致的資料正確（accuracy）、精準或「顆粒」粗細

（precision）、即時（updatedness）、與完整（completeness）才是更大的挑戰。

「資料量的多寡只要是夠用就好，只要數據大到夠用、夠完整，可以拿來決策分析或較有說服力的資訊即可」（受訪者 B3）

「政府在運用數據分析上，首先要考慮的是資料完備度是否足夠...其次是資料的顆粒度要足夠，有很多釋放出來的資料顆粒度是較低的，如果顆粒度太大，就沒辦法往下進行深入的研究與分析」（受訪者 B5）

「針對資料本身的可用度、正確性、格式以及更新頻率來添加補充，這樣比較能夠反應資料的好壞與否，也可以提供要往哪一個角度去作優化的參考」（受訪者 B6）

「資料格式很多時要怎麼去整併，如果要根據我的需求去分析，我可能不是只有我這個部門的資料，可能需要與其他部門的資料交互分析，...包括資料的來源不一樣，我怎麼把這些東西轉換成同樣的處理基準」（受訪者 B5）

「有很多機會可以介接到很多單位的資料，我們在介接資料最麻煩的部分，真的不是量跟速度，而是資料格式跟標準化的部分」（受訪者 B6）

受訪者的意見也呼應了前文中資料治理的組成元素與其生命週期，組織中的資料資產（data assets）必須從源頭起即有妥善的蒐集方法以確保其品質，通常基層業務主管單位在此肩負此蒐集與保管的重責，而中游通常由資訊單位主責的資料處理也難免面臨必須再加工確保資料的一致、正確、完整、與即時程度，以貢獻於下游統計單位或

業務主管及決策單位，這整個週期的資料價值鍊（value chain），不但牽涉可能牽涉同一個政府機關中的業務/資訊/統計/法規等單位，也頗有可能需要跨政府機關間的協力合作。

「大數據是一種跨領域的合作，通常會分成幾個步驟，首先是 data collection（資料蒐集），收到資料以後我們會有一個專門做 data processing（資料處理）的部分，將資料處理到可以被分析的狀態，才會進到 data analytics（資料分析），分析出的結果才能夠去支持某些特定決策與產生價值」（受訪者 B5）

第三節 政府機關導入巨量資料的搭配資源與因素

也由於前述巨量資料分析的特質，政府機關必須搭配權責（authority）、資源（resource）、與能力（capability）以實現其備受期待的預期效益，也須同時關注於可能出現的風險。

一、問題與資料導向的巨量資料導入模式

首先，作為政府善用當代資通訊科技的數位治理方案之一，巨量資料分析除了硬軟體技術之外，首先應關注的是如何與政策與決策妥善連動（policy and decision aligned），這也凸顯了巨量資料分析兼具問題與資料導向的雙重路線，由於政府機關的業務特質、技術整備度（readiness）、與既定制度文化，可能在應用模式上就可能有所差異。

「政府的大數據分析應用，必須是由上而下。在政府部門真正掌管資料的是業務主管，可是在推動 open data 與 big data 則是資訊主管... 如果不懂得業務問題，就不會知道應

該要以議題導向來準備資料給外界」(受訪者 B1)

「使用 big data 上，我們的邏輯應該翻轉過來，也就是先有問題，然後再去想 data 是什麼」(受訪者 B2)

「應該是要採取兼容的角度，一方面是問題導向，另一方面也可以作一些議題的預測，但大數據並非是什麼問題都能解決，有時候也許改用其他傳統的方式會更有效率」(受訪者 B3)

由於傳統的決策支援概念需要相對清晰的決策問題陳述或學理基礎，屬於理論/問題導向的資料分析模式，而巨量資料被寄予厚望的原因之一，也包括透過資料挖掘 (mining)、類屬 (classification)、與關連 (association) 等由下而上的途徑，顯現與釐清跨組織跨領域資料背後的癥結與理性。在本計畫子題的參與實作過程中的經驗，預期也會以此作為經驗省思的角度，如有可能也能以資料導向 (data-driven) 的相關演算分析 (如決策樹，decision tree) 作為資料分析的方法。

二、政府機關內部的跨單位分工與協力

提到巨量資料分析所需的基本條件，應該就是基本的硬軟體資訊資源了，其中特定業務主管的資訊系統與資料庫 (database)、或是跨系統與資料庫的資料倉儲 (data warehouse) 可說是基本所需，即有受訪者提到「準備的工作非常重要，如果沒有前面的資料倉儲、資料超市，或 data mining 等，不可能會有今天」(受訪者 B1)，而這些基礎資訊資源的準備通常需要政府機關內部資訊單位的協助取得與部署，不過值得注意的是，僅仰賴資訊單位其實也無法妥善規劃執行巨量資料分析，如前述資料治理的組成元素即可看出，業務問題與決策需求才有助於界定巨量資料所需的資料與分析規格。即使是單一政府機關

內跨業務的需求，基於資訊系統與資料的正確／一致／精確／即時／完整，也必須組成跨單位的任務編組或專案團隊，以單一政府機關內部為基礎，尤其是業務、資訊、統計、法規、與研究發展（綜合規劃）單位間的橫向合作，可說是成功導入與應用大數據的絕佳組合；而且如能以個別機關的政策或行政服務為起點，其實也更有助於引發跨機關資料的連結需求。

「在資料週期的運作中，資料的使用者與決策制定者也需要加入，也就是所謂的 domain knowledge（領域知識），才有辦法使資料被完善運用... 需要透過說明會、交流會或教育訓練，否則業務單位與資訊單位間會產生很大的隔閡」（受訪者 B5）

「大多會把業務單位與資訊單位找來開會，彼此提出需求與互相支援的方法，大部分不會使用由上而下的交辦方式，而是採取成立專案型組織執行這樣的業務」（受訪者 B5）

「最常看到的方式是各個部門自己去盤點自己有什麼樣的資料...這樣會造成很多重疊的狀況，如果沒有一個整合性或統合性的方式或流程，去讓所有資料作整合，其實這也只能形成一個倉庫而已，而且這個倉庫的東西也不是那麼有結構的可以被利用與加值」（受訪者 B6）

其次，更為細緻的團隊運作則應進一步考量各單位的分工，例如有受訪者指出：「首先是流程的問題，不管是由上而下或由下而上，我們必須去找出內部資訊流通的可能發起點為何，資料庫的掌管者不必挖空心思去想資料去要如何建置，我把這些定位在資料的提供端，重要的是要去回應提問者，瞭解哪些東西是可以使用的，在現有的東西中，我要怎麼要去提供給對方是可以省去麻煩的」（受訪者 B2）。

甚至有受訪者主張業務單位應主導此專案團隊的運作，資訊單位僅在技術規格與資安方面予以支援協助，此主張頗不同於我國實務上通常由資訊單位。另有受訪者指出，我國許多政府機關的資訊單位由於過度仰賴廠商建置維護其資訊系統，已經喪失了應用機關本身業務資料的規劃與執行能力，反而是業務單位如能有充足的資訊素養，如能與委外廠商妥善搭配，對於機關本身的大數據導入應用將有所助益。

「負責 IT 部分的人應該建立一個完善的平台，去支持資料蒐集的人，讓資料能夠在這個平台穩定的維護與運用，這應該是 IT 部門與資料蒐集部門擁有的一個明確界限，這也可以視為是一個生命週期，不論是否為問題導向，從蒐集資料開始，IT 部門就應專心於提供一個良好環境進行資料的彙整，而非是以資料蒐集者的角度進行蒐集」（受訪者 B5）

「在執行業務的過程中，完全以業務單位為主導，資訊與統計為後勤，有很多準備工作完成之後，資訊單位就可以退出，由業務單位單獨執行」（受訪者 B5）

此外，即時是資訊人員也可能對於巨量資料分析所需且日新月異的統計方法不甚瞭解，因此統計單位也將成為此巨量資料專案團隊中的一員，或是專案團隊人員必須具備一定程度的統計素養與技能。即有受訪者表示：「統計方法與資料探勘的面向，資訊人員對於處理資料後，如何形成模型的能力較為欠缺，需要資訊、業務與統計的人才相互配合，才有辦法靈活應用，因此政府針對大數據應用的人才培育與訓練仍有進步的空間」（受訪者 B4）。

最後如同大部分涉及行政與政策的深度方案，法規也是巨量資料

分析所需的基石，尤其是可能設計敏感與隱私資料的巨量資料應用，即時不牽涉開放政府資料（open government data），也可能成為難以跨越的阻礙，例如有受訪者抱怨：「我們也一直被質疑我們蒐集的資料顆粒度過大，我們只能提供統計資料，細部的資料恐怕會觸碰到個資法的部分」（受訪者 B4），或是「法規因素讓資訊單位執行資料蒐集有所疑慮與擔憂，且法務單位將法規解釋的功能丟回給資訊與業務等單位，形成許多模糊的灰色地帶，讓這些單位執行業務上皆採取保守的態度，不敢進行突破性的改變」（受訪者 B4）；也有受訪者表示將以專責單位釐清修訂跨業務的資料治理相關法規與對應的管理方案：「針對法規制度、組織管理面進行研議相關配套措施，思考是否成立專責機關負責此項業務，以及公開這些資料需要何種法源配合等議題」（受訪者 A1），這對於以下所整理歸納的跨機關或跨業務的巨量資料分析更具有迫切性。

三、 跨不同政府機關間的分工與協力

在跨政府機關間的巨量資料應用中，相對「單純」的可能是單一業務體系的案例，例如考試院銓敘部、行政院人事行政總處、與地方政府人事單位間，或是中央與地方政府的地政業務體系；不過即使它們具有其共同的特定業務功能與流程（如廣義人事法規與政策），仍不免有不同的管轄範圍與職責，也可能導致在資料治理上的殊異價值觀與主張。

「資料的正確性、權責、維護等...由於目前的資料整合與正確度不足，相關主管機關所整理之資料有落差...資料正確性定義不一致，資料編碼需要整合」（受訪者 A2）

「由於資料的來源與特性五花八門，且會有漏答的情形，無法進行比對與分析。且上級機關的資料格式無法符合現

況，使得下級機關在整理上需要多花功夫進行考核。建議上級機關在蒐集資料時，能與下級機關進行溝通，瞭解各自的需求，並進行協調與調整，或制定相對的配套措施，如此才能確保資料的品質」（受訪者 B3）

「針對正確性設計有激勵的措施，但反應其實有限，原因在於上級機關與基層機關的需求有落差，基層機關的需求比較偏向作業層次，因此無法產生共識」（受訪者 B4）

如果跨機關間同時也涵蓋了不同的業務體系，對於妥善規劃執行巨量資料分析應用就更具有挑戰性了，不過這也正是巨量資料有別於傳統決策支援系統與最被期待的預期效益。如以整段資料治理的生命週期來看，最初的資料蒐集與定義即可能出現頗為不一致的情境，而以上所述在法規與資源上的配套要求也勢必更加嚴格，也更需要高階主管的推動決心與穩定支持。

「最近有一個案例就是新住民與其第二代...衍生出一串問題跨越許多部會，但沒有人知道要準備什麼資料，到底要怎麼作資料，格式是什麼？這凸顯出一個問題，雖然是由上而下的議題導向，但負責的單位都沒有一個因應的計畫，讓輔助的單位去依循並幫助」（受訪者 B1）

「專業不同，需要準備的資料也會不同，但不見得符合其他單位的需求，專業解讀不同，其蒐集資料的管道也會不同，會變成議題導向的資訊需要重新蒐集。政府機關的資料蒐集各自領政，各自的角度與專業，不但是格式不一致，資料的品質與內容也不一致，資料治理最重要的問題不是那些格式，而是內容與品質」（受訪者 B1）

「在資料的即時更新上，如果能與其他機關合作，也會使

我們的資料有效程度提升，但其他機關也會有疑慮，因為我們不是法律上的權責機關，它們提供資料上也會有風險」（受訪者 B4）

「有許多跨機關、跨領域的資料數據分析應用，例如電子發票與健康管理的配合應用，或是花粉熱、行道樹與健保資料的相互應用，這些都可以提供作為公共政策的建議。業務單位在推動上也面臨了許多挫折，政府部門在運作上有很多限制，如果沒有高層主管的推動決心，很難讓跨部門或跨領域的資料分析順利成功」（受訪者 B1）

第四節 跨政府機關導入巨量資料的相關議題

如同政府機關大力推動的變革方案，各國與我國政府也期待巨量資料分析可以逐漸普及運用於各領域中，就有受訪者提到其機關首長肯定巨量資料分析對於循證政府治理的貢獻：「我們首長對於資訊使用也有專業性在，他也很難想像現在的政府竟然不是依靠證據來作事，而是利用經驗」（受訪者 B4）。然而過於急切地將巨量資料分析導入政府機關中，卻可能導致不適當的管考評估方法，例如有受訪者觀察到：「政府常使用許多 KPI 指標進行評比，但未必能符合現實，也不能代表這些資料的價值」（受訪者 B6）。

由於我國政府機關的巨量資料應用仍屬起步階段，本計劃據以主張在績效管考上應以鼓勵試辦為主，並且搭配相關管理措施鼓勵行動研究（action research）與經驗分享，相信也更能有助於以上跨機關、跨業務領域、以及跨業務/資訊/統計/法規單位間的經驗累積與深度融合於既有的行政與政策流程中。最後也有受訪者建議巨量資料分析也

應考量透過開放政府資料（open government data, OGD）廣納群眾智慧，以創新應用回饋刺激政府內部的資料治理流程；另外由於許多業務推動由基層地方政府負責，因此地方政府作為資料提供與品質確保的重要性也值得關注。

「政府部門運用 big data 時，有一個更重要的前提，就是各政府部門在推動 open data 如果作不到一定的程度，big data 是不可能成行」（受訪者 B1）

另外，為了解決跨政府機關的資料管理與整合問題，隸屬科技部國家實驗研究院的國家高速網路與計算中心（以下簡稱國網中心）目前已提供大數據分析平台，提供橫跨中央與地方政府機關、學術機構、與產業的應用，不論是對於大數據或開放政府資料而言，透過國網中心將跨政府機關資料有所管理整合，並提供足夠資安水準的數據分析與開放平台，的確能夠因應上述發現中所提到的有關跨業務資料格式、法規、與資訊安全的疑慮，雖然有受訪者提到擅長於技術服務的國網中心可能較無法發揮資料應用價值的整體規劃效益，但是如果妥善搭配業務與應用需求導向的政府、學術、與產業夥伴，將是頗值得期待的跨政府機關的應用推廣模式。

政府巨量資料分析與政策端應用效能提升之研析

第五章 網路輿情分析平台與群眾外包之實驗成果

本章說明本團隊以群眾外包方式進行輿情分析所進行的實驗以及實驗結果分析。

第一節 群眾外包輿情分析之實驗流程以及平台架設過程

如期中報告的說明，本團隊秉持「網路的問題用網路的方式解決」，規畫應用群眾外包（Crowdsourcing）（Howe, 2008）的作法來實驗人工判讀輿情的可行性，將大量的網路輿情資料的主觀判斷問題，經過分工設計後，讓廣大的群眾同時參與大量的輿情立場判定。

一、政大輿情分析平台與 PyBossa 之設計與應用

首先，就如何應用群眾外包進行網路輿情分析，必須先提出一些基本需求，接著是進行平台的挑選。經過一番評估後，我們選取 PyBossa¹¹這個以Python語言開發的開放源碼群眾外包平台，並以它建制了一個以群眾外包進行輿情分析的雛形系統，暫時稱為「政大輿情分析平台」。如圖 5所示，我們可以將網路輿情資料匯入此系統，讓群眾就一則一則的輿情資料進行判讀，這裡使用的就是上年度（2014年）就「自由經濟示範區」主題所獲致的網路輿情資料。

¹¹ PyBossa 的網址：<http://pybossa.com/>，事後發現（Borromeo and Toyama, 2015）也選了 PyBossa。



圖 5：群眾外包之網路輿情分析雛形系統

資料來源：本計畫。

目前我們正在進行實驗規劃與設計的步驟，主要將就以下四個面向評估實驗變項的設計，並持續研析如何對PyBossa進行進一步的客製化與調整，以利實驗之執行。

（一） 群眾的範圍

所有的專案都一律對外公開徵求群眾，還是要視專案的性質，有的專案會限縮群眾的範圍，甚至限於特定的對象？

（二） 群眾的資格限制

因為有些輿情的主題可能需要一定的專業背景方能有效判讀，

所以要如何對群眾設定資格限制（qualification）？

（三）輿情判讀的程序規範

個別群眾在進行輿情判讀時，是否要依循一定的程序規範？例如：每則輿情資料必須至少需經 3 位（奇數個）群眾判讀、每位群眾判讀的輿情則數是否要有限制等。

（四）輿情判讀結果的規則

輿情判讀結果是否直接以多數決來定案？還是要考慮個別群眾作答的可靠度，對群眾施以不同權重方式來加總後決定結果？

在PyBossa方面，我們針對以上幾個面向，整理出以下的機制與客製化功能。首先，PyBossa以專案（project）為單位來製定需要群眾協助的工作，並提供簡易的介面，供各專案工作資料的匯入。以我們的需求而言，工作資料就是搜集來的網路輿情資料。系統的使用者分為三種角色：匿名使用者，一般認證使用者和管理者。一般的專案的預設狀況是可以接受匿名使用者，也就是說對外公開徵求的匿名群眾。如果要限縮群眾的範圍，PyBossa有兩個機制，一是要求使用者註冊，以自選密碼登入；二是管理者可以替專案設定關關密碼，使用者必須輸入該密碼後，才能執行該專案的工作。因為使用者可以自行註冊成認證使用者，這兩種機制可以搭配使用，以限縮可以來執行判讀工作的使用者。

其次，因為參與的群眾是一群不特定的自願者，工作品質不易掌握。針對這點，Pybossa有教學功能，可以透過專案建立者設計的教程，一方面訓練受測者，另一方面也可測試讓他們的工作品質，在專案結束後，除了可以把工作結果匯出為CSV檔外，每個專案都有自己的統計頁面，方便該專案管理者檢視使用者答案的分佈的類型資訊，以及他們花多長的時間來完成該專案所有工作。藉此可以滿足部分群眾資格限制或群眾篩選的需求。

在輿情判讀的程序規範方面，PyBossa也有一些機制可以運用。例如：專案管理員可以設定每則輿情資料必須至少經過 n 位使用者判讀後，該項工作才算完成。至於若要限制每位使用者判讀的輿情則數，PyBossa目前無法直接支援，但PyBossa是開放原始碼專案，我們可以透過適度修改程式碼，滿足我們要的客製化需求，例如：提示使用者最少要判讀一定數量的輿情才能離開系統。

在輿情判讀結果的規則設定方面，我們可以運用PyBossa的兩種機制來協助。一是將專案的結果匯出為CVS檔，我們自行以程式進行後製，依自選規則，例如多數決，計算出專案中各則輿情資料的結果。另一種是支援不同權重的結果計算方式，這可透過PyBossa內稱之為校準（Calibration）的機制來完成。專案管理員可以安排一些已知答案的判讀工作，將其設定為校準任務，用來研判該答題人所提供答案的準確度。例如，如果一個使用者正確地回答了所有的校準任務，就可以給1點的權重，而如果使用者只答對50%，則可以給定為0.5的權重，藉此可達到不同使用者給予不同加權的效果。

接下來，我們要就以上四個實驗變項的面向，進一步設定可能的變項值，並依據不同的變項值制定不同的實驗情境，在進行網路輿情分析系統的客制化，以符合實驗之要求。最後就可以進行各項實驗，搜集結果後進行整體評估。

二、實驗設計之變數規劃

為求輿情分析平台的設置與實驗設計變數規劃得以相互配合，在實驗設計的規劃上，也將針對1.群眾的範圍、2.群眾的資格限制3.輿情判讀的程序規範4.輿情判讀結果的規則，等四大面向進行設計。

另外，根據Howe（2008）的建議，未來的群眾外包機制建立時，應該注意以下十點規則：1.選擇正確的模型（pick the right model）；2.選擇適合的群眾（pick the right crowd）；3.提供適當的誘因（offer the right incentive）；4.仍需持續招募既有員工（keep employing people）；

5.尋找善意的社群領導者(find benevolent dictators)；6.簡化任務(keep thing simple)；7.預設處置差疵不齊的資訊(be prepared for fluff)；8.試著在砂礫中找尋珍珠(Look for diamonds in the rough)；9.以群眾的意見為優先，可試著引導群眾，但避免過度的控制(The community's always right)；10.給予參與群眾合適的回饋(Give the crowd something)。這些規則與建議對於計劃團隊在平台建置與實驗變數設計上，也將具有引導性與啟發性。

再者，參考Borromeo與Toyama (2015)所進行的自動化與群眾外包的情緒分析實驗設計，在Borromeo與Toyama的實驗設計中，主要僅針對人工編碼、電腦自動化判定、以及有償與自願的群眾外包四種判讀機制進行區別與比較。但觀察其實驗設計的描述，仍有許多可操作的而未操作實驗變數可供參考。例如：在輿情判讀結果的規則上，除了以多數決為標準外，藉由增加多數決的人數，是否可能提升群眾對於輿論立場判讀的正確性？或者，對於具備特定背景或條件的群眾，其判讀的可信度的比重是否能有差異？另外，Borromeo與Toyama (2015)是以人工編碼者的判讀結果為絕對標準，雖然本計畫預設規劃也是如此，但群眾的立場判讀是否具絕對的標準，人工編碼者的判定結果又是否絕對正確，或者如何確定人工編碼者的判定結果是絕對正確的，這也是可再探討的問題。然而，若未來是希望以群眾外包機制為主要的輿情判定機制，那麼如何提升群眾判定的準確性，依然是我們在此應該追求的目標。

舉例來說，若是針對在群眾的範圍與資格限制上，除了以未預先設定的群眾為目標外，根據輿情資料的特性，我們也應可就理論或經驗上相信對該領域知識較熟悉與否，針對以特定背景的群眾為主要的徵求目標進行比較，或在事後區別參與群眾的個人背景而比較不同背景群眾的判讀結果。而可嘗試規劃問卷讓參與群眾在協助判讀後進行填答，而在問卷的設計上，也可增加詢問參與的參與動機與誘因，以更確實的掌握參與群眾的參與原因。

或者，在任務簡化上，平台的建立是否具備一定的便利性，或也可藉由前測或設置不同規格的平台進行比較。而針對任務本身，分配群眾特定數量的任務，或對其任務執执行程序規範的差異，是否會影響其參與動機與判讀結果，也將是可考慮的變數。

除上述外，根據政府對於政策議題輿論立場辨認的需求本質進行思考，民眾對於政府的信任程度、支持與否、政黨傾向，是否可能影響其參與動機以及參與行為。從政府與民眾之間的「供給—需求」關係互動進行思考，政府對輿情立場判讀的需求，對於在群眾外包機制下的民眾，某種程度上的服務供給者而言，這樣的需求與要求對於民眾而言是否具正當性以及合理性，其接受度又將如何，這些也都應是本計畫可深入探究的。而從這樣的角度思考，若政府以群眾外包為一新的治理途徑或工具時，該如何給予適當的論述，其徵求群眾的需求與立場是否具正當性，這都將涉及能否給予參與群眾合適的誘因與回饋、以及能否吸引到我們真正所需的群眾、匯集群眾之智慧。畢竟除了單純以提供財物為誘因外，任務本身符合群眾的興趣以及獲得群眾的認同，以及完成此任務後對於群眾的正向反饋，這些或許更是能吸引群眾參與的高層次誘因。而進一步思考，這過程也將影響群眾外包作為政府治理的新途徑時，除了在解決問題之外，能否也同時擴大民眾參與、深化民主、並提升政府與人民之間的信任，而可能避免將群眾外包的平台變成其特定「網軍」的戰場。

最後，在實驗設計的規劃上，如上所述，縱然有許多可深究之處，但仍必須以本計畫的目的為焦點，以提升群眾外包機制效能為出發點而進行規劃與取舍。此外，也仍需視技術上能否配合，即在平台建置能否符合實驗設計的需求而定。因此，最終的實驗設計將聚焦在：

- (一) 群眾外包與助理人工的輿情立場判定比較；
- (二) 說明與操作介面上的機構差異對於群眾判定是否具影響性；
- (三) 群眾進行立場判定時，是否會受到從眾效應的影響。

三、計畫進度

本計畫子題預定之研析時程將以10個月為度，並詳細各時程工作重點規劃如下表 10所示。目前已完成分析平台之比較優劣以及平台的架設測試，並規劃以Pybossa為我們主要的分析平台。後續將詳細規劃導入機關與其實作的公共服務與政策議題的內容與範圍，包括搭配此實作過程的資料蒐集活動，並預期在8月完成需求分析，而預計於11月初開始輿情分析平台之實驗，俾利在12月底產出結案報告初稿。另外，視資料蒐集狀況，平台將持續開放至1~2月間，以求盡可能的蒐集更多資訊，而可在最終結案報告時呈現更多的分析成果。

表 10：網路輿情分析平台實驗研析時程規劃表

預定時程 研析工作	104 年 6-7 月	104 年 7-8 月	104 年 9-10 月	104 年 11-12 月	105 年 1-2 月
研析分析平台（包括平台比較、平台架設測試）					
期中報告撰寫					
分析平台架設（包括需求分析）					
執行輿情分析實驗（包括群體測驗、數據分析）					
結案報告初稿撰寫					
最終結案報告撰寫					

資料來源：本計畫。

四、實驗平台架設與客製化運用

我們選取PyBossa¹²這個以Python語言開發的開放源碼群眾外包平台，來建構本次輿情分析平台之群眾外包實驗。PyBossa自2011年底推出以來，已廣為各界使用，包括美國太空總署（NASA）與歐洲高能物理研究中心（CERN）都曾以它建構citizen science的群眾外包平台。就在我們以它建制以群眾外包進行輿情分析的雛形平台時，我們也得知了日本學者Borromeo & Toyama最近也以PyBossa建構的群眾外包平台進行輿情分析，這也增加了我們不採用商業平台，而選用此一開放原始碼平台的信心。

PyBossa平台雖然免費，但它的架設與客製化運用需要資訊科技專業人員的投入。以本計畫而言，平台建置和管理上主要由一名資訊科學專長的協同主持人帶領兩位資訊科學與資訊管理的學生助理參與，其中一位助理熟悉系統工程的學生負責平台（包含搭配的Web伺服器與資料庫系統）的架設，另一位具備應用系統開發經驗的學生則負責平台的客製化擴充。此外，並由一名公共行政學系的博士生助理針對平台功能建置與實驗設計的需求配合上進行溝通，並參與平台建置、介面設計、與功能規劃。以下分別就本次的實驗平台架構、PyBossa主要功能設定、與客製化擴充作說明。

（一）實驗平台架設

本次實驗以PyBossa平台為主，包含Web伺服器（Nginx）與Python/Web介面元件uWSGI¹³以及PostgreSQL資料庫系統，此外，我們也應用了BaaS服務商Parse.com的服務來儲存參與群眾的基本資料（姓名與地址），以利發放完成任務的贈品。以下為本次實驗平台的系統架構圖：

¹² PyBossa 的官方網址：<http://pybossa.com/>

¹³ uWSGI 這個程式元件是為了讓 Python 所撰寫的 Web 應用程式能夠跟 Nginx、Apache 等等的 Web 伺服器溝通而設計的介面協定，類似標準的 CGI 介面。

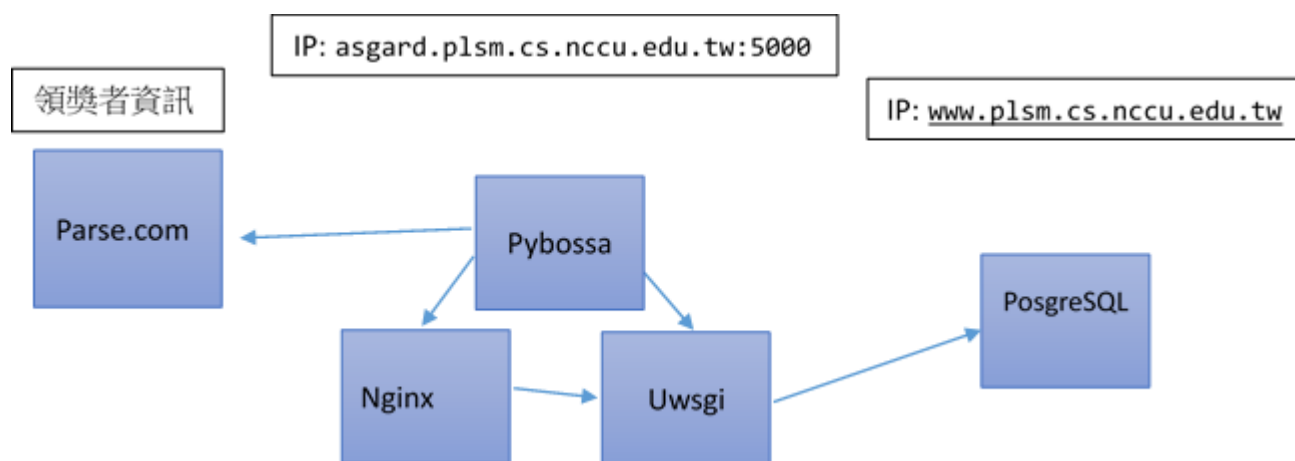


圖 6：實驗平台之系統架構圖

資料來源：本計畫。

PyBossa的功能相當完整，除了提供使用者相當多的自行設定功能外，也可以透過程式介面進行客製化與功能性擴充。以下條列我們認為PyBossa所具備的重要的可設定功能，並說明本次實驗所採用的設定。

- 1、**群眾的選擇**：PyBossa 的「群眾」（使用者）分為匿名者（anonymous）與以電子郵件帳號註冊過的使用者。每個群眾外包的專案都可以設定其實驗對象是匿名的就可以，還是必須是註冊過的使用者。此外，專案的管理員還可以替專案設定密碼，只接受知道密碼的群眾來進行專案實驗。對本計畫的專案而言，由於我們需要確定使用者身分，以利統計結果並發放禮物，所以我們要求使用者必須先進行註冊。
- 2、**群眾的資格限制**：此次實驗參與的群眾是一群不特定的自願者，並無設定參與資格。但就非特定背景之群眾的工作品質不易掌握此點，Pybossa 設有教學功能，專案建立者可設計教學流程，一方面訓練受測者，另一方面也可測試讓他們的工作品質。在專案結束後，除了可以把工作結果匯出為 CSV 檔外，每個專案

都有自己的統計頁面，方便該專案管理者檢視使用者答案的分佈的類型資訊，以及他們花多長的時間來完成該專案所有工作。藉此可以滿足部分群眾資格限制或群眾篩選的需求。但此次實驗本就是希望以不特定之參與群眾進行立場判定，雖然對於判定輿情主題「自由經濟示範區」的政策有一定的理解應該能提升其立場判定的效度，但並非必要條件。唯計畫團隊還是在任務開始前的說明頁面中，提供國發會網站上的「自由經濟示範區」的政策概要說明（懶人包），以利群眾事前參考。

- 3、**群眾進行輿情判讀的程序規範：**個別群眾在進行輿情判讀時，系統在顯示輿情資料的順序上是否要依循一定的程序規範？例如：輿情資料的顯示順序，是一則一則依序顯示，還是先顯示最少人已作答的輿情資料？又或是隨機顯示？第一種可視為深度優先，第二種則為廣度優先，PyBossa 三種都可支援。下圖揭示當作答題數至 15 題時，兩者之差異。

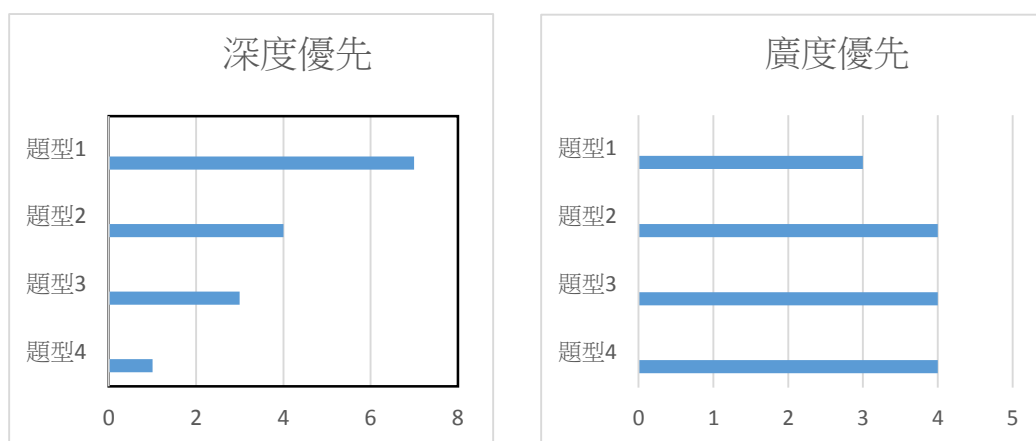


圖 7：群眾判讀輿情資料的可能順序

資料來源：本計畫。

本次實驗是採取以廣度優先為主的設定，其目的是為了平均每則輿情資料被判定的次數，以避免不同資料被判定的次數間出現明顯落差。其次，專案管理員也可以設定每則輿情資料必須至少經過數名使用者判讀後，該項任務才算完成，惟本次實驗並尚未應用此項設定。而本次實驗設定每位群眾必須完成所有被系統分配的35筆輿情資料的判讀才算完成任務，並獲得領取贈品的資格。但未完成全部任務的判定資料，也將被系統保存紀錄。

- 4、**群眾輿情判讀結果**：我們可以運用 PyBossa 的兩種機制來協助決定判讀結果。一是將專案的結果匯出為 CVS 檔，我們自行以程式進行後製，依自選規則，例如多數決，計算出專案中各則輿情資料的結果。另一種是支援不同權重的結果計算方式，這可透過 PyBossa 的校準（Calibration）機制完成。即專案管理員可以安排一些已知答案的判讀工作，將其設定為校準任務，

用來研判該答題人所提供答案的準確度。例如，如果一個使用者正確地回答了所有的校準任務，就可以給 1 點的權重，而如果使用者只答對 50%，則可以給定為 0.5 的權重，藉此可達到不同使用者給予不同加權的效果。本次專案則選擇前者，以程式採多數決的方式決定群眾判定的結果。

（二）實驗平台之客製化擴充

圖 8 展示本實驗進行之主要流程。民眾在接收到本次實驗的廣告或推廣資訊，並點選連結進入我們的平台首頁後，必須先以電子郵件帳號註冊為平台的使用者。平台系統在接到註冊要求後，會自動發出電子郵件進行使用者認證。使用者點選認證電子郵件中的連結後，將被系統隨機分配至四個不同情境介面的專案之中的任一專案進行輿情立場判讀任務。將所有的任務完成後，平台會要求使用者填寫簡單的個人資訊，以利後續寄送贈品給完成任務的使用者。

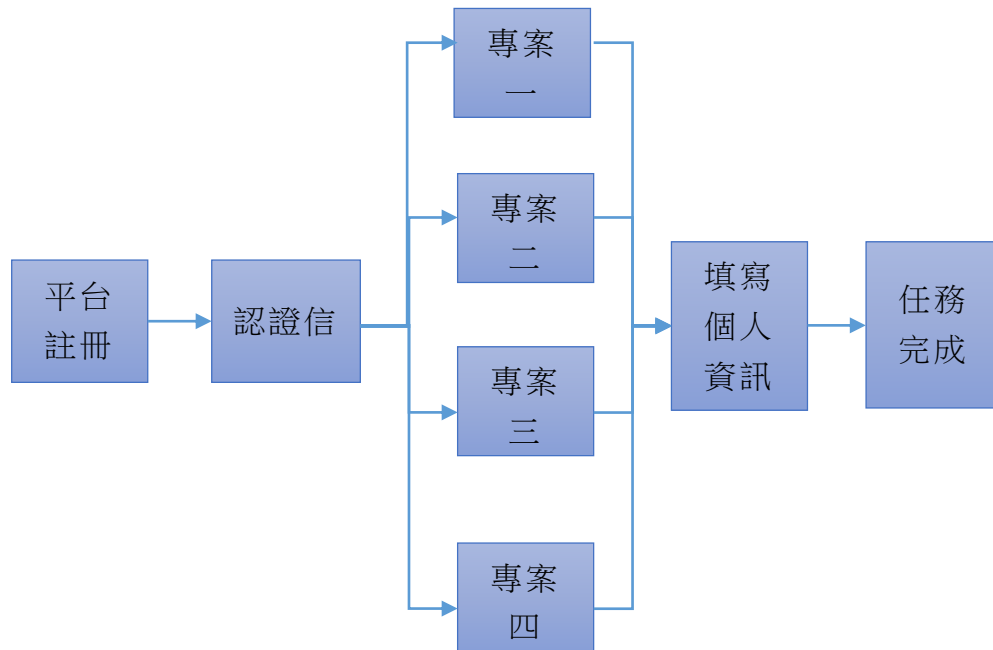


圖 8：本次實驗之主要流程圖

資料來源：本計畫。

基於實驗設計說明所提到的理由，本次實驗以PyBossa制訂了四個專案，不同專案的提供給使用者判定的輿情資料都是一樣的，但根據機構效應、以及從眾效應的實驗設計，在一開始的任務說明、以及介面配置上有所差異。不過輿情文章立場判定的功能介面則是一致的，如下圖所示，每則輿情文章的立場判定主要分為三大步驟：1.閱讀訊息內容、2.挑選幫助判定立場的關鍵語句、以及3.確定立場判定結果。

自由經濟示範區之網路輿情的立場判定實驗

1. 請閱讀下列關於自經區的網路輿情，並判斷下段文字是否支持自經區的政策

教育部一再宣稱首要目標在於「標準學習」，但卻看不到開放對於當前高等教育亂象可能製造的問題，避重就輕地回應文中提出的根本疑問，讓公民團體、教師工會、家長、老師、學生以及更多鄉民參與決策過程，才能避免自經區淪為下一個黑箱作業的服貿協定。

1. 首先，請您閱讀下面黑色字體的文章，並判斷「文章作者」對於自經區政策的立場為何

2. 當滑鼠移到上方文章時，您所指到的句子便會變成紅色字體並有底線。請您點選您認為該篇文章中，最能幫助你判斷作者立場的關鍵句。

2. 點選上段文章中，讓您下此判斷的關鍵句子(請選擇三句話)

讓公民團體、教師工會、家長、老師、學生以及更多鄉民參與決策過程 x

才能避免自經區淪為下一個黑箱作業的服貿協定 x

您所點選的句子將會自動出現在白色對話框內，如同上方的綠色文字與方塊。若需要修改，請按下綠色文字方塊右側的X，消除該文字方塊後，重新選擇關鍵句。

3. 您認為上述言論的立場是支持或反對自經區的設立?

