

2015 年

犯罪防治研究

大數據運用在親密關係暴力
犯罪防治分析之探討

研究實錄

發現趨勢、
創見未來

中華民國104年12月

目錄

1.大數據運用在親密關係暴力犯罪防治訪談紀錄(王珮玲教授)	3
2.大數據運用在親密關係暴力犯罪防治訪談紀錄(廖美鈴主任)	6
3.大數據運用在親密關係暴力犯罪防治訪談紀錄(李御璽教授)	10
4.大數據運用在親密關係暴力犯罪防治訪談紀錄(蔡俊輝副主任)	20
5.大數據運用在親密關係暴力犯罪防治訪談紀錄(陳泉錫處長)	24
6.大數據運用在親密關係暴力犯罪防治訪談紀錄(張秀鴛司長)	30
7.第一場焦點座談	34
8.第二場焦點座談	41
9.第三場焦點座談	51
10.期中報告審查會議紀錄	66
11.期末報告審查會議紀錄.....	69
12.2015 年犯罪防治研究學術發表會--發現趨勢、創見未來	73
13.研究案活動照片	80

1. 大數據運用在親密關係暴力犯罪防治訪談紀錄(王珮玲教授)

訪談對象:暨南大學社會政策與社會工作學系王珮玲教授

訪談時間:104 年 4 月 17 日 14 時

訪談地點:內政部警政署秘書室專門委員室

從親密關係暴力防治現況開始，坦白講現在親密關係暴力案件的特性是甚麼，是從有限資料中去了解，所謂有限資料是指透過通報資料去了解，可是通報資料在我們全部發生案件中它還是少數，大部分案件還是沒有出來，所以從現在有限資料來看的話，大概會看到一些特別在某些族群上是有比較高的狀況，比如說在新住民這塊比較高，女性受暴比較高，年齡層大概就是 20 幾歲到 40 歲期間，所占比例最高，這是就特殊性。可是有些在文獻有提到，但資訊未顯現，比如區域的差異，因為國外發現到城鄉會有差異，因為城市的狀況與鄉村狀況不太一樣，包括固守傳統的價值觀念，也可能家族會聚在一起，加上司法人員包括警察，在鄉村地區相對保守，喜歡介入或不喜歡介入，所以在國外城鄉差異，有些人在做比較鄉村地區家暴案件求助困境及司法反應、還有資源到位，這都是存在問題，但現在在台灣還是看不出來，只能從人的發生率、當地人口數來處理，但這些是不精準，因人是移動的，台灣這麼小，不見得在當地通報，他的戶籍就在當地，這就是有很多的問題，有沒有這狀況不知道。再來就是說，台灣還有二個不知道的地方，其中一個對目前，坦白說這次修法是比較好的，它將範圍擴大到情人之間、青少年，從英國全國調查資料，就可看到青少年親密關係比率不低於成年人，我們國內調查已出現從國中開始與異性交往比率已經蠻高的，到高中更高。到大學那就更不用講了。這些部分從青少年他們未成年，他們與異性交往就有親密關係可能性，他們的嚴重性、狀態…我們也是不知。另外不知的是老年人那一塊，老年人出來通報的比率更低，原因就是認為已經這麼久了已經習慣了。以時代因素他們在那年代工作比率相對少，所以他們在經濟獨立沒那麼高，所以對施暴者依賴就會更多，再加上不要給子女難堪，所以對於這一塊，成年人 20 到 40 歲是服務對象比較多，但這二端的在現況了解上是不足的。

國外在講說親密關係有不同類型，不是每個案子都是長得同一個型，所以對加害人有不同分類。對於案件有不同的分類，以案件的特性，如一般的爭吵，或親密恐怖，或者是甚麼，對於施暴者他有不同分類，各式各樣不同標準所做出分類，也有對案件做出不同分類，這部分我們在台灣還沒有辦法掌握。可能這些未知的與大數據分析也許有連結，但現在我沒辦法想到如何從這裡到這一步，可是這些是我們在做這些工作時所不知道的，但是是很重要的。加害人的分類大家都知道，但分類比率怎樣，我們也不知道。不同類型有不同特性，比如說脅迫控制，它的特性在權力控制就是一個核心，他的暴力行為就有控制性行動、跟蹤騷擾的

