

行政院主計總處「政府資料開放諮詢小組」 106年度第2次會議紀錄

壹、時間：106年12月22日(星期五)下午2時30分

貳、地點：行政院第3會議室

參、主席：蔡召集人兼委員鴻坤

紀錄：蔡宏宜

肆、出席人員：林委員金龍、林委員恭正、徐委員偉初(請假)、張委員源俊、梁委員德馨、黃委員叔娟、張委員惟明(戴專門委員美英代)、李委員國興、呂委員秋香、許委員碧蘭(請假)、葉委員滿足(徐專門委員健中代)、陳委員憫、潘委員城武、蔡委員孟哲(請假)

列席人員：尹副處長慧珍、李副執行長佳航、林高級分析師文遠、莊高級分析師雄鈞、黃科長慶裕、陳分析師婷妤、陳分析師信成、蔡設計師宏宜、龍研究員運濤

伍、報告事項

一、上次會議決議事項辦理情形，報請鑒察。

決定：洽悉。

二、本總處106年7至12月資料開放業務辦理情形，報請鑒察。

(一)張委員源俊：

請問107年資料開放應用獎比賽會如何公布，可以鼓勵不同的人來參加，請問有怎樣的規定，比賽結果會公布在哪？

(二)黃科長慶裕：

有關資料開放應用獎評獎作業是由各機關主動推薦，國發會預計在明年4-5月左右會函請各機關提報參獎申請，6-7月由機關提報優質資料應用案例，9月左右由評審委員評獎，其中20%分數是由民眾票選，詳細作業方式可參考本會議資料第14頁。

決定：洽悉。

三、本總處開放資料集之相關應用案例報告

(一)林委員恭正：

有關零時政府 gov 所製作之中央政府總預算視覺化資料內容，該資料經過彙整處理後可能會與原資料內容略有差異，引用該資料時應審慎評估。

決定：洽悉。

四、本總處大數據研究案研究成果報告。

(一)潘委員城武：

本研究案所選擇的三個主題均是將來可能的應用，如由應用既有歲計會計資料進行未來預測、數據輿情提供與資料去識別化等。目前研究案已告一段落，後續將藉由這些研究經驗成果進行應用。

(二)林委員金龍：

輿情資料來源是由各媒體、PTT 及臉書而來，但如部分媒體有防制機制，並未開放資料庫，且抓資料時間較短無法獲得歷史性資料，請問如何因應？

(三)黃科長慶裕：

本研究案輿情分析部分是由研究團隊提供的工具雛型，並已與工具內可搜尋的媒體來源擁有者洽談授權方式以獲得資料內容。

(四)梁委員德馨：

臉書資料可設定瀏覽權限，需擁有者設定為開放才可瀏覽，並非任何資料均提供他人瀏覽。

(五)張委員源俊：

1. 很多機關用大數據來進行決策，提醒因工具是隨手可得，只要能取得資料，每個人可以作自己的模型，如何去評斷及審查分析結果的正確性。尤其是將結果提供政府部門進行決策的時候更要特別小心。
2. 另外在 Text Mining 部分，因為現在的報章雜誌都明顯的有立場，尤其是臉書，會有很多有特定立場的人會踴躍發表意見，所以會有很嚴重的選擇性誤差，這些資訊要怎樣參考，參考到怎樣程度，應該要有另外一個評估的機制，才能知道可信度有多高。

3. 有關去識別化問題，內部必須要有一個評估的機制評估去識別化是否有效。

(六) 蔡召集人兼委員鴻坤：

主計總處是幕僚機關，並非決策機關，且主計資訊處人力有限，很多系統都要委外，有關大數據系統要先考慮要不要做，再來考慮怎麼做，如果根本不要做，就不用再多花錢或人力時間。