☒ 支持

☐ 反對

☐ 中立

☐ 無法判斷

3. 最後，點選您對以上文章內容的判斷選項。
若您認為文章立場偏向支持自經區，便點選第1個圓圈；
若為反對，請點選第2個圓圈。
文章立場中立或無法判斷，則分別點選第3與第4個圓圈。

已判定文章: 3

剩餘文章: 32

您可以在這邊看到您目前的判斷進度。

4. 點選以下方塊前往下一題，並繼續以上步驟

下一題

圖 9：輿情立場判定之任務執行介面

資料來源：本計畫。

為配合本次實驗設計，我們進行了一些客製化擴充，主要包含四個主要部分：電子郵件認證、隨機指定專案、首頁與專案網頁介面與群眾個資蒐集，分別概述如後。

- 1、**電子郵件認證**：我們利用 PyBossa 的內建機制，自行開發一個電子郵件認證的程式模組，一方面可以藉此確認使用者註

冊時填寫的是真實的電子郵件帳號，另一方面也可藉此引導使用者隨機進入由系統分配的輿情立場判定專案。若任由使用者就四個專案挑選，無法遂行我們的實驗目地。

- 2、**隨機指定專案**：如前述說明，為了要測試機構效應與群眾效應兩大變數，設計了四個不同介面的專案進行實驗，因此計畫團隊開發了一個隨機指定專案的模組，將使用者隨機分配到其中一個專案。
- 3、**首頁與專案網頁介面**：為了強化可能影響機構效應與群眾效應的變數，除了首頁介面的說明必須客製外，四個專案的操作介面也有設計差異。例如下圖 10 顯示為有顯示其他人作答結果統計的畫面，與上圖 9 的主要差異在於左下角顯示了此篇輿情目前所有群眾的判定結果統計（圖 10 尚未有群眾開始判定的情況，因此各項數值均為 0，各選項的數值將隨著群眾的判定結果即時地變動）。計畫團隊便以此差異營造是否可能產生從眾效應的變數。對此預設的假設是，能夠看到其他群眾判定結果的使用者，會更傾向於和多數人的判定結果而使得該專案的整體判定結果，比起看不到其他群眾判定結果的專案，會具有較高的判定結果一致性。
- 4、**群眾個資蒐集**：為了取得正確的寄送贈品資訊，我們也開發了一個使用者簡單個資蒐集模組，在使用者完成所有輿情訊息的立場判讀後執行，要求使用者填入姓名與地址。此部分為求簡化，我們並沒有去修改擴充 PyBossa 的資料庫表格，而是另外使用第三方服務提供者 Parse.com 所提供資料存取服務，將資料存放在那裡。

自由經濟示範區之網路輿情的立場判定實驗

1. 請閱讀下列關於自經區的網路輿情，並判斷下段文字是否支持自經區的政策

教育部一再宣稱首要目標在於「標準學習」，但卻看不到開放對於當前高等教育亂象可能製造的問題，避重就輕地回應文中提出的根本疑問。讓公民團體、教師工會、家長、老師、學生以及更多鄉民參與決策過程，才能避免自經區淪為下一個黑箱作業的服貿協定。

1. 首先，請您閱讀下面黑色字體的文章，並判斷「文章作者」對於自經區政策的立場為何

2. 點選上段文章中，讓您下此判斷的關鍵句子(請選擇三句話)

讓公民團體、教師工會、家長、老師、學生以及更多鄉民參與決策過程 x

才能避免自經區淪為下一個黑箱作業的服貿協定 x

您所點選的句子將會自動出現在白色對話框內，如同上方的綠色文字與方塊。若需要修改，請按下綠色文字方塊右側的X，消除該文字方塊後，重新選擇關鍵句。

2. 當滑鼠移到上方文章時，您所指到的句子便會變成紅色字體並有底線。請您點選您認為該篇文章中，最能幫助你判斷作者立場的關鍵句。

3. 您認為上述言論的立場是支持或反對自經區的設立?

☐ 支持
☒ 反對
☐ 中立
☐ 無法判斷

3. 最後，點選您對以上文章內容的判斷選項。
若您認為文章立場偏向支持自經區，便點選第 1 個圓圈；
若為反對，請點選第 2 個圓圈。
文章立場中立或無法判斷，則分別點選第 3 與第 4 個圓圈。

4. 點選以下方塊前往下一題，並繼續以上步驟

下一題

目前投票狀況

支持	反對	中立	無法判斷
0	0	0	0

已判定文章: 4 剩餘文章: 31

您可以在這邊看到此篇文章目前其他人的判斷狀況，以及您目前的判斷進度。

圖 10：提供判讀結果統計的專案執行介面

資料來源：本計畫。

第二節 網路輿情平台之活動推廣與運作概況

本計畫子題除了建置輿情平台於 PyBossa 之外，為了加強推廣以便邀請不特定的群眾參與活動，並於 Facebook 平台上建立粉絲專頁，進行推廣以及和參與群眾進行互動。以「全民參與、齊力鍛金：政大輿情平台」為名，建立粉絲專頁，並突破 FB 所認可按讚人數限制而取得設定固定網址的權限¹⁴。

為了有效推廣本次活動，計畫團隊並同時採取付費以及免費的多重管道同時進行宣傳。關於付費管道，在初步的了解與規劃後主要設定兩種方案，第一：向 google 或 yahoo 購買關鍵字。第二：在 facebook 刊登「網站點擊次數廣告」。這兩種方式都是以點擊數的多寡來決定廣告的預算，差別在於一個是以搜尋關鍵字，另外一個是從 FB 的廣告放送直接連到我們的網站。

而考量本次輿情分析的活動並不希望在一開始便限定針對特殊群體進行宣傳，並希望能盡可能地讓訊息接觸到各種不同背景和興趣的網路使用者。因此，決定配合 FB 粉絲團的設立，並在 FB 刊登網站點擊次數廣告，在除設定在臺灣地區發送廣告外，並不另外設定其他廣告受眾的限制。

另一方面，在免費的網路宣傳上，計畫團隊在事前調查臺灣主要的網路論壇，例如台大批踢踢實業坊 BBS (PTT)¹⁵、Dcard¹⁶、伊莉討論區¹⁷等。並在符合各大論壇各自的規定下，發布活動資訊與網站連結。

¹⁴全民參與、齊力鍛金：政大輿情平台粉絲專頁之固定網址為：
<https://www.facebook.com/nccusdap>

¹⁵ 台大批踢踢實業坊 telnet://ptt.cc

¹⁶ 網址：<https://www.dcard.tw/>

¹⁷ <http://www03.eyny.com/index.php>

表 11：不同之常見網路付費廣告比較

不同之常見網路付費廣告比較	
一、Google關鍵字	
Google關鍵字之刊登手續	需填寫廣告委刊單，內含基本資料、每日預算、廣告網址、購買廣告總金額。填寫完列印之後請簽章，以傳真及郵寄方式回覆，待確認刊登資料完整後，即為您安排專屬企劃人員，進行關鍵字廣告製作事宜(關鍵字挑選、廣告文案內容設計)，約需三~五個工作天，於確認付款後即可安排廣告上線。
申請表格填寫	表格內容包括：公司名稱、連絡人/職稱、辦公電話、手機電話、E-mail、廣告目標網址(官網)、目標市場、刊登需求選擇、廣告預算金額、方便連絡時間、與備註資訊
計費方式	Google AdWords 關鍵字廣告授權代理商會向廣告客戶收取帳戶開戶費NT3,000元以及廣告費的20%為服務費。
二、Yahoo奇摩關鍵字	
計費方式	(一) 點選計費制：廣告採點選計費，只有當網路使用者實際點選廣告並進如您的網站時，才會開始收費。
	(二) 關鍵字起標：針對特定關鍵字設定標價，初始設定之起標價約為NT\$3上下。不過同一關鍵字可能會同時有不同廣告主進行競標，因此起標價主要取決於：

	<p>1. 該關鍵字的熱門程度，由競價相同關鍵字的廣告主數量，以及他們所出價的金額同時決定，因此越熱門的關鍵字起標價就會相對較高；</p> <p>2. 關鍵字與廣告的品質，與同業相比，若網路使用者越偏好點選您的廣告，代表廣告的品質相對較高，起標價也會相對較低。</p>
三、Facebook廣告	
FB廣告形式	廣告可出現在行動版和桌面版的FB動態消息及／或右方欄位。並配合FB的使用者資料庫，可自行設定廣告受眾。
廣告費用	在Facebook 上刊登廣告與您於當地報章發佈廣告一樣需要收費。廣告主可以為廣告設定「單日預算」或「總經費」來自由控制費用。除了廣告主自行設定的費用外，並不會產生其他額外費用。
廣告成效管理	<p>廣告會被多少人看見或點擊，則取決於諸多因素，包括您的預算、您選擇的投標類型（是單次點擊成本（CPC）還是每千次廣告曝光成本（CPM）），以及您鎖定的廣告受眾（是廣泛的受眾，還是具有特定興趣、行為和人口統計資料的受眾）等等。</p> <p>FB並會提供廣告管理介面，讓廣告主可隨時管理、修改、終止廣告、或是選擇特定的時間點、發送廣告，並可即時觀看目前廣告的發放成效，以及廣告受眾分析。</p>

資料來源：本計畫。

以本次計畫團隊使用FB付費廣告為例，廣告成效如下表所示。從2015/11/5日正式發放廣告，至2015/12/4之間，廣告的總處及人數達101,109，其表示有機會看到我們所發送的廣告人數。不過實際點擊廣告並進入平台首頁的人數，為3,073人。在此便可明顯發現有機會看到廣告的人數以及實際會點擊廣告進入平台首頁人數之間的落差。在所有理論上可接觸到廣告訊息的受眾中，約僅有0.03%的廣告受眾真正留意到廣告資訊，並受到廣告內容吸引而實際進入平台首頁。然而，在進入平台首頁後，是否願意註冊活動帳號、嘗試進行任務、以及實際完成所有任務並留下個人資料，在這些步驟之間，人數將會持續的遞減流失。這些資訊的取得，能讓我們更精確地了解到活動的成效，以及可能大量減少參與者的階段為何，而有機會及早發現問題已進行調整。因此，本次計畫並已將Google Analytic實際導入於本次輿情平台的網站，在後續的其他章節中，本計畫子題也將以本次經營平台網站的實際經驗為例，展現相關的網站使用者行為資訊分析能力。

表 12：輿情平台 FB 廣告成效

廣告專案名稱	政大輿情平台連結點擊次數
分析報告開始時間	2015/11/5
分析報告結束時間	2015/12/4
廣告投放狀態	持續中
廣告資訊觸及人數	101,109
實際點擊總人數	3,073
每次點擊的成效成本（TWD）	1.43
總支出金額（TWD）	4409

資料來源：本計畫。

FB的廣告管理本身也提供了多樣化的廣告受眾資訊分析給廣告主。其主要的優勢在於結合其FB使用者的資訊，而可有效掌握廣告受眾的相關資訊。例如基本的年齡、性別、所處國家、地區、連線裝置等等。以下表的廣告受眾年齡分析為例，FB的廣告管理便提供了各個年齡層或是更進一步地細分不同年齡與性別的廣告觸及效果、實際點擊網站的次數分布、並針對支出金額的效益進行估算。如此，我們便可清楚地了解到針對不同背景的廣告受眾之廣告發送成本效益，而得以循證地調整廣告行動方案。

表 13：輿情平台 FB 廣告受眾年齡與成效分析

廣告受眾年齡	網站點擊次數 成果	觸及人數	每次成效的成本 (TWD)	支出金額 (TWD)
18-24	457	15798	1.376368	629
25-34	382	16021	1.570681	600
35-44	475	20279	1.557895	740
45-54	806	26829	1.395782	1125
55-64	661	16479	1.373676	908
65+	292	5706	1.393836	407
總計	3073	101113	1.434754	4409

資料來源：本計畫。

表 14：輿情平台 FB 廣告受眾年齡/性別之成效分析

廣告受眾年齡與性別	點擊次數成果	廣告觸及人數	支出金額
18-24女性	263	9,318	NT\$362
18-24男性	194	6,464	NT\$267
18-24不明	-	20	-
25-34女性	216	9,310	NT\$312
25-34男性	165	6,627	NT\$279
25-34不明	1	100	NT\$9
35-44女性	278	11,428	NT\$423
35-44男性	196	8,793	NT\$316
35-44不明	1	80	NT\$2
45-54女性	382	12,311	NT\$514
45-54男性	422	14,477	NT\$607
45-54不明	2	52	NT\$3
55-64女性	233	6,321	NT\$363
55-64男性	427	10,150	NT\$544
55-64不明	1	24	NT\$2

廣告受眾年齡與性別	點擊次數成果	廣告觸及人數	支出金額
65+女性	97	2,344	NT\$152
65+男性	195	3,336	NT\$255
65+不明	-	32	-
總計	3070	101,075	4,404

資料來源：本計畫。

除上述展現的廣告受眾年齡、性別之外，FB的廣告管理並可提供國家、國家／地區、廣告版位、廣告版位與連線裝置等不同的廣告成效與廣告受眾之分析。並提供多種分析欄位的交叉樞紐表以及評估指標的篩選。因此，網站分析所能提供的資訊，相較過去傳統的網站成效評估、使用者經驗意見調查相比，所能提供的資訊是更為客觀且即時性的。並隨著網站管理者的設定，而可事先記錄各種不同的特定資訊。以此次計畫在FB廣告管理上的成果分析為例，便展現了網站分析與管理的基本圖像、和傳統偏重主觀意見的網站成效調查的差異、以及透過網站分析技術，可能如何及時地對行動方案成效進行監控與調整，並有效地進行事後成效評估

第三節 群眾與人工之輿情立場判定結果比較分析

回到本次輿情平台實驗的分析重點，即藉由一般群眾的分工模式進行大量網路輿情立場判定之結果，與透過事先培訓的少數助理進行人工立場判定結果之間的比較分析。本次實驗採用政府應用巨量資料精進公共服務與政策分析之可行性研究（蕭乃沂、陳敦源、廖洲棚，2014）中所使用的自由經濟貿易區網路輿情分析資料，以及該委託研究案中助理進行人工判定結果為比較基準，在此次輿情平台中使用

相同的350篇輿情文章讓一般不特定的群眾進行立場判定，而最終群眾判定的結果則以簡單多數決的方式計算，來決定不特定多數群眾對於特定輿情文章的立場判定。

從活動正式開始的2015/11/2號開始，本次研析擷取至2015/11/26時所有的群眾判定結果進行比較分析。在完成所有任務的人次上，達223人。針對350篇輿情文章，總共取得8752筆判定結果。每篇輿情文章被不同群眾判定的次數為23~28人次之間，平均每篇輿情文章的判定人次達25次，在數量上已遠遠超越原本進行人工立場判定助理人數3位。

此次平台的設計，讓群眾針對每篇文章的立場可決定為支持自經區設立、反對自經區設立、持中立意見、以及無法判斷四種。為此部分與去年助理人工判斷有些微差異。去年助理判斷的結果主要分為，支持、反對、無法判斷、以及同時皆有支持與反對意見。因此在群眾的判定的中立、無法判斷，以及助理判定的無法判斷、以及同時皆有支持與反對意見之間在判斷思考的邏輯上會有所落差，而可能造成兩造意見之間具有較大的偏誤。

除此之外，由於群眾判定的最終結果是以多數決的方式進行，因此可能出現最高票選項不只一個的情況，而出現一種無法判斷最終結果的狀況，團隊將其稱為模稜兩可。因此，本次研析將針對群眾和助理判定為中立、無法判斷、或同時皆有支持與反對意見、以及模稜兩可的狀況（設為遺漏值）試著進行不同的編碼結果分析，並在最後主要以群眾和助理判定為支持與反對的結果進行比較。

一、 全體群眾判斷與助理判斷立場之比較分析

如前段所述，首先，在一開始的分析我們將先保持原始的編碼結果，但將群眾呈現模稜兩可而無法決定最終立場的結果設為遺漏值。如下表所示，在350篇文章中，共有6篇文章為目前群眾仍無法有優勢意見而無法決定判斷。

表 15：觀察值處理摘要-群眾判斷結果 * 助理判斷結果

	觀察值					
	有效		遺漏		總計	
	N	百分比	N	百分比	N	百分比
群眾判斷結果 * 助理判斷結果	344	98.3%	6	1.7%	350	100.0%

資料來源：本計畫。

針對群眾以及助理的判定結果，除了支持、反對沒有問題外，我們先將助理判斷之正反皆有和群眾的無法判斷設為同一結果，而將助理無法判斷的結果，和群眾判斷為中立的結果視為相同。其針對 344 篇輿情文章判定結果的交叉表如下。

在 344 篇文章中，助理共判斷出 60 篇反對立場文章、而判斷 127 篇支持立場的文章。但群眾但判斷出 108 篇的反對文章、以及 155 篇的支持文章，並主要從助理判定為無法判斷的文章中所增加。因此，群眾在助理所認為無法判斷的文章之中，更多地判斷出立場為支持或反對的文章。因而在結果上具有相當大的差異。

造成如此結果的可能有二，一為助理對於立場判定的標準較為保守，二則可能是群眾對於支持或反對立場的判斷更為敏感。但目前我們尚無法得知實際原因為何。不過，若僅針對助理判定為反對或支持的結果來看。助理判定為反對的 60 篇文章中，群眾也做出相同判斷的數量為 51 篇，比例高達 85%；而助理判斷立場為支持的 127 篇文章中，群眾也做出一致判斷的數量達 100 篇，一致比例為 78.7%。以助理判斷出支持或反對立場的文章上，兩者判定結果的一致性比例超過 80%。從此可初步發現助理和群眾對於立場可能較為明確的文章，在判斷上的差異並不大；此外，群眾對於助理們較為保守判斷為無法判斷的文章中，可能具有更敏感的能力去辨認其立場，或者說是較積極

地去判斷文章立場是屬於支持或反對。

表 16：群眾判斷結果 * 助理判斷結果交叉表（原始）

			群眾判斷				總計
			反對	支持	正反皆有 /無法判斷	無法判斷 /neutral	
助理判斷	反對	計數	51	8	0	1	60
		助理判斷的 %	85.0%	13.3%	0.0%	1.7%	100.0%
		群眾判斷的 %	47.2%	5.2%	0.0%	1.5%	17.4%
		佔總計的百分比	14.8%	2.3%	0.0%	0.3%	17.4%
	支持	計數	15	100	0	12	127
		助理判斷的 %	11.8%	78.7%	0.0%	9.4%	100.0%
		群眾判斷的 %	13.9%	64.5%	0.0%	18.2%	36.9%
		佔總計的百分比	4.4%	29.1%	0.0%	3.5%	36.9%
	正反皆有/ 無法判斷	計數	8	6	0	4	18
		助理判斷的 %	44.4%	33.3%	0.0%	22.2%	100.0%
		群眾判斷內的 %	7.4%	3.9%	0.0%	6.1%	5.2%
		佔總計的百分比	2.3%	1.7%	0.0%	1.2%	5.2%
	無法判斷 /neutral	計數	34	41	15	49	139
		助理判斷的 %	24.5%	29.5%	10.8%	35.3%	100.0%
		群眾判斷的 %	31.5%	26.5%	100.0%	74.2%	40.4%
		佔總計的百分比	9.9%	11.9%	4.4%	14.2%	40.4%
總計	計數		108	155	15	66	344
	助理判斷 內的 %		31.4%	45.1%	4.4%	19.2%	100.0%
	群眾判斷內的 %		100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百分比		31.4%	45.1%	4.4%	19.2%	100.0%

資料來源：本計畫。

不過，若以Kappa一致性係數進行檢定，發現整體群眾判斷和助理判斷的結果其Kappa係數僅為0.401，p值為0.000為高度顯著。因此，雖然在助理判斷為支持或反對的文章上，群眾具有一致判斷的比率高達80%，但整體來說，兩者在全體文章立場的判斷上一致性僅為低度的0.401。

但是這樣的結果主要在於前述所說，在一開始的結果選向上，群眾和助理判定為中立、無法判斷、或同時皆有支持與反對意見之間具有邏輯上的偏誤。而就上述交叉表的結果來看，誤差也主要來自於上述非支持或反對的判定結果上。因此，我們將試著調整一下對於支持和反對以外的編碼結果，觀察兩者之間的一致性是否能有所提升。

表 17：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（原始）

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.401	.035	12.196	.000
有效觀察值個數	344			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

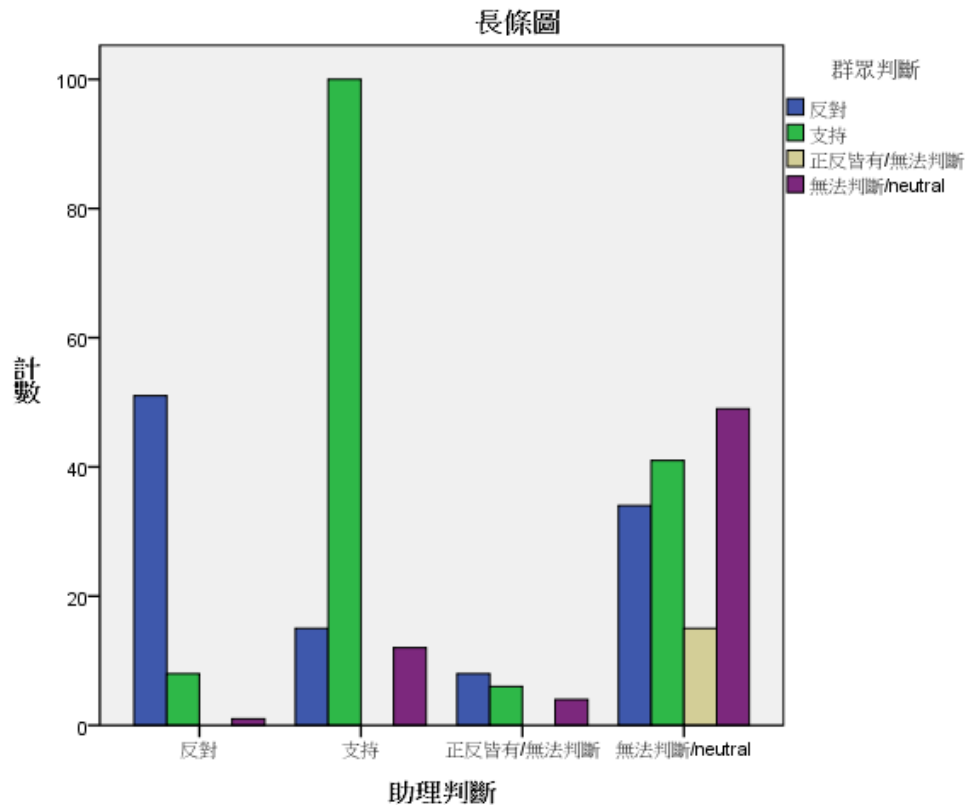


圖 11：群眾判斷結果 * 助理判斷之長條圖

資料來源：本計畫。

二、全體群眾判斷與助理判斷立場之比較分析（合併無法判斷資料）

接著，我們同時將群眾和助理判定為支持或反對以外的結果，皆編碼為無法判斷，並再次進行結果的交叉比對與Kappa一致性分析。如下表所示，兩者對於支持與反對的判定結果如預期般並沒有變化，不過在助理判斷為無法判斷的一致性上因為編碼的整併而從35.3%上升為43.3%。

表 18：群眾判斷結果 * 助理判斷結果交叉表（合併無法判斷資料）

			群眾判斷_6			總計
			反對	支持	無法判斷	
助理判斷_6	反對	計數	51	8	1	60
		助理判斷_6 內的 %	85.0%	13.3%	1.7%	100.0%
		群眾判斷_6 內的 %	47.2%	5.2%	1.2%	17.4%
		佔總計的百分比	14.8%	2.3%	0.3%	17.4%
	支持	計數	15	100	12	127
		助理判斷_6 內的 %	11.8%	78.7%	9.4%	100.0%
		群眾判斷_6 內的 %	13.9%	64.5%	14.8%	36.9%
		佔總計的百分比	4.4%	29.1%	3.5%	36.9%
	無法判斷	計數	42	47	68	157
		助理判斷 內的 %	26.8%	29.9%	43.3%	100.0%
		群眾判斷 內的 %	38.9%	30.3%	84.0%	45.6%
		佔總計的百分比	12.2%	13.7%	19.8%	45.6%
總計	計數	108	155	81	344	
	助理判斷_6 內的 %	31.4%	45.1%	23.5%	100.0%	
	群眾判斷_6 內的 %	100.0%	100.0%	100.0%	100.0%	
	佔總計的百分比	31.4%	45.1%	23.5%	100.0%	

資料來源：本計畫。

不過，就整體判定結果的一致性來說，其Kappa係數僅些微上升為0.459，p值為0.000具高度顯度，整體結果僅接近中度的一致性。因為這樣的編碼調整並無法影響群眾更大量的在助理無法判定的文章中更積極地判斷其立場為支持或反對的結果。

表 19：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（合併無法判斷資料）

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.459	.036	12.806	.000
有效觀察值個數	344			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

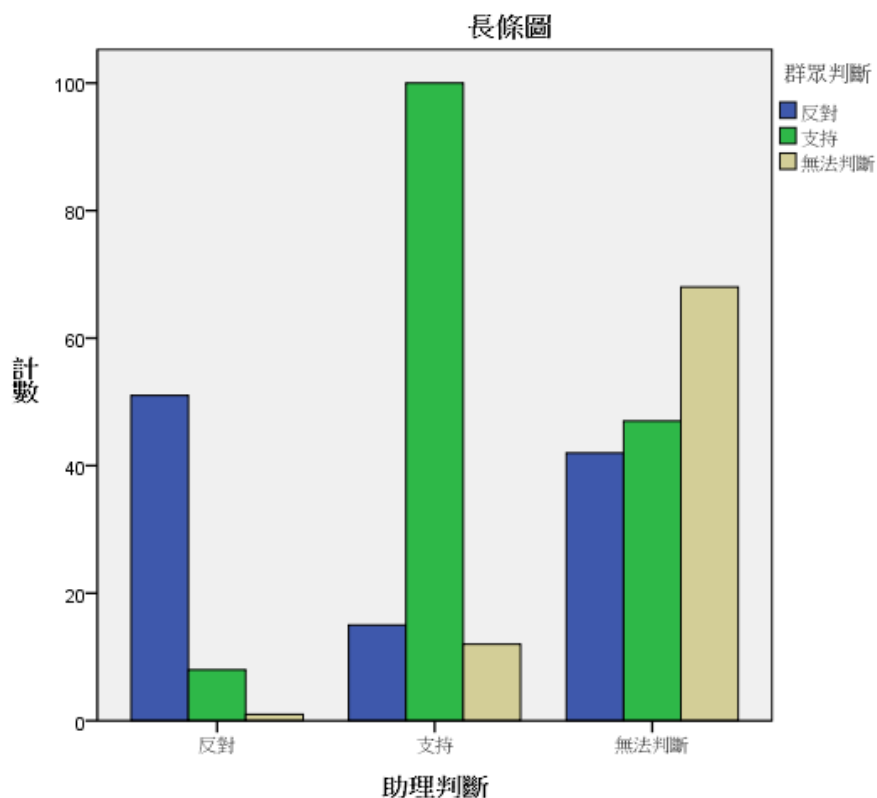


圖 12：群眾判斷結果 * 助理判斷結果長條圖（合併無法判斷資料）

資料來源：本計畫。

三、 群眾判斷結果 * 助理判斷結果交叉表（只看判斷為支持與反對立場）

最後，我們嘗試排除所有除了被判定為支持或反對立場的文章，觀察群眾和助理在判定上的一致性。總共針對174篇助理和群眾皆有判斷立場為支持或反對的文章結果進行比較。

根據以下交叉表所示，助理判斷為反對立場，且群眾有做出支持或反對的判定結果的59篇文章中，群眾也做出一致反對立場判斷的數量為51篇，一致性比例達86.4%；判斷成相反支持立場的數量為8篇，

比例為13.6%。至於助理判斷為支持立場，且群眾有做出支持或反對立場的115篇文章中，兩者具相同意見的數量為100篇，一致性比例高達87%，不一致的數量為15篇，比例為13%。

因此，從助理的角度來看，群眾與他們判斷的結果一致性比例為86.7%，不一致比例為13.3%。而反過來從群眾的角度來看，助理們所判斷為反對立場的結果的66文章中，助理判定為一致的數量51篇，比例為77.3%；而群眾判定為支持立場的108篇文章中，助理和群眾判定結果一致的數量為100篇，一致性比例為92.6%。

表 20：群眾判斷結果 * 助理判斷結果交叉表（只看判斷為支持與反對立場）

			群眾判斷		總計
			反對	支持	
助理判斷	反對	計數	51	8	59
		助理判斷 內的 %	86.4%	13.6%	100.0%
		群眾判斷 內的 %	77.3%	7.4%	33.9%
		佔總計的百分比	29.3%	4.6%	33.9%
	支持	計數	15	100	115
		助理判斷 內的 %	13.0%	87.0%	100.0%
		群眾判斷 內的 %	22.7%	92.6%	66.1%
		佔總計的百分比	8.6%	57.5%	66.1%
總計	計數		66	108	174
	助理判斷 內的 %		37.9%	62.1%	100.0%
	群眾判斷 內的 %		100.0%	100.0%	100.0%
	佔總計的百分比		37.9%	62.1%	100.0%

資料來源：本計畫。

至於從Kappa一致性係數的結果來看，群眾和助理在各自皆辨認為支持立場和反對立場的結果一致性，其Kappa係數高達0.713，p值為0.000的高度顯著。顯示兩者在此種情況下的判定結果具有相當高度的一致性。

表 21：群眾判斷結果 * 助理判斷之 Kappa 一致性結果（只看判斷為支持與反對立場）

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.713	.055	9.446	.000
有效觀察值個數	174			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

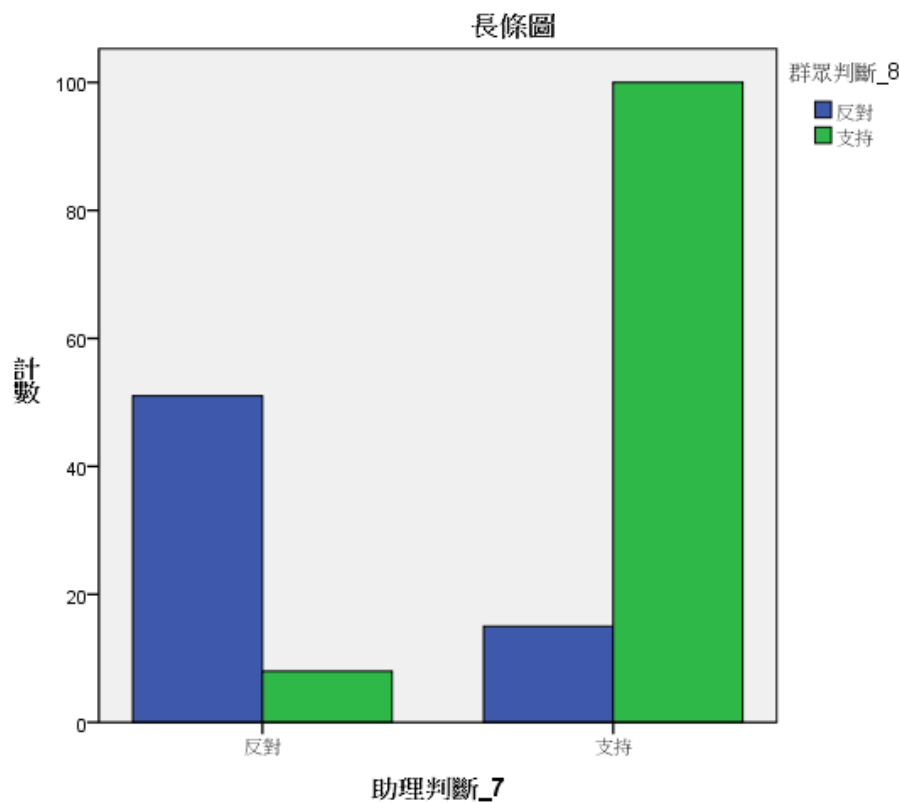


圖 13：群眾判斷結果 * 助理判斷結果長條圖（只看判斷為支持與反對立場）

資料來源：本計畫。

因此，就上述三種的分析結果來說，此次實驗設計因為先前助理對於支持與反對立場結果以外的判定方式，以及群眾的判定邏輯因為在判定機制的設計上而具有差異，因此大幅影響了整體群眾和助理在輿情立場的判定結果。

但仔細觀察其結果也可發現，群眾在此次的實驗結果上，相較於助理對於新聞文章的立場判定較為保守來說，展現出群眾是更有可能更精確地或更敏感地反映出潛藏在輿情文章中的支持或反對立場。此

外，試著排除判斷不明的結果之後，我們發現其實群眾跟助理之間對於支持立場和反對立場的輿情文章判定具有非常高的一致性，其 Kappa 係數高達 0.713。這樣的結果也再次加強了我們對於以群眾外包進行網路輿情立場判定的準確性與可行性之信心。

第四節 群眾輿情判定之機構效應與從眾效應分析

除了比較群眾和助理判斷立場結果之間的一致程度如何外，本次平台在操作介面上也根據機構效應和從眾效應所設計的實驗變數而設計了 4 種不一樣的平台介面，如下表所示。接下來將試著分析群眾分別在四種不同操作介面進行輿情判定的結果是否具有差異，觀察此次實驗情境對於判定結果的影響如何。

表 22：輿情平台實驗設計之 4 種專案介面

	介面顯示與任務說明之代表機構差異	
是否顯示其他人的判定結果	A1 電子治理研究中心／ 顯示其他人的判定	B1 國家發展委員會／ 顯示其他人的判定
	A2 電子治理研究中心／ 無顯示其他人的判定	B2 國家發展委員會／ 無顯示其他人的判定

資料來源：本計畫。

在機構效應的情境變數上，我們以政治大學公行系為執行團隊，對外設計兩套不同的委託機關的名稱和計畫目的邀請群眾參與我們的

輿情判定活動。其中兩套任務介面為強調本團隊是受電子治理研究中心所委託，希望針對群眾和少數人工在輿情判定的比較進行學術研析；另外兩套則為強調本計畫是受國家發展委員會委託而進行，並以期待群眾能共同參與輿情立場判定，而可能有助於政府政策規劃之效能提升。除此之外，並在不同的介面上皆持續分別顯示電子治理研究中心與國家發展委員會的標題圖像，以試著強化群眾對於兩委託機關對於此活動的意象。最後，希望觀察不同機關圖像介面的判定結果是否會具有明顯的差異，以檢驗此次研析設計的機構效應是否對群眾的判定上具明顯影響。

其次，在從眾效應的實驗設計，我們則設計兩套立場判定介面是能夠即時看到其他群眾對於每篇文章的判定結果，另外兩套介面則並沒有此設計。因此，根據從眾效應所提出的假設是，當準備對文章進行判讀的民眾，在下決定前看到其他人的判定結果，將可能傾向和其他人判定的結果一致。所以，檢測看的到其他人判定結果的群眾所判定的立場結果，其一致性是否比另一套介面的群眾更高，將是我們判斷群眾效應程度的方式。

一、不同專案介面的判定結果一致性

本次群眾判定在完成所有任務的人次上，達223人。針對350篇輿情文章，總共取得8752筆判定結果。每篇輿情文章被不同群眾判定的次數為23~28人次之間，平均每篇輿情文章的判定人次達25次。而四種專案介面的實際群眾判定概況如下表所示。

表 23：四種專案介面的群眾判定概況

專案介面	總判定 次數	每篇文章被 判定的次數	與全體群眾判 定結果比較之 Kappa係數	與助理判定 結果比較之 Kappa係數
A1 電子治理研究中心/ 顯示其他人的判定	2550	7~8	0.814	0.416
A2 電子治理研究中心/ 無顯示其他人的判定	2201	5~8	0.785	0.396
B1 國家發展委員會/ 顯示其他人的判定	1864	5~7	0.705	0.343
B2 國家發展委員會/ 無顯示其他人的判定	2137	5~7	0.729	0.365

資料來源：本計畫。

從上表觀察四個平台專案的判定結果和全體群眾的判定結果以及助理判定結果之Kappa一致性係數以及總判定次數，可以明顯發現兩者之間具正向之相關。但除此之外，並觀察不到四組群眾的判定結果是否因介面情境差別而具有明顯差異。

二、 各專案介面的組內判定一致性

接著，試著計算四組群眾在判定結果上各自的一致性程度。計算方法是將各組群眾對每篇文章判定選項中的最高票除以該篇文章的判定總次數，作為該組群眾在多數決機制後的判定結果一致性或者產出之立場可信度。

表 24：群眾之組內和全體判定結果一致性

專案介面	組內判定結果一致性
A1 電子治理研究中心/ 顯示其他人的判定	0.7682483
A2 電子治理研究中心/ 無顯示其他人的判定	0.678465986
B1 國家發展委員會/ 顯示其他人的判定	0.786884354
B2 國家發展委員會/ 無顯示其他人的判定	0.698952381
All 全體群眾的判定	0.674976557

資料來源：本計畫。

根據上表四組群眾判定以及全體群眾判定結果之一致性，可明顯發現，可以看到其他人判定結果的A1（0.768）和B1（0.787）兩組介面，其群眾判定的一致性明顯高於全體群眾（0.675）判定和A2（0.678）、B2（0.699）兩組的判定結果一致性。從這樣的結果可以推論，可以看

到其他人判定結果的A1和B1兩組專案介面的群眾，在立場的判定上確實呈現了從眾效應的結果。

第五節 輿情平台之使用者瀏覽行為分析

為了解參與本實驗之群眾在我們的平台上的一些瀏覽行為，我們特別在系統網頁中加入蒐集使用者瀏覽行為的 Google Analytics(GA) 程式碼，搜集使用者的基本瀏覽行為。透過 GA 的統計結果，我們可以得知如瀏覽者從什麼樣的管道進入我們網站、網站到訪者的個人資訊，像是性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統。也可以統計網站到訪者實際使用網站服務的情況，分析網站系統的回應速度。以下條列說明為根據 GA 提供的瀏覽紀錄所可得知的基本資訊，試著展現應用 GA 可提供的多元化資料，以及本次輿情平台網站的使用者行為概況。

第一、瀏覽者從什麼樣的管道進入我們網站：由下圖可知大部分皆從 Facebook 網站進入平台。而其來源應包含（1）粉絲專頁的推廣、（2）Facebook 的廣告宣傳、（3）計畫團隊成員利用個人 Facebook 網絡的介紹、（4）接受到以上訊息的 Facebook 使用者之連結分享。但在數量上應為（2）藉由廣告宣傳而進入網站的使用者為大宗。

政府巨量資料分析與政策端應用效能提升之研析

來源/簡介	客戶開發			行為			轉換		
	工作階段	% 新工作階段	新使用者	跳出率	單次工作階段頁數	平均工作階段時間長度	目標轉換率	目標達成	目標價值
	2,470 % 總計: 100.00% (2,470)	69.72% 資料檢視平均 值: 68.80% (0.17%)	1,722 % 總計: 100.17% (1,719)	63.08% 資料檢視平均 值: 63.08% (0.00%)	3.37 資料檢視平均 值: 3.37 (0.00%)	00:02:47 資料檢視平均 值: 00:02:47 (0.00%)	0.00% 資料檢視平均 值: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
1. (direct) / (none)	1,228 (49.72%)	63.19%	776 (45.08%)	53.58%	4.04	00:03:30	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
2. lm.facebook.com / referral	633 (25.63%)	70.46%	446 (25.90%)	90.68%	1.14	00:00:49	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
3. facebook.com / referral	235 (9.51%)	68.09%	160 (9.29%)	28.09%	6.80	00:07:14	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
4. m.facebook.com / referral	102 (4.13%)	95.10%	97 (5.63%)	48.04%	3.53	00:01:09	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
5. get-your-social-buttons.info / referral	45 (1.82%)	100.00%	45 (2.61%)	100.00%	1.00	00:00:00	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
6. site63783351.snip.to / referral	30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
7. site68358751.snip.to / referral	30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
8. site13989387.snip.to / referral	24 (0.97%)	100.00%	24 (1.39%)	100.00%	1.00	00:00:00	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
9. l.facebook.com / referral	23 (0.93%)	100.00%	23 (1.34%)	56.52%	1.70	00:00:43	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)
10. site65137379.snip.to / referral	19 (0.77%)	100.00%	19 (1.10%)	100.00%	1.00	00:00:00	0.00% 0 (0.00%)	0 (0.00%)	\$0.00 (0.00%)

圖 14：使用者進入平台網站之管道

資料來源：本計畫。

第二、性別分析：由下圖可知參與者為女性居多，1304 位能辨認出性別的平台使用者之中，有將近 55% 的使用者為女性。顯示群眾外包的公共活動，未必是男性的天下，反而有可能吸引女性的參與。但實際的因果關係仍有待確認。