行動、或性暴力的行動，這些都會與類型有關係，這些東西的掌握，對您的處理是有幫助的。我們前面有很多東西是未知，很多未知就會影響後面的防治處理。所以那些因素在親密關係暴力的防治上是重要、關鍵的？我很難告訴你少數的，我覺得很多都很重要，沒有辦法說只有幾個比較重要。

現在對案件的族群、性別、年齡掌握都是靠通報資料，年齡會影響通報、社經地位會影響通報，因為我們資料都在有限通報資料當中，衛福部現有一科技計畫，要去作台灣大規模的 survey 調查，如果那東西開始做的話，我覺得對於台灣整體現象掌握有幫助。

恐怖情人之類的親密關係暴力遽增現象，很難講是遽增，也不太相信是遽增，是我們有沒有去了解或不知道，恐怖情人是各式各項的都有存在，只是現在沒有納入通報系統裡面，所以這個發生率未知，就無從比較，就無法確定是否遽增，但這事情本來就存在，國外文獻指出，只要有交往就有暴力可能，這是個很自然的事情，只是我們不知道它的狀況是什麼，所以這本來就存在的。雖然新聞媒體曝光量有增加，但我不認為可解讀為遽增，因為本來是多少不知道，就像每多人講說，家暴案件一直在成長，從來沒有掉下來過，那能說家暴工作越做越糟糕嗎？當然不是，在做調查時會發現黑數是一大堆，可是有事件發生去促發大家意識到這是個問題，這個不求助會更有問題，所以每次發生事件時，通報就急速增加，那不是現在才有這種事情，而是社會以前沒意識到、關注到，也就沒有通報。不能以通報數字增加就說狀況是越來越糟，正向解讀是大家越來越願意出來求助。這不是個壞現象。

對事前的訊息可以預防被害發生而言，若這些訊息大家事前知道，但沒有警覺而錯失了。這是對此議題了解認知的問題，大家都已知道訊息，但沒有意識到問題的嚴重性，沒當一回事，這就要回到怎樣讓社會大眾知道，當你的朋友、親密伴侶出現這樣時就是個危險徵兆。是一種認知教育部分。我要談另外一部分知的權利，在英國試辦中，因曾發生一個案子，它的未婚夫(或同居人)對她施暴，她也申請保護令，但未婚夫違反保護令強暴她，她都有向警察報案，但最後仍被殺。女方家人非常不諒解警察，她多次報案時警察都有掌握到這位未婚夫曾對前幾任太太、女朋友、同居人有傷害、甚至嚴重傷害紀錄，被告判刑服監數年，警察從紀錄上知道這個人是危險的，但警察沒有將這資訊告訴她，他是危險人物。家屬非常質疑為何都沒有給她提示、警告，讓她一步步陷入險境。此事後經檢討也經倡議，以後民眾對這種安全有知的權利，有二種管道，如果我交往對象有暴力傾向，可去警察局申請調查他以前有無家暴紀錄，但基於個資法，是要經討論認為是高危險的話，就可將訊息告知申請人，這是人民有查詢的權利。另一管道是，她不知要申請，警察可將資訊提報討論後認為有高危險，就可以主動告知她。這方案自 2013 年開始試辦。所以，這是種事前訊息的告知，讓她去預防被害。