(七) 潘委員城武：

目前在下一階段，我們要做的是歲計會計管理報表的視覺化的呈現，也就是先把報表格式轉成視覺化，可以有更活潑的呈現方式。

決定：洽悉。

五、美國「證據基礎決策委員會」總結報告。

(一) 徐專門委員健中：

1. 美國也面臨因法令規章限制，造成機關間甚至機關內資料無法串接與後續處理的問題。機敏資料使用是以統計為目的，而且是要在安全的環境之下才能夠進行，因此報告內有一個核心重點，就是要建置國家安全防護資料服務(National Secure Data Service, NSDS)機構，並建議修法賦予權力與豁免，主要是做服務而不是建一個倉儲系統，而且是在授權且核准的計畫下，進行暫時性串接相關資料，串接後在一定期限之內完成後資料也不保留。另也針對 OMB、聯邦政府提出組織功能調整建議。主要目的是要提升機敏資料取得連結、分析與後續應用的效益。

2. 整個報告提出規劃藍圖，後續實際作業，則有待華府智庫兩黨政策中心來落實及推動。

(二) 蔡召集人兼委員鴻坤：

1. 美國從1980年開始建立書面文件減量化，所有統計調查及政府要民眾填的任何資料上面都要寫需要花多少時間，每年檢討成果並提交至國會。

2. 另外報告中提到2014年決定進行盤點，建立聯邦政府現有資料清單

及背景資料，用一個清單讓大家知道政府所有資料庫的格式、欄位及儲存格式以方便有需要的人使用。

3. 建議我國政府建立類似 NSDS 機關，評估計畫是否需要這麼多預算，還有預期效果，如果評估效果不理想就不給預算，所以該委員會的出發點就是為了預算，後來發現只要統計做的好自然可以解決預算問題

(三)梁委員德馨：

非常認同不再成立倉儲系統，而是有意義的處理與協助，在期間內完成資料介接不進行儲存。

(四)林委員金龍：

1. 對於不同格式與來源的資料，如何彙總、互相勾稽，個別資訊的使用權限、隱私權及資安疑慮等問題均是屬於資訊科學領域。
2. 最重要的是如何利用資料。同樣資料經由不同使用者可得到不同結果，要由資料出發去彰顯公共行政政策執行效果並予以量化，牽涉到分析基本理論與架構，難度頗大。
3. 我國目前已建置的政府資料開放平臺，受到很多人肯定，建議可尋求資料科學方面或統計方面的學會合作，召開專門會議，探討及推廣政府開放資料平臺的應用。
4. 目前已有人以 R 軟體建置搜尋介面，協助使用者搜尋資料平臺資料集，並將蒐尋到的資料匯出，對開放資料的推廣應用很有幫助。
5. 網路上有學術文獻搜尋資料庫(如 DBLP)，由學者收集後放置於網路提供查詢，每日更會將所有索引資料壓縮後提供下載。網站上亦提供用多種程式語言寫的相關程式，對於分析索引資料庫有極大的助益。因目前政府資料開放平臺資料量多，但並未提供整合的搜尋工具，建議可在平臺上提供使用者上傳相關程式碼與分析成果，以方便其他有需要的人使用。

(五)張委員源俊：

製作資料清單是一件很重要的事情，如果使用者可以取得政府部會資料庫建立的資料清單，就可以大約知道政府已蒐集哪些資料以及該如

何應用。

(六)陳委員憫：

本案簡報介紹的 NSDS 運作流程，目前衛福部資料科學中心已經完成大約八成到九成，包括資料去識別、清理、揭露及審閱等機制，只有在最後發布的部分，有一些考量所以作法比較保守。當初未建置倉儲系統的原因除了成本因素外，亦有資安考量以避免各單位有一些疑慮，目前已成功整合該部內各單位及所屬機關資料庫，也開放某些跨部會的重大計畫使用，相關做法很值得參考。

決定：

(一)請主計資訊處潘處長將本報告適時提供國家發展委員會參考。

(二)洽悉。

陸、討論事項

一、修正「行政院主計總處資料開放管理要點」(草案)，提請討論。

決議：

(一)請將本計畫第十三項，有關「績效卓越」文字酌作修正。

(二)文字修改後通過。

二、修正「行政院主計總處資料開放具體行動計畫」(草案)，提請討論。

決議：照案通過。

柒、散會：下午4時20分