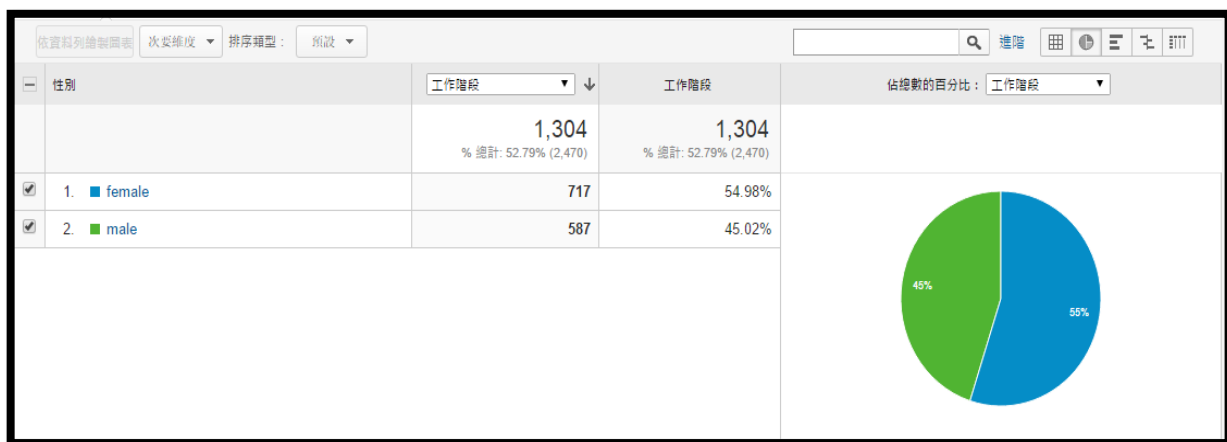


圖 15：使用者之性別數量與比例分析

資料來源：本計畫。

第三、使用者所處網域地區分析：下圖顯示參與者幾乎都在台灣地區。從下圖可得知，參與者高達 2276 人，佔所有人數的 92.16% 是在台灣地區的網域與平台進行連線。不過，也是有從美國（50／2.23%）、中國（16／0.65%）、日本（10／0.4%）、南韓（8／0.32%）、德國（6／0.24%）、英國（4／0.16%）、俄羅斯（4／0.16%）、巴西（3／0.12%）等，從世界各國網域連線至平台的情形。

這樣的結果也確實反映了本次平台宣傳與推廣的主要範圍僅針對臺灣民眾，但少量卻遍布世界各國的其他網域使用者，也顯示了國際網路在資訊傳遞範圍上的無遠弗屆。

政府巨量資料分析與政策端應用效能提升之研析

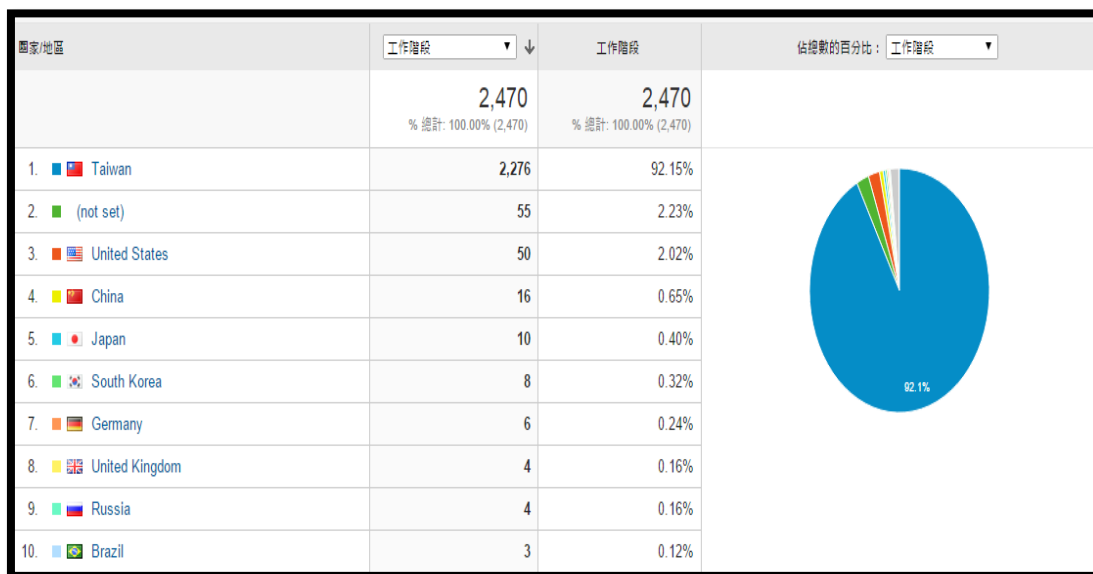


圖 16：使用者所處網域之地區分布

資料來源：本計畫。

第四、紀錄使用者停留時間以及跳出率：下圖顯示總共 3172 的瀏覽量中，具有 2050 的不重覆連線量，而使用者平均停留在平台首頁的時間為 71 秒。這樣的結果顯示平台使用者對於閱讀首頁資訊所願意的花費時間，從這樣的平均時間長度推測，幾乎所有使用者都願意詳細閱讀平台活動說明。

不過從使用者的跳出率為 63%來看，有超過 6 成的使用者在看完平台首頁的資訊跳出頁面，而離開率為 49%。顯示約為半數的使用者在此階段後並未進一步進行活動註冊。這樣的結果對於網站管理者來說是一大警訊，而應留意調整平台首頁的說明，或是在這個階段強化誘因，以留住使用者接續下一階段行為。



圖 17：使用者停留時間以及跳出率

資料來源：本計畫。

第五、統計參與者使用的瀏覽器：下圖顯示大部分為 **chrome** 的使用者，高達 70% 以上，其次則為 **Android** 系統之瀏覽器，但僅約 10%。不過，雖然 **Android** 系統瀏覽器的使用者數量為第 2 多，但從平均資料檢視時間上僅為 49 秒，遠低於其他瀏覽器使用者的時間（其餘平均檢視時間為 2~6 分鐘以上）。可能必須留意本平台首頁在 **Android** 系統瀏覽器上的相容性是否較低，而可能影響使用者的檢視，換句話說，群眾外包的運作平台，要如何同時滿足桌電與行動裝置的介面需求，也是未來應該注意的問題。

瀏覽器	裝置效能			行速			耗電		
	工作階段 ↓	% 工作階段 ?	新增裝置 ?	輸出率 ?	單次工作階段 段數 ?	平均工作階段時間 段數 ?	日總耗電率 ?	日總耗電 ?	日總耗電 值 ?
	2,470 % 總計: 100.00% (2,470)	69.72% 資料總量平均 值: 69.60% (0.17%)	1,722 % 總計: 100.17% (1,719)	63.08% 資料總量平均 值: 63.08% (0.00%)	3.37 資料總量平均 值: 3.37 (0.00%)	00:02:47 資料總量平均 值: 00:02:47 (0.00%)	0.00%	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
1. Chrome	1,841 (74.53%)	71.05%	1,308 (75.96%)	64.91%	3.37	00:02:45	0.00%	0 (0.00%)	\$0.00 (0.00%)
2. Android Browser	253 (10.24%)	71.94%	182 (10.57%)	88.54%	1.34	00:00:49	0.00%	0 (0.00%)	\$0.00 (0.00%)
3. Safari	137 (5.55%)	38.69%	53 (3.08%)	44.53%	4.36	00:04:48	0.00%	0 (0.00%)	\$0.00 (0.00%)
4. Internet Explorer	95 (3.85%)	76.84%	73 (4.24%)	27.37%	4.67	00:03:41	0.00%	0 (0.00%)	\$0.00 (0.00%)
5. Safari (in-app)	68 (2.75%)	85.29%	58 (3.37%)	42.65%	4.25	00:02:24	0.00%	0 (0.00%)	\$0.00 (0.00%)
6. Firefox	57 (2.31%)	59.65%	34 (1.97%)	33.33%	5.96	00:05:45	0.00%	0 (0.00%)	\$0.00 (0.00%)
7. Opera	10 (0.40%)	80.00%	8 (0.46%)	20.00%	5.70	00:05:54	0.00%	0 (0.00%)	\$0.00 (0.00%)
8. Edge	9 (0.36%)	66.67%	6 (0.35%)	22.22%	6.33	00:06:02	0.00%	0 (0.00%)	\$0.00 (0.00%)

圖 18：使用者連線之網路瀏覽器分析

資料來源：本計畫。

第六、分析網站到訪者實際使用網站服務的情況：下圖顯示大部分的訪客都只有進到首頁，之後就離開了。符合前述使用者跳出率所發現的狀況。

網站類型	服務紀錄			訪問			使用			轉換		
	工作區	↓	% 工作區	使用時間	↓	% 使用時間	每頁瀏覽時間	↓	% 每頁瀏覽時間	每頁跳出率	↓	% 每頁跳出率
	2,470		69.72%	1,722		63.08%	3.37		00:02:47	0.00%		\$0.00
	% 總計: 100.00% (2,470)		69.69% (0.17%)	% 總計: 100.17% (1,719)		63.08% (0.00%)	3.37 (0.00%)		00:02:47 (0.00%)	0.00% (0.00%)		% 總計: 0.00% (\$0.00)
1. /	1,971 (79.80%)		71.64%	1,412 (82.00%)		63.83%	3.26		00:02:46	0.00%		\$0.00 (0.00%)
2. /account/signin	52 (2.11%)		38.46%	20 (1.63%)		67.31%	3.15		00:01:19	0.00%		\$0.00 (0.00%)
3. /?from=http://site63783351.snip.to/	30 (1.21%)		100.00%	30 (1.74%)		100.00%	1.00		00:00:00	0.00%		\$0.00 (0.00%)
4. /?from=http://site68388751.snip.to/	30 (1.21%)		100.00%	30 (1.74%)		100.00%	1.00		00:00:00	0.00%		\$0.00 (0.00%)
5. /project/Economic-Areas-of-PA/	30 (1.21%)		46.67%	14 (0.81%)		40.00%	3.17		00:01:55	0.00%		\$0.00 (0.00%)
6. /?from=http://get-your-social-buttons.info/?ref=36573101	28 (1.13%)		100.00%	28 (1.63%)		100.00%	1.00		00:00:00	0.00%		\$0.00 (0.00%)
7. /?from=http://site13989387.snip.to/	24 (0.97%)		100.00%	24 (1.38%)		100.00%	1.00		00:00:00	0.00%		\$0.00 (0.00%)
8. /project/Economic-Areas-of-NOCA/	24 (0.97%)		83.33%	20 (1.16%)		29.17%	4.75		00:04:10	0.00%		\$0.00 (0.00%)
9. /?from=http://site65137379.snip.to/	19 (0.77%)		100.00%	19 (1.10%)		100.00%	1.00		00:00:00	0.00%		\$0.00 (0.00%)
10. /project/Economic-Areas-of-PA/	19 (0.77%)		73.68%	14 (0.81%)		26.32%	4.58		00:04:37	0.00%		\$0.00 (0.00%)

圖 19：使用者利用網站服務之紀錄

資料來源：本計畫。

第七、新的網站訪客比率：下圖顯示有將近七成的人為新進的訪客。

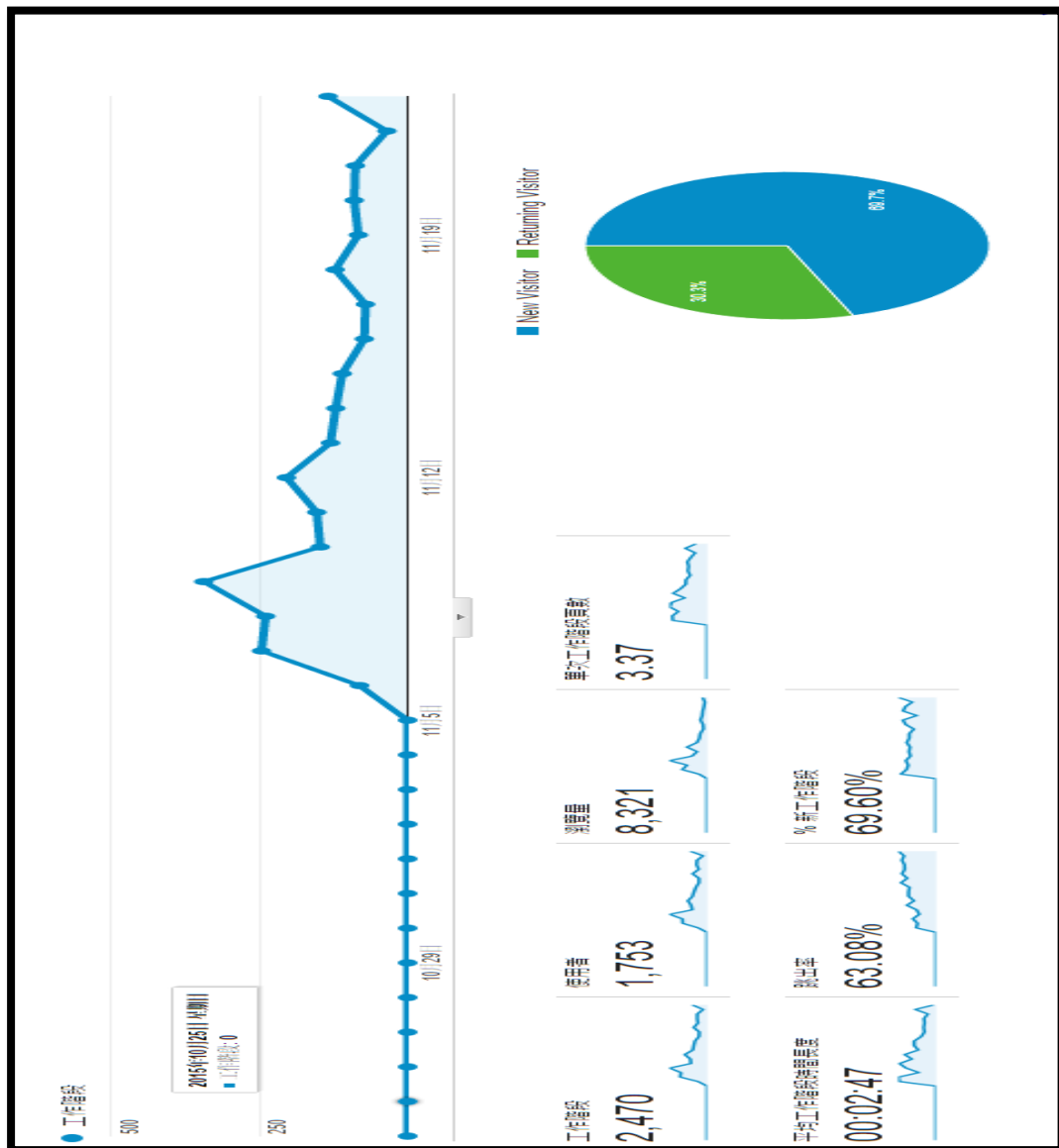


圖 20：不重複之新網站訪客比率

資料來源：本計畫。

第八、分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程：

下圖顯示首頁相當多人進入，但依序在註冊頁以及正式的活動進行頁面的進程中，使用者逐漸的流失。符合團隊在平台設計時所擔心的問題，即流程階段越多，將大幅影響使用者繼續使用的意願。為此次平台系統的架設由於需配合實驗情境的設計，以及行政需要。因此已是目前能想出最精簡的流程。而由此次經驗顯示，未來若欲進行類似的活動與平台架設、或是網站資訊的提供上，流程、階段或是網頁連結次數的簡化將會是影響使用者使用意願的一大關鍵。



圖 21：網站使用者操作網站之過程記錄

資料來源：本計畫。

綜上所述，透過這些 GA 所蒐集的網站使用者瀏覽行為資訊，以及所提供的分析結果，我們可以對進入網站的訪客的瀏覽行為有非常精確的了解，而可以作為活動宣傳效益與網頁易用性的判斷依據。此外，GA 所提供的分析項目，可作為我們評估使用者瀏覽行為上的基本要素。而對網站的管理者來說，便應策略性的根據網站特性與目標而指認出重要項目，並進行有效的管理與評估。

政府巨量資料分析與政策端應用效能提升之研析

第六章 政府服務網之網站分析技術應用研析成果

第一節 政府為民服務網站深度訪談分析

一、深度訪談設計

為瞭解政府大型為民服務網站的經營模式以及比較政府為民服務網站與民間顧客導向服務網站間的差異，本計畫子題透過深度訪談方式，來瞭解各網站經營者的網站服務理念、營運重點以及營運績效，並藉此進一步探索各網站經營者應用對於網站日誌資料協助分析的想法與實際作為。為此，深度訪談採半結構化訪綱的方式進行，以便讓各個訪談結果有共同的比較基準，預擬的訪談題綱分為兩個構面進行設計，分別是網站預期管理作為以及網站實際管理作為，茲就訪談分析重點以及訪談題綱內容整理如下表 25 所示：

表 25：公私部門服務網站經營者深度訪談重點及訪談題綱對照表

訪談構面	訪談分析重點	訪談題綱
(一) 網站預期管理作為	1. 網站營運目標：網站設立背景、資金來源、經營目標、目前成就以及技術團隊和技術。	1. 請您先簡單介紹您的網站，包括：設立背景、資金來源、經營目標、目前成就以及技術團隊和技術？
	2. 網站績效要求：網站資源分配者對於網站營運的績效要求。	2. 請問您的網站的財務資助者對網站經營管理的績效要求為何？
	3. 網站服務對象：網站提供服務對象什麼服務，	3. 請問您的網站主要的服務對象為何？

訪談構面	訪談分析重點	訪談題綱
	服務對象的特質（目標市場）。	
	4. 網站經營的問題：特別關注網站服務和使用 者互動之間產生的問題，以及經營者對網站 服務品質持續提昇的 自我期許。	4. 請問您認為網站的使用者分析 重不重要？為什麼？ 5. 請問您對於網站服務品質的未 來期許為何？ 6. 要達到您的期許，您規劃從事 哪些改善工作？ 7. 使用者行為分析對於您從事這 些改善工作的幫助在哪裡？
（二）網站實 際管理 作為	8. 目前做成網站服務改 善決策的依據。	8. 請問您目前如何進行網站的經 營決策分析？（用什麼分析工 具、哪些是重要決策資訊、這 些資訊代表什麼意義且影響哪 些決策？）
	9. 目前分析網站使用者 行為的方法。	9. 請問您目前是透過何種方法來 瞭解網站使用者行為？
	10. 網站維運資源分配者 對於目前網站經營所 關切的重點。	10. 請問您的網站的財務資助者 目前最關切的績效目標為何？

資料來源：本計畫。

本計畫子題的訪談對象選擇採用立意抽樣的方法，主要考量的重點在於網站服務的代表性以及服務對象的可接近性。因此，經計劃團

隊評估後，在民間網站部分選擇近年來在國內受到愈來愈多使用者歡迎的熱門服務網站「愛料理（iCook）食譜社群網」的主要網站營運規劃者作為訪談對象；在公部門網站部分，則由國發會推薦中央政府機關四個瀏覽人次最多的為民服務網站，分別是：財政部電子申報繳稅服務網、內政部不動產實價登錄網、法務部全國法規資料庫以及國發會 e 管家 Plus 等四個網站的網站營運規劃者作為訪談對象，並分別於 2015 年 7 月 28 日至 9 月 15 日間進行訪談。茲就各受訪者背景、訪談日期、逐字稿分析代碼以及訪談者等資訊整理如下表 26 所示：

表 26：公私部門服務網站經營者背景

網站名稱（網址）	受訪者職級	性別	訪談日期	代碼	訪談者
愛料理（iCook）食譜社群網 （ https://icook.tw ）	網站技術長	男	2015/7/28	A1	廖洲棚老師 陳恭老師
財政部電子申報繳稅服務網 （ http://tax.nat.gov.tw ）	科長	男	2015/8/06	A2	廖洲棚老師 陳恭老師
	科員	女	2015/8/06	A3	
	專案經理 （委外廠商）	男	2015/8/06	A4	
內政部不動產實價登錄網 （ http://lvr.land.moi.gov.tw ）	科長	男	2015/9/02	A5	廖洲棚老師
	科員	男	2015/9/02	A6	

網站名稱（網址）	受訪者職級	性別	訪談日期	代碼	訪談者
法務部全國法規資料庫 (http://law.moj.gov.tw/)	科長	男	2015/9/04	A7	廖洲棚老師
	科員	男	2015/9/04	A8	
國發會 e 管家 Plus (https://msg.nat.gov.tw)	專案經理 (委外廠商)	女	2015/9/15	A9	廖洲棚老師

資料來源：本計畫。

二、受訪網站簡介及訪談結果分析

本計畫子題共計訪談五個公私部門設立之服務網站，為讓讀者瞭解各網站營運宗旨、營運目標以及訪談結果，故分別就各受訪者意見以及本團隊上網搜尋彙整結果整理受訪網站的簡介以及訪談結果，並按照訪談時間先後次序分別說明如下。

（一）愛料理（iCook）食譜社群網

- 1、成立時間：2010 年 8 月。
- 2、主責單位/委外單位：寶利拾股份有限公司。
- 3、網站經營目標：愛料理團隊依照「看見需求、滿足需求、擁抱用戶」的做法，且善用社群與技術，隨時調整與開發最符合使用者的服務；目前食譜總數量已突破 60,000 道，2014 月 12 月起在 Alexa 排名最高躍升至第 51 名，

2015 年的重要目標是食譜內容的「一源多用」¹⁸。

- 4、績效指標：穩定成長（站穩業界第一名）、網站使用排名、影響力排名、訪客數、與網友互動性。
- 5、營運內容：開發並且營運食譜社群網站愛料理以及其多平臺的手機應用軟體。
- 6、訪談結果分析：本計畫整理愛料理（iCook）食譜社群網經營者的訪談結果，如表 27 所示，相關結果的引述依據，則另行摘錄於附錄一，請讀者再行對照參考。



圖 22：愛料理網站首頁

資料來源：愛料理情報室，網址：<https://newsroom.icook.tw/presses>。

¹⁸ 資料引用自愛料理情報室，網址：<https://newsroom.icook.tw/presses>，最後檢閱日期：2015 年 12 月 1 日。

表 27：愛料理（iCook）食譜社群網經營者訪談重點整理

訪談焦點	關切議題	經營現況
1. 網站營運目標	(1) 設立背景	<p>A. 網站設立緣起來自於幾位夥伴的創業動機。</p> <p>B. 經營目標聚焦在滿足民眾自行料理的需求。</p> <p>C. 網站成立時剛好趕上社群網站及 APP 發展風潮。</p>
	(2) 資金來源	<p>A. 個人投資者。</p> <p>B. 國際創投公司。</p>
	(3) 經營目標	<p>A. 擴大網路影響力。</p> <p>B. 追求網站瀏覽及服務使用人次成長。</p> <p>C. 瀏覽人次成長達極限後，開始嘗試跨域經營，從網路跨入實體書籍及服務。</p>
	(4) 網站經營方式	<p>A. 內聘技術人力，網站開發主要使用 Ruby on Rails, JavaScript 技術，伺服器則架在亞馬遜的平台。</p> <p>B. 網站同時提供 APP 服務。</p>
2. 網站績效要求	(1) 財務資助者要求	<p>A. 網站瀏覽人次的成長。</p> <p>B. 民間網站特性為贏者全拿，投資者會著眼於未來的大幅收益，可承擔暫時沒有獲利的風險。</p>
	(2) 經營管理重點	<p>A. 從追求瀏覽人次轉為擴大願意分享資源的使用者佔比，以及增加網友與網站的互動性。</p> <p>B. 區隔不同使用者分群，找出和經營目標最相關的目標市場，再將行銷資源加重投資於主要的目標市場。</p> <p>C. 隨時關切目標達成度，並據以檢視和修</p>

訪談焦點	關切議題	經營現況
		正網站的規劃。
3. 網站服務對象	(1) 服務對象特質	A. 依據性別、年齡區隔不同的服務需求。 B. 持續分析不同目標市場對於網站目標的貢獻度。
4. 網站使用者行為分析方法	(1) 分析架構	A. 導入網站創業公司常採用的 AARRR 模式作為分析架構，內容包含 Acquisition, Activation, Retention, Referral, Revenue。
	(2) 分析重點	A. 經營者需先行確認網站的績效目標與各目標的績效指標。 B. 運用 AARRR 架構定義不同面向的績效目標，再觀察不同目標下的績效指標達成度。 C. 預先設定網站績效分析的框架，如 AARRR，再決定網站使用者行為數據的分析工具，如 GA (Google Analytics) 或網站日誌 (web log) 分析
	(3) 決策經驗	A. 運用社群媒體攬客。 B. 設計使用者行為誘因，鼓勵使用者推薦網站服務。
	(4) 資料蒐集方法	A. 深度訪談瞭解不同客群偏好。 B. 運用網路問卷，調查使用者意見。 C. 設定 GA，分析使用者行為及預期經營目標達成度。 D. 在網站技術面分析上則透過蒐集網路日誌，進行日誌分析。
	(5) GA 實作經驗	A. 區辨使用者來源：運用 GA 中的 New Visitor, Returning User，如有使用社群

訪談焦點	關切議題	經營現況
		<p>網站攬客，可進一步使用 UTM tag，但並不一定需要在所有網頁上進行設定，未設定是常態。</p> <p>B. 分析網站目標達成程度：運用 GA 的 Goal Setting 來瞭解使用者使用行為與目標達成程度的關係。</p>
	(6)使用 GA 的限制	<p>A. GA 分析具有樣本數的限制（免費版 50 萬筆，付費版 75 萬筆）。</p> <p>B. 網站經營的微觀細節需依靠 log 分析補足，但多數的使用者行為分析可透過 GA 的 Goal Setting 設定取得相關資訊。</p> <p>C. 在實際經營上，log 分析多用來進行網站架設及管理技術等面向的分析。</p>

資料來源：本計畫。

（二）財政部電子申報繳稅服務網

- 1、成立時間：於 1997 年配合行政院推動電子化政府的政策籌備，1999 年正式上線；1999-2004 年為第一階段；2005-2010 為第二階段；2011-2015 為第三階段；暫訂 2016 年起為第四階段。
- 2、主責單位/委外單位：財政部財政資料中心／關貿網路股份有限公司
- 3、網站經營目標：提供給納稅義務人多元化的申報的管道，達成政府便民化服務的電子化政府政策目標，並且落實節能減碳。
- 4、績效指標：網路報稅率、系統穩定性、回應性。
- 5、營運內容：於網站放置報稅軟體供民眾下載。提供報稅後

第六章 政府服務網之網站分析技術應用研析成果

的查詢服務。

訪談結果分析：本計畫整理財政部電子申報繳稅服務網經營者的訪談結果，如表 28 所示，相關結果的引述依據，則另行摘錄於附錄一，請讀者再行對照參考。



圖 23：財政部電子申報繳稅服務網首頁

資料來源：財政部網站。

表 28：財政部電子申報繳稅服務網經營者訪談重點整理

訪談焦點	關切議題	經營現況
1. 網站營運目標	(1)設立背景	<p>A. 配合行政院電子化政府政策推動，由財政部財稅資訊中心國稅組負責建置與營運。</p> <p>B. 86 年規劃，88 年上線服務，至 93 年為第一期計畫。</p> <p>C. 第二期計畫始於 94 年終於 99 年；第三期計畫始於 100 年終於 104 年，預計在 105 年啟動第四期計畫。</p> <p>D. 為提供民眾網路報稅管道，本案以提供網路報稅軟體開始，逐步設立網站、電子郵件信箱及客服專線服務納稅義務人，目前提供稅額試算服務、綜所稅結算申報、執行業務所得、外僑綜所稅、遺產稅與贈與稅等五種個人以及營業稅、各類所得（含信託）、營所稅結（決）算（含機關團體、清算）、營所稅暫繳、貨物稅、菸酒稅、證券交易稅、期貨交易稅以及扣除額單位電子資料交換等九種非個人的國稅繳納服務。</p>
	(2)資金來源	公務預算。
	(3)經營目標	報稅軟體穩定度、回覆民眾使用軟體問題、報稅進度查詢、FAQ 彙整提供。
	(4)網站經營方式	委外經營，網站採用架 html 網站技術，有 ap server, web server, DB 等。
2. 網站績效要求	(1)財務資助者要求	無特別要求。

訪談焦點	關切議題	經營現況
	(2)經營管理重點	軟體、網站穩定度，提供民眾 FAQ。
3. 網站服務對象	(1)服務對象特質	國稅的納稅義務人、法人。
4. 網站使用者行為分析方法	(1)分析架構	使用者意見的質化資料分析。
	(2)分析重點	整理常見問答資訊，人工整理數據資料後，回覆長官、記者詢問民眾申報數統計。
	(3)決策經驗	經驗判斷。
	(4)資料蒐集方法	以機關的經驗判斷為主，委外單位收受民眾 email，電話意見的蒐集為輔。

資料來源：本計畫。

(三) 內政部不動產實價登錄網

- 1、成立時間：配合不動產交易透明化政策，於 2012 年 10 月上線，後續 2013-2015 年每年已計畫編列經費增修功能以及系統維護。
- 2、主責單位/委外單位：內政部地政資訊中心／接收資訊端：康大資訊股份有限公司、查詢服務端：光特資訊科技股份有限公司。
- 3、網站經營目標：達成不動產交易透明化的政策目標。
- 4、績效指標：網站瀏覽人次、open data 下載次數。
- 5、營運內容：建立地政資料建置系統，提供一個交易人申報的網路平台。後台提供檢核機制，以利後續地政事務所稽核是否為合理交易。負責統整不動產交易資料庫資料，以

政府巨量資料分析與政策端應用效能提升之研析

建置系統供民眾查詢。

- 6、訪談結果分析：本計畫整理內政部不動產實價登錄網經營者的訪談結果，如表 29 所示，相關結果的引述依據，則另行摘錄於附錄一，請讀者再行對照參考。



圖 24：內政部不動產實價登錄網首頁

資料來源：內政部不動產實價登錄網。

表 29：內政部不動產實價登錄網經營者訪談重點整理

訪談焦點	關切議題	經營現況
1. 網 站 營 運目標	(1)設立背景	A. 配合內政部不動產交易透明化政策。 B. 100 年規劃，自 101 年 10 月起上線服務。
	(2)資金來源	A. 公務預算。

訪談焦點	關切議題	經營現況
		B. 100 年以標餘款委請廠商開發。 C. 102 年起由內政部編列預算支應。
	(3)經營目標	達成不動產交易透明化政策目標。
	(4)網站經營方式	委外經營，網站使用技術未詳細說明。
2. 網站績效要求	(1)財務資助者要求	網站瀏覽人次和 open data 下載次數，目前每月瀏覽人次約 100 萬人次，open data 下載次數每月約 7000~8000 次。
	(2)經營管理重點	A. 資料建置系統著重在資料登錄及資料正確性檢核功能。 B. 所有系統開發都將友善度和使用量列為績效重點。
3. 網站服務對象	(1)服務對象特質	對不動產交易行情有興趣或需求之業者或人士
4. 網站使用者行為分析方法	(1)分析架構	缺乏使用者行為紀錄，目前尚未分析。
	(2)分析重點	尚未分析。
	(3)決策經驗	尚未分析。
	(4)資料蒐集方法	網路問卷調查使用者行為。
	(5)GA 實作經驗	廠商管理時使用但未實際應用於地政司業務。

資料來源：本計畫。

(四) 法務部全國法規資料庫

政府巨量資料分析與政策端應用效能提升之研析

- 1、成立時間：配合行政院「電子化／網路化政府中程推動計畫」，於 1997 年開始籌備，並於 2001 年 4 月上線。
- 2、主責單位/委外單位：法務部資訊處/法源資訊股份有限公司。
- 3、網站經營目標：第一個目的：提供一個及時、完整、專業的官方法律資料庫。為了專業人士所設置。第二個目的：提供一個親和、易用屬於全民的法規資料庫，拉近全民法律感情。
- 4、績效指標：使用量／親民度（舉辦活動）。
- 5、營運內容：建置一個可服務大量民眾連線以及完整收錄國內法規的前後台系統。
- 6、訪談結果分析：本計畫整理法務部全國法規資料庫經營者的訪談結果，如表 30 所示，相關結果的引述依據，則另行摘錄於附錄一，請讀者再行對照參考。

第六章 政府服務網之網站分析技術應用研析成果



圖 25：法務部全國法規資料庫首頁

資料來源：法務部全國法規資料庫網站。

表 30：法務部全國法規資料庫經營者訪談重點整理

訪談焦點	關切議題	經營現況
1. 網站營運目標	(1)設立背景	A. 配合行政院電子化政府計畫，將電子法規納入電子化網路化中程推動計畫。 B. 86 年規劃，90 年 4 月正式上線。
	(2)資金來源	公務預算。
	(3)經營目標	A. 設置一個即時完整又專業的法規資料庫。 B. 建立一個親和屬於全民的法規資料庫。
	(4)網站經營方式	委外經營，網站使用技術未詳細說明。

訪談焦點	關切議題	經營現況
2. 網站績效要求	(1) 財務資助者要求	國會無特別要求。
	(2) 經營管理重點	A. 持續提高使用量。 B. 讓使用者以親和易用的方式取得及時、完整、專業之法律資訊與常識。 C. 協助深化全民之法治素養，並拉近民眾與法律之感情。
3. 網站服務對象	(1) 服務對象特質	A. 專業人士（包括法律專業人士及公務人員）。 B. 一般民眾。
4. 網站使用者行為分析方法	(1) 分析架構	無
	(2) 分析重點	法規關鍵字、法規檢索的整合、熱門詞彙的法規查詢、通俗用語查找、情境式智慧查找、協助教育部推動法治教育。
	(3) 決策經驗	A. 透過各電子信箱及問卷調查得知使用者對網站功能的反應。 B. GA 分析應用於使用者查找資料之統計，幫助調整網站新增內容。
	(4) 資料蒐集方法	A. 電子信箱 B. 問卷調查 C. GA D. 專案會議
	(5) GA 實作經驗	A. 熱門網頁點閱率。 B. 通俗用語分析。 C. 智慧查找用量分析。

資料來源：本計畫。

(五) 國發會 e 管家 Plus 網站

- 1、 成立時間：配合第四階段電子化政府進程，於 2008 年上線。
- 2、 主責單位/委外單位： 國家發展委員會/聯合報系聯經數位股份有限公司
- 3、 網站經營目標：期待透過電子化政府建置，使政府得以站在第一線服務民眾的需求，提供民眾正確完整的政策資訊，以及透過訂閱服務以及條件查詢提供民眾切身的政策資訊。
- 4、 績效指標：會員成長量、會員使用黏著度、APP 下載數。
- 5、 營運內容：在民眾端，提供全民快訊、福利自己查、電子刊物、生活地圖等服務，利用訂閱服務以及條件查詢等方式俾使民眾可以搜尋到與自身最相關的政策或是生活資訊，例如婚育補助、國考訊息、附近加油站以及餐廳資訊。在機關端，提供與協助各縣市政府機關網站服務與 e 管家服務平台，協助各機關編輯政令文宣文字，並重新調整成較親民的文章。
- 6、 訪談結果分析：本計畫整理國發會 e 管家 Plus 網站經營者的訪談結果，如表 31 所示，至於相關結果的引述依據，則另行摘錄於附錄一，請讀者再行對照參考。



圖 26：國發會 e 管家 Plus 首頁

資料來源：國發會 e 管家 Plus 網站首頁。

表 31：國發會 e 管家 Plus 網站經營者訪談重點整理

訪談焦點	關切議題	經營現況
1. 網 站 營 運目標	(1)設立背景	A. 第四階段電子化政府服務。 B. 電子化政府服務的便民服務應用。
	(2)資金來源	公務預算
	(3)經營目標	A. 服務符合民眾需求。 B. 民眾具有使用黏著度。
	(4)網站經營 方式	委外經營，網站使用技術未詳細說明。
2. 網 站 績 效要求	(1)財務資助 者要求	會員成長數、黏著度（包括登入的頻率，功能使用點擊數、登入停留時間）、APP 下載數
	(2)經營管理 重點	A. 以合約規範為主，但會因委託機關長官意見而略有不同。 B. 民眾使用度有多高。 C. 機關接受度有多高。 D. 機關端介接積極度。
3. 網 站 服 務對象	(1)服務對象 特質	A. 一般民眾（重視分眾化需求的滿足）。 B. 公務機關同仁（提供資源協助建立資料、介接平台和行銷）。
4. 網 站 使 用 者 行 為 分 析 方法	(1)分析架構	使用者行為、使用者流量、預期目標轉換率
	(2)分析重點	分析使用者情況、使用行為、服務利用率、目標轉換率、訂閱服務使用率等等。
	(3)決策經驗	A. 提供日報、雙週報、月報等資料供主管機關經營決策參考。 B. 在過程中發現有不理想的部分，經營團

訪談焦點	關切議題	經營現況
		隊會於檢討作法後，再往上做呈報調整。
	(4)資料蒐集方法	A. 自行開發 log 分析工具。 B. 運用 GA 分析網站目標轉換情形。
	(5)GA 實作經驗	報表資訊以自行開發 log 分析工具為主，GA 用於輔助參考使用。

資料來源：本計畫。

三、小結

綜合前述的分析，本計畫整理公私部門服務網站的經營落差分析如表 32 所示，作為本節的小結。

表 32：公私部門服務網站的經營落差分析

訪談構面	訪談分析重點	經營落差分析
(一) 網站預期管理作為	1. 網站營運目標。	私部門：贏得最高市佔率，獲得高報酬。
		公部門：達成網站設立的政策目標。
	2. 網站績效要求。	私部門：使用人數及影響力成長。
		公部門：大部分關心使用量、服務滿意度，少數關心使用黏著度。
	3. 網站服務對象。	私部門：不斷發掘能符合經營目標的特定目標市場，並聚焦在特定目標市場的經營。

訪談構面	訪談分析重點	經營落差分析
		公部門：政策決定，除E管家外，較少深入探求不同服務對象的需求差異，以及鎖定特定目標市場經營。
	4. 網站經營的問題。	私部門：如何隨時因應市場變化彈性調整，不斷適應市場及經營目標需求調整服務內容。
		公部門：專注在網站功能的精進，與服務項目的增加，且會關心使用者問題的解決，但受限於委外合約及預算程序，網站服務內容的調整彈性比私部門低。
(二) 網站實際管理作為	5. 網站服務改善決策的依據。	私部門：專注發掘網路使用者行為資料所隱藏偏好，以及透過網路排名調查瞭解自己的網路影響力。運用自己的開發團隊開發服務內容。
		公部門：主要透過機關同仁專業經驗、電子郵件、客服專線等途徑蒐集網站服務改善的資訊，再透過與廠商的專案會議討論解決方式。少部分會透過網站流量分析軟體，如GA，來蒐集和分析民眾使用網站服務情況。
	6. 分析網站使用者行為的方法。	私部門： 1. 有具體的分析架構AARRR。

訪談構面	訪談分析重點	經營落差分析
		2. 蒐集網站使用者日誌檔，並運用自行開發的log分析工具以及GA進行使用者行為分析。
		3. 搭配使用深度訪談及網路問卷調查。
	7. 網站維運資源分配者關切的重點。	公部門 1. 各網站無具體分析架構，且分析方法及重點皆不同。 2. 主要使用電子信箱蒐集民眾意見，偶爾使用問卷調查。 3. 除全國法規資料庫（僅用部分GA功能）以及E管家等網站（使用自行開發log分析工具及部分GA功能）外，其他網站沒有蒐集及分析使用者行為的網站日誌檔。
		私部門：主要資金來源為投資者，能承受較高風險，短期關心市佔率的擴大，以追求未來的高報酬。 公部門：國會多無意見。委託廠商專注在配合委託機關委外需求，若訂有服務指標的話，則追求服務指標的達成程度。

資料來源：本計畫。

第二節 中央政府機關網站服務調查分析

為瞭解政府機關目前分析網站使用者行為的作法，以及網站經營決策考量的重點，本計畫規劃透過網路問卷方式，調查中央政府各機關（一至三級機關），適用國家發展委員會「政府網站版型與內容管理規範」的機關入口網（官網），以及各機關另行設置且瀏覽人次最多之為民服務網站（若無另行設置則免填）的經營現況¹⁹。

本調查的分析單元為「網站」，調查工具係採用 SurveyMonkey 網路問卷調查平台，調查內容包含政府機關網站的基本資料（包括瀏覽人次、績效指標、服務對象、建置及維運費用等等）、網站的營運模式、網站經營考量重點、網站行銷方式以及瞭解使用者瀏覽網站行為的分析需求。另外，本計畫參考 104 年「政府網站營運績效檢核計畫」、「中央所屬機關、22 縣市政府旅遊網站營運績效檢核報告」以及「中央部會及各直轄市及縣市政府第 2 次網站營運績效檢核報告」等資料，於資料庫中增加「是否為檢核對象」、「104 年網站營運績效檢核成績」等二個變數。本計畫期望透過本調查結果的分析釐清下列問題：