這是否與大數據有所關連，目前尚不知曉。

現行在親密關係暴力防治有效途徑，範圍是很大的，教育部現在正在推動情感教育，因青少年到一定年齡就會有情感交往的事情，這是很自然的，所以老師都認為應該教，但家長卻持反對意見，但我們認為這是很自然、擋不住的，應該教，從預防性的觀點這種情感教育就是要進來，這是一級預防，二級預防對一些有危險可能的個案，去發現他們，讓他們得到協助、支持、輔導及幫忙，三級預防是對已經有暴力行為的人，由學校、家長、社區網絡一起合作對已經有此行為的人要做什麼事情。另外教育部也在推動校園內危險評估工具。

2. 大數據運用在親密關係暴力犯罪防治訪談紀錄(廖美鈴主任)

訪談對象:內政部警政署資訊中心廖美鈴主任

訪談時間:104年4月17日14時30分

訪談地點:內政部警政署資訊中心會議室

訪談回應內容:

本來覺得我們的東西是否符合這研究的需要,後來覺得即使做這樣研究也會與我們所做資料庫給我們一些想法,從學者的角度,與從我們內部的調查的角度,我們要收集什麼資料,我們有想過,那時候請外縣市同仁來的時候,他們提列了二十八項資料庫,建議我們以署的身分去向誰要,然後一起納進來作分析。這些納進來的資料是他們在偵辦過程所需要的,要去拜託人家的,不管是醫療資料、勞健保資料...,他們都要以個案一一去拜託人家。所以,假設我們有能力將這些資料都放到我這邊來,那我們也會被質疑是個恐怖國家、警察國家,他會認為你警察的權力太大了,現在還是要透過一些機制的。那大數據是這一二年用在bussiness比較多,它試圖是找出原來結構式資料庫所不能做到的分析,它有很大好處,原來我們是從結構式資料庫分析,受限於欄位,但有很多資料它是非結構性的,它可能是開放的填寫,我們現在受理報案,有所謂發生事實,它就是開放式填寫,而開放式填寫有很大的缺點是,每一個人表達的方式不一樣,有些人時間寫前面地點寫後面,有些人顛倒相反,表達方式、用語也不太一樣,這些東西對於未來要做分析,會是個挑戰。中文是個非常豐富的東西,當您要去用電腦做個分析,電腦無法去判斷,您講的真正意思是什麼,還有很多那些隱藏或隱射的字眼,那電腦無法去做這些判斷。但是,我們還是有需要的,因用固定格式去做分析仍有不足的地方。刑事局曾做過一小小的test model,他們覺得原來的刑案資料庫就是固定式的,雖然它有data mining的概念,仍受限來源是固定格式,所以,他們也覺得不夠用。因為現有大數據技術,所以他們做個test model,把移送書資料與原來資料庫內資料、筆錄開放式問答放進去做測試,發現筆錄裡所記的人,比如聯絡過的人、毒品案相關人,卻未必出現在移送書上,他們就去思考說,這是很奇怪的事情,為什麼這些人可以重複在筆錄上出現而沒有被移送,他們應該是很重要關係人,否則怎麼沒被移送,他們就有個想法,那些人以後被分析出來,就是他們要去找的人。雖然不知道哪些人代表甚麼,但是他不斷在筆錄出現,maybe他就是個組織、上源,只是這些人他應該很了解法令、很懂得去保護自己,所以他沒有被發現證據違法的資料。刑事局已經開始做這個的想法。那我們做大數據都是偏重說,在犯罪偵查部分,我們從這麼多資料裡面,把這個人跟其有關的找出來,如生活動態、生活習慣。在過去我們是單筆單筆去查我們沒有看到這現象。我們偵查角度就是將與人有關的、與車有關的,都把它找出來,