第一、瞭解中央政府各機關官網及為民服務網站目前營運情況以及及關注使用者需求的現況。

第二、瞭解中央政府各機關為民服務網站推動使用者導向服務的具體作法及未來需求。

第三、瞭解中央政府各機關便民服務網站使用 Google Analytics（GA）分析使用者行為的現況及分析需求。

在計劃團隊內部完成問卷設計後（詳細問卷內容附錄五），由國立政治大學於 104 年 11 月 18 日函請行政院 148 個一至三級機關，協助邀請各機關網站管理人員，就其管理網站的實際狀況填答問卷。計劃團隊在 11 月 18 日至 12 月 3 日止，總計回收 238 份問卷，經扣除測試問卷、重複填答問卷以及非調查範圍機關填答問卷後，共獲得 127 份

¹⁹以國發會為例，受訪者需分別就機關官網和我的 E 政府網站的營運現況，各填寫一份問卷；若是交通部中央氣象局僅設置官網，則僅需填寫一份問卷。

有效問卷，其中有 27 個二級機關及 83 個三級機關管理人員填答問卷，各機關總答覆率為 74.32%，一級機關答覆率為 0%，二級機關答覆率為 81.82%，三級機關答覆率為 72.81%。

一、問卷調查敘述統計

首先，計劃團隊先就中央政府各機關網站的描述性統計資料整理如表 33 所示。從表 33 可知，有 82.7%的網站透過全部委外或部分委外方式進行網站建置，超過一半以上的機關網站未設有網站績效指標，最近一次改版至今的平均使用年限約為 2 至 3 年。在管理人力部分，二級和三級機關的專任人力數量差異不大，平均皆不到 2 人，但三級機關間的兼任人力數差異較大，不過平均而言高於二級機關，所有機關平均皆低於 4 人。值得一提的是，計劃團隊以變異數分析比較人力數和網站建置方式間的關係，發現不論二級機關或三級機關，其專、兼任管理人力的任用數與建置方式無顯著關係²⁰。有 74.8%的網站除提供全國民眾使用外，另設有特定的服務對象。從網站每月平均瀏覽人次、建置及維運費用來看，各網站的瀏覽人次及建置維運費用的差異很大，但整體來講二級機關普遍高於三級機關。有 77.2%的受訪者表示其管理之網站適用國發會「政府網站版型與內容管理規範」²¹，但計劃團隊另比較「政府網站營運績效檢核計畫」²²的檢核對象規定，並

²⁰ 本計畫採用單因子變異數分析進行檢測，結果發現二級機關在專、兼任人力數和網站建置方式之 F 值分別為 1.709 和 .390，其 p 值都大於 .05；同樣方式檢測三級機關，所得之 F 值分別為 .715 和 .383，其 p 值也都大於 .05。

²¹ 依據國發會 104 年 3 月發布之「政府網站版型與內容管理規範」規定，該規範適用以下四種網站類別的政府網站，但不適用短期的行銷活動網頁。第一，依網站功能分類，包括各機關入口網站，各機關主題網站（含各機關建置之專業網站、線上申辦服務、觀光文化、互動社群等各類網站）；第二，依網站使用者需求分類，包括各機關分眾網站，包括外語網，如英文、日文等，以及各機關行動版網站（Mobile Web）（國發會，2015：14-15）。

²² 計畫團隊依據國發會於 104 年 1 月發布之「政府網站營運績效檢核計畫」規定

在資料庫中增設一個變項，以比較兩者的差異。計劃團隊進行卡方檢定後，發現大部分認知自己適用該管理規範規定的網站管理者，其所轄網站也同時是計畫的檢核對象，僅有 7 個網站的管理者認為所轄網站不適用該管理規範（但其實是適用該規範），但也有 6 個網站的管理者未表示意見，由於檢核計畫的檢核指標和該管理規範的規定並不完全一致，也可能因此造成相關機關管理者的混淆。

其次，網站規劃時，對相關利害關係者意見重視程度由高至低依序是：網站主要使用者、機關首長、業務單位、網站管考單位、委外廠商、民意代表、財政及主計單位、審計單位。進一步分析發現，有 52% 的網站在規劃時有調查使用者需求，使用最多的調查方法為網站流量分析（33.9%）、其他依序為問卷調查（26%）、深度訪談（17.3%）、行動觀察（11%）、實驗研究（3.1%），另有 10.2% 的管理者選擇包括同事討論、學者專家意見、同仁意見、參考網站作業規範、民意信箱意見等等方法。

再其次，僅有 34.6% 的網站有辦理網站行銷活動，目前最常用來行銷網站的方式依序為：FB 粉絲頁（26%）、機關網站廣告（15%）、網路抽獎（15%）、電子郵件（13.4%）、搜尋引擎優化設定（12.6%）、FB 廣告（8.7%）、Google 關鍵字廣告與其他關鍵字廣告（7.1%）、LINE 貼圖（3.1%）、商業網站廣告（1.6%），另有 2.4% 的管理者選擇其他方法，包括 Youtube、機關公文或電子報等。此外，諸如 FB 粉絲頁、Youtube、LINE、Google+ 等社群媒體以及電子郵件，是最常用來方便使用者推薦機關網站的途徑。

進行調查，總計有 33 個中央部會機關網站、118 個中央部會所屬機關網站、22 個直轄市與縣市政府網站以及 22 個直轄市及縣市政府觀光旅遊網站被列為檢核對象。該計畫預期達成的目標有三，分別是：1. 促進各政府機關營運網站內容資料豐富、時效與正確性；2. 為協助各機關持續提升網站經營績效，檢核結果將提供各機關精進參考；3. 整體檢核執行完畢，將針對檢核所蒐集之資料進行彙整，並加以分析，進而提出政府網站營運績效改善建議報告，以做為未來政府網站營運改善之參考依據（國發會網站，網址：<http://goo.gl/gyvkOy>，最後檢閱時間：2015/12/05）。

表 33：中央政府各機關網站描述性統計

調查項目	答項	次數	百分比
網站所屬機關層級？	一級	0	0%
	二級	31	24.4%
	三級	96	75.6%
網站是否為機關官網？	是	109	85.8%
	否	18	14.2%
機關是否設有網站績效指標？	是	54	42.5%
	否	73	57.5%
除服務全國民眾外，網站是否有設定其他的服務對象？	是	95	74.8%
	否	32	25.2%
網站規劃及建置方法？	委外規劃及建置	48	37.8%
	自行規劃但委外建置	57	44.9%
	自行規劃及建置	6	4.7%
	其他	10	7.9%
網站是否適用國發會「政府網站版型與內容管理規範」？	是	98	77.2%
	否	7	5.5%
	不知道	16	12.6%
網站是否為國發會 104 年政府網站營運績效檢核計畫檢核對象？	是	92	72.4%
	否	35	27.6%
規劃時是否有調查使用者需求？	是	66	52%
	否	54	42.5%
	漏答	7	5.5%
是否辦理行銷活動推廣網	是	44	34.6%

站？	否	76	59.8%
	漏答	7	5.5%
網站目前有使用 Google Analytics (GA) 分析使用者行為嗎？	有	63	49.6%
	沒有	56	44.1%
	沒聽過 GA	1	0.8%
	漏答	7	5.5%
調查項目	答項	平均數	標準差
這一年來每月平均瀏覽人次？	二級機關	584,807	1207656.13
	三級機關	382,108	1041348.43
最近一次改版建置費用？ (千元)	二級機關	2,165.39	2,219.93
	三級機關	1,415.25	1,522.74
改版後年平均維運費用？ (千元)	二級機關	1,145.96	1,654.15
	三級機關	550.52	743.72
改版迄今使用年限？	二級機關	2.17	2.29
	三級機關	2.78	3.16
網站管理專任人力？ (位)	二級機關	1.07	1.03
	三級機關	1.25	1.87
網站管理兼任人力？ (位)	二級機關	1.59	1.72
	三級機關	3.99	12.12
104 年度網站營運績效檢核成績？	二級機關	95.54	3.19
	三級機關	84.95	10.43

資料來源：本計畫。

二、問卷調查推論統計

為瞭解各網站遵循國發會「政府網站版型與內容管理規範」的程度、「政府網站檢核營運績效」成績，和管理者自評所轄網站在使用者親和性 (user friendly) 與網站可用性 (usability) 表現間的關係，本計畫依據調查結果重新轉換出三個變項進行檢測，分別是遵守網站管理規範程度 (R)、網站親和度 (F) 以及網站可用性 (U)，變項轉換方式

則是原問卷各相關題組得分的加總，至於檢核成績則是採用國發會報告的原始分數。另外，為求比較結果的一致性，本計畫僅選擇被列為 104 年度網站營運績效檢核的網站作為分析對象，總計共分析 92 個網站，其中屬中央部會機關（二級機關）網站者有 26 個；中央部會所屬機關（三級機關）網站者有 66 個。

為瞭解人工檢核成績與管理者自評相關指標（R、F、U）得分以及網站每月平均瀏覽人次等使用者導向指標的關係，本計畫採用 Pearson 相關分析方法進行檢測，分析結果如表 34 所示。從表 34 得知，104 年政府網站營運績效檢核成績，與中央部會機關（二級機關）網站管理者自評相關指標（R、F、U）間沒有顯著關係；與中央部會所屬機關（三級機關）網站管理者自評相關指標（R、F、U）有顯著正相關，惟網站檢核成績和各機關網站的每月瀏覽人次之間沒有顯著關係。

表 34：Pearson 相關係數分析結果

	機關層級	每月瀏覽 人次	合規範圍 (R)	親和度 (F)	可用性 (U)
104 年政府 網站營 運績效檢 核成績	所有機關	.134	.516**	.396**	.386**
	二級機關	-.288	.328	.328	.362
	三級機關	.212	.601**	.479**	.475**

說明：**p<.01

資料來源：本計畫。

本計畫認為上述分析結果喻含兩種意義，首先，即使政府機關網站已獲得網站管理者以及機關外部檢核者評價符合相關使用者導向的建置規範、使用者親和度以及網站可用性，但並不會因此增加或減少網站的瀏覽人次。換言之，政府機關網站每月瀏覽人次的高低，無法僅從以人工檢視為主的主觀評估指標的數據高低中推測出兩者的關聯性。當然，本計畫受限於資料取得之限制，無法得知各機關網站標的

服務對象瀏覽人次和檢核分數以及主觀評估指標（R、F、U）的相關性，不過從總體來看，各政府機關若想要瞭解網站使用者的實際使用行為，僅以上述主觀指標的達成度做為網站營運管理績效的思考，顯然是不足的。

其次，上述分析結果顯示中央部會機關（二級機關）網站的營運績效檢核成績，與網站管理者自評相關指標（R、F、U）得分的高低無關。本計畫推論此一資訊隱含的問題是二級機關的網站管理者和外部績效檢核者之間，對政府網站營運績效的認知有差異。在假定問卷填答者係依真實營運狀態填答問卷的情況下，兩者的不一致將造成網站維運的困難，顯然未來需要更客觀的評價機制來評估中央部會機關網站的實際營運績效。

此外，為釐清政府機關建置網站時，有無事先調查使用者需求，對於網站營運績效檢核成績、每月平均瀏覽人次以及網站管理者自評相關指標（R、F、U）等項目的表現差異，本計畫進一步採用獨立樣本 T 檢定進行分析，分析結果如表 35 所示。從表 35 可知，各政府機關在網站建置時是否有事先調查使用者需求，和網站營運檢核績效成績以及網站管理者自評相關指標（R、F、U）得分的高低無關。這個分析結果顯示，在政府網站目前的績效評估體制下，網站規劃時對於使用者需求資訊的瞭解程度，顯然還未成為能影響網站經營績效表現的重要資訊。

表 35：網站經營績效相關指標與使用者需求調查之獨立樣本 T 檢定結果

調查項目	有調查	無調查	T 值
網站營運績效檢核成績	88.23	88.34	-.052
每月平均瀏覽人次	343166.94	411523.59	-.378
合規程度 (R)	101.84	98.75	1.192
親和度 (F)	54.32	53.51	.616
可用性 (U)	61.75	60.79	.655

說明：* $p < .05$

資料來源：本計畫。

綜合上述的討論，本計劃發現網站規劃建置前的使用者的需求以及建置後的實際使用行為和滿意度，尚未能成為檢核政府網站經營績效的依據，依賴專家的外部檢核和管理者主觀判斷，仍是目前政府網站經營的主要決策依據。若進一步對照前節的深度訪談結果，吾人可知政府機關網站的經營者，尚未能充分利用網站日誌所記錄的大量網站使用者行為資料於網站經營決策。因此，本計畫建議政府機關未來能參考前節指出之私部門網站的決策分析模式，嘗試運用網站流量的巨量資料分析，並進一步網站使用者的瀏覽行為資訊納入網站營運決策的參考資訊，以改善目前決策僅依賴專家判斷的不足。

第三節 應用GA於政府服務網站能提出哪些問題解方

網站流量分析可彌補傳統運用網路問卷調查與實驗設計法，來分析網站使用者行為的不足，主因在於網站流量分析是以網頁日誌或紀錄檔 (web page logs) 或網頁標記 (page tagging) 來分析網站使用者的網站足跡，所得資訊具有客觀性，可排除受觀察者主觀意識導致的調查偏誤，故稱為「非干擾式網路行為觀測」(鄭江宇、張佳榮，2015)。

在眾多網路流量分析工具中，以 Google Analytics (GA) 最為各界所青睞²³。

GA 為 Google 企業所推出，大多數功能為免費的網頁日誌或紀錄檔分析工具。一般大眾或企業使用 GA 的目的在於分析所屬網站的使用者瀏覽行為，以改善網站的經營決策。GA 可支援網站分析師進行分析工作，例如：追蹤許多標準的網站指標，如造訪數、個別訪客數、跳離率、放棄率以及追蹤業務成果等等；追蹤不同類型的行銷活動以及協助區隔不同客戶的網站使用行為；提供豐富的報表方便閱讀分析結果（胡為君譯，2011：4-7）。特別是以目前在企業網站經營的分析架構 AARRR 為例，GA 可配合此一分析架構蒐集足夠且重要的使用者行為資訊，供網站經營者參考。AARRR 架構是美國矽谷知名育成公司 500 Startups 的創辦人 Dave McClure 提出的網站經營分析指標，其認為網站經營者應該考慮以下的問題²⁴：

- (一) Acquisition (獲得使用者)：如何讓使用者從各種管道得知並連結到網站？
- (二) Activation (成為有效的使用者)：如何讓已瀏覽網站的使用者使用網站提供的主要功能、服務或閱讀重要資訊，而不是隨意瀏覽即離開？
- (三) Retention (留住使用者)：如何讓使用者願意重複造訪網站？
- (四) Referral (使用者推薦他人)：如何讓使用者願意推薦他人使用網站？
- (五) Revenue (使用者付費或執行預期行動)：如何讓使用者增加網站的造訪率、停留時間或服務利用率？

上述問題，可以透過網路流量分析取得重要的網站使用者行為資

²³ 截至 2013 年為止，美國商業雜誌「財星」所列出的世界排名 500 大企業中，約有 63.2% 的企業採用 GA 做為網站流量分析工具（鄭江宇、張佳榮，2015：19）。

²⁴ 資料參考本計畫受訪者 A1 意見以及

<https://blog.alphacamp.co/2014/12/01/growth-hacking-aarr-metrics/>，最後檢閱日期：2015 年 12 月 5 日。

訊，以協助網站經營者改善網站的服務，而 GA 所提供的相關功能，正好可以配合 AARRR 分析架構，來幫助管理者取得相關的網站瀏覽行為資訊。

一、 GA 功能需求分析

本計畫的問卷調查發現，目前已有 63 個中央政府機關網站(49.6%)有使用 Google Analytics (GA) 分析使用者行為。至於，目前未使用 GA 的 57 個網站中，管理者回答其未使用的理由，主要是因為未列在委外合約內(16.5%)、沒有需要(11.8%)、不會使用(6.3%)以及其他理由(10.2)，受訪者回答其他理由的內容包括，資訊廠商未使用、使用其他分析工具、尚在評估中、預計明年度導入等原因。由於 GA 功能眾多，若無一個明確的分析架構和分析目的，管理者將迷失於茫茫如大海般的龐大資料中，計劃團隊為協助政府機關釐清分析需求，以設定相對應的 GA 功能，故依據 AARRR 架構所需資訊設計調查問題，其相關對應整理如表 36 所示，各政府網站管理者可依據問題需求，自行評估對於 GA 功能設定的需求。換言之，政府網站管理者可視經營管理需要設定適當的 GA 功能，並不需要完全遵守 AARRR 分析架構。

表 36：AARRR 分析架構與 GA 功能關聯一覽表

AARRR 分析 架構	對應之 GA 功能概述	自我評估問題
Acquisition (獲得使用者)	1. 設定目標對象 (Audience) 分析來觀察新訪客。 2. 設定攬客 (Acquisition) 分析來觀察網路行銷活動、AdWord 廣告活動、Google 站外搜尋和網站訪客到訪行為的關係。	1. 瞭解網路民眾從什麼管道得知我們的網站。 2. 釐清網路民眾來源管道和網路行銷活動的關係。 3. 蒐集我們的網站到

AARRR 分析 架構	對應之 GA 功能概述	自我評估問題
		<p>訪者的個人資訊(如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊)。</p> <p>4. 瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站(訪客站外搜尋關鍵字分析)。</p>
<p>Activation (成為有效的使用者)</p>	<p>1. 設定行為 (Behavior) 分析來觀察訪客行為流程、網站內容、網站速度、站內搜尋、觸發事件 AdSense 廣告成效、網頁活動以及不同版本網頁實驗情況等相關資訊。</p> <p>2. 設定轉換 (Conversion) 分析來設定網站服務目標。</p> <p>3. 設定網站速度分析 (Site Speed)，可取得資訊包含網站操作時間、網頁載入平均需時、網頁下載平均需時、伺服器反應平均需時等等多項網站管理的技術資訊。</p>	<p>1. 分析我們的網站到訪者實際使用網站服務的情況(如網站最熱門網頁、服務或最常被下載的檔案)。</p> <p>2. 分析網站訪客曾在我們的網站搜尋哪些關鍵字(訪客站內搜尋關鍵字分析)。</p> <p>3. 分析網站訪客搜尋關鍵字(包括站內與站外搜尋)和網站服務的對應關係。</p> <p>4. 分析網站系統的回應速度。</p>

AARRR 分析 架構	對應之 GA 功能概述	自我評估問題
Retention (留住使用者)	1. 設定目標對象 (Audience) 分析來觀察舊訪客再訪率、造訪頻率、再訪問隔時間、瀏覽量、參與度、裝置、作業平台等資訊，並可進一步設定興趣分析來區辨訪客感興趣的廣告類型、有意願使用網站服務的訪客類型。	1. 分析網站訪客的再訪頻率。 2. 釐清網站訪客再訪率和行銷活動的關係。 3. 區隔網站訪客類型 (如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型)。
Referral (使用者推薦他人)	1. 設定攬客 (Acquisition) 的推薦連結分析來觀察訪客到訪管道 (Channels) 以及所有推薦連結流量分析。 2. 設定行為 (Behavior) 分析來觀察新訪客到訪和原訪客推薦連結的關係。 3. 設定轉換 (Conversion) 分析來觀察社交網路推薦連結的參訪流量。	1. 瞭解網站訪客推薦他人使用我們的網站的情況。
Revenue (使用者付費或執行預期行動)	1. 設定行為 (Behavior) 分析來觀察訪客的瀏覽流程。 2. 設定轉換 (Conversion) 分析來觀察服務被利用前的瀏覽歷程 (含熱門瀏覽歷程)、網站服務利用次數、	1. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程 (訪客的進入點、停滯點和離去點)。 2. 分析網站預期績效

AARRR 分析 架構	對應之 GA 功能概述	自我評估問題
	網路行銷功勞歸屬等資訊。	指標的達成率。

資料來源：本計畫。

本計畫在問卷調查中，請政府網站管理者自行評估對於網站使用者相關瀏覽行為資訊的需求程度，各受訪者需在 1 到 10 分的範圍內給予各題分數（例如：1 分表示完全不需要，10 分表示非常需要，5-6 分代表中間意見，5 分稍偏不需要，6 分稍偏需要），整體意見整理如表 36 所示，至於不同機關層級（二級機關 29 位、三級機關 90 位）的意見，則另意見整理如表 37 和表 38 所示。

表 37：中央政府各機關網站經營者的網路流量分析資訊需求

調查問題 (n=119)	平均數	標準差
1. 瞭解網路民眾從什麼管道得知我們的網站。	7.69	1.934
2. 釐清網路民眾來源管道和網路行銷活動的關係。	7.23	2.133
3. 蒐集網站到訪者的個人資訊（如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊）。	6.75	2.627
4. 瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站（訪客站外搜尋關鍵字分析）。	7.53	2.267
5. 分析網站到訪者實際使用網站服務的情況（如網站最熱門網頁、服務或最常被下載的檔案）。	8.25	1.723
6. 分析網站訪客曾在我們的網站搜尋哪些關鍵字（訪客站內搜尋關鍵字分析）。	8.00	1.855

調查問題 (n=119)	平均數	標準差
7. 分析網站訪客搜尋的關鍵字（包括站內與站外搜尋）和網站服務的對應關係。	7.88	1.927
8. 分析網站系統的回應速度。	7.96	1.744
9. 分析網站訪客的再訪頻率。	7.71	1.931
10. 釐清網站訪客再訪率和行銷活動的關係。	7.06	2.290
11. 區隔網站訪客類型（如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型）。	7.28	2.258
12. 瞭解網站訪客推薦他人使用我們的網站情況。	6.98	2.358
13. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程（訪客的進入點、停滯點和離去點）。	7.49	2.066
14. 分析網站預期績效指標的達成率。	7.45	2.189

資料來源：本計畫。

表 38：中央政府二級機關網站經營者的網路流量分析資訊需求

調查問題 (n=29)	平均數	標準差
1. 瞭解網路民眾從什麼管道得知我們的網站。	7.21	2.144
2. 釐清網路民眾來源管道和網路行銷活動的關係。	7.21	1.916
3. 蒐集網站到訪者的個人資訊（如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊）。	5.86	2.669

調查問題 (n=29)	平均數	標準差
4. 瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站(訪客站外搜尋關鍵字分析)。	6.82	2.342
5. 分析網站到訪者實際使用網站服務的情況(如網站最熱門網頁、服務或最常被下載的檔案)。	8.07	1.907
6. 分析網站訪客曾在我們的網站搜尋哪些關鍵字(訪客站內搜尋關鍵字分析)。	7.72	2.136
7. 分析網站訪客搜尋的關鍵字(包括站內與站外搜尋)和網站服務的對應關係。	7.34	2.240
8. 分析網站系統的回應速度。	7.38	2.243
9. 分析網站訪客的再訪頻率。	7.10	2.093
10. 釐清網站訪客再訪率和行銷活動的關係。	6.45	2.414
11. 區隔網站訪客類型(如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型)。	6.72	2.313
12. 瞭解網站訪客推薦他人使用我們的網站情況。	6.24	2.400
13. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程(訪客的進入點、停滯點和離去點)。	7.14	2.189
14. 分析網站預期績效指標的達成率。	7.03	2.413

資料來源：本計畫。

表 39：中央政府三級機關網站經營者的網路流量分析資訊需求

調查問題 (n=90)	平均數	標準差
1. 瞭解網路民眾從什麼管道得知我們的網站。	7.84	1.848
2. 釐清網路民眾來源管道和網路行銷活動的關係。	7.23	2.209
3. 蒐集網站到訪者的個人資訊(如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊)。	7.03	2.564
4. 瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站(訪客站外搜尋關鍵字分析)。	7.74	2.211
5. 分析網站到訪者實際使用網站服務的情況(如網站最熱門網頁、服務或最常被下載的檔案)。	8.31	1.667
6. 分析網站訪客曾在我們的網站搜尋哪些關鍵字(訪客站內搜尋關鍵字分析)。	8.09	1.759
7. 分析網站訪客搜尋的關鍵字(包括站內與站外搜尋)和網站服務的對應關係。	8.06	1.792
8. 分析網站系統的回應速度。	8.14	1.518
9. 分析網站訪客的再訪頻率。	7.91	1.844
10. 釐清網站訪客再訪率和行銷活動的關係。	7.26	2.226
11. 區隔網站訪客類型(如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型)。	7.46	2.224
12. 瞭解網站訪客推薦他人使用我們的網站情況。	7.22	2.307
13. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程(訪客的進入點、停滯點和離去	7.60	2.027

調查問題 (n=90)	平均數	標準差
點)。		
14. 分析網站預期績效指標的達成率。	7.59	2.109

資料來源：本計畫。

整體而言，各政府機關管理者對 AARRR 架構的個別構面的需求雖大多為正向，但中央政府二級機關的網站管理者，重視「成為有效的使用者」和「使用者付費或達成預期行動」等構面相關資訊（各題平均分數皆在 7 分以上，部分落在在 8 分以上）的程度，會高於其他構面；中央政府三級機關的網站管理者，則是各構面都重視（各構面題目的平均分數皆在 7 分以上）。

二、 GA 設定實例

計劃團隊搭配本計畫群眾外包網站平台的架設，一併在系統後台設定使用 GA 進行網站流量分析。計劃團隊先在該平台的標題頁面埋入 GA 工具所提供的一小段 JavaScript 代碼，讓網站使用者在每次執行平台的頁面時都能傳送通知，例如：瀏覽者從什麼樣的管道進入網站、蒐集網站到訪者的個人資訊，像是性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統，分析網站到訪者實際使用網站服務的情況，分析網站系統的回應速度，透過 GA，整合成易讀的資訊給平台管理者知道。底下，本計畫依據前述各調查問題可蒐集到的範例型態，依序以本計畫在群眾外包網站平台蒐集到的 GA 資訊，做為示範。

政府巨量資料分析與政策端應用效能提升之研析

1、問題一：瞭解網路民眾從什麼管道得知我們的網站。

來源/媒介	客戶開發			行為			轉換		
	工作階段	% 新工作階段	新使用者	跳出率	單次工作階段頁數	平均工作階段時間長度	目標轉換率	目標達成	目標價值
	2,470 % 總計: 100.00% (2,470)	69.72% 資料檢視平均值: 69.80% (0.17%)	1,722 % 總計: 100.17% (1,719)	63.08% 資料檢視平均值: 63.08% (0.00%)	3.37 資料檢視平均值: 3.37 (0.00%)	00:02:47 資料檢視平均值: 00:02:47 (0.00%)	0.00% 資料檢視平均值: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
1. (direct) / (none)	1,228 (49.72%)	63.19%	776 (45.08%)	53.58%	4.04	00:03:30	0.00%	0 (0.00%)	\$0.00 (0.00%)
2. lm.facebook.com / referral	633 (25.63%)	70.46%	446 (25.90%)	90.68%	1.14	00:00:49	0.00%	0 (0.00%)	\$0.00 (0.00%)
3. facebook.com / referral	235 (9.51%)	68.09%	160 (9.29%)	28.09%	6.80	00:07:14	0.00%	0 (0.00%)	\$0.00 (0.00%)
4. m.facebook.com / referral	102 (4.13%)	95.10%	97 (5.63%)	48.04%	3.53	00:01:09	0.00%	0 (0.00%)	\$0.00 (0.00%)
5. get-your-social-buttons.info / referral	45 (1.82%)	100.00%	45 (2.61%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
6. site63783351.snip.to / referral	30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
7. site68358751.snip.to / referral	30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
8. site13989387.snip.to / referral	24 (0.97%)	100.00%	24 (1.39%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
9. l.facebook.com / referral	23 (0.93%)	100.00%	23 (1.34%)	56.52%	1.70	00:00:43	0.00%	0 (0.00%)	\$0.00 (0.00%)
10. site65137379.snip.to / referral	19 (0.77%)	100.00%	19 (1.10%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)

圖 27：訪客到站來源範例

資料來源：本計畫。

- 問題二：釐清網路民眾來源管道和網路行銷活動的關係。
本計畫未購買 Google 的關鍵字廣告，所以沒有相關資訊。
- 問題三：蒐集網站到訪者的個人資訊（如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、

作業系統等等資訊)。

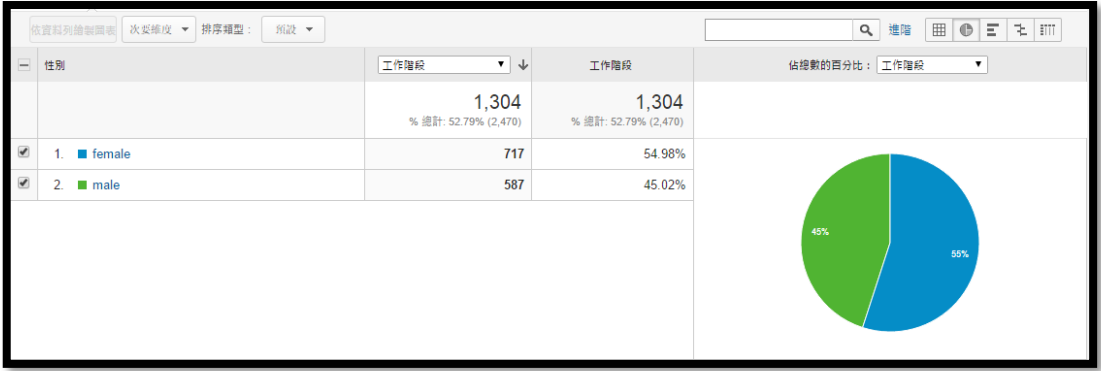


圖 28：訪客性別範例

資料來源：本計畫。

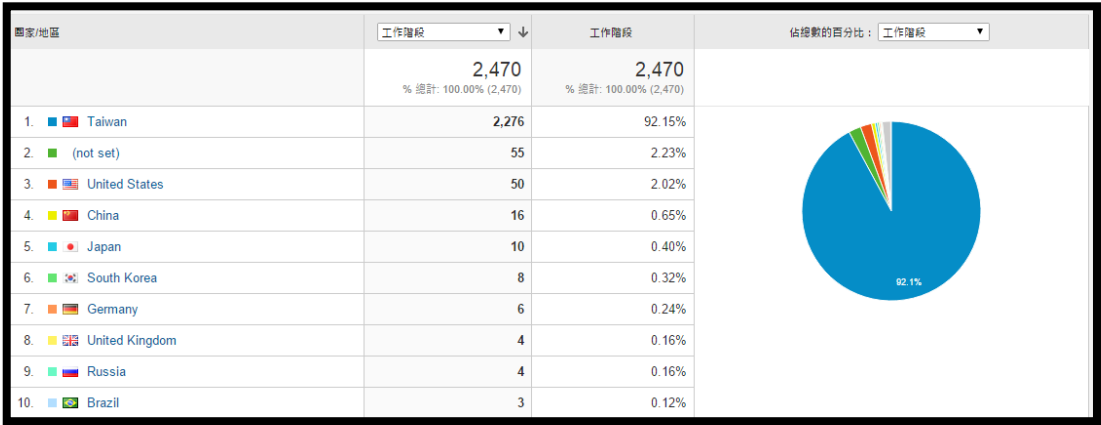


圖 29：訪客居住地範例

資料來源：本計畫。

政府巨量資料分析與政策端應用效能提升之研析



圖 30：訪客停留時間範例

資料來源：本計畫。

	瀏覽器 ?	客戶開發			行為			轉換		
		工作階段 ? ↓	% 新工作階段 ?	新使用者 ?	跳出率 ?	單次工作階段頁數 ?	平均工作階段停留時間 ?	目標轉換率 ?	目標達成 ?	目標價值 ?
		2,470 % 總計: 100.00% (2,470)	69.72% 資料檢視平均值: 69.60% (0.17%)	1,722 % 總計: 100.17% (1,719)	63.08% 資料檢視平均值: 63.08% (0.00%)	3.37 資料檢視平均值: 3.37 (0.00%)	00:02:47 資料檢視平均值: 00:02:47 (0.00%)	0.00% 資料檢視平均值: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
<input type="checkbox"/>	1. Chrome	1,841 (74.53%)	71.05%	1,308 (75.96%)	64.91%	3.37	00:02:45	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	2. Android Browser	253 (10.24%)	71.94%	182 (10.57%)	88.54%	1.34	00:00:49	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	3. Safari	137 (5.55%)	38.69%	53 (3.08%)	44.53%	4.36	00:04:48	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	4. Internet Explorer	95 (3.85%)	76.84%	73 (4.24%)	27.37%	4.67	00:03:41	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	5. Safari (in-app)	68 (2.75%)	85.29%	58 (3.37%)	42.65%	4.25	00:02:24	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	6. Firefox	57 (2.31%)	59.65%	34 (1.97%)	33.33%	5.96	00:05:45	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	7. Opera	10 (0.40%)	80.00%	8 (0.46%)	20.00%	5.70	00:05:54	0.00%	0 (0.00%)	\$0.00 (0.00%)
<input type="checkbox"/>	8. Edge	9 (0.36%)	66.67%	6 (0.35%)	22.22%	6.33	00:06:02	0.00%	0 (0.00%)	\$0.00 (0.00%)

圖 31：訪客使用瀏覽器類型分析範例

資料來源：本計畫。

第六章 政府服務網之網站分析技術應用研析成果

- 4、問題四：瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站（訪客站外搜尋關鍵字分析）。

本計畫未購買 Google 的關鍵字廣告，所以沒有相關資訊。

- 5、問題五：分析網站到訪者實際使用網站服務的情況（如網站最熱門網頁、服務或最常被下載的檔案）。

	到達網頁 ?	客戶開發			行為			轉換		
		工作階段 ? ↓	% 新工作階段 ?	新使用費 ?	跳出率 ?	單次工作階段貢獻 ?	平均工作階段時間長度 ?	目標轉換率 ?	目標達成 ?	目標價值 ?
		2,470 % 總計: 100.00% (2,470)	69.72% 資料抽樣平均值: 69.60% (0.17%)	1,722 % 總計: 100.17% (1,719)	63.08% 資料抽樣平均值: 63.08% (0.00%)	3.37 資料抽樣平均值: 3.37 (0.00%)	00:02:47 資料抽樣平均值: 00:02:47 (0.00%)	0.00% 資料抽樣平均值: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
1. /		1,971 (79.80%)	71.64%	1,412 (82.00%)	63.83%	3.26	00:02:46	0.00%	0 (0.00%)	\$0.00 (0.00%)
2. /account/signin		52 (2.11%)	38.46%	20 (1.16%)	67.31%	3.15	00:01:19	0.00%	0 (0.00%)	\$0.00 (0.00%)
3. /?from=http://site63783351.snip.to/		30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
4. /?from=http://site68358751.snip.to/		30 (1.21%)	100.00%	30 (1.74%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
5. /project/EconomicAreaForPA/		30 (1.21%)	46.67%	14 (0.81%)	40.00%	3.17	00:01:55	0.00%	0 (0.00%)	\$0.00 (0.00%)
6. /?from=http://get-your-social-buttons.info/?ref=36573101		28 (1.13%)	100.00%	28 (1.63%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
7. /?from=http://site13989387.snip.to/		24 (0.97%)	100.00%	24 (1.39%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
8. /project/EconomicAreaForNDCA/		24 (0.97%)	83.33%	20 (1.16%)	29.17%	4.75	00:04:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
9. /?from=http://site65137379.snip.to/		19 (0.77%)	100.00%	19 (1.10%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
10. /project/EconomicAreaForPAB/		19 (0.77%)	73.68%	14 (0.81%)	26.32%	4.58	00:04:37	0.00%	0 (0.00%)	\$0.00 (0.00%)

圖 32：最熱門網頁分析範例

資料來源：本計畫。

- 6、問題六：分析網站訪客曾在我們的網站搜尋哪些關鍵字（訪客站內搜尋關鍵字分析）。

本計畫的網站平台未設站內搜尋，故無相關資訊。不過，站內搜尋服務的使用情況應是政府服務性網站重要的 GA 分析

項目，建議政府網站經營者能注意本項服務產生的資訊。

7、問題七：分析網站訪客搜尋的關鍵字（包括站內與站外搜尋）和網站服務的對應關係。

本計畫的網站平台未設站內搜尋，加上網站外公開且未購買關鍵字廣告，故無相關資訊。但是這可能也是政府服務性網站重要的 GA 分析項目。

8、問題八：分析網站系統的回應速度。



圖 33：網站回應速度分析範例

資料來源：本計畫。

9、問題九：分析網站訪客的再訪頻率。

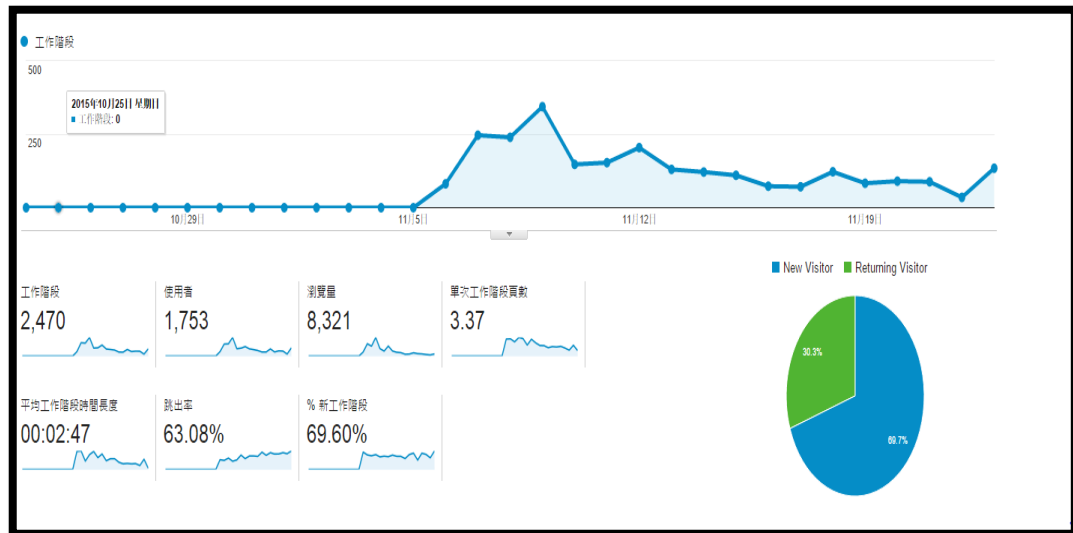


圖 34：訪客再訪率分析範例

資料來源：本計畫。

10、問題十：釐清網站訪客再訪率和行銷活動的關係。

本計畫購買 Facebook 廣告，故需從 Facebook 後台得知相關資訊，相關分析如前第五章第二節所呈現。

11、問題十一：區隔網站訪客類型（如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型）。

政府巨量資料分析與政策端應用效能提升之研析





國家/地區	工作階段	% 新工作階段	新使用者	跳失率	單次工作階段貢獻	平均工作階段時間長度	目標轉換率	目標達成	目標價值
所有工作階段	2,470 % 總計: 100.00% (2,470)	69.72% 資料檢視平均: 69.72% (0.00%)	1,722 % 總計: 100.00% (1,722)	63.08% 資料檢視平均: 63.08% (0.00%)	3.37 資料檢視平均: 3.37 (0.00%)	00:02:47 資料檢視平均: 00:02:47 (0.00%)	0.00% 資料檢視平均: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
回訪者	748 % 總計: 30.28% (2,470)	0.00% 資料檢視平均: 69.72% (-100.00%)	0 % 總計: 0.00% (1,722)	61.36% 資料檢視平均: 63.08% (-2.72%)	3.32 資料檢視平均: 3.37 (-1.34%)	00:03:01 資料檢視平均: 00:02:47 (7.88%)	0.00% 資料檢視平均: 0.00% (0.00%)	0 % 總計: 0.00% (0)	\$0.00 % 總計: 0.00% (\$0.00)
1.  Taiwan									
所有工作階段	2,276 (92.15%)	67.27%	1,531 (88.91%)	60.02%	3.57	00:03:02	0.00%	0 (0.00%)	\$0.00 (0.00%)
回訪者	745 (99.60%)	0.00%	0 (0.00%)	61.21%	3.33	00:03:01	0.00%	0 (0.00%)	\$0.00 (0.00%)
2. (not set)									
所有工作階段	55 (2.23%)	100.00%	55 (3.19%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
回訪者	0 (0.00%)	0.00%	0 (0.00%)	0.00%	0.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
3.  United States									
所有工作階段	50 (2.02%)	100.00%	50 (2.90%)	98.00%	1.02	< 00:00:01	0.00%	0 (0.00%)	\$0.00 (0.00%)
回訪者	0 (0.00%)	0.00%	0 (0.00%)	0.00%	0.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
4.  China									
所有工作階段	16 (0.65%)	100.00%	16 (0.93%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
回訪者	0 (0.00%)	0.00%	0 (0.00%)	0.00%	0.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
5.  Japan									
所有工作階段	10 (0.40%)	90.00%	9 (0.52%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)
回訪者	1 (0.13%)	0.00%	0 (0.00%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	\$0.00 (0.00%)

圖 35：訪客類型區隔分析範例

資料來源：本計畫。

12、問題十二：瞭解網站訪客推薦他人使用我們的網站情況。
本計畫的網站平台未使用使用 Facebook 推薦功能，故無範例資訊。

13、問題十三：分析網站到訪者瀏覽資訊或使用網站特定服務

的操作過程（訪客的進入點、停滯點和離去點）。

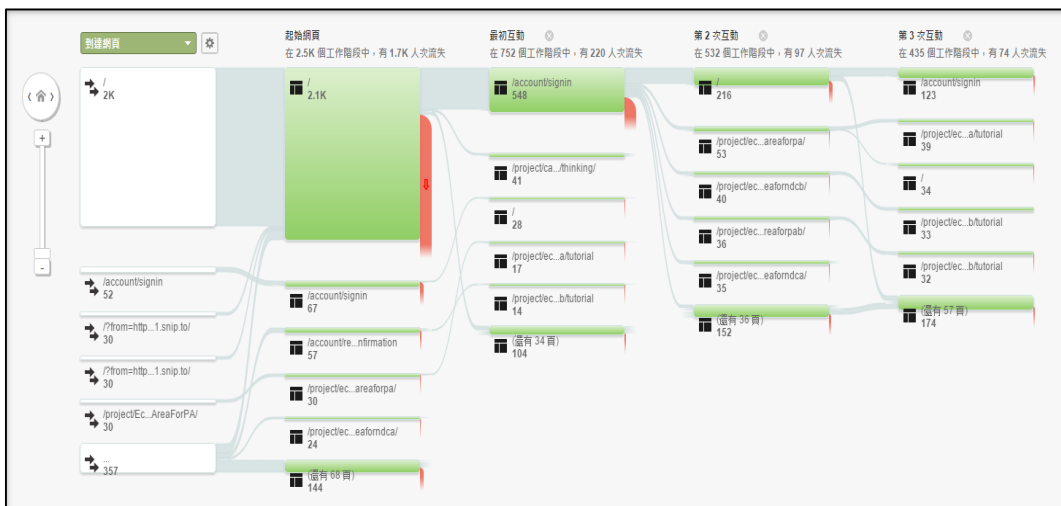


圖 36：訪客到訪流程分析範例

資料來源：本計畫。

14、問題十四：分析網站預期績效指標的達成率。

本計畫為實驗性質的群眾外包網站，並未特別設定預期指標資訊，故無範例資訊。

上述 14 點即為網站管理者使用 GA 進行網站分析時，能應用的 14 項網站分析技術及可回應的關鍵問題。本計畫建議政府機關未來能嘗試運用網站流量的巨量資料分析，將網站使用者的瀏覽行為資訊納入網站營運決策的參考資訊，以改善目前相關之決策與管理可能偏重賴專家判斷或主觀性意見而可能的不足之處，因此計劃團隊依據前述調查得知中央政府機關各網站管理者所關切的網站使用者行為資訊偏好次序(依據平均分數高低排序)，並整理出一份政府機關網站常用(或需求性較高)之 GA 服務的服務設定需求表(如表 39 所示)，供政府機關參考運用，並可依據管理需求自行或委託管理廠商協助設定 GA 報表。

表 40：中央政府機關網站使用 GA 服務需求表

受訪者 需求排序	調查問題	GA 功能概述
1	5.分析網站到訪者實際使用網站服務的情況（如網站最熱門網頁、服務或最常被下載的檔案）。（8.25 分）	設定行為（Behavior）分析來觀察訪客行為流程、網站內容、網站速度、站內搜尋、觸發事件 AdSense 廣告成效、網頁活動以及不同版本網頁實驗情況等相關資訊。
2	6.分析網站訪客曾在我們的網站搜尋哪些關鍵字（訪客站內搜尋關鍵字分析）。（8.00 分）	設定行為（Behavior）分析來觀察訪客行為流程、網站內容、網站速度、站內搜尋、觸發事件 AdSense 廣告成效、網頁活動以及不同版本網頁實驗情況等相關資訊。
3	8.分析網站系統的回應速度。（7.96 分）	設定網站速度分析（Site Speed），可取得資訊包含網站操作時間、網頁載入平均需時、網頁下載平均需時、伺服器反應平均需時等等多項網站管理的技術資訊。
4	7.分析網站訪客搜尋的關鍵字（包括站內與站外搜尋）和網站服務的對應關係。（7.88 分）	設定轉換（Conversion）分析來設定網站服務目標。
5	9.分析網站訪客的再訪頻率。（7.71 分）	設定目標對象（Audience）分析來觀察舊訪客再訪率、造訪頻率、再訪問隔時間、瀏覽量、參與度、裝置、作業平台等資訊，並可進一步設定興

受訪者 需求排序	調查問題	GA 功能概述
		趣分析來區辨訪客感興趣的廣告類型、有意願使用網站服務的訪客類型。
6	1.瞭解網路民眾從什麼管道得知我們的網站。 (7.69 分)	設定目標對象 (Audience) 分析來觀察新訪客。
7	4.瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站 (訪客站外搜尋關鍵字分析)。 (7.53 分)	設定攬客 (Acquisition) 分析來觀察網路行銷活動、AdWord 廣告活動、Google 站外搜尋和網站訪客到訪行為的關係。
8	13.分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程 (訪客的進入點、停滯點和離去點)。(7.49 分)	設定行為 (Behavior) 分析來觀察訪客的瀏覽流程。
9	14.分析網站預期績效指標的達成率。(7.45 分)	設定轉換 (Consersion) 分析來觀察服務被利用前的瀏覽歷程 (含熱門瀏覽歷程)、網站服務利用次數、網路行銷功勞歸屬等資訊。
10	11.區隔網站訪客類型 (如依據資訊瀏覽或服務使用行為、訪客來源地	設定目標對象 (Audience) 分析來觀察舊訪客再訪率、造訪頻率、再訪問隔時間、瀏覽量、參與度、裝置、作

受訪者 需求排序	調查問題	GA 功能概述
	點、再訪頻率、來訪時間等資訊區隔訪客類型)。(7.28 分)	業平台等資訊，並可進一步設定興趣分析來區辨訪客感興趣的廣告類型、有意願使用網站服務的訪客類型。
11	2.釐清網路民眾來源管道和網路行銷活動的關係。(7.23 分)	設定攬客 (Acquisition) 分析來觀察網路行銷活動、AdWord 廣告活動、Google 站外搜尋和網站訪客到訪行為的關係。
12	10.釐清網站訪客再訪率和行銷活動的關係。(7.06 分)	設定目標對象 (Audience) 分析來觀察舊訪客再訪率、造訪頻率、再訪問隔時間、瀏覽量、參與度、裝置、作業平台等資訊，並可進一步設定興趣分析來區辨訪客感興趣的廣告類型、有意願使用網站服務的訪客類型。
13	12.瞭解網站訪客推薦他人使用我們的網站情況。(6.98 分)	<p>1. 設定攬客 (Acquisition) 的推薦連結分析來觀察訪客到訪管道 (Channels) 以及所有推薦連結流量分析。</p> <p>2. 設定行為 (Behavior) 分析來觀察新訪客到訪和原訪客推薦連結的關係。</p> <p>設定轉換 (Conversion) 分析來觀察社交網路推薦連結的參訪流量。</p>

受訪者 需求排序	調查問題	GA 功能概述
14	3.蒐集網站到訪者的個人資訊(如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊)。(6.75 分)	設定目標對象(Audience)分析來觀察訪客個人資訊。

資料來源：本計畫。

政府巨量資料分析與政策端應用效能提升之研析

第七章 政策建議與結論

第一節 政府機關巨量資料分析應用的未來

透過本計畫三個包括「跨部門巨量資料分析」、「群眾外包應用在網路輿情分析」、以及「政府服務網站的網站 log 分析」等研析發現，本計畫可以歸納整理出三個面向的結論，此三個彼此獨立卻也相互關連的研析結論包括（1）政府巨量資料應用落差與（2）技術導入的分析結果。

一、政府運用巨量資料技術的「落差分析」：

依循本計畫方法中落差分析（gap analysis）的引導，從前面三個研析結果可以發現，如果從巨量資料時代政府引進電子化政府改革方案時，其「情境」（situation）的基礎上，發現下面三個落差，而這三個落差的彌平，是巨量資料時代電子化政府研究，建構起「循證的公共治理」（evidence-based public governance）的一個絕佳的契機：

首先，**決策應用的落差**：對於政府機關而言，與其將巨量資料視為另一個電子化政府甚至新公共管理的「流行運動」，不如將它穩健紮實地鑲嵌融入於以資料與資訊為基礎的「決策支援系統」（decision support systems, DSS）以因應「循證的公共治理」訴求。以本計畫的實證資料來看，巨量資料分析的基礎除了當代統計與資訊科學的操作平台與軟體演算法之外，例如資料與文字探勘（data/text mining）、資料視覺化（data visualization）等技術面因素之外，對於公部門而言更是因為它一方面不排除公共治理的政治特質，更強調其邏輯與實證的交互啟發與檢證，不論是應用巨量資料技術於跨部門資料運用、群眾外包輿情分析、或是網站 log 分析，都是為了提供政府決策單位，更多關於「政府/民眾」介面的循證資訊（evidence-based），以利於作出

強化該介面能協助執政者同時滿足民主回應與問題解決績效的目標，有趣的是，雖然本計畫內容大多是從技術的焦點出發（跨部門資料整合、群眾外包、log 分析），但是這些技術問題的解決，都有前述強化「政府/民眾」溝通介面的意涵，當然，如何強化政府應用巨量資料的決策功能，是本計畫發現最重要的一個應用落差。

再者，**資料治理的落差**：本計畫的受訪者主張在巨量資料原始的 3V（volume, velocity, variety）特質中，資料量（volume）雖然是巨量資料的命名基礎，但並非其最大的挑戰。仰賴日新月異的技術平台，除了在少數的即時應用情境之外，龐大資料量的蒐集與處理通常足以因應，尤其在限縮業務領域後的資料量可能也沒如此巨量；反而因為資料的更新與傳播速度（velocity）與多元格式與來源（variety），所導致的資料正確（accuracy）、精準或「顆粒」粗細（precision）、即時（updatedness）、與完整（completeness）才是更大的挑戰，期中報告審查委員對於「強化政府所管理主要資料庫及網站，在社會、經濟、科技面向之治理，包括所需要的資料擷取規範、儲存架構、分析工具及法律限制等內涵」，就是這個落差的問題；因此，這個落差也同時出現在群眾外包研析的資料判讀與後續應用問題，以及 log 分析政府服務網站研析中，政府網站對於 log 資料的建置與應用的需求與實際差異都可明顯看出，政府單位缺乏資料治理架構的引導，進行技術與實務的同步行政改革，是十分值得重視的。

最後，**議題導向的落差**：本計畫各國的巨量資料相關文獻與實務中，存在著以公共問題/決策導向（problem/decision/domain-driven）與事實/探索導向（fact/pattern-driven, exploratory）的兩種不同的分析應用路線，前者較為接近傳統決策支援學理所闡釋的以組織所面臨的策略、決策問題、與領域知識為範圍，導引後續的資料蒐集處理與分析所需的方法；而後者則是以探索性的資料分析方法先從資料中找到特定的關聯規則或型態，再透過領域知識探討背後的邏輯予以探討。實作經驗的政府機關而言，本計畫認為以「公共問題議題導向的導入模式」應該更為適用，不論是跨域的資料整合（特定議題的專業需求）、

還是群眾外包所要面對的網路輿情問題（特定議題的意見分析）、還是政府服務網站 log 分析的應用等看似技術性的問題（特定議題的網站應用），在電子化政府的應用上，都與特定議題（issue）的管理脫不了關係，前述「公共問題/決策」導向的巨量資料應用，也必須以特定議題為管理核心，這與企業的資料應用大多聚焦單一「產品/議題」有很大的不同，因為政府的公共問題處理，是「多議題」與「跨議題」的，如何面對議題與公共問題的區隔，是政府應用巨量資料不能迴避的工作。

二、 技術導入的分析結果

承接上述研析發現之決策應用、資料治理、以及議題導向等落差，接下來本計畫結論將分三部分討論，政府應用巨量資料導入跨域、群眾外包、以及網站 log 分析的結果發現。

（一） 跨部門巨量資料分析：

本計畫發現，問題/決策導向的巨量資料分析難免遇到無法滿足的需求，例如資料正確性與精確度的落差、資料欄位定義不清與格式轉換等，然而這些不足之處也正好刺激其分析應用需求，重新檢討、研擬與修正所需資料的蒐集（例如人工輸入與介接既有系統）與處理（例如格式定義、清理、勘誤、轉換）規格與程序，這對於欠缺先前經驗的政府機關而言，也正是導入巨量資料分析所需的基礎條件。

即使是單一政府機關中的特定業務領域，例如台北市政府人事處針對其人力資源管理議題，例如彈性上班時段的安排，以上所需資料的蒐集與處理也可能設計業務、資訊、統計、與法規單位間的討論；如果巨量資料分析的應用領域及資料來源又涉及跨政府機關（如部會、局處、中央/地方）或跨業務領域（如健保、社福、人事），為達成以上基礎條件所需的溝通協調與整合需求勢必更為繁複。本計畫透過國內外相關文獻與實證資料也得知，透過任務編組的專案團隊堪稱頗為適

合的運作模式，待業務分析單位瞭解其問題與資料需求，資訊單位也能適度支援其規格，統計單位補強其所需的資料分析預測方法及模型，法規單位釐清其資料蒐集處理利用的權責適法性，就可以再交由業務單位主導將巨量資料逐步融入其常態業務流程中。

（二）群眾外包應用在網路輿情分析：

本次研析是基於過去兩年電子治理研究中心對於網路輿情分析發現所建構的實驗設計，這個研析的過程中，假設政府未來可以應用群眾外包的網路機制，提供一般巨量資料分析的「情緒分析」所不能提供的「立場分析」資訊；這樣假設的主要原因是由電腦關鍵字組所篩選出的「正/負」向情緒的工具，在公共領域面臨網民「說反話（用負面的詞支持某些言論，或是用正面的詞反對某些言論）」的干擾，必須用人工的方式進行「立場分析」，以拉近巨量資料分析與決策者之間對於可用性的認知差異。當然，這個建構的過程與結果，對於未來政府建構網路輿情機制，是一樣地深具啟發性。

首先，在群眾外包平台的設計與行銷上，本計畫自行應用特定網站工具（PyBossa）進行從零開始的群眾外包平台設計，中間所花的跨專業溝通與修改的資源，超乎原先的想像，加上過程中因為實驗的需要，加入「機構效應」與「從眾效應」的實驗元素，讓整個網站的設計工作拖到期中報告時還未能上線；另外，除了設計的問題之外，還有上線行銷的問題，一個群眾外包的新網站，如果不能想出善於流傳的口號與 logo，以及足夠的網路廣告資源進行推動，政府運用群眾外包的可能性會大大降低，這些應用關鍵技術的能力，對政府來說是一個新的挑戰。

再者，本實驗的結果顯示，第一，分析群眾外包結果與去年研究助理團隊結果差異，如果從助理角度排除不好判斷的資訊之後，不論是正面與負面的立場判斷，網民與助理的判斷約有接近七成五的相容率，顯示兩種方法有一定的類同率，因此，成本效益與時效因素可能成選擇哪一種方法的重要考量，當然，兩種方法處理難以分析的資訊

的問題，還需要更多的分析；第二，「機構效應」對參與群眾外包者判斷資訊的立場影響不大，但是「從眾效應」卻有顯著的影響，這個部分的影響可能來自網民上網看似自主但卻非常容易受到網路大多數人意見所影響的結果，這個結果讓未來群眾外包的網路設計，可能還是要採取較為「封閉」的判斷網路輿情的環境，讓每一位上來協助的網民，都可以有獨立決策的機會，事實上，這是民主理論個人意見整合成為人民意志的關鍵問題，以是為何選罷法規定選舉投票前十天開始不得公佈民調的原因。第三，在人口分佈上，女性超過男性、以台灣地區的網民為主、以及約有 60%左右的「離開率」，都可當作未來設計政府群眾外包網頁「使用者」可以參考的重點。

最後，關於「情緒分析（純機器）」、「助理編碼（純人工）」與「群眾外包（人工加上網路協助）」等三種方法的成本效益分析，時效上，情緒分析當然是最快的，群眾外包次之，助理編碼最慢，成本上來看，情緒分析統整在一整個網路輿情分析的合約包裹內，助理編碼的成本大多是工資，而群眾外包是有償的工作，再加上行銷費用，需要進一步確認成本與效益的差異，不過一般來說，群眾外包可以改變參與的密度與設計不同的報償機制，在時間的控制上也比助理編碼來得有彈性，成本效益的價值並不差，可能端看如何設計到最適的水平，應該可以得到更好的 CP 值，只不過行銷的成本效益較難分析，這也可能是群眾外包方法的成敗關鍵所在。

（三）政府服務網站的網站 log 分析：

首先，本計畫的網站經營者訪談資料顯示，政府的網站設立與私部門網站設立最大的差異，就是公部門並沒有一個客觀的「獲利」標準可以當作營運與績效管理的標的，如果網站還有網路購物的功能，一般而言私部門會有更大的誘因去改善網民瀏覽經驗，以增加消費者打開皮夾花錢的可能性；因此，政府對於應用網民瀏覽 log 分析來改善政府網路使用經驗的誘因是不足的，包括政府服務品質獎項下的政府網路應用的相關指標，也沒有要求政府各單位進行網路網民瀏覽 log

的分析，以作為網頁改善的決策之用。本計畫認為，過去研究顯示，網民對於政府資料擷取與表達意見的興趣是高於一切的，但是，政府的網站經營是否真有確實處理這兩個基本需求，是有很大疑慮的，因此，政府網站的瀏覽率低一方面可能是公共議題比起演藝圈花絮有其先天的劣勢，但是另一方面，政府網站沒有基於網民瀏覽 log 資料進行網路設計的改善作為，或許也是網民對政府網站沒有興趣的重要原因之一，因此，這個可以用巨量資料分析進行網頁瀏覽的舒適度提升。

再者，即使政府機關網站已獲得網站管理者以及機關外部檢核者評價符合相關使用者導向的建置規範、使用者親和度以及網站可用性，但並不會因此增加或減少網站的瀏覽人次。換言之，政府機關網站每月瀏覽人次的高低，無法僅從以人工檢視為主的主觀評估指標的數據高低中推測出兩者的關聯性。各政府機關若想要瞭解網站使用者的實際使用行為，僅以上述主觀指標的達成度做為網站營運管理績效的思考，顯然是不足的。

最後，網站流量分析可彌補傳統運用網路問卷調查與實驗設計法，來分析網站使用者行為，主因在於網站流量分析是以網頁日誌或紀錄檔（web page logs）或網頁標記（page tagging）來分析網站使用者的網站足跡，所得資訊具有客觀性，可排除受觀察者主觀意識導致的調查偏誤，在眾多網路流量分析工具中，以 GA 最為各界所青睞。在 AARRR 分析架構下，整體而言各政府機關管理的在個別構面的需求雖大多為正向，但中央政府二級機關的網站管理者，重視「成為有效的使用者」和「使用者付費或達成預期行動」等構面相關資訊（各題平均分數皆在 7 分以上，部分落在在 8 分以上）的程度，會高於其他構面；中央政府三級機關的網站管理者，則是各構面都重視，值得關心的問題是，這種新的 GA 分析能力，除了外包一途之外，公部門內部人員的相關能力品質，以及能夠連接 GA 分析與政府網站改進決策的人才何處覓等人力資源的問題，也都值得關心。

第二節 政府機關巨量資料分析應用的實務建議

一、跨部門巨量資料分析：

本計畫以下表 41 整理了我國政府機關巨量資料分析的推動建議，其中主協辦政府機關是指巨量資料應用分析的範圍，有可能是單一政府機關（如行政院特定部會、地方政府特定局處、或整體行政院、台北市政府等），或是跨政府機關（如衛福部主辦、財政部協辦，或行政院人事總處主辦、各縣市政府人事處協辦），而業務領域是指巨量資料應用的業務範圍（如健保、社福、人事）。

如此交叉則而成為四個政府機關推動巨量資料推動建議與進程，其中類型 A 代表單一政府機關中將巨量資料應用於單一特定領域，中央政府如財政部將綜合所得稅申報與後續查核補徵結果資料作為後續查核樣態的預測分析，地方政府如台北市政府人事處運用歷年差勤記錄研析彈性上下班的可能性。類型 B 則為單一業務領域但跨機關主協辦，中央政府如行政院衛福部整合本身的食物及原物料資料與財政部的電子發票資料，作為追蹤可疑食材來源；地方政府如新北市政府教育局整合新北市社會局的通報資料作為中輟學生的事前預測與事後分析；跨中央與地方政府如行政院人事總處彙整各縣市政府人事處的公務人力訓練資料，瞭解全國公務人力的訓練樣態。

表 41：政府機關巨量資料推動進程

主協辦機關 (如部會、局處、中央/地方)	業務領域 (如健保、社福、人事)	單一 業務領域	跨 業務領域
單一機關		A	C
跨機關		B	D

資料來源：本計畫。

類型 C 則是單一主協辦機關但卻跨越業務領域的巨量資料應用，例如台北市政府研考會整合了 1999、市長電子信箱等民眾意見（通常為文字資料）、與各局處跨業務領域的決算與績效資料，以瞭解台北市政府整體經費配置於各業務領域的適當性。最後，類型 D 代表巨量資料的應用同時跨越了政府機關與業務領域的界線，例如中央與地方的社會福利、戶政、與教育政策機關，共同整合相關資料以瞭解新住民的相關議題。

首先，本計畫主張：現階段我國應先鼓勵試辦類型 A 的巨量資料應用案例，使得單一主管機關內部的業務/資訊/統計/法規/研展綜規單位，都能透過特定業務領域的應用案例，以問題與決策主導資料治理的各項需求，包括資料的蒐集、保管、更新、定義、清理、勘誤、轉換、分析、與分享等，磨合跨單位相關法規、技術、政治、與行政管理的立場與調整，並累積經驗以提升巨量資料分析應用所需的素養與技能，更重要的是，單一主管機關的大數據應用頗有可能引發跨政府間與跨業務的資料需求，也是後續幾種大數據推動類型的經驗基礎。接著類型 B 的巨量資料應用也值得儘早累積相關經驗，例如行政院人事行政總處、考試院銓敘部、各縣市政府人事處可以各自需求研商公

務人力資料的通用格式或轉換規格，同時調整所需法規以有助於全國公務人力巨量資料應用所需的資料統整。

相對而言，類型 C 與 D 牽涉更為複雜的技術與法規調適，例如各業務專屬法規可能會界定民眾資料的利用範圍限制於該業務中，或是不同業務領域的資訊系統硬軟體與資料規格也可能頗為殊異，且其巨量資料分析與應用也頗為多樣與異質，因此本計畫建議更應以較為策略層次的問題與決策需求（相對於資料探勘導向）予以引導巨量資料專案的規劃與執行，以避免範圍發散而易生挫折，尤其是類型 D 勢必更為利害關係人的關鍵問題導向，更需要有類型 A 與 B 的經驗作為基礎。

其實不論是以上 ABCD 哪一種巨量資料應用類型，資料的蒐集、盤點、定義、維護、交換等機制都必須妥善規劃執行，本計畫也發現目前許多政府機關由於多數資訊系統均已委外建置維運，其業務單位甚至資訊單位都必須仰賴外部廠商規劃執行以上的資料管理方案，而不同廠商之間的合作成效其實就決定了跨機關業務的資料整合交換的成效，因此審慎規劃並發揮國網中心的技術服務角色應該是可行的跨機關跨業務資料管理與分析應用方案，縱使龐大的政府資料不可能完集中存放於國網中心，但是如果能跨資料集串連的共享機制（如 API），輔以妥善的個資與安全標準（如經濟部標準檢驗局的資訊安全認證），就能打通大數據應用最基本的資料與法規障礙，對於被課以保管民眾資料的個別公務人員而言，就能更無後顧之憂地依據其政策與行政需求規劃執行巨量資料應用。

二、 群眾外包應用在網路輿情分析：

群眾外包的平台是巨量資料時代政府輿情分析的一種方法的選擇，當然，網路輿情分析對目前各級政府是非常重要的，群眾外包的平台設計、運作與應用，都可以在本計畫的基礎上，繼續前進開發。然而，本計劃所得到的結果，在實務上可以提出下面三點改革的策略性建議。

（一） 平台的建立策略：

政府群眾外包的網路平台設計，由於初步的實驗顯示還有許多可以改善的空間，建議可以將這個設計的工作當作一個網路問題解決競賽的推廣作法，上網邀請熟悉網路平台設計的團隊，共同來參與設計比賽，並且從中找出可以合作的夥伴，最後定出政府各單位需要群眾外包平台的標準模式，作為推廣應用的技術性基礎。

（二） 平台的運作：

群眾外包在輿情分析上的應用雖然實務上是可行的，仍然需要在三個點上繼續進行技術的創新，首先，如何讓網民進入網頁時降低「離開率」，應該要有更好的設計策略；再者，網民對於付費服務的反應如何，以及是否存在更有利於網民參與的付費方式，應該再加強，當然，從政府這一方來看，這也是一個成本效益的問題，如何精算網路輿情分析各種途徑的 CP 值，也可以投注資源進行研究；最後，平台的行銷策略該如何擬定，也是平台運作成功的重要因素，也應該設法進行策略分析與建構。

（三） 平台的應用策略：

目前本計畫的應用範圍是在網路輿情的立場判斷，可否以更為廣泛的群眾外包應用，比方說，結合「電子化規則設定」(e-rulemaking)或是其他解讀民眾意見反映的整體與個體資料等工作，都可以繼續投注資源開發；另外，網路輿情分析除了立場分析之外，還有下一階段的「議題論述」的分析，是否可以結合群眾外包進行更多的解讀提供政策決策單位進行決策參考，都是可以繼續努力的地方。

三、政府服務網站的網站 log 分析：

政府網頁的 log 分析，最重要的目的就是應用巨量資料分析的技術，分析並歸納政府網站效能提升的決策支援資料，這個部份因為公私部門的差異關係，過去政府部門服務品質提升上，並沒有受到應有的重視，因此，本計畫建議的改革策略一共有三點。

（一）網站服務效能評估：

建議政府資訊管理單位，除了建構相關的網站設置標準之外，也可以配合巨量資料時代所產生新的可能性，對政府單位的網站進行服務效能提升的標準制定與輔導作為，過去相關文件中加入 log 資料收集的底線、GA 分析的標準、以及如何用這些分析出來的資料改善網站運作效應。

（二）網站服務的品質提升策略：

建議在國發會政府服務品質獎的電子化政府項目中，加入「能夠應用收集到的網頁 log 資訊，進行網路使用行為的分析，並且從中提出網站營運的改善意見」的項目，從品質獎參獎的角度切入，推動政府網站依照 log 資訊回饋優化的政府網站改革。

（三）網站營運的策略：

目前政府網頁的策略分析並不是網站經營的重要工作，未來如果 log 收及與 GA 分析成為標準作業程序之後，就可以開始針對特定的服務性網站進行策略分析與建構，包括網頁與 FB 的交互應用、目標群體的需求定位分析、以及建構虛實整合的「政府／民眾」互動接面的策略，以改善民眾與政府之間的關係。

第三節 政府機關巨量資料分析的後續研析建議

本計劃針對這一年來的研析成果，提出下面兩點後續研析建議，第一是操作理論與實務連接的「機制」部份，第二則是後續研析「議題」的部份，分別說明。

首先，在機制的問題上，本計劃有兩點後續研析建議，第一，增加政府電子治理改革業務與產學協力的質與量。雖然目前科技部已有鼓勵政府實務界協同學術團隊共同規劃執行巨量資料應用分析補助計畫，但其數量與推廣仍頗為受限，為了更快累積實務與學術界的實作經驗，本計劃建議政府機關如決定推動巨量資料分析，則應視其應用案例主動尋求與相關學術領域的研究團隊合作，透過行動學習與行動研究（action learning and action research）途徑累積實作經驗，並同步累積研究成果以回饋修正於巨量資料專案，相信此產官學合作模式必能提升我國政府機關運用巨量資料分析的能量，如也能透過適當的經驗交流分享機制，也將吸引更多政府機關投入並逐步融入於常態行政與政策流程中，相信也終能逐步檢驗與發揮巨量資料功效與限制，使得巨量資料分析得以貢獻於循證公共治理。

第二，本計劃也認為在應用「績效管理」工具加速或強化政府推動巨量資料應用的過程中，目前仍屬於初步導入階段，因此，國發會的資訊處可以以鼓勵試辦與經驗（成功與失敗案例同等重要）的分享與推廣為主軸，至少未來兩年內不適合直接訂定績效目標作（例如一機關每年至少一個巨量資料應用案例）為管考手段，也更能有助於以上跨機關、群眾外包、以及網頁 log 分析經驗累積與深度融合於既有的行政與政策流程中，以外部研究團隊的角度來看，目前我國政府機關推動巨量資料分析應用的模式，仍有加強與外部技術與學術團隊合作交流的空間，甚至有受訪者建議巨量資料分析也應考量透過開放政府資料（Open Government Data, OGD）徵求群眾智慧、過程中以 log 分析確認政府與民眾的連接，並且最後應用回饋刺激政府內部的資料治理流程不斷成熟。。

再者，未來研析的方向建議上，本計劃提出三個可能方向。第一，政府資料治理的跨議題與跨部門機制原型（**prototype**），從傳統政府組織分工後跨域管理的介面，以及傳統公共政策議題設定多議題的環境出發，找出政府推動巨量資料「無界限」（**borderless**）的資訊架構，這個部份絕對會涉及到政府組織、資訊自由、個人資料保護等相關法規的相容性問題，或許，政府目前積極推動「資訊基本法」的當下，除了從產業面出發以外，政府無界限的資訊架構的基礎法制建設，也應該是其中重要的一環，這部份的議題值得投資資源進行相關研究；第二，網路輿情分析的最後一哩路－「輿情論述萃取機制」（**Internet Arguments Extraction Mechanism, ITEAM**）的研究，從網路輿情的資料萃取、聲量-情緒-立場分析等技術的開發與應用，最終的目的還是在於篩選出來的資料，可否真正協助政府進行對應的政策與公關作為，本計劃認為，以本中心過去三年的相關研究顯示，目前應該到了要探索網路輿情「正反論述萃取」的階段，需要設計能將公共議題之正反意見演化流程與論述內容的方法，與政府現有的政策行銷與公關部門的工作內容接軌。第三，「全觀式外部輿情資訊系統」（**Holistic External Opinion System, HEROS**）架構之研究，由於目前政府收集外部輿情資訊的管道十分多元，包括民意調查、剪報資訊、質化資訊（會議、公聽會、首長信箱、行政訴訟...）、以及網路輿情等，要如何整合這些多元管道，協助政府建構一個「全觀式的輿情系統」，也是可以投注資源的所在。

政府巨量資料分析與政策端應用效能提升之研析

參考文獻

中文部分

- 方映鈞（2014）。「全國文官雲端資料庫」人事資料管理的下一步。
公務人員月刊，**212**，65-71。
- 方映鈞、林進富（2015）。全國公務人力資料雲端服務平台—雲端及大數據環境下公務人力資料的整合、管理及運用。**人事行政**，**190**（4），2-10。
- 王聖銘（2015）。群眾外包災情資訊之評估與檢證—以莫拉克颱風為例。
災害防救科技與管理學刊，**4**（1），31-48。
- 王豐勝、黃彥文（2013）。Big Data 時代啟示錄—雲端巨量資料的國際潮流與產業需求。**經濟前瞻**，**146**，123-128。
- 吳宜蓁（1998）。**議題管理**。台北：正中。
- 吳嘉川（2010）。2009 年紐約「主資料管理高峰會議」考察紀要。**金融聯合徵信雙月刊**，**11**，55-59。
- 宋方捷（2012）。政府跨領域資料管理與應用—以高雄市社會經濟資料庫決策應用系統為例。**國土及公共治理**，**3**（1），121-130。
- 宋餘俠、李國田（2012）。政府部門資料加值推動策略與挑戰。**研考雙月刊**，**36**（4），10-21。
- 林巧敏（2011）。政府資訊公開與申請應用網站內容分析：以臺灣與美英兩國政府機關為例。**大學圖書館**，**15**（2），78-98。
- 林俊宏（譯）（2013）。**大數據**（Mayer-Schönberger, V., and Cukier, K. 原著）。臺北：天下文化。
- 林珊如（2002）。網路使用者特性與資訊行為研究趨勢之探討。**圖書資訊學刊**，**17**，35-47。
- 林蔚君、詹雅慧（2013）。巨量資料分析與 GIS 應用價值創造。**國土資訊系統通訊**，**88**，12-20。
- 邱耀漢（2010）。公部門資訊治理評估模式。**國家菁英季刊**，**5**（2），

37-60。

胡為君（譯）（2011）。網頁分析：**Google Analytics**（原作：Cutroni, J.）。臺北市：基峯資訊股份有限公司。

國發會（2014）。個人家戶數位機會調查報告書報告書（103 年度）。2015 年 08 月 08 日，取自：
<http://www.ndc.gov.tw/cp.aspx?n=55C8164714DFD9E9>。

莊素宜（2010）。從雲端運算看公部門人事行政資訊服務整合。人事月刊，299，37-41。

陳玫真、黃營杉（2009）。網站可用性對消費者行為意圖影響之研究。東吳經濟商學學報，67，79-106。

陳恆鈞（2011）。議題管理初探。T&D 飛訊，111，1-14。

陳祥、林明童（2002）。我國「電子化政府整合型入口網站」使用者行為分析。圖書館學與資訊科學，28（2），1-10。

陳敦源（2009）。民主治理：公共行政與民主政治的制度性調和。台北市：五南圖書公司。

陳敦源、李仲彬、張鎧如、陳玉豐、林錦鈺、簡鈺琿（2014）。邁向循證基礎的人力資源管理：以主客觀資料整合建立「臺灣文官調查資料庫」的經驗分享。人事月刊，346，11-28。

陳敦源、李仲彬、黃東益（2007）。推動電子化政府可以改善公眾接觸嗎？台灣個案的分析〉。東吳政治學報，25（3），25-92。

陳敦源、潘競恆（2011）。政府就是「我們」：Web 2.0 時代民主治理的希望或幻夢？研考雙月刊，35（4），10-22。

黃彥棻（2011）。資訊部門主動發掘使用者需求 銓敘部用 IT 加速績效評量作業，2011 年 9 月 23 日，取自：
<http://www.ithome.com.tw/node/69906>。

詹大千、金傳春（2013）。地理資訊系統在傳染流行病學上的應用。國土資訊系統通訊，88，21-31。

- 廖洲棚、陳敦源、蕭乃沂、廖興中（2013）。運用巨量資料實踐良善治理：網路民意導入政府決策分析之可行性研究。行政院研究發展考核委員會（編號：RDEC-MIS-102-003），未出版。
- 劉龍龍、葉乃嘉、何志宏、余孝先（2013）。各國政府之雲端發展策略與推動現況。公共治理季刊，1（3），22-34。
- 潘競恆、王千文、陳敦源（2011）。電子治理與 Web2.0 社群網站使用行為之實驗研究。台灣公共行政與公共事務系所聯合會（TASPAA）年會暨學術研討會，台北市。2011 年 06 月 29 日。
- 蔡良文（2014）。人事行政學：論現行考銓制度（修訂六版）。台北：五南。
- 鄭江宇、張佳榮（2015）。掌握行銷新趨勢：你不可不知的網站流量分析。新北市：新頁圖書股份有限公司。
- 蕭乃沂、陳敦源、廖洲棚（2014）。政府應用巨量資料精進公共服務與政策分析之可行性研究。行政院研究發展考核委員會（編號：RDEC-MIS-103-003），未出版。
- 蕭元哲、葉上葆（2003）。電子化政府之使用分析—檔案分析法。電子商務研究，1（2），207-224。
- 羅晉（2008）。電子化參與的效益與風險：民眾對政府網站的認知與使用行為之初探。資訊社會研究，15，181-208。

英文部分

- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. **Wired Magazine**, http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory.
- Ansoff, H. I. (1980). Strategic Issue Management. **Strategic Management Journal**, 1 (2), 131-148.
- Axelsson, K., U. Melin., & I. Lindgren. (2010). Exploring the Importance of Citizen Participation and Involvement in e-government Projects:

- Practice, Incentives, and Organization. **Transforming Government: People, Process and Policy**, 4(4), 299-321.
- Balm, G. J. (1996). Benchmarking and gap analysis: what is the next milestone? **Benchmarking for Quality Management & Technology**, 3 (4), 28-33.
- Behn, R. D. (2014). **The PerformanceStat Potential: A Leadership Strategy for Producing Results**. Washington, D. C.: Brooking Institution Press.
- Brabham, D. C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. **Convergence: The International Journal of Research into New Media Technologies**, 14 (1), 75-90.
- Buie, E. & D. Murray. (2012). **Usability in Government Systems**. Waltham, MA: Elsevier Press.
- Chun, S. A., S. Shulman., R. Sandoval., & E. Hovy. (2010). Government 2.0: Making Connections between Citizens, Data and Government. **Information Polity**, 15 (1/2), 1-9.
- Clark, W. R. & M. Golder. (2015). Big Data, Causal Inference, and Formal Theory: Contradictory Trends in Political Science? **PS: Political Science and Politics**, January, 65-70.
- Cobb, W. N. W. (2015). Trending Now: Using Big Data to Examine Public Opinion of Space. **Space Policy**, 1-6. (article in press).
- Coleman, R., P. Lieber., A. L. Mendelson., & D. D. Kurpius. (2008). Public Life and the Internet: If You Build a Better Website, Will Citizens Become Engaged? **New Media & Society**, 10 (2), 179-201.
- Criado, J. I., R. Sandoval-Almazan., & J. R. Gil-Garcia. (2014). Government Innovation through Social Media. **Government Information Quarterly**, 30, 319-326.
- Davis, P. M. (2004). Information-Seeking Behavior of Chemists: A Transaction Log Analysis of Referral URLs. **Journal of the**

- American Society for Information Science and Technology**, **55** (4), 326-332.
- Demirkan, H. & D. Delen. (2013). Leveraging the Capabilities of Service-oriented Decision Support Systems: Putting Analytics and Big Data in Cloud. **Decision Support Systems**, **55**, 412-421.
- Desouza, K. C. (2013). Implementing Open Innovation in the Public Sector: The Case of Challenge.gov. **Public Administration Review**, **73** (6), 882-890.
- Desouza, K. C. & B. Jacob (2014). Big Data in the Public Sector: Lessons for Practitioners and Scholars. **Administration & Society**, (article in press).
- Dumbill, E. (2012). What is Big Data? An Introduction to the Big Data Landscape. Retrieved May 2, 2015 from <https://beta.oreilly.com/ideas/what-is-big-data>
- Eaton, C., D. Deroos., T. Deutsch., G. Lapis., & P. Zikopoulos. (2012). **Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data**. New York: McGraw-Hill Press.
- Estevez, E., C. Chesnevar., Maguitman, A., & R. Brena. (2012). DECIDE 2.0 – A Framework for Intelligent Processing of Citizens' Opinion in Social Media. The Proceedings of the 13th Annual International Conference on Digital Government Research, accessed from: http://delivery.acm.org/10.1145/2310000/2307778/p266-estevez.pdf?ip=140.119.45.41&id=2307778&acc=ACTIVE%20SERVICE&key=AF37130DAFA4998B%2EBBA1407E900E193F%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=685680900&CFTOKEN=41099447&__acm__=1434940647_6c09cd350ecfd063821b6739e69f6160.
- Fang, W. (2007). Using Google Analytics for Improving Library Website Content and Design: A Case Study. **Library Philosophy and Practice**, **9** (3), 1-17.