讓以後我們同仁去釐清線索的時候，他很快就可以切入原來這個人經常有活動是那些人，而且我們會把一些動態資料放進來，這樣就有助於我們同仁加速，他不需要自 28 個資料庫去萃取資料，28 個資料庫有的東西我都已經倒出來了，那他就可以從這個大的 pool 哪面去，下個關鍵字去搜尋分析。同樣的，國外有很多的研究，犯罪預防的部分，從現有發生時間、地點、範圍、犯罪習性，他可以預測說下次犯罪可能什麼時間出現，如洛杉磯警方就想拿這判斷結果作勤務佈署，因為警力很有限，根據此分析結果他就把警力佈署在那幾個犯罪可能出現的地方，而且把可能犯罪的這些人通知轄區分局，就是我們所稱治安顧慮人口，說這些人已到轄區活動了，根據我們資料庫分析，他應該會在這段時間跑到你這邊有犯罪行為，然後這些警察就瞪大眼睛來找這些人，來看所來是不是這個人，會不會在預期的時間去做犯罪，這樣犯罪率真的是下降。如果我們資料庫有這樣效果是非常好的，所以我說是偵防並重的資料庫，因為我們從這麼多資料庫去把資料收集起來。同樣的，現在除犯罪偵查之外，另外一塊是很多人在做的，從網路上的臉書、社群...，將近 5000 個網站、粉絲團等等，這種資料是更散的，無法要到的，但有很多工具去搜尋，現市面上有廠商專門去搜尋這些資料之後去做分析。分析什麼，今天很多人關心此事，他是正面或是負面，列個表，提供主官去判斷這件事情嚴不嚴重，要不要回應或堅持等政策決定。所以，我們也希望我們的大數據的未來，從網路上搜尋資料，但現在年輕人語言甚難聚焦分析，但如果有業者已能分析相當成熟，我們就可以拿來跟我們的資料庫再作 match，這樣我們得到的預測或分析會更準確。現在在做是這三大方面，一是偵查、一是預防，再來是未來智慧分析。

現在最不能預防的是那些不特定對象，您也不知道他有沒有犯罪動機，也沒有犯罪前科的那些，我們都知道百分之八十的犯罪集中在百分之二十人口，警察偵查犯罪大多從這百分之二十找起，至於說不能預期的，則不能事先知曉，你要預防這些人犯罪、有沒有犯罪動機，要做此類分析，或可以在資料庫去找試試看，但效益有待評估。你的 data 來源，要去從不同資料庫去找出它的關聯性，什麼人會來做什麼犯罪是有關連性的，要將這些潛在的人找出來，這是大家要用智慧。搞不好要從臉書資料找，有些人的情緒表達是比較激動的，或常酗酒的、或失業致經濟有狀況的、或有精神疾病病史的，或從毒品矯治裡面，將關聯性建立起來，那是很豐富的，或可歸納集中出某部分人。第一類是有前科的，第二類是去找的，從有歷史的紀錄不一定是犯罪紀錄，去找出來關聯性。就像大數據之前也是找購物行為，提升銷售率。例如男生去買尿布時會想買啤酒喝，東西就可擺一起，提升銷售率，這是以前所未想到的，這就是大數據要去找的。所以說，親密暴力的界定要在哪裡，要找的資料庫可從哪裡去找，或從要得到關聯性是什麼再去反推。但要去找出這些資料就要花相當大功夫，連我們警察要去找都會碰到相當的困難。現在資料是非常多，像悠遊卡公司要將悠遊卡改成記名，去找出人的活動軌跡，但有隱私問題，最好善用現有的。現有的數位 data 太多了，大數據就是要從中

找出更有用的，有更大價值的。對有關聯繫資料有哪些、放在那邊要做調查，再去向其要資料，但現在個人資料保護法實施後，向其要原始資料是有困難的，不會將整份資料抄給您，只允許你單筆單筆跟他查。

要做大數據分析用整個倒入統合資料庫是比較好的，若個別式就要考慮你如何去連，要什麼資料，那系統設計會更複雜。現行政院針對學術研究部分，由國發會主導，將去識別化的資料統一放在國家實驗研究院的國網中心，提供學術研究申請用，就不用到個別機關去要了，現在正進行中。