- Gayo-Avello, D. (2011). Don't Turn Social Media into Another 'Literary Digest Poll'. **Communication of the ACM**, **54** (10), 121-128.
- George, G., M. R. Haas & A. Pentland. (2014). From the Editors: Big Data and Management. **Academy of Management Journal**, **57** (2), 321-326.
- Gill, K. (2012). Usability of Public Web Sites. In Elizabeth Buie and Dianne Murray (Eds.), **Usability in Government Systems User Experience Design for Citizens and Public Servants** (pp. 21-39). Waltham, MA: Elsevier, Inc.
- Groves, K. (2007). **The Limitations of Server Log Files for Usability Analysis**. Retrieved July 22, 2015 from <http://web.elastic.org/~fche/mirrors/www.cryptome.org/2014/01/boxesandarrows-log-spy.pdf>.
- Heath, R. L. & K. R. Cousino. (1990). Issues Management: End of First Decade Progress Report. **Public Relations Review**, **16** (1), 6-18.
- Howe, J. (2006a). **Crowdsourcing: A Definition**, Crowdsourcing: Tracking the Rise of the Amateur (weblog, 2 June) , Accessed 20 July 2015:
http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html
- Howe, J. (2006b). The Rise of Crowdsourcing. **Wired**, **14** (6), Accessed 20 July 2015:
<http://www.wired.com/wired/archive/14.06/crowds.html>
- Howe, J. (2008). **Crowdsourcing: How the Power of the Crowd is Driving the Future of Business**. Random House.
- Jang, S. M. & P. S. Hart. (2015). Polarized Frames on 'Climate Change' and 'Global Warming' across Countries and States: Evidence from Twitter Big Data. **Global Environmental Change**, **32**, 11-17.

- Jansen, B. J. (2006). Search Log Analysis: What It Is, What's Been Done, How to Do It. **Library & Information Science Research**, **28**, 407-432.
- Kim, G. H., S. Trimi., & J. H. Chung. (2014). Big-Data Applications in the Government Sector. **Communications of the ACM**, **57** (3), 77-85.
- Lazer, D., R. Kennedy., G. King., & A. Vespignani. (2014). The Parable of Google Flu: Traps in Big Data Analysis. **Science**, **343**(6176), 1203–1205.
- Lewis, S. C., R. Zamith., & A. Hermida. (2013). Content Analysis in an Era of Big Data: A Hybrid Approach to Computational and Manual Methods. **Journal of Broadcasting & Electronic Media**, **57**(1), 34-52.
- Lodge, M. & K. Wegrich. (2014). Crowdsourcing and Regulatory Reviews: A New Way of Challenging Red Tape in British Government? **Regulation & Governance**, **9**(1), 34-46.
- Mahrt, M. & M. Scharkow. (2013). The Value of Big Data in Digital Media Research. **Journal of Broadcasting & Electronic Media**, **57**(1), 20-33.
- Manovich, L. (2012). Trending: The Promises and the Challenges of Big Social Data. In M. K. Cold(ed.)**Debates in the Digital Humanities**, pp.460-475. Minneapolis, MN: University Of Minnesota Press.
- McMillan, S. J. (2000). The Microscope and the Moving Target: The Challenge of Applying Content Analysis to the World Wide Web. **Journalism and Mass Communication Quarterly**, **Spring 77** (1), 80-98.
- Milakovich, M. (2013). **Anticipatory Government: Integrating Big Data for Smaller Government**. In Policy, Internet, and Politics Conference, Oxford, UK: Oxford University, accessed from:

- <http://ipp.oii.ox.ac.uk/sites/ipp/files/documents/Anticipatory%20Government%20Oxford.pdf>.
- Neuman Russell. W., L. Guggenheim., S. Mo Jang., & S. Y. Bae. (2014). The Dynamics of Public Attention: Agenda-setting Theory Meets Big Data. **Journal of Communication**, **64**, 193-214.
- Nicholas, D., P. Huntington., N. Lievesley., & R. Withey. (1999). Cracking the Code: Web Log Analysis, **Online and CD-Rom Review**, **23**(5): 263-269.
- Obi, T. & N. Iwasaki. (2015). **A Decade of World e-Government Rankings**. Nieuwe Hemweg, Netherlands: IOS Press.
- Parasuraman, A., V.A. Zeithaml., & L. L. Berry. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. **Journal of Marketing**, **49**(4), 41-50.
- Perrott, B. E. (1996). Managing Strategic Issues in the Public Service. **Long Range Planning**, **29**(3), 337-345.
- Perrott, B. E. (2011). Strategic Issue Management as Change Catalyst. **Strategy & Leadership**, **39**(5), 20-29.
- Poetz, M. K., & M. Schreier. (2012). The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas? **Journal of Product Innovation Management**, **29**(2), 245-256.
- Borromeo, R. M. & M. Toyama. (2015). **Automatic vs. Crowdsourced Sentiment Analysis**. Proceedings of the 19th International Database Engineering & Applications Symposium, 90-95.
- Roloff, J. (2007). Learning from Multi-Stakeholder Networks: Issue-Focussed Stakeholder Management. **Journal of Business Ethics**, **82**, 233-250.
- Shi, Q. & M. Abdel-Aty. (2015). Big Data Applications in Real-time Traffic Operation and Safety Monitoring and Improvement on Urban Expressways. **Transportation Research Part C**, (article in press).

- Silvestro, R. (2005). Applying gap analysis in the health service to inform the service improvement agenda **International Journal of Quality & Reliability Management**, **22**(3), 215 – 233.
- Sjovaag, H. & E. Stavelin. (2012). Web Media and the Quantitative Content Analysis: Methodological Challenges in Measuring On-line New Content. **Convergence: The International Journal of Research into New Media Technologies**, **18**(2), 215-229.
- Suri, P.K. and Sushil. (2012). Planning and Implementation of e-governance Projects: A SAP-LAP Based Gap Analysis. **Electronic Government: An International Journal**, **9**(2), 178-199.
- Sushil. (2000). SAP-LAP Models of Inquiry. **Management Decision**, **38**(5), 347-353.
- Surowiecki, J. (2004). **The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations**. New York: Doubleday.
- U.S. Dept. of Health and Human Services. (2006). **The Research-Based Web Design & Usability Guidelines**, Enlarged/Expanded edition. Washington: U.S. Government Printing Office, <http://www.usability.gov/>, 2015.08.10 last visited.
- Wang, H. J. & J. Lo. (2012). Determinants of Citizens' Intent to Use Government Websites in Taiwan. **Information Development**, **29**(2), 123–137.
- Wood, C. (2014). **Managing and Organizing Data is the Next Phase of State and Local Government's Hopeful Metamorphosis**. Retrieved April 29, 2014, from <http://www.govtech.com/data/Data-Governance.html>
- Xiang, Z., Z. Schwartz., J. H. Gerdes., & M. Uysal. (2015). What Can Big Data and Text Analytics Tell Us about Hotel Guest Experience and

- Satisfaction? **International Journal of Hospitality Management**, **44**, 120-130.
- Yiu, C. (2012). **The Big Data Opportunity: Making Government Faster, Smarter and More Personal**. Policy Exchange
- Zamith, R. & S. C. Lewis. (2015). Content Analysis and the Algorithmic Coder: What Computational Social Science Means for Traditional Modes of Media Analysis. **Annals, AAPSS**, **659**, 307-318.
- Zikopoulos, P., D. DeRoos., K. Parasuraman., T. Deutsch., D. Corrigan., & J. Giles. (2013). **Harness the Power of Big Data: The IBM big data platform**. McGraw-Hill.

附錄

附錄一、政府為民服務網站深度訪談結果摘要

（一）愛料理（iCook）食譜社群網

1. 網站營運目標：

- (1) 設立背景：如果可以的話當然也想做一個像是Facebook或是Line這樣一個大的產品，不過回歸現實，我們當然在資金還有經驗上面也比較不足，所以我們想說可以做一個跟在地文化比較相關的。看看去餐廳的很多人做了嘛，所以就覺得說到吃的，在家吃可能是一個很好的方向。所以就成立了icook網站，還有手機的這些應用程式。我覺得我們運氣很好，剛好有搭上兩波列車，一個是Facebook，像我們現在FB上面的粉絲團有120多萬個粉絲，那是因為我們在比較前期就比其他競爭者或是網路從業人員更早的投入。那另外一個我們搭上的列車就是APP，那APP就是在這幾年就是如日中天，我們很早很早就開始做的話，其實就是上面卡個位啊，那卡到位之後啊，其實這樣兩三年下來其實我們也累積了200多萬的下載數量這樣子。（A1，P2：18-27）
- (2) 資金來源：資金來源的話主要是，就第一年的話XXX有投資，第二年的話就是得到日本的XXX投資²⁵。（A1，P2：28-30）
- (3) 經營目標：我認為我們在網路上面的影響力，大概我覺得這是一個比較合理的位置，再往前大概前進不了多遠。因為終究還是有一個分眾大小的問題，……所以我認為我們在這個領域已經算是很Peak

²⁵ 資金來源主要為國內個人投資以及日本的創投公司投資，考量商業競爭的保密原則，故將詳細資料改以 XXX 代替。

很頂端，所以說在今年的話，我們就是著去拓展更多跨領域的，比如說我們今年出了兩三本書，在食譜這個領域上面也都是做得很好。（A1，P3：2-9）（B1：請問會員數目前有多少？）有註冊在案的會員數是有破百萬的，那沒註冊的話，每個月不重複訪客可以到500萬。（A1，P3：27-28）

- (4) 網站經營方式：（B1：請問目前的技術團隊？）內聘，人數約莫就是10個左右。（A1，P2：31）（B1：主用採用的技術是？）網站技術主要是用Ruby on Rails。然後配合我們整個伺服器都是放在亞馬遜的平台上，當然還會混搭一些小的東西，像是前端就有用JavaScript，這些當然都是要有的。各平台的話當然就是用他們原生的東西，像是APP的話就用像是IOS之類的。（A1，P3：33-P4：2）。

2. 網站績效要求：

- (1) 財務資助者要求：網路世代是一個Start-up，那他們就是會要求要有成長，……所以成長是很重要的……那績效其實就是成長，對於我們的網站而言，訪客數當然是一個重要的成長指標。（A1，P4：3-8）在網路時代是贏者通拿。就是最大的那一家拿全部，大家只會用Google，只會用Youtube。所以他們擔心的其實是你們不能成為那個贏者……所以他們寧願你不會被拉下來之後再去思考獲利的问题……對於投資者來說，他們要的是高額的return，不是短期的小錢……要投資就是要買風險。所以投資者就是可以承擔3到5年沒有return的風險。（A1，P4：24-33）
- (2) 經營管理重點：我們的使用者其實分成很多環節，網路上面的這些食譜其實是由少數的使用者分享的，就跟FB一樣，Content Provider其實是少的，這個比重在我們的網站上面是非常懸殊的，幾乎是九成跟一成。前幾年對我們來說當然是思考如何養大我們的基數，到了今年我們開始思考如何把那個九成的人變成那個一成，變成願意去分享，……我們的一個新的方向，就是如何去增加一個網友的互動性。（A1，P4：8-21）最終你要做的事情都一樣，就是你把目

標訂出來，指標訂出來，然後切不同的source，比較好的source加重投資，然後你會發現有些source不管在那些地方就是會比較差，那你可能就要去思考說是不是網站規劃出了什麼問題。（A1，P15：28-30）

3.網站服務對象

- (1)服務對象特質：年齡層也許大家可以理解就是25到35歲的女性居多，可是中間就是性質不太一樣。我們現在也做很多使用者訪談，那25歲到30歲這一段其實會做菜的就是for fun，覺得很有趣啊，在外租屋的周末找朋友一起來玩一下這樣。過了30歲以後其實是因為結婚了，家裡需要（煮飯），而這個就不是core的使用族群。比較contribution的使用者就很兩極喔，分布在20到25歲以及35歲以上的。啊中間的就結婚啦生小孩，生小孩哪有空在那邊寫食譜啊。所以說婚前婚後是我們很重要的讀者這樣。那男生的部分看起來都是for fun啦，年紀大的或是年紀小的都是。（A1，P5：4-11）

4.網站使用者行為分析方法

- (1)分析架構：通常在Internet Start-up上面，會有一個term叫做「AARRR」。這個代表的是使用者從進入這個網站到最後回饋的這一段過程中，他可能會面對的每一個環節。好，所以我們從第一個A開始，第一個A是Acquisition，也就是如何去得到這個使用者。第二個A是Activation，怎麼讓他真正變成我們的使用者，以我們的情境來說，Acquisition可能是透過粉絲團或是廣告來獲取或知道我們的網站，那在第二個A，他是怎麼樣註冊成你的會員，或是他是怎麼involve進你的網站，又或許是不一定要註冊成會員，我們可能會對他們的使用有一個期待。我們可能期待他的停留時間，或是期待他能翻超過三頁，我們就相信這是成功的，成功的把使用者留在這裡。第三個R其實是最困難的，叫做Retention，怎麼樣子讓他可以常常地回來，那我不知道以E-Government需不需要常常回來，這個我覺得每個Buesinese不太一樣。以我們的情況來講的話，當然我們

是希望使用者是可以回流的，最好是針對那些Contributer，最好他們是可以常常回來的。最後兩個R分別是Referral，他怎麼樣把這個東西介紹給其他人，以及最後一個R是Revenue，怎麼樣在這最後產生一些獲利。所以AARRR他是一個有關聯的，前面三個Acquisition、Activation、Retention是接在一起的，那最後Referral跟Revenue則是前面三個發生之後接下來的事情。（A1，P5：14-31）（B1：對使用者屬性的瞭解除了深度訪談外，是否有其他方法？）其實在很micro的數據主要還是靠FB，因為如果你有左FB粉絲團的話，他會告訴你producer他的profile大概會在這裡這樣。（A1，P6：26-27）

- (2) 分析重點：在這每一個環節當中，我們要的指標是什麼？像是在Activation方面，我們認為我們的食譜搜尋功能是很好的，因為我們想做的是解決人們生命中會遇到的問題，那搜尋食譜是我們observe最多人使用我們網站的方式，也是我們認為我們可以創造價值的地方。所以當我們用各種方法Acquisition使用者之後，我們如何使他們嘗試看看我們很棒的搜尋功能，這不管是在UI上面的調整，或是給他們很明確的說明給一個示範。那之後這些人在使用之後，有沒有辦法留下來，那留下來可能用各種方式啊，像是說你可以註冊一個會員啊，我們每一周會把精選的食譜或是新的東西告訴你，這個過程當中就有可能再把他釣回來，或者是最差的我們可能會請他們在FB粉絲團上面按個讚，這樣至少他們以後會持續接收到我們的東西，人才有辦法再帶回來。在這樣的一個大框架背後，工具我們反而覺得是比較次要的，就是說我們這個AARRR個別的分析是什麼，可能目標都定義出來之後，那不管說我們要用GA（Google Analytics）或者是Log分析，其實就是只是輔助把這樣子的數據算出來。（A1，P5：32-36；P6：1-9）

- (3) 決策經驗：以Acquisition而言，我們自己的經驗告訴我們還是透過FB啊，那像現在的話我們也有在是用Line啊或是用其他的工具之類的。Referral其實是難在你要怎麼去設計一些誘因啊。像是大家知道Dropbox嘛，你邀請一個朋友可以加幾G這樣，他們就是在這個環節

上面做得很好這樣。但坦白講，像是愛料理這樣比較偏媒體性質的網站上面，其實能做的事情相當有限。除非說就是我們就是拿錢去發啊，不然你要怎麼鼓勵他花時間註冊？很難很難。但是我們會有一些活動或是比賽，就是針對食譜的一些作者，他們用食譜去參加比賽，然後會有一些獎品，其中得獎的依據其中一項，就是他們的食譜被多少人看過啊，他可能就會去拉票，請大家去愛料理幫他看這個食譜，幫忙按個讚之類的。（A1，P6：11-23）

- (4) 資料蒐集方法：（B1：對使用者屬性的瞭解除了深度訪談外，是否有其他方法？）其實在很micro的數據主要還是靠FB，因為如果你有左FB粉絲團的話，他會告訴你producer他的profile大概會在這裡這樣。（A1，P6：26-27）（B1：有發網路問卷嗎？）對，同一時間也去搜尋資訊這樣。我覺得網路問卷要做得好，不能是太明確的問卷，就是你要綁一些點。比方說像我們的點可能是說，「喜歡看我們影片的這一些朋友們，可不可以告訴我們接下來要拍什麼？」（B1：問卷發放方式？）註冊後可能會有使用者的E-mail，那你可以針對這些人在發。我想可能我們也有做一點點的E-commerce，那可能還會在網站上面有消費行為的人來做發送。（A1，P7：1-10）（B1：Google Analytics這一個工具在協助你們去了解使用者行為，他大概扮演了哪樣的角色？）剛才講到的AARRR全部都可以在GA上面去實做出來。（A1，P7：16-17）（B1：所以使用者分析還是仰賴GA嗎？）對，除非說當初沒埋到的，後來發現說不行這東西很重要，就要再把log翻出來再去做。（A1，P12：28-29）
- (5) GA實作經驗：在**Acquisition**裡面，你會看到一個叫做，New Visitor跟Returning User，你就可以看到那些使用者是新Acquire進來的，GA也有提供一個UTM的Tag機制，UTM的部分，譬如你有Line，然後有加一些粉絲團，一些商業的粉絲團貼的網址你可以發現第一個他一定是縮網址，你去看都會發現後面會加一個UTM source，UTM media是什麼。就是當使用者使用這些連結的時候，他在GA裡面就會標記說這個使用者是從哪裡來的。剛剛說的AARRR只是一個分

析方式，那分析出來之後，你就要切各種channel去try，比方說Line進來的AARRR狀況是什麼，電子報進來的AARRR狀況是什麼？當你發現某一個區塊的直視特別好的，你就可以去加重那一塊的比重。（A1，P7：17-26）（B2：如果你沒有特別作tagging的話，是不是就會廣泛的歸類成social？）對，但是他那個是看你的HTTP refer，如果他是從FB來的他就會掛social，如果他是從Google或是從Yahoo來的他就會掛search。（A1，P7：29-30）**Activation**其實跟後面的這些不論是**Revenue**、**Referrl**這些都在GA裡面有一個重要的功能叫做Goal Setting設定目標。那設定目標的話一班不外乎三種方式去設定，一個是說他有沒有照你的想法走到某一個頁面去，舉例來說，跟我們註冊網站時，頁面會有一個表單，要送出的時候我們會謝謝他並會盡快跟他聯絡，假設我是政府網站，我可能會希望他們透過這個表單去問問題之類的，所以一個好的標單應該是可以讓他們成功地去送出這個表單，在網頁上面我可能會設計說在最後一頁會有一個感謝的頁面，回到GA的話，他應該要先進到表單頁面然後到感謝頁面。代表他有成功送出這個表單，GA旁邊會看到有一個goal可以選，設定完成之後你就可以看到不同channel的進來這個表單到ending page，你就可以看到完成的人比率有多少。（A1，P8：8-17）我覺得GA的重點是網址，裡面的內容是還好，他主要是認網址。只要網址是分開的，他在GA裡面就會是separate的。（A1，P9：19-20）

- (6) 使用GA的限制：GA要做大數據，好像是要做一個很複雜的運算，那為甚麼它可以很快做出來呢？是因為這邊要注意的是它有data sampling的問題，如果沒有額外付錢的話，GA他最多就是ssample 50萬人這樣。這個50萬筆其實滿容易遇到的，比如說你設定五天，那可能一天十萬人的話就到了。（A1，P12：11-15）付費版本也就只是從50萬到75萬而已，所以一樣。大數據沒這麼簡單，所以.....Log就是要拿來補這種洞的，像是這些很micro data的這些，就是要來補這些洞。（A1，P12：21-23）（B1：所以除了GA外，你們也有在做網站log分析？）有，但大多是在做網頁DOM這一類

的，就是主要是網站技術面上的問題，多數的東西其實大部分用GA goal setting就可以。（A1，P12：25-26）

（二）財政部電子申報繳稅服務網

1.網站營運目標：

- (1) 設立背景：86年的時候配合行政院推動電子化政府的政策.....提供給納稅義務人多元化的一個申報的管道.....，所以我們那時候把叫做網際網路報繳稅作業列為主要的便民服務措施。.....在第一階段的時候我們就是針對五個主要的稅務項目開始，綜所稅、營所稅還有各類所得、營業稅還有營所稅暫繳。.... 在第二期規劃到第三期規劃之間，我們又增加很多稅目了，增加了九個稅目，像是遺贈稅、菸酒、稅額試算、外僑這些與國際接軌的服務....。我們一路走來都是配合政府那邊的電子化政策，也是慢慢的在循序漸進.....（A3，P2：15-30）。
- (2) 資金來源：我們的資金呢就是正常的一個公預算編列（A3，P2：31）。
- (3) 經營目標：我們的目標是希望說可以提供民眾一個方便的申報管道，還有資訊安全，....提供給民眾全天候24小時的服務，.....當然也是配合政府的一個政策，多用網路少走馬路，節省了不少社會成本這樣（A3，P3：12-16）。
- (4) 網站經營方式：成立這個案子是配合我們政府的政策，然後整體委外。（A3，P3：11）（B1：請問委託哪間公司？）關貿網路公司（A2，P3：27）。（B1：在經營這些網站和報稅軟體，大概是使用什麼樣的技術？）網站的部分比較單純，大部分是靜態網路，所以在我們機房那邊就是web server，AP server還有DB.....。所以在網頁那邊看的話，大部分的網路就是html的。那個其實也就沒什麼技術，就是一個網頁。（A4，P3：29-31～P4：1-7）。（B2：稍微釐清一下，

你們的使用率是軟體本身，但是網站本身就是比較少啦，就是方便大家可以下載軟體而已。）對沒錯（A4，P5：2）。

2.網站績效要求：

- (1)財務資助者要求：（B1：國會或者上級長官對網站經營有無績效要求？）對這一點到沒有特別要求或是訊息（A3，P4：23）。（B1：委外合約有無績效目標要求？）我們的績效要求大部分是在民眾使用上不會有效能上的問題啦，.....目前我們要求廠商就是在申報的時候要提供一個比較穩定的一個服務這樣子（A3，P4：26-29）。
- (2)經營管理重點：我們的這個網站主要是放置我們報稅的軟體，那其實除了下載這個報稅的資訊外，我們也可以去查詢報完稅之後也可以去查訊是否有成功報稅，基本上我們服務的項目主要是在這裡（A3，P3：1-3）。（B1：所以這個網站目前的管理是誰在負責？）也是我們在負責，需求的部分則是由中心這邊來提（A4，P7：19-20）。

3.網站服務對象

- (1)服務對象特質：就是納稅義務人啊，.....有一些稅目是給國人報稅，又有一些是法人，公司行號。那報稅的自然人的話，可能就是本國國人嘛，後來又因為有外僑，所以服務的對象又擴增了。還有有憑證的人，像是金融憑證也可以透過網路來申報。另外還有公司行號使用的工商憑證，那些其實都是可以用的（A3，P6：13-17）。

4.網站使用者行為分析方法

- (1)分析架構：我們現在可能是用人工的方式，用mail把資訊給我們部裡的長官，譬如說到今天有多少人用這個申報.....（A3，P7：16-17）。一方面來說就是我們的經驗嘛，比如說像是稅法更動嘛，或是像是今年APP的上線，.....我們可能事先知道這些問題，民眾可能就不需要再打這些電話了（A3，P7：26-30）。（B1：所以你們主要是從客服專線蒐集民眾問題嗎？）也不一定，還有Email（A2，P7：32）。客服專線是全年的，就算不是申報月份還是會有陸陸續續一些電話

進來，那客服中心會去統計資訊，然後再把這些資訊給我們。那另外還有國稅局，國稅局那邊也會有民眾提問，那我們會把資訊彙整起來後，把程式更新後就把FAQ都一起丟上去（A4，P8：4-7）。另一方面我剛剛提到的那個HOT的地方，是我們中心整理的。我們也會去蒐集各網站裡頭大家提的一些問題，我們這邊有人會去匯集這些資料再去做整理（A2，P8：10-12）。

（三）內政部不動產實價登錄網

1.網站營運目標：

- (1) 設立背景：配合不動產交易透明化政策，自101年10月上線。有關不動產實價的運作……大概有四道程序，第一個就是接受交易人的呈報資料，第二個是作檢核，第三個則是結合我們資料庫的資料，把它變成比較完整的資料，最後給人民去做查詢。然後實價登錄量，從100年開始大概是117萬件，每個月網站查詢的部分則是大概130萬左右。opendata下載數大概一個月8、9000筆（A5，P2：14-24）。
- (2) 資金來源：公務預算，設置初期（100年）以標餘款委請廠商開發，自102年開始至104年，內政部在相關計畫編列預算辦理功能增修及系統維護。
- (3) 網站經營方式：採用委外方式，資料建置系統是康大資訊公司，查詢網站是由光特資訊公司建置（A5，P2：25-31）。（B1：網站維護費一年大概是多少？）我們有兩個系統嘛，所以申報的那個平均差不多200萬，查詢網差不多100萬。這個預算編到105年（A5，P4：1-2）。
- (4) 經營目標：達成不動產交易透明化政策目標，自101年至今已提供110餘萬筆不動產交易實價資訊供各界查詢（引自受訪者提供資料）。

2.網站績效要求：

- (1) 財務資助者要求：網站的績效要求為網站瀏覽人次和 open data 下載次數，目前每月瀏覽人次約 100 萬人次，open data 下載次數每月約 7000~8000 次（引自受訪者提供資料）。
- (2) 經營管理重點：我們那個有關不動產實價的運作，大概有分兩個系統。第一個是資料的建置系統，也就是交易人交易的時候要去申報的系統，也就是在交易的後時候要去把交易的價值用這個系統建置上去。那資料上去之後，我們地政事務所就會上去看他所呈報的資料合不合理，我們有一個檢核機制。那之後才會把我們這邊有的資料，就是地質資料補上去，之後才會產說所謂的實價登錄資料。所以大概有四道程序，第一個就是接受交易人的呈報資料，第二個是作檢核，第三個則是結合我們資料庫的資料，把它變成比較完整的資料，最後給人民去做查詢（A5，P2：14-21）。（檢核機制）目前還是以人工判斷，就是說看他有沒有異常的一個交易（A6，P3：16）。申報端因為我們有提供一個客服電話，也可以解決一些在申報過程的疑義（A6，P4：3-4）。有關操作的東西當然是放在操作的友善度啦，所以才要有客服....。在技術這邊，要用憑證會是一個很大的關卡（A5，P4：8-10）。
- (3) 內部的績效要求：對我們自己來說，最 normal 的還是剛剛提到的那兩個，就是友善度跟使用量。其實對我們來講內部最大的挑戰還是在資料分析。他資料分析可能有兩個方向，就是一個 Big Data 的方向，那另外一個可能就是決策分析的方向。那從 Big Data 可能就是從很多資料去建立分析模型嘛，那對我們這邊還有很多行政院、財政部這邊可能比較偏決策分析這個方向，從確定的關係裡面，我們能不能從裡面去做一些決策（A5，P4：21-26）。

3. 網站服務對象

- (1) 服務對象特質：查詢網站主要服務對象為對不動產交易行情有興趣或需求之業者或人士（受訪者提供資料）。（資料建置網站）都

是法律規定的人，第一順位是地政士，然後是交易人，還有不動產經理人。只是說在法律規範上，地政士是主體啦（A5，P3：6-7）。

4.網站使用者行為分析方法

- (1)分析架構：我們明年有編一些錢準備來做資料分析這一塊。那Big Data不會做，也不是和行政機關來做這個，那應該由學術機關來做比較適合（A5，P4：28-29）。
- (2)分析重點：現在對我們來說最困難的是我們沒有辦法做分析，現在open data有一個議題，就是它的效益是怎樣？但是open data的先天條件是說，我不能去記錄你是什麼人來用我的這個，所以這是最大的限制，因為我要記錄你的行為，就是我需要你是我的會員嘛，但問題是我們根本沒有這個機制存在，所以我很難去分析使用者行為（A5，P5：2-7）。（B1：請問有蒐集使用者行為的資料嗎？）。
- (3)決策經驗：沒有會員的話，它會很片段。沒有會員就是沒有資料可以分析，那就要回到傳統面的蒐集，但那是很單向的，你就很難有數據可以去分析（A5，P7：13-14）。
- (4)資料蒐集方法：老師有提到的那個服務品質獎，那次評分的老師委員有提到說要做問卷（網路問卷），然後我們有做，期間大概是兩三個月....針對他造訪網站的行為做一些使用上的提問（A6，P5：9-14）

GA實作經驗：廠商有處理，但實際業務未應用（受訪者提供資料）。

（四）法務部全國法規資料庫

1.網站營運目標：

- (1)設立背景：全國法規資料庫的建置是在民國86年，配合行政院將電子法規納入電子化網路化中程推動計畫，由法務部這邊在同年九月召集行政院、立法院相關單位還有學術單位組成一個全國法規資料

庫的推動委員會，開始做一些推動的規劃。那這個網站就在90年的四月正式上線。提供全民一個即時、完整還有兼顧專業與親和力的一個電子法規資料庫（A7，P2：13-18）。

- (2) 資金來源：整個預算來源就是公務預算。我們現在的預算規模，因為我們同時有包含本部主管法規的網站以及行政院下面的主管法規共用系統和檢察官的檢索系統，那每年都會有個維護費用差不多是九百多萬，因為它不只是全國法規資料庫在用，那現在是一個給23個機關主管法規共用的系統，那個也包含在裡面。還有一個是本部業務通報的系統，所以是整個包起來的（A7，P2：18-24）。
- (3) 經營目標：在全國法規資料庫入口網有兩個設置的目的，當然就是設置一個即時完整又專業的法規資料庫，那以免各機關投入相同的資源去這裡，就會不知道要看哪一個，所以就由官方去設置一個entry point。那另外一個就是建立一個親和屬於全民的法規資料庫，希望透過這個平台去拉近全民法律生活的感情（A7，P2：29-35）。
- (4) 網站經營方式：（B1：請問目前的技術團隊？）法務部經營管理，並經由委外服務方式，委託兼具法律及資訊專業之廠商維運 目前是交給法源（A7，P2：27）。

2. 網站績效要求：

- (1) 財務資助者要求：國會那邊對我們呈報的數據沒有太大的意見（A7，P5：29-30）。
- (2) 經營管理重點：提供更多之使用者，能以親和易用的方式取得及時、完整、專業之法律資訊與常識，以深化全民之法治素養，並拉近民眾與法律之感情（受訪者提供資料）。使用者分析可協助設計親和易用的網站。目前提供服務信箱及時蒐集和處理使用者建議，另處理自行政院院長信箱、各部會首長信箱及各機關函轉之問題。每年會辦理使用者問卷調查，以及使用GA進行使用者查找資料之統計……（受訪者提供資料）。

- (3) 目前經營成就：目前全國法規資料庫的一個使用情形，累積的使用次數已經到達2億2500萬人次，這是截至今年8月30日之前。那1月到8月的月平均點閱數，大概一個月有500多萬點閱人次。那電子報的訂閱戶則是14萬5千多人次。那全國法規資料庫的法規筆數，可以參考手上的資料，這是統計至今年6月底為止，那法規總共有1萬1千480筆的法規，那英譯的法規有2181筆，司法判決有1萬1千筆，這是指法規，比方說刑法民法都各算一筆。那如果以條數來看，現行的法規是從90年開始蒐集，那現行法規有18萬6千餘條法規，那英譯法規目前有6萬5千多（A7，P3：4-11）。目前我們的通俗用語統計到8/7為止，目前有1萬8千270筆資料，對應的法律專門用語詞彙建置則有27393筆……另有有55則智慧查找的案例（A7，P3：26-31）。

3.網站服務對象

- (1) 服務對象特質：一個是給專業人士使用，就是法律專業人士或是公務人員使用；另外一個當然就是給全民使用……民眾在全國法規資料庫這邊找到的資料，我們可以確定說是最新的也是最完整即時的一個資料（A7，P2：33-35；P3：1-4）。

4.網站使用者行為分析方法

- (1) 分析重點：全國法規除了提供法規的關鍵字，讓專業人士去檢索查詢之外，另外有一個比較特別的地方，它提供對於法律用語比較不懂或陌生的民眾一個查詢的方式，另外一個則是法規檢索的整合，如果大家有進入我們的法規資料庫看，裡面有幾個檢索的方式，你只要進去打某個關鍵字，或者是部分的關鍵字，我們就會去把它導引出來，就可以找到需要的法規。那還有一個是熱門詞彙的法規查詢，以及關鍵字詞彙推薦的法規查詢，如果找出來是相關、接近的法規，我們會在旁邊推薦這樣。另外有一個叫做通俗用語，一般對法律用語不是那麼熟悉的民眾，……一般民眾可能無法這麼precisely地找到它所需要的法條或法規，比方說他如果打小偷，但

他不知道要打竊盜。或者是說他打小三、酒駕，.....現在有建立這個通俗用語之後，我們會把這些用語的相關法條列出來。那即便他沒有找到，假設出現新的通俗用語，我們會用GA把他抓出來，我們會檢討說這些通俗用語要不要建立條目讓民眾去做檢索這樣子。.....一般比較沒有法律專業素養的民眾，透過這樣的方式可以讓他們比較容易接近這樣子的資料庫。另外我們還有一個功能叫做智慧查找，一般我們法律專業人員或是公務人員，都是先去找法條，再來判斷說跟情境是不是符合構成要件，但因為一般民眾不知道哪個法條更適合這個情境，所以我們就把這個使用情境顛倒過來，從情境去導引他找到法條。目前已經有55則智慧查找的案例，它有一些插圖跟分類讓大家比較容易去找.....（A7，P3：11-31）。（B1：在網站介面友善度方面有何改善？）我們現在使用的還是比較傳統的這種三欄式的，一些歷史因素啦，專業人士在使用這個網站已經有他的習慣，那所以我們也沒有冒然的去做修改....（A7，P9：27-29）。這個網站在使用上其實已經頗獲好評了，在台灣只要你搜尋法條的話，第一名其實應該都是全國法規資料庫。那進一步我們是想說怎麼去improve這個網站的價值，就是說如何在法治教育這一塊有所增進..... 因為行政院在285號院會有函示說法治教育已經是院頒的命令，教育部這邊要配合，就是說學生在不同的階段應該要接受多少的法治教育，那這就是未來我們全國法規資料庫大概會進行的作業（A7，P10：2-14）。

- (2) 資料蒐集方法及決策經驗：透過各信箱及問卷調查可得知使用者對網站功能的反應，GA即使用者查找資料之統計，則可幫助調整網站新增內容，例如智慧查找點閱率分布；增加通俗用語（受訪者提供資料）。通俗用語的產生是依照法律用語本身它比較有類似的地方。比方說一般民眾可能只知道酒駕，這種比較口語的，而不會去知道說它背後的法律條文。那我們明顯知道酒駕是指這種的話，我們就會去把他們對應起來，就是說他如果去搜尋酒駕，他就可以在資料庫裡面找到對應的條目。一開始像是小偷拉、小三這種用語是

不會出現在法規裡面的，小三可能跟通姦罪有關，像這種就會去用人工把它對應起來。另外一種是，因為我們的資料量是滿龐大的，那GA會把有找到的跟沒找它的分開，那針對沒找到的我們就會去做分析，因為這可能是新找到的通俗用語，那我們會跟法律專業人員討論適不適合，如果適合，我們就會把它對應起來（A7，P4：4-13）。另外，每年我們有辦一個教育訓練，然後會做問卷調查，上面有回饋表可以給他們去做建議，就是可能有哪些通俗用語是上面還找不到的，我們也會透過這樣子的管道去知道說民間還有哪些通俗用語是會用到的（A7，P4：16-19）。每個月都有跟廠商開一個專案會議，那會針對這一個月以來營運的方式做一個報告，那假如說有一些建議或是回饋的事項，就可以透過剛剛那些管道一併報告，那如果有額外的需求，我們這邊會要求把它放上去。如果有緊急的事件，我們這邊也隨時可以跟廠商那邊做一個修改。那因為我們的瀏覽量在中央機關來說算是很高的，那可能部長有一些訊息，可能也會放在上面做一個公布，所以如果上面有一些要求的話，我們也會要求廠商盡快做一些修改（A7，P7：18-24）。

- (3) GA實作經驗：第一個是GA的用法，就是在網頁上，哪一個link是點選率最高的，它可能代表是最被需要或是比較顯眼比較上面的位置，那邊的功能可能是要加強的，另外一個就是使用的location這邊，目前大概都還是國內在使用。針對這一個點閱率的部分，我們的這一個智慧查找的方式，因為它有很多案例，所以我們可以知道說哪一些案例是比較多人點的，.....我們會做分類，假設我們分成十類，假設某幾類像是勞工問題或是家庭問題，是點閱率比較多的，那我們就會去現有的案例裡面看說是不是足夠，有沒有再增加的空間，這邊就有助於我們把資源放在我們應該要關注的位置（A8，P7：7-15）。針對一般民眾，有一些新的通俗用語我們會透過GA去cue下來之後來做分析使用。那剛也有提到就是智慧查找，....我們會針對各種case的使用量去做分析，....這就是我們在GA上找到的一些

數據，幫助我們做一些判讀這樣，那對於我們後續的服務也會有幫助（A7，P8：16-26）。

（五）國發會 e 管家 Plus 網站

1. 網站營運目標：

- (1) 設立背景：E管家這個服務，是國發會那邊有個第四階段電子化政府服務推展開來的，那有一項就是要做便民服務的應用.....。八年前他們開始在談的時候，是有三個脈絡。一個就是針對公務人員有一個e公務服務，方便他們公務人員在內部做一個方便的溝通。另外一個是e幫手，幫助企業創業做一些串聯，那針對民眾的部份的話，就是我們e管家。.....e管家的部分，大概這八年來也經歷過一兩次比較大的轉變，早期的話比較單純，就是國發會居中協調各個部會以及各級地方政府，提供一個個人化訂閱的服務，那可能包括個人化停車服務、繳稅還有藝文活動這些，按照你的興趣做的訂閱，另外一個走向的話就是全民快訊，就是提供一個不是那麼針對個人的資訊提供給你，服務規劃就以這兩個為主。我們接手e管家這個案子之後，會發現兩個很大的問題，就是跟民眾的需求會比較遠一點...他在會員的部分其實綁訂了e政府的帳號，.....e政府這樣一個帳密走遍天下的概念需要改變，實務上面民眾在申請這個長號的時候會覺得有一點點麻煩....所以在我們這邊的規劃是把open ID拉進來。然後在服務面也可以做更貼近他們的需求，所以除了原先規劃的需求之外，....我們希望給民眾的一個體驗是，你不需要知道這些知識背景、關鍵字，你也可以找到你要的資訊。所以我們做了一個就是福利自己查的服務，你不用思考太多，你只要按照自己的需求點點點，你可以很輕鬆地蒐集到適合自己的資訊（A9，P2：12-36；P3：1-22）。

- (2) 資金來源：國發會公務預算（A9，P5：22）。

- (3) 經營目標：最重要的當然是希望這個東西可以符合到民眾需求，然後也希望有黏著度，而不要只是會員數增加，可是實際用的人很少（A9，P4：12-13）。（委託機關有績效指標要求嗎？）
- (4) 網站經營方式：(B1：技術團隊是自己還是再做分包？) 要看開發的東西，主要當然是我們這邊。那額外可能有些活動要架額外的網站，那一塊可能會切出去做（A9，P4：17-18）。（B1：如何和其他縣市資料介接？）最直接的辦法就是我用技術介接的方式來做，那也許這些東西直接跟你們的網站直接接就可以了...。但.....不是每一個縣市政府都有一個成熟的平台....，從實務面上來看，最可行的方式是，只要我們這邊還有資源，只要你告訴我你們的網站上面已經有類似的資料了，我們就可以幫你們建一個基礎的資料...（B1：大部分由你們人工看到之後就把它放進去？）對.....介接的話，反而是不可行的，它沒有經費去開發這個東西（A9，P6：7-17）。