親密關係暴力若是在情侶之間，則這些大多是無前科的、或學生、上班族，比較難掌握。比較能掌握的是家暴、兒少性侵，因家暴有些特徵，如毒品、酒等等，另弑父弑母，也可以根據其就業情況、菸酒、交通違規，就比較能掌握。但家暴案件在刑事部分有些表述是經過隱蔽的，不像一般刑案，但如果其已經殺人，就是殺人案件，但性侵害案件就會用代碼。如要找性侵殺人，則與刑案和在一起，要去找也是一個難題。有些資料太雜了而且來源、格式也不一致。若從E化報案開始就是勾稽它的案類，或也是解決方法。另，恐怖情人資料分析是比較難的，除非有上年紀、有前科紀錄，此與家庭暴力、兒少性侵、弑父弑母這些都可以抓在一起，就是從他前科資料、交通狀況、報稅、結離婚紀錄等等去看，現社會上有些案例是單親媽媽，又有吸毒、與男朋友一起住，女兒被侵犯，還可以找；但單親父親自己帶小孩性侵女兒的，就比較難找，除非這個父親有酗酒的問題等等。另外針對高風險家庭，曾看過報導指出，幼童時期曾遭受家暴的，比較容易於長大後有家暴傾向，如果這樣歷程可在哪資料找到的話，就可列入高風險有暴力傾向者，除刑事資料外還可將家暴中心的資料到入，大數據不怕資料多，也不怕資料格式不同。但要找實際處理經驗的人，才知道他的判斷是甚麼，需要那些資料。而每個每個的判斷模式不同，把它放進去就會變成 general 的，判斷的通則就會比較準確。這種 know-how 在大數據是很重要的。

政府不同機關間資料連結不是那麼容易外，法令(個資法)限制、資訊安全的問題還要面對，另外資料正確性問題，會影響到將來分析的結果。此外，要將要來資料找個地方放置分析，還需經費來建置。

美國紐約市前局長轉往加州警察局任職八年，期間做了些犯罪預測的 model，後又回任紐約警局局長，也正推動這些的犯罪預測。他強調自從加州警局用此套預測，犯罪率有下降。目前在台灣我們是第一個起步的，但在商業已做二年了，商業是最需要去比對你的購物行為、你的喜好，據此，從擺架位置、price 都要去調整，很多東西放在那没人要，就會淘汰掉，因涉及店面使用率及利潤。所以，商業的 model 是最多的，分析購物行為是比較容易的。家暴案件大多是隱藏性，今天吵架到派出所，隔天就又到和好，如有通報至家暴中心，才能就現有資料庫內做分析，現家暴中心是有很多資料，但是要想到與其他資料庫的關聯性，可從一

筆一筆的關聯性蒐集起來，再來重新設定關聯性，一步一步來，而且大數據可不限格式資料，就可將訪談等資料納進來。

3. 大數據運用在親密關係暴力犯罪防治訪談紀錄(李御璽教授)

訪談對象：私立銘傳大學資訊工程學系暨研究所李御璽專任教授

訪談時間：104 年 4 月 20 日 10 時 10 分至 11 時 50 分

訪談地點：銘傳大學資訊大樓會議室

訪談者：范國勇

一、犯罪防治

面對政府要求資料開放、大數據等政策及國際發展趨勢之下，您認為大數據運用在犯罪防治的公共議題上可行性如何？發展的方向？

政府不同機關間的資料庫連結(或資料釋出)，您認為要注意哪些重點？可能會面臨哪些困難？該如何解決？

您認為政府或是社會上有哪些資料(或資訊、資料庫)是與犯罪或親密關係暴力有關的？或是對犯罪預防或親密關係暴力防治有用的？有哪些是基礎的條件(例如資料的建置程度、資料內容、個資保護…)？又需要哪些配套措施？