2. 網站績效要求：

- (1) 財務資助者要求：去年的話，在專案上面有要求幾個指標，一個是會員成長、黏著度還有一個APP的下載數（A9，P4：15-16）。（B1：請問黏著度如何計算？）看登入的頻率，然後看他每個功能使用上點擊的次數（A9，P4：18）。（B1：有包含登入時間嗎？）也有（A9，P4：20）。
- (2) 經營管理重點：（B1：國發會關心什麼？）民眾的使用度有多高，就是真的有沒有幫上民眾的忙....另外以介接來說的話，他們也會去關心說機關的接受度有多高，配合度有多高（A9，P5：25-28）。我們會定期兩個禮拜跟國發會召開專案會議，主要就是跟長官就專案執行的狀況做一個回顧（A9，P8：2-3）。其實這個專案有點坎坷，換了好幾個承辦的長官，那每一個長官都會有對這個專案一點點出入的想像，那至少每一個長官近來最關心的就是使用情況跟與機關端的介接積不積極（A9，P8：6-8）。

3. 網站服務對象

- (1) 服務對象特質：一般民眾和公務機關同仁（訪談者於訪談過程中歸納）。

4. 網站使用者行為分析方法

- (1) 分析架構：一個是我們系統自己紀錄的Log，我們自己記錄下來然後去做使用者行為分析，那另外一個就是最普遍使用的GA，可以看到使用者的流量，跟我們期待的目標轉換這樣子（A9，P6：26-28）。
- (2) 分析重點：Log部分的話，依照我們的專案執行，都會有提供日報跟月報，那日報跟月報都會針對我們使用者的情況去做分析。GA的部分主要是我們Web的使用情況，你可以知道網站的流量、停留的時間，跳出率還有使用者能不能達到目標的轉換率（A9，P7：16-17）。基本上做完識別化之後，我們會在我們的後台針對服務的主項包括全民快訊、福利自己查跟生活地圖去抓他們被使用的情形，然後後臺你可以看到的是，以全民快訊來說，我們有分九個頻道，那九個頻道大概佔有比率的分布是多少，還有一些福利自己查因為有個人化的設定，那在這個項目上面，我們就可以知道男女、年齡分布等狀況，因為有個人資料的關係，所以我們可以比較去抓到使用者的輪廓（A9，P7：7-12）。
- (3) 決策經驗：（B1：日報內容如何？）會比較routine的報表。比如說會員數、登入數、APP的下載量、各管道會員增加的比率等等，然後還有我們的訂閱服務....我們會追蹤使用的情況，最近一年有登入或訂閱等等（A9，P7：20-29）。（B1：日報和月報差在時間趨勢的長度不同嗎？）對（A9，P7：32）。我們也會有一個雙週報，包括系統面以及社群推廣面等等項目執行的情況。那除此之外也會把客服端回饋的資訊一併作報告。那如果在過程中發現有不理想的部分，我們會去檢討作法，再往上做呈報（A9，P8：3-6）。
- (4) 資料蒐集方法：自行開發的log分析工具以及GA（訪問者歸納）

- (5) GA實作經驗：在GA的部分，主要的成分是在參考，它所能撈出來的資料對我們最後要產出的報表功用會比較小一點。舉例來說，目標轉換這一塊，過去因為只能抓到Web的資料，APP是沒辦法抓到，所以在過去我們是沒有非常著重在這一塊上面，那反倒是未來我們在考慮到未來服務的發展，這些資料收容在Web端的時候，我們會開始去累積一些可以參考的資料質，所以我們現在會開始去看說每個項目被點擊的情況，比如說它點進全民快訊這個內容的情況是怎麼樣，為了是說後續我們在APP的功能上要慢慢去做收斂的時候，希望把流量都導到我們的Web時，可以有一個參考值（A9，P7：1-3）。

政府巨量資料分析與政策端應用效能提升之研析

附錄二、輿情平台群眾分析結果表

一、文章結果之觀察值處理摘要表

觀察值處理摘要

	觀察值					
	有效		遺漏		總計	
	N	百分比	N	百分比	N	百分比
群眾判定（全部） *電子治理研究中心A1(O)	321	91.7%	29	8.3%	350	100.0%
群眾判定（全部） *電子治理研究中心A2(X)	303	86.6%	47	13.4%	350	100.0%
群眾判定（全部） *國發會B1(O)	310	88.6%	40	11.4%	350	100.0%
群眾判定（全部） *國發會B2(X)	311	88.9%	39	11.1%	350	100.0%

資料來源：本計畫。

二、群眾判定 * 電子治理研究中心 A1

交叉表

			電子治理研究中心A1				總計
			反對	無法判斷	支持	中立	
群眾判定 反對	計數		92	1	0	9	102
（全部）	All結果2 內的		90.2%	1.0%	0.0%	8.8%	100.0%
	%						

政府巨量資料分析與政策端應用效能提升之研析

	無法判斷	Test1_2 內的 %	92.0%	6.7%	0.0%	12.3%	31.8%
		佔總計的百分比	28.7%	0.3%	0.0%	2.8%	31.8%
	支持	計數	1	7	0	4	12
		All結果2 內的 %	8.3%	58.3%	0.0%	33.3%	100.0%
		Test1_2 內的 %	1.0%	46.7%	0.0%	5.5%	3.7%
		佔總計的百分比	0.3%	2.2%	0.0%	1.2%	3.7%
	中立	計數	4	1	131	9	145
		All結果2 內的 %	2.8%	0.7%	90.3%	6.2%	100.0%
		Test1_2 內的 %	4.0%	6.7%	98.5%	12.3%	45.2%
		佔總計的百分比	1.2%	0.3%	40.8%	2.8%	45.2%
	總計	計數	3	6	2	51	62
		All結果2 內的 %	4.8%	9.7%	3.2%	82.3%	100.0%
		Test1_2 內的 %	3.0%	40.0%	1.5%	69.9%	19.3%
		佔總計的百分比	0.9%	1.9%	0.6%	15.9%	19.3%
	總計	計數	100	15	133	73	321
		All結果2 內的 %	31.2%	4.7%	41.4%	22.7%	100.0%
		Test1_2 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
		佔總計的百分比	31.2%	4.7%	41.4%	22.7%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
--	----	---------------------	-------------------	--------

附錄二、輿情平台群眾分析結果表

合約的測量 卡帕 (Kappa)	.814	.027	21.836	.000
有效觀察值個數	321			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

三、 群眾判定 * 電子治理研究中心 A2

交叉表

			電子治理研究中心 A2				總計
			反對	無法判斷	支持	中立	
群眾判定 反對 (全部)	計數		89	3	4	5	101
	All結果2 內的 %		88.1%	3.0%	4.0%	5.0%	100.0%
	Test2_2 內的 %		94.7%	18.8%	2.7%	10.6%	33.3%
	佔總計的百分比		29.4%	1.0%	1.3%	1.7%	33.3%
無法判斷	計數		0	8	2	1	11
	All結果2 內的 %		0.0%	72.7%	18.2%	9.1%	100.0%
	Test2_2 內的 %		0.0%	50.0%	1.4%	2.1%	3.6%
	佔總計的百分比		0.0%	2.6%	0.7%	0.3%	3.6%
支持	計數		1	2	131	8	142
	All結果2 內的 %		0.7%	1.4%	92.3%	5.6%	100.0%
	Test2_2 內的 %		1.1%	12.5%	89.7%	17.0%	46.9%
	佔總計的百分比		0.3%	0.7%	43.2%	2.6%	46.9%
中立	計數		4	3	9	33	49

政府巨量資料分析與政策端應用效能提升之研析

	All結果2 內的 %	8.2%	6.1%	18.4%	67.3%	100.0%
	Test2_2 內的 %	4.3%	18.8%	6.2%	70.2%	16.2%
	佔總計的百分比	1.3%	1.0%	3.0%	10.9%	16.2%
總計	計數	94	16	146	47	303
	All結果2 內的 %	31.0%	5.3%	48.2%	15.5%	100.0%
	Test2_2 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百分比	31.0%	5.3%	48.2%	15.5%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.785	.030	19.973	.000
有效觀察值個數	303			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

四、 群眾判定 * 國發會 B1

交叉表

			國發會 B1				總計
			反對	無法判斷	支持	中立	
群眾判定	反對	計數	90	2	6	4	102

附錄二、輿情平台群眾分析結果表

定（全部）	All結果2 內的 %	88.2%	2.0%	5.9%	3.9%	100.0%
	Test3_2 內的 %	82.6%	10.0%	4.4%	9.1%	32.9%
	佔總計的百分比	29.0%	0.6%	1.9%	1.3%	32.9%
	無法判斷	計數	2	9	1	1
	All結果2 內的 %	15.4%	69.2%	7.7%	7.7%	100.0%
	Test3_2 內的 %	1.8%	45.0%	0.7%	2.3%	4.2%
	佔總計的百分比	0.6%	2.9%	0.3%	0.3%	4.2%
	支持	計數	6	2	122	10
	All結果2 內的 %	4.3%	1.4%	87.1%	7.1%	100.0%
	Test3_2 內的 %	5.5%	10.0%	89.1%	22.7%	45.2%
	佔總計的百分比	1.9%	0.6%	39.4%	3.2%	45.2%
	中立	計數	11	7	8	29
	All結果2 內的 %	20.0%	12.7%	14.5%	52.7%	100.0%
	Test3_2 內的 %	10.1%	35.0%	5.8%	65.9%	17.7%
	佔總計的百分比	3.5%	2.3%	2.6%	9.4%	17.7%
總計	計數	109	20	137	44	310

政府巨量資料分析與政策端應用效能提升之研析

All結果2 內的 %	35.2%	6.5%	44.2%	14.2%	100.0%
Test3_2 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
佔總計的百分比	35.2%	6.5%	44.2%	14.2%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.705	.033	18.448	.000
有效觀察值個數	310			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

五、 助理判斷 * 國發會 B1

交叉表

			國發會B2				總計
			反對	無法判斷	支持	中立	
群眾判定	反對	計數	79	4	2	10	95
(全部)	All結果2 內的 %		83.2%	4.2%	2.1%	10.5%	100.0%
	Test4_2 內的 %		87.8%	20.0%	1.5%	15.6%	30.5%
	佔總計的百分比		25.4%	1.3%	0.6%	3.2%	30.5%

附錄二、輿情平台群眾分析結果表

無法判斷	計數	1	9	1	1	12
	All結果2 內的 %	8.3%	75.0%	8.3%	8.3%	100.0%
	Test4_2 內的 %	1.1%	45.0%	0.7%	1.6%	3.9%
	佔總計的百分比	0.3%	2.9%	0.3%	0.3%	3.9%
支持	計數	5	1	127	13	146
	All結果2 內的 %	3.4%	0.7%	87.0%	8.9%	100.0%
	Test4_2 內的 %	5.6%	5.0%	92.7%	20.3%	46.9%
	佔總計的百分比	1.6%	0.3%	40.8%	4.2%	46.9%
中立	計數	5	6	7	40	58
	All結果2 內的 %	8.6%	10.3%	12.1%	69.0%	100.0%
	Test4_2 內的 %	5.6%	30.0%	5.1%	62.5%	18.6%
	佔總計的百分比	1.6%	1.9%	2.3%	12.9%	18.6%
總計	計數	90	20	137	64	311
	All結果2 內的 %	28.9%	6.4%	44.1%	20.6%	100.0%
	Test4_2 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百分比	28.9%	6.4%	44.1%	20.6%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.729	.032	19.428	.000
有效觀察值個數	311			

政府巨量資料分析與政策端應用效能提升之研析

- a. 未使用虛無假設。
- b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

六、 助理判斷 * 電子治理研究中心 A1

交叉表

			電子治理研究中心A1				總計
			反對	支持	正反皆有/ 無法判斷	無法判斷 /neutral	
助理判斷	反對	計數	50	6	1	1	58
		助理判斷_5 內的 %	86.2%	10.3%	1.7%	1.7%	100.0%
		test1_5 內的 %	50.0%	4.5%	6.3%	1.3%	17.8%
		佔總計的百 分比	15.3%	1.8%	0.3%	0.3%	17.8%
	支持	計數	17	86	1	15	119
		助理判斷_5 內的 %	14.3%	72.3%	0.8%	12.6%	100.0%
		test1_5 內的 %	17.0%	64.2%	6.3%	19.7%	36.5%
		佔總計的百 分比	5.2%	26.4%	0.3%	4.6%	36.5%
	正反皆有/無 法判斷	計數	7	7	0	3	17
		助理判斷_5 內的 %	41.2%	41.2%	0.0%	17.6%	100.0%

附錄二、輿情平台群眾分析結果表

		test1_5 內的 %	7.0%	5.2%	0.0%	3.9%	5.2%
		佔總計的百分比	2.1%	2.1%	0.0%	0.9%	5.2%
	無法判斷 /neutral	計數	26	35	14	57	132
		助理判斷_5 內的 %	19.7%	26.5%	10.6%	43.2%	100.0%
		test1_5 內的 %	26.0%	26.1%	87.5%	75.0%	40.5%
		佔總計的百分比	8.0%	10.7%	4.3%	17.5%	40.5%
	總計	計數	100	134	16	76	326
		助理判斷_5 內的 %	30.7%	41.1%	4.9%	23.3%	100.0%
		test1_5 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
		佔總計的百分比	30.7%	41.1%	4.9%	23.3%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.416	.036	12.058	.000
有效觀察值個數	326			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

七、 助理判斷 * 電子治理研究中心 A2

交叉表

			電子治理研究中心A2				總計
			反對	支持	正反皆有/ 無法判斷	無法判斷 /neutral	
助理判斷	反對	計數	47	10	1	0	58
		助理判斷_5 內的 %	81.0%	17.2%	1.7%	0.0%	100.0%
		test2_5 內的 %	49.0%	6.8%	6.3%	0.0%	18.8%
		佔總計的百 分比	15.3%	3.2%	0.3%	0.0%	18.8%
	支持	計數	13	97	0	11	121
		助理判斷_5 內的 %	10.7%	80.2%	0.0%	9.1%	100.0%
		test2_5 內的 %	13.5%	66.0%	0.0%	22.4%	39.3%
		佔總計的百 分比	4.2%	31.5%	0.0%	3.6%	39.3%
	正反皆有/無 法判斷	計數	8	4	1	4	17
		助理判斷_5 內的 %	47.1%	23.5%	5.9%	23.5%	100.0%
		test2_5 內的 %	8.3%	2.7%	6.3%	8.2%	5.5%
		佔總計的百 分比	2.6%	1.3%	0.3%	1.3%	5.5%
	無法判斷	計數	28	36	14	34	112

附錄二、輿情平台群眾分析結果表

/neutral	助理判斷_5 內的 %	25.0%	32.1%	12.5%	30.4%	100.0%
	test2_5 內的 %	29.2%	24.5%	87.5%	69.4%	36.4%
	佔總計的百 分比	9.1%	11.7%	4.5%	11.0%	36.4%
總計	計數	96	147	16	49	308
	助理判斷_5 內的 %	31.2%	47.7%	5.2%	15.9%	100.0%
	test2_5 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百 分比	31.2%	47.7%	5.2%	15.9%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.396	.037	11.406	.000
有效觀察值個數	308			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

八、 助理判斷 * 國發會 B1

交叉表

			國發會 B1				總計
			反對	支持	正反皆有/ 無法判斷	無法判斷 /neutral	
助理判斷	反對	計數	48	5	0	2	55
		助理判斷_5 內的 %	87.3%	9.1%	0.0%	3.6%	100.0%
		test3_5 內的 %	42.5%	3.6%	0.0%	4.5%	17.5%
		佔總計的百 分比	15.3%	1.6%	0.0%	0.6%	17.5%
	支持	計數	17	89	3	10	119
		助理判斷_5 內的 %	14.3%	74.8%	2.5%	8.4%	100.0%
		test3_5 內的 %	15.0%	65.0%	15.0%	22.7%	37.9%
		佔總計的百 分比	5.4%	28.3%	1.0%	3.2%	37.9%
	正反皆有/無 法判斷	計數	7	6	0	2	15
		助理判斷_5 內的 %	46.7%	40.0%	0.0%	13.3%	100.0%
		test3_5 內的 %	6.2%	4.4%	0.0%	4.5%	4.8%
		佔總計的百 分比	2.2%	1.9%	0.0%	0.6%	4.8%
	無法判斷	計數	41	37	17	30	125

附錄二、輿情平台群眾分析結果表

/neutral	助理判斷_5 內的 %	32.8%	29.6%	13.6%	24.0%	100.0%
	test3_5 內的 %	36.3%	27.0%	85.0%	68.2%	39.8%
	佔總計的百 分比	13.1%	11.8%	5.4%	9.6%	39.8%
總計	計數	113	137	20	44	314
	助理判斷_5 內的 %	36.0%	43.6%	6.4%	14.0%	100.0%
	test3_5 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百 分比	36.0%	43.6%	6.4%	14.0%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.343	.034	10.489	.000
有效觀察值個數	314			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

九、 助理判斷 * 國發會 B2

交叉表

			國發會 B2				總計
			反對	支持	正反皆有/ 無法判斷	無法判斷 /neutral	
助理判斷	反對	計數	41	6	2	6	55
		助理判斷_5 內的 %	74.5%	10.9%	3.6%	10.9%	100.0%
		test4_5 內的 %	45.1%	4.3%	10.0%	9.1%	17.5%
		佔總計的百 分比	13.0%	1.9%	0.6%	1.9%	17.5%
	支持	計數	17	93	0	12	122
		助理判斷_5 內的 %	13.9%	76.2%	0.0%	9.8%	100.0%
		test4_5 內的 %	18.7%	67.4%	0.0%	18.2%	38.7%
		佔總計的百 分比	5.4%	29.5%	0.0%	3.8%	38.7%
	正反皆有/無 法判斷	計數	5	4	0	6	15
		助理判斷_5 內的 %	33.3%	26.7%	0.0%	40.0%	100.0%
		test4_5 內的 %	5.5%	2.9%	0.0%	9.1%	4.8%
		佔總計的百 分比	1.6%	1.3%	0.0%	1.9%	4.8%
	無法判斷	計數	28	35	18	42	123

附錄二、輿情平台群眾分析結果表

/neutral	助理判斷_5 內的 %	22.8%	28.5%	14.6%	34.1%	100.0%
	test4_5 內的 %	30.8%	25.4%	90.0%	63.6%	39.0%
	佔總計的百 分比	8.9%	11.1%	5.7%	13.3%	39.0%
總計	計數	91	138	20	66	315
	助理判斷_5 內的 %	28.9%	43.8%	6.3%	21.0%	100.0%
	test4_5 內的 %	100.0%	100.0%	100.0%	100.0%	100.0%
	佔總計的百 分比	28.9%	43.8%	6.3%	21.0%	100.0%

資料來源：本計畫。

對稱的測量

	數值	漸近標準錯誤 ^a	大約 T ^b	大約 顯著性
合約的測量 卡帕 (Kappa)	.365	.037	10.469	.000
有效觀察值個數	315			

a. 未使用虛無假設。

b. 正在使用具有虛無假設的漸近標準誤。

資料來源：本計畫。

政府巨量資料分析與政策端應用效能提升之研析

附錄三、跨部門巨量資料分析之訪談結果摘要

一、跨部門巨量資料之應用

(一) 巨量資料使用經驗

這兩個資料庫的內容欄位有些具有重疊，但資料內容與性質會有差異，主要是因為兩者的蒐集過程與方式有所不同（銓敘資料 vs. 派令資料），因為資料是經過銓敘審核過，因此銓敘部的資料正確程度較高。如果需要調查公務人員的基本資料，基本上會整合這兩個資料庫，因此資料庫的整合是在正在進行當中。銓敘部近兩三年來推動「全國公務人力雲端平台」，將銓敘、考績等資料皆上傳至此平台。(A1)

資料庫進行整合後，全國各政府行政機關皆會提出資料運用之業務需求，例如傑出貢獻獎、保訓會人力資源規劃、地方政府選委會等，故目前資料庫的應用相當多元化，服務的對象也以政府機關為主。(A1)

公務人力資料庫的現況：銓敘部與保訓會的資料會回流至公務人力資料庫進行整合。(A2)

(二) 政府機關巨量資料應用之未來規劃與方向

有關於未來近期目標方面，由於行政院推動 Open Data 的緣故，銓敘部也朝著資料開放之目的前進，但面臨僅擁有人事資料，缺乏金錢資料的困境，包括人事費用的資料。銓敘部近期會與主計總處進行研議，與其「待遇福利系統」整合與分析，短期目標是整合人事費用的系統，一旦人事、組織與費用資料系統建構完成，期待能有更完善的利用。

銓敘部規劃建立我國的「公務人力資料倉儲」，將會把統計的資料，定期公布以提供相關主管機關運用或進行追蹤。銓敘部亦規劃將「全國公務人力雲端平台」的服務範圍擴大，將原屬保訓會的訓練進修項目，以及考選部的相關資訊整合納入此平台。長期

政府巨量資料分析與政策端應用效能提升之研析

目標方面，考試院要求銓敘部研議「公務人力運用之公共課責制度」，將特定的公務人員資料公開，使民眾加以進行監督。(A1)

我國政府的資料運用面向非常多元，人事資料庫應用之可研究方向就包括了：公務人力評估、國旅卡使用之分析、完整退休資料與考試院年金制度的連結、資料治理、全體公務職系的流動率分析比較、生活津貼待遇系統，以及升遷之因素分析等議題，皆可應用巨量資料與進行分析。(A2)

二、 專家學者焦點座談

(一) 巨量資料分析於政府機關之應用

1、 哪些業務的應用符合巨量資料分析之 3V 基本特質

資料量大小取決於一種相對的概念，如果要作主題式或未來有價值的分析一定是要切割領域。就健保領域而言，健保署每天蒐集所有醫療院所的就醫資料，蒐集到的資料是屬於較為定型化、結構完整的資料，並不像散置的、影像化的非結構化資料。會有這樣的資料特性，是因為健保開辦初期時，資料處理的能力不大，因此對於申報格式的精簡上相當鑽研，因而產生出具有結構化的格式。

許多跨機關、跨領域的資料數據分析應用，例如電子發票與健康管理的配合應用，或是花粉熱、行道樹與健保資料的相互應用，這些都可以提供作為公共政策的建議。業務單位在推動上也面臨了許多挫折，政府部門在運作上有很多限制，如果沒有高層主管的推動決心，很難讓跨部門或跨領域的資料分析順利成功。(B1)

我們機關目前對於大數據的應用也只是剛起步而已，現在設立了一個大數據工作小組才要開始運作，因此資料倉儲還沒開始建置，目前還停留在構想階段。我們在思考：資料量的大小對於分析的結果是否真的有影響？就市府而言，員工的人數大約六七萬人，跟氣象分析比起來真的是小巫見大巫。其實我們的觀點跟剛剛葉組長相當類似，就是資料量的多寡只要是夠用就好，只要數據大到夠用、夠完整，可以拿來決策分析或較有說服力的資訊即可。(B3)

在即時更新上，我們也遇到了困難，可能是牽扯到權責的問題，因為總處負責蒐集資料，為了要符合即時更新，我們也建立了 WEB HR 共用系統讓大家來用，但這些都是從資訊角度出發，但在權責角度上人事總處並沒有處理的權責，權責的

執行單位而是銓敘部。資料多元化的部分，我們沒有想到去蒐集輿情或語音資料來作分析，目前還停留在結構性的量化資料，即時性與正確性一直是很大的問題，我們也一直被質疑我們蒐集的資料顆粒度過大，我們只能提供統計資料，細部的資料恐怕會觸碰到個資法的部分。(B4)

我們比較不屬於產生資料的單位，所以我是從使用者的角度分享。我們當初在講大數據有幾個 V，對我們來說並非需要同時符合這些 V 才屬於大數據，它只要有某一些類型的特性存在的時候，就會有一些挑戰或需要被處理的事情，就需要產生大數據的技術，所以我們不認為所有的 V 都要存在才叫大數據。(B5)

我覺得幾個 V 真的不是重點，重點其實是現在大家講的 Value(價值)，對於更新速度有多快，資料數量有多大，這些討論其實都有點過於鑽牛角尖了，最重要反而真的要從需求導向去看，因為資料量真的太多了，蒐集再多拿來作比較，就我們看來真的只是噓頭而已。(B6)

2、何種業務類型可利用巨量資料的分析方法

首先我們在作大數據的分析應該是議題導向；其次，政府的大數據分析應用，必須是由上而下。在政府部門真正掌管資料的是業務主管，可是在推動 open data 與 big data 則是資訊主管，以健保署的角度來看，他們只懂技術問題，不懂業務問題，他如果不懂得業務問題，就不會知道應該要以議題導向來準備資料給外界。(B1)

為何要使用大數據，先不用期待它可以帶來績效的提升或政策的協助，我們現在最大的問題點是在於我們處在什麼樣的情境？我們並不知道我們的現況為何。譬如說，現在的戶籍概念其實跟居住的概念是不一樣的，但我們很多的政策是用戶籍的概念出發，所以我們應該先弄清我們現在的狀況，以及

位處的情境。

其次，使用 big data 上，我們的邏輯應該翻轉過來，也就是先有問題，然後再去想 data 是什麼；換句話說，針對問題的性質為何，先去尋找手中現有的資料，瞭解是否能夠分析問題，如果無法運用的話，再向外尋求協助。就這樣的邏輯下，我們有兩個問題可以進行思考：首先是流程的問題，不管是由上而下或由下而上，我們必須去找出內部資訊流通的可能發起點為何，資料庫的掌管者不必挖空心思去想資料去要如何建置，我把這些定位在資料的提供端，重要的是要如何去回應提問者，瞭解哪些東西是可以使用的，在現有的東西中，我要怎麼要去提供給對方是可以省去麻煩的？(B2)

大數據的應用，應該是要採取兼容的角度，一方面是問題導向，另一方面也可以作一些議題的預測，但大數據並非是什麼問題都能解決，有時候也許改用其他傳統的方式會更有效率，針對問題的性質選擇不同方式解決會比較恰當。(B3)

從我們的角度而言，政府在運用數據分析上，首先要考慮的是資料完備度是否足夠，因為政府釋放資料的意義是讓使用者進行運用，一旦完備度夠了，分析才能產生價值；其次是資料的顆粒度要足夠，有很多釋放出來的資料顆粒度是較低的，如果顆粒度太大，就沒辦法往下進行深入的研究與分析；因此我們也同意資料只要夠用，價值就會產生，而非是量大才有價值。(B5)

因為政府在推動 open data 的方式，雖然是由上而下推動，但執行方式卻是由下而上，我們最常看到的方式是各個部門自己去盤點自己有什麼樣的資料，依照可開放程度決定是否開放，這樣會造成很多重疊的狀況，如果沒有一個整合性或統合性的方式或流程，去讓所有資料作整合，其實這也只能形成一個倉庫而已，而且這個倉庫的東西也不是那麼有結構的可以被利用與加值。如果可以去作一個由上到下的整合，資料的可用度會比現在來得好很多，我們真的建議有些東西是適

合由上而下的去執行。(B6)

(二) 資料治理於政府機關之應用

1、資料治理在政府機關應用的障礙

雖然是由上而下的議題導向，但負責的單位都沒有一個因應的計畫，讓輔助的單位去依循並幫助，進而累積出成果。

專業不同，需要準備的資料也會不同，但不見得符合其他單位的需求，專業解讀不同，其蒐集資料的管道也會不同，會變成議題導向的資訊需要重新蒐集。政府機關的資料蒐集各自領政，各自的角度與專業，不但是格式不一致，資料的品質與內容也不一致，資料治理最重要的問題不是那些格式，而是內容與品質。(B1)

資料的格式無法反映現實狀況，由於資料的來源與特性五花八門，且會有漏答的情形，無法進行比對與分析。且上級機關的資料格式無法符合現況，使得下級機關在整理上需要多花功夫進行考核。(B3)

不同的機關所擁有的資料不一樣，使得資料以片段的方式呈現，無法瞭解整體資料的特性；且不同機關維護資料的方式不一，資料的編碼方式也不同，使得資料統整與整合的難度倍增，凸顯出政府資訊整合有很大的問題。(B4)

2、分析政府資料應用障礙之原因與解決之道

資訊單位的角色應該是在前期進行準備，後期應該由業務單位執行，但目前行政院與國發會的角度仍然是由資訊單位來推動大數據的業務，其實是有點本末倒置的，從顧客導向的觀點來看，不應該由資訊單位來扛起這樣的責任。(B1)

建議電子治理研究中心可以向行政院提案，由副院長或一位政務委員成立一個計畫，先完成一至兩個標竿的專案，藉由

此專案排除我們討論過，在大數據應用上會面臨到的問題。會這樣做的原因在於大家對大數據並不瞭解，且執行上有疑慮，藉由先導型的專案，一方面排除困難，二方面告訴大家這個是不錯的，引起各部會的興趣。(B2)

建議上級機關在蒐集資料時，能與下級機關進行溝通，瞭解各自的需求，並進行協調與調整，或制定相對的配套措施，如此才能確保資料的品質。(B3)

法規因素讓資訊單位執行資料蒐集有所疑慮與擔憂，且法務單位將法規解釋的功能丟回給資訊與業務等單位，形成許多模糊的灰色地帶，讓這些單位執行業務上皆採取保守的態度，不敢進行突破性的改變。其次，權責問題使得人事行政總處在資料蒐集上遇到相當多困難。最後，欠缺應用 big data 的能力，使人事行政總處只知道如何處理資料，欠缺循證的方法，換句話說就是統計方法與資料探勘的面向，資訊人員對於處理資料後，如何形成模型的能力較為欠缺，需要資訊、業務與統計的人才相互配合，才有辦法靈活應用，因此政府針對大數據應用的人才培育與訓練仍有進步的空間。(B4)

在資料格式的部分，很多資料進到平台時都各有各的欄位與格式，並且以代號的形式呈現，這個東西其實是提供者才會知道，使用者並不知道，所以要怎麼為這些資料定義並定好欄位，這是必須要清楚說明，讓需要分析的人拿到資料之後，可以應用與加值。因此，在資料週期的運作中，資料的使用者與決策制定者也需要加入，也就是所謂的 domain knowledge(領域知識)，才有辦法使資料被完善運用。在以上的過程中，都需要被明確的分工與賦予責任，這些都是可以在管理面與法規政策方面被執行的。

在這個過程中常常會出現一個問題，資料提供者與平台提供者各自執行其業務時，會出現想法、觀念與作法上的落差，此時就需要透過說明會、交流會或教育訓練，否則業務單位與資訊單位間會產生很大的隔閡，難以溝通。在業界方面處理此

政府巨量資料分析與政策端應用效能提升之研析

類的問題，大多會把業務單位與資訊單位找來開會，彼此提出需求與互相支援的方法，大部分不會使用由上而下的交辦方式，而是採取成立專案型組織執行這樣的業務。(B5)

附錄四、 PyBossa 平台架設之系統文件

一、 專案程式碼

專案程式碼位置: <https://github.com/Skytim/nccuTEG>

(一) 安裝 nccuTEG

nccuTEG 是基於 pybossa 所建立的專案，目的在於幫助 TEG(電子治理研究中心)分析政府想關的資料。以下的教學將以本計畫建立的 nccuTEG5 專案平台為例，展示如何透過 Vagrant 建立系統環境。Vagrant 是一個輕量級的套件，可以讓你快速地透過 virtualbox 部署專案設計的系統環境。

(二) 需求環境

作業系統：

ubuntu 14.04LTS；

Mac OSX Yosemite；

並安裝 git。

如果是 Ubuntu 請在 terminal 輸入 apt-get install git

如果是 Mac 請下載 git 安裝

<http://sourceforge.net/projects/git-osx-installer/>

二、 設置 PyBossa 使用 Vagrant

你需要安裝以下套件

- 1、 VirtualBox (min version 4.2.10)
- 2、 Vagrant (min version 1.2.1)
- 3、 Vagrant Plugin (vagrant-ansible-local)

先安裝完上述 1、2 套件後，在安裝套件 3 時，請在終端機輸入：`vagrant plugin install vagrant-ansible-local`

`clone nccuTEG`

`git clone https://github.com/Skytim/nccuTEG`

`clone`

輸入後，切換到此目錄

`cd nccuTEG`

並且在 terminal 輸入指令 `vagrant up`

之後，需等待數十分鐘，讓 `vagrant` 下載 Ubuntu 作業系統環境完成後進入到虛擬機環境 `vagrant ssh`，並且運行 `python run.py`。最後，開啟 browser 並輸入 `http://127.0.0.1:5000`（平台網址），便可以看到 `nccuTEG` 的起始畫面。

三、 **Pybossa** 功能簡介

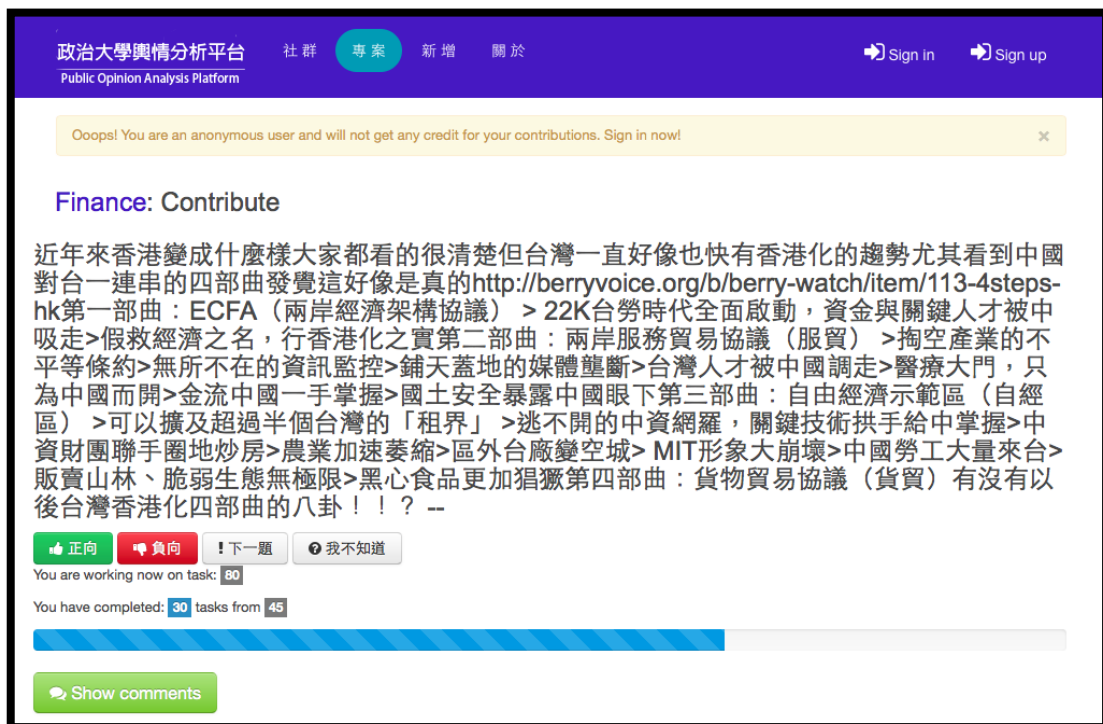
（一） 前台

- 1、 使用者收到 **Pybossa** 寄來的信，裡面有專案連結與專案密碼。
- 2、 使用者透過內附連結先登入 **Pybossa** 的使用者，之後透過信裡的密碼進入專案。
- 3、 該專案在使用者測驗前會有提示教學，告知使用者其專案的測驗方法。
- 4、 使用者開始問答，選擇文章的正負向程度，並可填入原因

至舉證區。

(二) 後台

- 1、管理者可以設定整個專案每題任務將被測驗的次數，其優先順序可以設定為深度或廣度優先。
- 2、管理者可以設定其專案的專案密碼。
- 3、可以查看目前任務的統計圖表。
- 4、管理專案的任務。



資料來源：本計畫。

PyBossa 有三種使用者，分別為匿名使用者，一般使用者和管理者。在預設條件下，建立 PyBossa 專案的第一個用戶將成為管理者，

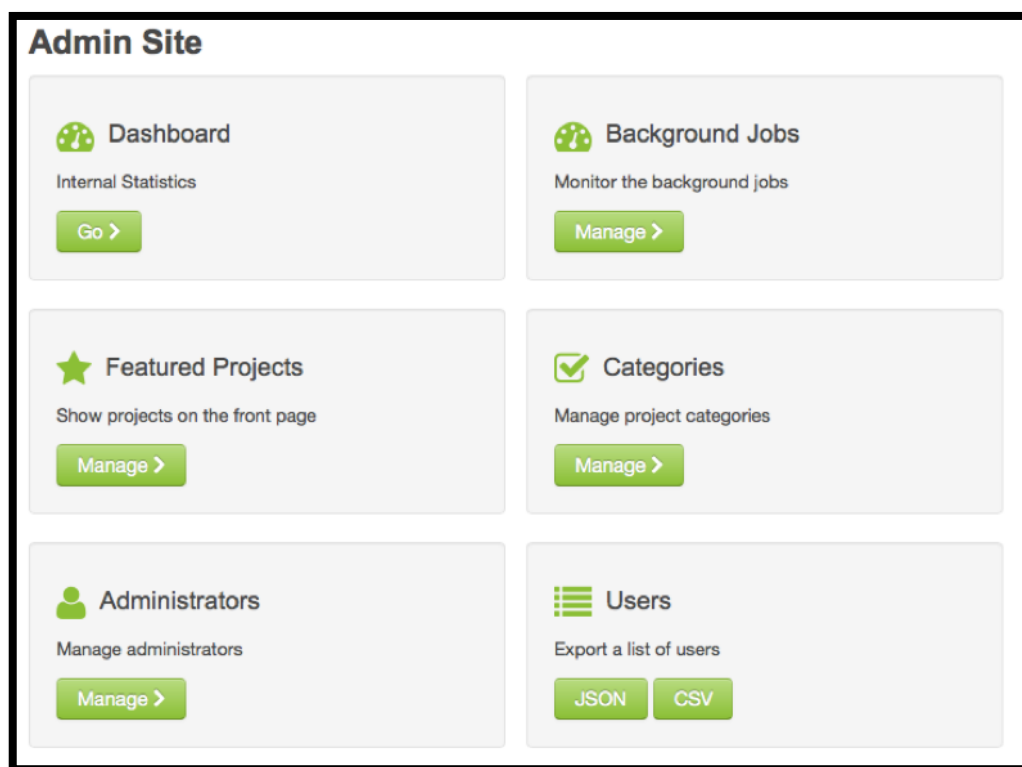
且擁有所有權限管理網站。管理者能夠透過點擊畫面右上角的用戶名稱，到網站管理者的管理介面。



資料來源：本計畫。

管理者在管理 Pybossa 時，擁有三個不同的功能：

- 1.受關注的專案
- 2.分類專案
- 3.管理者功能

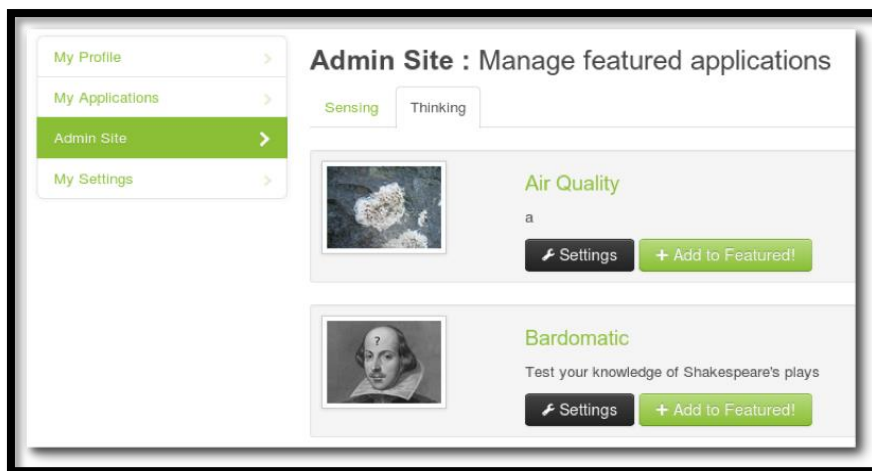


資料來源：本計畫。

管理者可以修改所有專案，也可以看到哪些專案被標記為草稿：專案中至少要建立一個任務，管理者可顯現任務頁面，讓其他使用者參與專案任務。

管理者也可在管理者頁面中獲得所有註冊過的 **PyBossa** 使用者列表，並能以 **JSON** 或 **csv** 格式下載該列表。此外，管理者還可以透過 **RQ** 儀表版來監控所有的後台作業，或者取消、重啓目前正進行中的任務。

（三） 顯示專案



資料來源：本計畫。

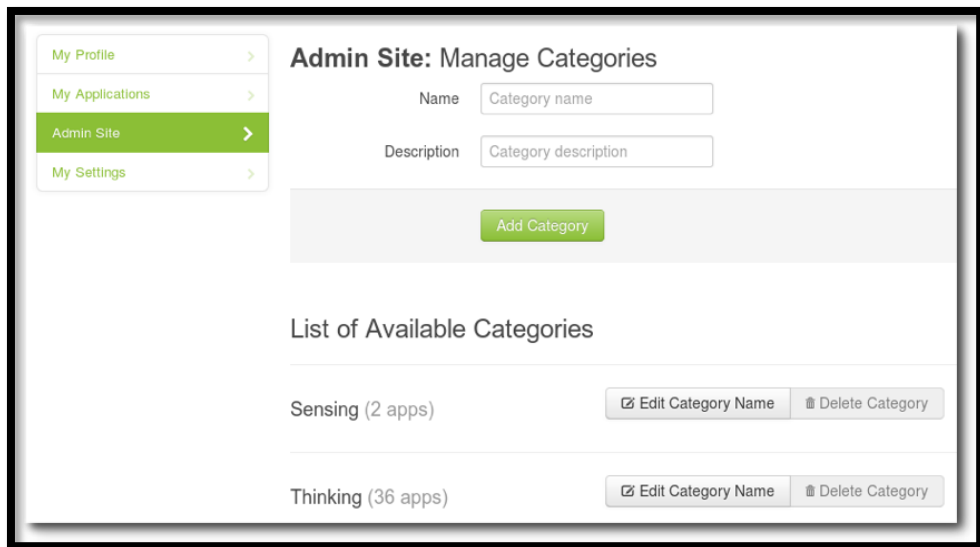
我們可以透過該頁面將一個專案加入到首頁，或是從中刪除。

（四） 分類

PyBossa 一開始有兩種專案分類：

- 1、 思考（Thinking）：以使用者可以用他們的技能來解決問題來分類（如圖像或聲音識別）。
- 2、 檢測（Sensing）：使用者可以透過像是 EpiCollect 工具，幫助專案管理者蒐集資料，然後回傳數據到 PyBossa。

管理者想要添加多個類別的話，只需輸入名稱和它的描述，即可添加專案。



資料來源：本計畫。

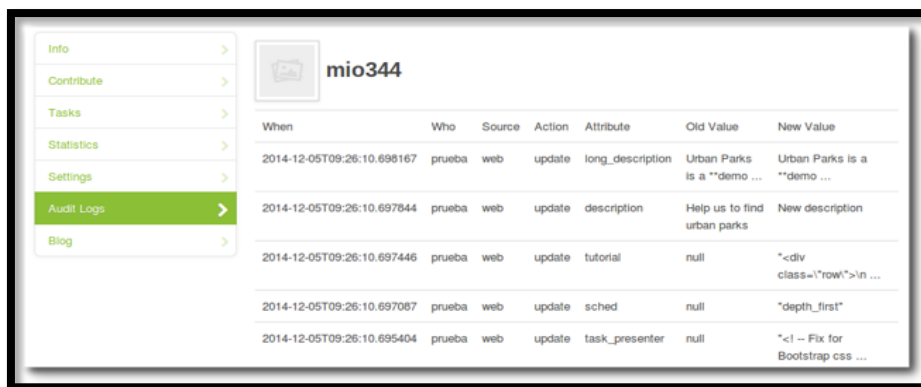
當有一個以上的專案在上面時，你不能刪除該類別。但是，當沒有相關專案連接到該類別時，你可以重新命名類別或刪除它。

（五） 管理者

在這一個頁面，管理者可以加入/刪除使用者的管理者的角色。基本上，你可以通過用戶別名進行搜索，並將他們加到管理者群組。跟分類功能很像，點擊按鈕即可允許或刪除將使用者的管理者權限。

（六） 審計日誌

當一個專案被建立時，刪除或更新，系統登記其在服務器操作。管理者將有機得在每個項目頁面中的所有記錄的行為，在一個名為部分審計日誌。



When	Who	Source	Action	Attribute	Old Value	New Value
2014-12-05T09:26:10.698167	prueba	web	update	long_description	Urban Parks is a "demo ...	Urban Parks is a "demo ...
2014-12-05T09:26:10.697844	prueba	web	update	description	Help us to find urban parks	New description
2014-12-05T09:26:10.697446	prueba	web	update	tutorial	null	"<div class="row">in ...
2014-12-05T09:26:10.697087	prueba	web	update	sched	null	"depth_first"
2014-12-05T09:26:10.695404	prueba	web	update	task_presenter	null	"<l - Fix for Bootstrap css ...

資料來源：本計畫。

只有管理員和被授權的使用者才可以看到審計日誌。透過該資訊你可以了解：

- 1、 **When**：記錄使用者的活動時間。
- 2、 **Action**：使用者的修刪改查記錄。
- 3、 **Source**：分別是透過 **API** 或 **Web** 頁面進行操作。
- 4、 **Attribute**：該專案的屬性已被更改。
- 5、 **Who**：哪一個使用者。
- 6、 **Old value**：修改前的值。
- 7、 **New value**：修改後的值。

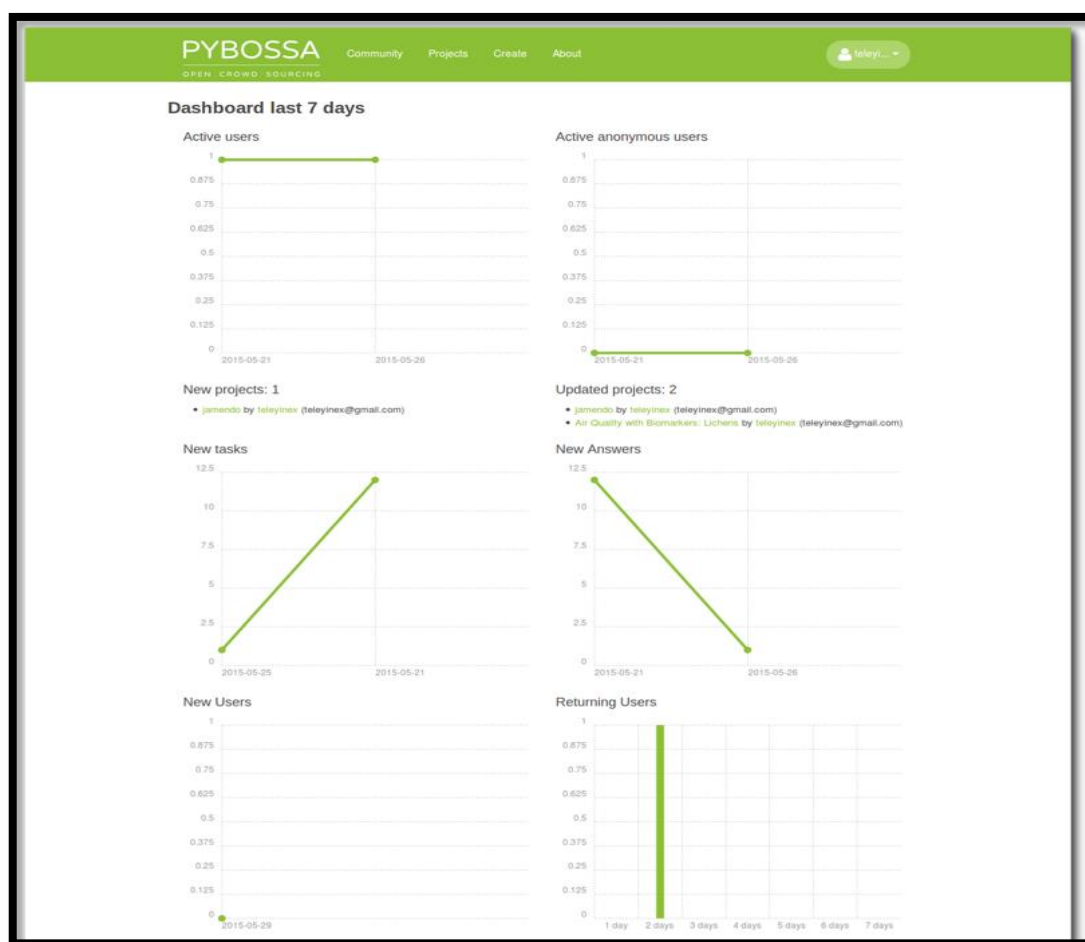
（七） 系統儀表板（**Dashboard**）

儀表板可以讓你監測 **PyBossa** 運作狀況，在儀表板上會顯示過去一個禮拜內的資訊，並 24 小時即時更新，其顯示資訊包含：

- 1、 活躍使用者：在過去的 7 天進行至少 1 個任務的使用者數。
- 2、 活躍匿名者：在過去的 7 天進行至少 1 個任務的匿名使用

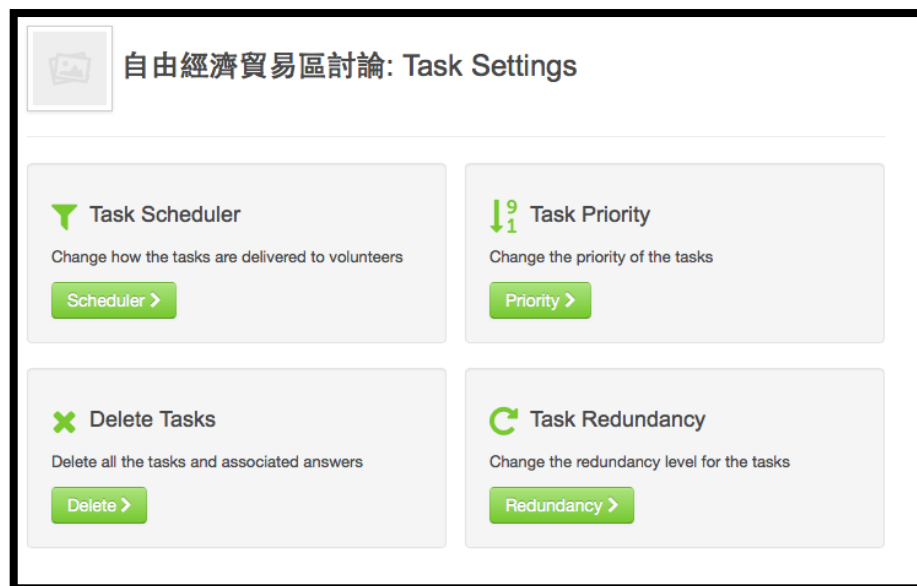
者數。

- 3、新專案：在過去的 7 天中創建的新專案。
- 4、更新專案：在過去的 7 天更新專案。
- 5、新使用者：在過去的 7 天註冊的新用戶數。
- 6、返站使用者數：返站使用者在過去 7 天內回來的數量。
- 7、近期活動：最後在 Pybossa 上處理的活動。




資料來源：本計畫。

(八) 任務設定及管理



資料來源：本計畫。

管理者可以設定整個專案每題任務可以被測驗幾遍,其優先順序,以及深廣度優先。



自由經濟貿易區討論: Browse tasks

This page shows all the available tasks for this project.

For each task, you can find the following information:

Task #0000 This number identifies the task for the project and it is unique

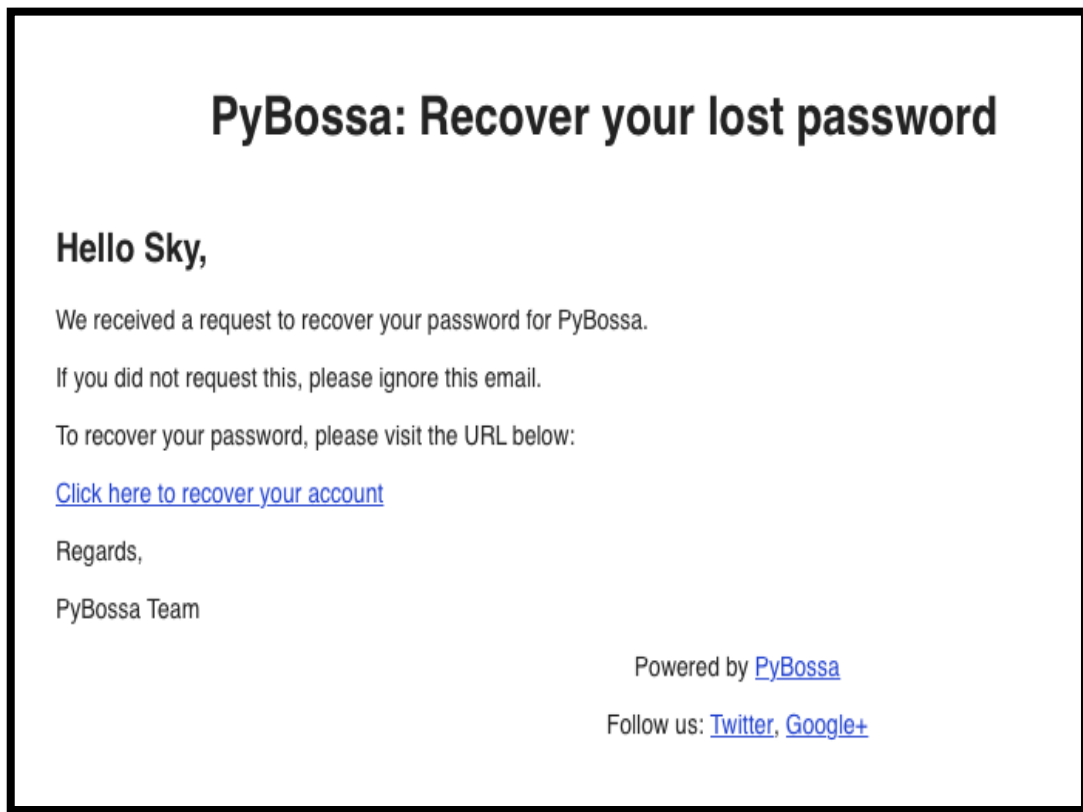
0 of 30: The first number shows how many answers have been submitted for the task and the **second number** how many need to be obtained to mark the task as **completed**.

Progress Bar: The progress bar will the percentage that has been completed for the task.

Task #266 3 of 5	<div><div></div></div>	Download Partial results
Task #267 3 of 5	<div><div></div></div>	Download Partial results
Task #268 3 of 5	<div><div></div></div>	Download Partial results
Task #269 3 of 5	<div><div></div></div>	Download Partial results
Task #270 3 of 5	<div><div></div></div>	Download Partial results

資料來源：本計畫。

(九) E-mail 驗證 (包括忘記密碼與驗證身份)



資料來源：本計畫。

當使用者申請新帳戶或是忘記密碼使用者使用認證信箱時，Pybossa 會自動寄一封確認信給使用者。

四、測驗流程

自由經濟貿易區示範教學

歡迎,以下是本專案的說明介紹

自由經濟貿易區討論: Contribute 1.報導題目

自由經濟示範區懶人包-同經濟起飛的序曲雖然技術上軌行的不錯但是這個論述也大...自我感覺良好吧-D-下面留言都有人說得很清楚,然後依然是各自針對不過也有難得順暢溝通的例子(然後又被黨工惡版)

舉證說明 2.報導文章

自我感覺良好吧

3.舉證說明

支持 反對 中立 不清楚 4.文章情感

You have completed 4 tasks from 4

Hide comments

1 Comment Crowdcrafting Login

Recommend Share Sort by Best

Join the discussion...

SKY · 40 minutes ago
我覺得這個議題有點疑問

5.討論區

畫面中的五個區域分別是

- 1.報導題目:
- 2.報導文章:每個專案都有它的子任務,你必須評斷該文章的情感,並在下方選擇
- 3.舉證說明:你可以填入評斷該文章情感的依據內文
- 4.文章情感:有四個選項 1.支持2.反對3.中立4.不清楚
- 5.討論區:可以跟他人討論該文章的內容

開始進行

資料來源：本計畫。

該專案在使用者測驗前會有提示教學,告知使用者其專案的測驗方法。

自由經濟貿易區討論: Contribute

自由經濟示範區 懶人包-悶經濟逆轉勝的序曲雖然技術上執行的不錯但是這個論述也太...自我感覺良好吧-D-下面留言都有人說得很清楚，然後依然是各自跳針不過也有難得順暢溝通的例子（然後又被黨工洗版...

舉證說明

請輸入舉證說明

支持

反對

中立

不清楚

You have completed: 0 tasks from 5

資料來源：本計畫。

使用者開始問答,選擇文章的正負向程度,並可填入原因至舉證區

task_run_calibration	task_run	task_run	task_run	task_run__info	task_run	task_run	task_run	task_run	task_run
None	2015-08	2015-08	549	No/自我感覺良好/然後又被黨工洗版	13	266	None	2	None
YES	2015-08	2015-08	550	Yes/台灣不能置身事外	13	267	None	2	None
None	2015-08	2015-08	551	Yes/精彩的對談 也都不會開放	13	268	None	2	None
YES	2015-08	2015-08	552	Yes/台灣人不用怕經濟整合，台灣人不是被嚇大的	13	269	None	2	None
None	2015-08	2015-08	553	Yes/示範區政策內涵公開、透明 香港與新加坡業者表達興趣 教育	13	270	None	2	None

資料來源：本計畫。

其蒐集資料結果有其答題人之答題選項，答題之答案以及作答時間。

附錄五、 中央政府各機關網站服務現況調查規劃

V4

2015 年 11 月 13 日

壹、調查說明：

- 一、 本項調查的對象為中央政府各機關（一～三級機關），適用國家發展委員會「政府網站版型與內容管理規範」的機關入口網（若不適用免填）以及該機關瀏覽人次最多之主題網站（若無免填）。填答者為該網站管理人（需要事先掌握有哪些網站）。以國發會為例，受訪者請分別就國發會的機關官網和我的 E 政府網站的營運現況，各填寫一份問卷。
- 二、 本調查分析單元為「網站」，調查工具為網路問卷（SurveyMonkey），調查內容包含政府機關網站的服務功能、行銷方式以及推動使用者導向的需求。
- 三、 調查有關政府機關網站回應使用者需求的需求，係參考 AARRR 分析架構設計，以結合 GA 設定需求。
 - （一） Acquisition（獲得使用者）：如何讓使用者從各種管道得知並連結到網站？
 - （二） Activation（成為有效的使用者）：如何讓已瀏覽網站的使用者使用網站提供的主要功能、服務或閱讀重要資訊，而不是隨意瀏覽即離開？
 - （三） Retention（留住使用者）：如何讓使用者願意重複造訪網站？
 - （四） Referral（使用者推薦他人）：如何讓使用者願意推薦他人使用網站？
 - （五） Revenue（使用者付費或執行預期行動）：如何讓使用者增加網站的造訪率、停留時間或服務利用率？（考量政府網站特質，

政府巨量資料分析與政策端應用效能提升之研析

故調整原始定義，改參照免費提供內容，但藉廣告獲利的商業網站模式，設計衡量重點及調查問題）

四、 預期調查成果：

- （一） 調查中央政府各機關官網及為民服務網站目前營運情況以及及關注使用者需求的現況。
- （二） 瞭解中央政府各機關為民服務網站推動顧客導向服務的具體作法及未來需求。
- （三） 瞭解中央政府各機關便民服務網站使用 Google Analytics（GA）協助分析使用者行為的需求。

貳、調查內容

一、調查說明

親愛的公務界先進您好：

我們正進行一項由國家發展委員會委託電子治理研究中心辦理的研究案，希望透過中央政府各機關網站管理者的問卷調查，讓我們對於政府回應民眾網路服務需求的公務運作實況有更深入的了解。

本項調查的對象為適用國家發展委員會「政府網站版型與內容管理規範」的機關入口網(官網)以及各機關瀏覽人次最多之主題網站，每一個網站請填答一份問卷，若機關未設有主題網站，則僅需就官網營運現況填寫一份問卷。以國發會為例，受訪者請分別就國發會的機關官網和我的 E 政府網站的營運現況，各填寫一份問卷(重新連至問卷網址即可重填)。每份問卷僅需花費您十分鐘的時間，希望您能依據個人從事網站管理的專業及工作經驗依序填答相關問題。您填答的資料僅作學術分析之用，個別問卷都將匿名以及保密處理，請您放心。非常感謝您提供寶貴的意見以及對於本研究的支持。

計畫主持人	政治大學公共行政學系教授	陳敦源
協同主持人	政治大學公共行政學系副教授	蕭乃沂
	政治大學資訊科學系教授	陳 恭
	空中大學公共行政學系助理教授	廖洲棚

敬上

2015 年 11 月 16 日

聯絡人：政治大學公共行政學系博士生 陳揚中
電 話：(02)2939-3091~51145
E-mail：103256503@nccu.edu.tw

二、網站基本資料

- (一) 請問 貴機關名稱：_____（開放題）
- (二) 請問 貴機關的機關代碼：_____（開放題）
- (三) 請問 貴機關網站網址：_____（開放題）
- (四) 請問 貴機關網站是否為機關官網（入口網站）？ 是 否
- (五) 請問 貴機關網站預設的主要服務對象為何？（如 65 歲以上老年人、婦女、12 歲以下孩童、所有納稅人、全國國民、關心房地交易者……等）_____（開放題，至多寫三種）
- (六) 請問 貴機關網站這一年來每月平均瀏覽人次為何？
（開放題）
- (七) 請問 貴機關網站是否有設定網站績效指標？是 否（跳答下一題）
- (八) 請問 貴機關網站設了哪些關鍵績效指標？_____（開放題）
- (九) 請問 貴機關網站最近一次改版的上線服務時間？民國_____年
月
- (十) 請問 貴機關網站最近一次改版的網站建置費用（不含硬體）約為
多少？
新台幣_____千元（開放題）
- (十一) 請問 貴機關網站從最近一次改版迄今，平均每年約需編列多
少維運費用（不含硬體）？ 新台幣_____千元（開放題）
- (十二) 請問 貴機關網站目前需投入多少人力來管理或維護？專任
人；兼任____人（開放題）

三、網站設計的方式

- (一) 請問 貴機關網站採用以下哪一種方法規劃和建置（單選）：
1. 委外規劃及建置
 2. 自行規劃但委外建置
 3. 自行規劃及建置。
 4. 其他：（請說明）。
- (二) 請自我評估 貴機關網站在規劃及建置時，對於下列利害關係者意

見的重視程度：(在 1 到 10 分的範圍內給予分數，例如：1 為非常不重視，10 為非常重視，5-6 代表中間意見，5 稍偏不重視，6 稍偏重視)

1. 網站主要使用者。
2. 機關首長。
3. 民意代表。
4. 業務單位。
5. 財政及主計單位。
6. 審計單位。
7. 網站管考單位。
8. 委外廠商。

(三)請問 貴機關網站是否為適用國家發展委員會「政府網站版型與內容管理規範」的網站？ 是 否(系統設定跳答下二題) 不知道(系統設定跳答下二題)

(四)請自我評估 貴機關網站符合「政府網站版型與內容管理規範」所列之必要要求的程度(若要求項目僅為建議項目則不計算)：(請就以下的陳述在 1 到 10 分的範圍內給予分數，例如：1 都完全不符合，10 完全符合，5-6 代表中間意見，5 稍偏不符合，6 稍偏符合)

1. 使用者的呈現裝置(共四項)。
2. 網頁組成要素(共五項)。
3. 導覽(共七項)。
4. 首頁設計(共三項)。
5. 文字樣式與連結(共七項)。
6. 圖片與多媒體(共五項)。
7. 表單(共五項)。
8. 搜尋(共兩項)。
9. 應提供內容(共六項)。
10. 內容呈現格式(共四項)。

11. 內容管理（共四項）。

(五)請自我評估 貴機關網站的使用者親和性（user friendly）表現：（請就以下的陳述在 1 到 10 分的範圍內給予分數，例如：1 完全不認同，10 完全認同，5-6 代表中間意見，5 稍偏不認同，6 稍偏認同）

1. 我們的機關網站能清楚傳達網站的設置目的，提供網站必要資訊，讓民眾產生良好印象。
2. 我們的機關網站能提供使用網站的必要提示，並提供方便民眾操作的網站功能。
3. 我們的機關網站能提供網站使用動線，方便民眾找到相關政府資訊。
4. 我們的機關網站能提供標準化的網頁內容格式以及一致性的視覺設計，讓民眾能順暢無阻地瀏覽網站。
5. 我們的機關網站有方便的站內搜尋功能，讓民眾輸入關鍵字，就能找到所需之服務或找到所需資訊來解決疑惑。
6. 我們的機關網站能提供最真實、正確的資訊或服務，讓民眾產生信賴感。

(六)請自我評估 貴機關網站的網站可用性（usability）表現：（請就以下的陳述在 1 到 10 分的範圍內給予分數，例如：1 完全不認同，10 完全認同，5-6 代表中間意見，5 稍偏不認同，6 稍偏認同）

1. 我們的機關網站內容更新速度很快，經常在完成內部簽核程序後就立即上網公告或更新。
2. 我們的機關網站設計符合使用者直覺，使用者很容易或不需費力就能操作網站提供的服務。
3. 我們的機關網站使用者可以憑直覺就學會網站提供的基本功能。
4. 我們的機關網站回應速度很快，使用者在操作時不會有延遲感。
5. 我們的機關網站內容容易記憶，使用者用後可以對內容留有

深刻印象。

6. 我們的機關網站從未發生網頁找不到（104 not found）或其他非預期的錯誤。
7. 我們的機關網站使用者對網站內容及服務感到滿意。

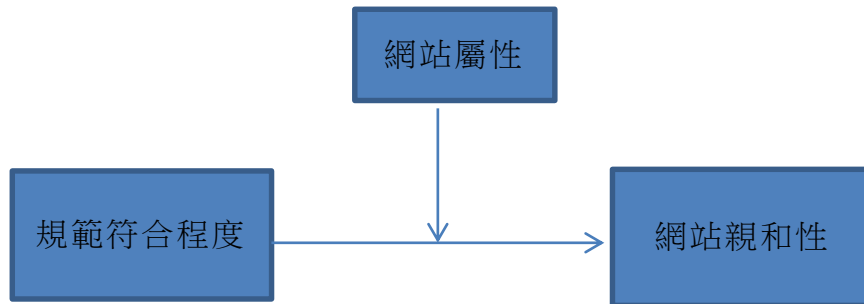
四、網站管理者對使用者需求的認識現況與未來期望

- (一) 請問貴機關在規劃本網站時是否有調查使用者需求？是 否（系統設定跳答下一題）
- (二) 請問您用了以下哪些方法調查網站使用者需求？（複選）問卷調查、深度訪談、行動觀察、實驗研究、網站流量分析、其他
- (三) 請問貴機關是否有辦理網路行銷活動來推廣本網站？是 否（系統設定跳答下兩題，到第六題）
- (四) 請問您目前透過以下哪些網路行銷管道來讓使用者知道您的網站？商業網站廣告、政府機關網站廣告、FB 廣告、FB 粉絲頁、Google 關鍵字廣告、搜尋引擎優化設定（SEO）、關鍵字廣告（SEM）、LINE 貼圖、網路抽獎、電子郵件、其他（請說明）
- (五) 請問您曾運用以下哪些方法方便使用者推薦他人造訪您的網站？（複選）FB 粉絲頁、Google+網頁、LINE 貼圖、YouTube 影片、電子郵件、其他（請說明）
- (六) 請問您的網站目前有使用 Google Analytics（GA）分析使用者行為嗎？有（跳答下一題）沒有 沒聽過這個分析工具（結束訪問）
- (七) 請問沒有使用 GA 的原因是？沒有需要 不會使用 未列在委外合約 其他（請說明）
- (八) 請問您對於下列網路流量分析資訊的需求程度為何：（請就以下的陳述在 1 到 10 分的範圍內給予分數，例如：1 完全不需要，10 非常需要，5-6 代表中間意見，5 稍偏不需要，6 稍偏需要）
 1. 瞭解網路民眾從什麼管道得知我們的網站。
 2. 釐清網路民眾來源管道和網路行銷活動的關係。
 3. 蒐集網站到訪者的個人資訊（如性別、IP 位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊）。

4. 瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站（訪客站外搜尋關鍵字分析）。
5. 分析網站到訪者實際使用網站服務的情況（如網站最熱門網頁、服務或最常被下載的檔案）。
6. 分析網站訪客曾在我們的網站搜尋哪些關鍵字（訪客站內搜尋關鍵字分析）。
7. 分析網站訪客搜尋的關鍵字（包括站內與站外搜尋）和網站服務的對應關係。
8. 分析網站系統的回應速度。
9. 分析網站訪客的再訪頻率。
10. 釐清網站訪客再訪率和行銷活動的關係。
11. 區隔網站訪客類型（如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型）。
12. 瞭解網站訪客推薦他人使用我們的網站情況。
13. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程（訪客的進入點、停滯點和離去點）。
14. 分析網站預期績效指標的達成率。

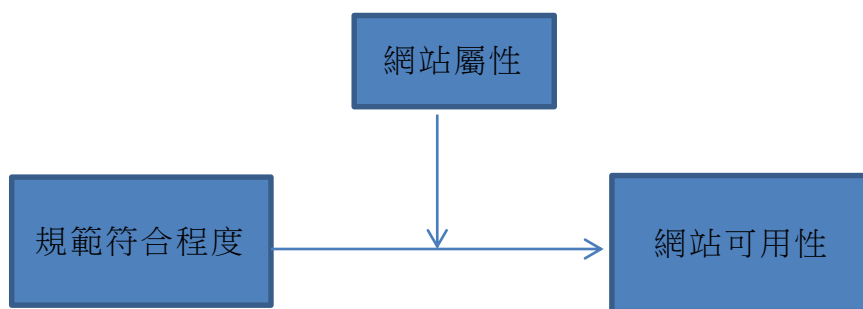
參、問卷設計分析架構

一、分析架構一



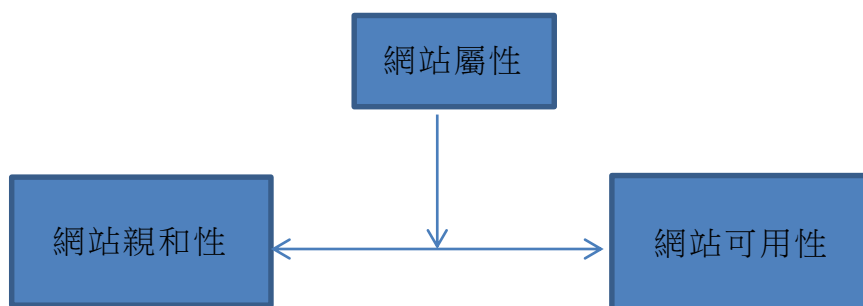
資料來源：本計畫。

二、分析架構二



資料來源：本計畫。

三、分析架構三



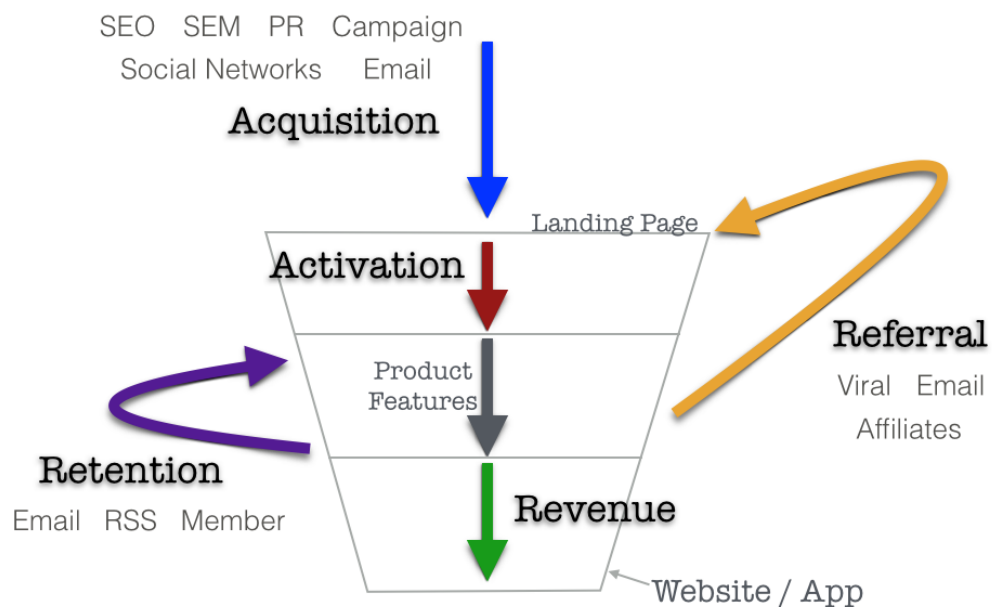
資料來源：本計畫。

四、分析架構四

著眼在描述性統計，內容包括：網站屬性、網站設計、對使用者需求的認識現況、對使用者需求的未來期許、對 GA 的資訊需求。

五、AARRR 架構、GA 設定需求與問卷題目

AARRR! metrics



by George Chang, modify from Dave McClure's Startup Metrics

資料來源：<https://blog.alphacamp.co/2014/12/01/growth-hacking-aarr-metrics/>，最後檢閱日期 2015/11/06。

題組：三、(八) 請問您對於下列網站使用資訊的需求程度為何？

AARRR 架構	問卷子題	對應之 GA 功能概述
Acquisition (獲得使用者)	1. 瞭解網路民眾從什麼管道得知我們的網站。	1. 設定目標對象 (Audience) 分析

附錄五、 中央政府各機關網站服務現況調查規劃

AARRR 架構	問卷子題	對應之 GA 功能概述
	<p>2.釐清網路民眾來源管道和網路行銷活動的關係。</p> <p>3.蒐集我們的網站到訪者的個人資訊(如性別、IP位址、居住地、到訪時間、停留時間、瀏覽器、電子裝置類型、作業系統等等資訊)。</p> <p>4.瞭解網路民眾在 google 搜尋了哪些關鍵字才找到我們的網站(訪客站外搜尋關鍵字分析)。</p>	<p>來觀察新訪客。</p> <p>2. 設 定 攬 客 (Acquisition) 分析來觀察網路行銷活動、AdWord 廣告活動、Google 站外搜尋和網站訪客到訪行為的關係。</p>
Activation(成為有效的使用者)	<p>5.分析我們的網站到訪者實際使用網站服務的情況(如網站最熱門網頁、服務或最常被下載的檔案)。</p> <p>6.分析網站訪客曾在我們的網站搜尋哪些關鍵字(訪客站內搜尋關鍵字分析)。</p> <p>7.分析網站訪客搜尋關鍵字(包括站內與站外搜尋)和網站服務的對應關係。</p> <p>8.分析網站系統的回應速</p>	<p>1. 設 定 行 為 (Behavior) 分析來觀察訪客行為流程、網站內容、網站速度、站內搜尋、觸發事件 AdSense 廣告成效、網頁活動以及不同版本網頁實驗情況等相關資訊。</p> <p>2. 設 定 轉 換 (Consersion) 分析來設定網站服務目標。</p>

AARRR 架構	問卷子題	對應之 GA 功能概述
	度。	
Retention (留住使用者)	<p>9. 分析網站訪客的再訪頻率。</p> <p>10. 釐清網站訪客再訪率和行銷活動的關係。</p> <p>11. 區隔網站訪客類型(如依據資訊瀏覽或服務使用行為、訪客來源地點、再訪頻率、來訪時間等資訊區隔訪客類型)。</p>	<p>2. 設定目標對象 (Audience) 分析來觀察舊訪客再訪率、造訪頻率、再訪問隔時間、瀏覽量、參與度、裝置、作業平台等資訊，並可進一步設定興趣分析來區辨訪客感興趣的廣告類型、有意願使用網站服務的訪客類型。</p>
Referral (使用者推薦他人)	<p>12. 瞭解網站訪客推薦他人使用我們的網站的情況。</p>	<p>4. 設定攬客 (Acquisition) 的推薦連結分析來觀察訪客到訪管道 (Channels) 以及所有推薦連結流量分析。</p> <p>5. 設定行為 (Behavior) 分析來觀察新訪客到訪和原訪客推薦連結的關係。</p> <p>6. 設定轉換 (Conversion) 分</p>

AARRR 架構	問卷子題	對應之 GA 功能概述
		析來觀察社交網路推薦連結的參訪流量。
Revenue(使用者付費或執行預期行動)	13. 分析網站到訪者瀏覽資訊或使用網站特定服務的操作過程(訪客的進入點、停滯點和離去點)。 14. 分析網站預期績效指標的達成率。	3. 設定行為 (Behavior) 分析來觀察訪客的瀏覽流程。 4. 設定轉換 (Consersion) 分析來觀察服務被利用前瀏覽歷程(含熱門瀏覽歷程)、網站服務利用次數、網路行銷功勞歸屬等資訊。

資料來源：本計畫。

六、問卷設計參考資料

(一) 自行設計

(二) 美國政府網站可用性規範 (<http://www.usability.gov/what-and-why/usability-evaluation.html>)

(三) 國發會「政府網站版型與內容管理規範」(網站親和性、規範要求) (<http://www.webguide.nat.gov.tw/index.php/ch/speci/>)

對此規範的評論(<http://buzzorange.com/techorange/2015/02/04/how-to-improve-gov-websites/>)

(四) 鄭江宇、張桂榮(2015) 掌握行銷新趨勢你不可不知的網站流量分析，新北市：新頁圖書股份有限公司。(GA 相關內容)

（五）AARRR 分析架構：Dave McClure “Product Marketing for Pirates: AARRR! (aka Startup Metrics for Internet Marketing & Product Management)”

<http://500hats.typepad.com/500blogs/2007/06/internet-market.html>

（六）新創事業一定要掌握的成長指標－AARRR Metrics

<https://blog.alphacamp.co/2014/12/01/growth-hacking-aarr-metrics/>

（七）GA for AARRR <http://blog.denivip.ru/index.php/2014/04/google-analytics-for-pirates/?lang=en>

（八）Coleman, R., P. Lieber, A. L. Mendelson & D. D. Kurpius (2008) Public Life and the Internet: If You Build a Better Website, Will Citizens Become Engaged? New Media & Society, Vol10(2):179-201.

七、問卷調查期間：2015 年 11 月 9 日至 11 月 20 日，調查工具為 surveymonkey 網路問卷平台，建議調查前先找 5 位機關網站負責人協助前測。

八、問卷發放方式：公文（內容請參考肆），請製作問卷網址以及網址 QR code 當附件。另外，中央機關一～三級機關名單請確認。

附錄六、政府電子治理內涵與品質之前瞻發展焦點座談相關資料

一、焦點座談邀請函

各位專家學者勛鑒：

電子治理研究中心研究團隊，刻正執行國家發展委員會委託「政府巨量資料分析與政策端應用效能提升之研析」乙案，為瞭解我國在巨量資料分析與應用的發展現狀與未來趨勢，誠摯邀請您參與專家學者座談會。

此次座談希望能廣邀我國公部門、實務界、以及學術界等各方先進參與，在此期盼您能參與此次座談，而得藉由您的意見分享和共同討論，協助研究團隊對於電子治理的發展趨勢，以及未來政府的電子治理改革應有的策略與方向能有更深入的理解和多元的想像。

本座談會預定於國立政治大學公共行政學系（台北市文山區指南路二段64號，綜合院館南棟11樓271149教室）召開，但為瞭解專家學者出席意願與出席時段，還請您就下列選項擇一回覆本案研究助理（聯絡資訊如下），研究團隊將安排最多專家學者方便出席的時段安排會議。。

- ☐ 105年2月19日(五)上午10點至12點
- ☐ 105年2月25日(四)上午10點至12點
- ☐ 上述兩個時段皆可
- ☐ 謝謝邀請，但不克出席。

二、座談會討論提綱：

- (一) 請討論未來十年，大數據、開放資料、與群眾外包、或其他等網路新興科技發展的趨勢。
- (二) 從管理的角度來看，公部門與私部門在推動網路新科技變革上有何異同？為甚麼？
- (三) 台灣政府近年所推動Web2.0時代電子治理方案所遇見的問題，其改革的方案為何？
- (四) 新政府上台後，台灣公部門電子治理改革，是否有哪些值得追求的新亮點？