近年來恐怖情人之類的親密關係暴力遽增現象，您認為運用大數據分析可行性如何？可能受到的限制或面臨的困難為何？如要運用大數據分析，進行的步驟應先從哪些方面著手？

(一)面對政府要求資料開放、大數據等政策及國際發展趨勢之下，您認為大數據運用在犯罪防治的公共議題上可行性如何？發展的方向？

1. 監視器資料蒐集是為大數據資料，並可運用於犯罪防治

犯罪防治，我是覺得像以往跟現在其實就有很大的不同，那我常聽我弟弟他在講，他就是說以往是靠地緣上的一些關係然後去破案，那他像現在不一樣，他現在很多都依賴那個監視器，以往就是沒有收集，那監視器收集下來其實那個量也非常可怕，那我自己知道說之前，您可能也知道，就是有一個蠻牛下毒案在一週之內就破案，因為他就看那個歹徒在哪個地方打的電話，其實這個電話的資料也是一個大數據的資料，在哪個地方打的電話確定他的地緣的位置、地理的位置，然後就去調超商的那個錄影帶，那個錄影帶的收集其實也是一個大數據……所以那他那時找我去是說，有沒有辦法，開發一種系統就是輸入錄影帶那輸出告訴我說哪幾個人可能是，比對的這種，那他說因為蠻牛是一個大的一個指標性的案件…

2. 使用車牌辨識系統提高破案率

…但是有這樣的一個大數據分析的系統，是不是我們以後就可以普遍用在各個地方，尤其現在地方到處都有攝影機，那如果有辦法做一些攝影機裡面的人物的、車的一些辨識，將來這個就很有幫助，那我知道警政署像這種車牌辨識，他這個應該已經比較成熟一點，現在雖然沒有收費站，很多地方他其實也再照，他照的目的其實將來就是，尤其現在你如果用的是 etag，你現在 etag

一上去，他馬上就知道你人在哪裡，那以後如果你車子不見了！但是我知道我 etag 號碼，他只要敢開出去，敢開到高速公路，馬上就知道那輛贓車現在在哪裡流動，從哪裡要走到哪裡去，那這個對於以後要抓這個贓車，我覺得現在所有東西如果電子化之後…

3. 運用網路社群媒體有助於犯罪偵防

…前一陣子幾次抓到要犯，都是他一上去 facebook 一登入，馬上人家就鎖定他的定位馬上就抓到，像西門町柯 p 那個他是上 facebook！馬上就知道，那 facebook 以前也不是以前的東西，他 facebook 上面的資料也是一個大數據！所以現在這種犯罪防治，其實我是覺得如果東西漸漸的數位化電子化，之後這個犯罪防治是無所遁形！

4. 記名悠遊卡及智慧型手機為大數據發展之方向，並可運用在犯罪防治

那像我弟弟最近跟我講說，他們現在好像警察好像如果用悠遊卡去記名的話，到便利超商好像有那個優惠！他說那我要不要去記名那，我就跟他說你最好不要去記名啦（台語）！就說我們都在做這個分析嘛！你這樣一記名，他就你怎麼這個時候跑去買咖啡，對不對馬上被人家逮到，我就跟你說最好不要去記名（台語），你不要貪這一點小便宜，你搞不好你只是去買個咖啡就回來，人家就給你冠上一個一個怠忽職守什麼的，所以現在也是一個大數據的引用，就是說你現在只要有使用智慧型手機，基本上你都被追蹤，因為你不要以為智慧型手機，因為我們分析智慧型手機的數據我們分析很多，你只要拿出來馬上被追蹤，尤其如果你又有定位那更糟糕，馬上可是現在很多應用就是需要，需要你定位啊…

5. 大數據應由政府整合不同單位來源的資料

根據那個罪犯資料發現一些現象就是說，那些罪犯通常都是他的樓下犯罪，樓上就犯罪就是那個有群聚效應，不是整個布魯布魯克林區每個地方都在犯罪，但是有群聚效應就是這個地方如果犯罪很多，那就是在那一區啊！布魯克林區裡面在在，在更幾個他們說熱點區啊！那幾個整棟樓可能都在犯罪，那他找到幾個熱點區之後，他在去分析這些熱點區附近政府到底為他們做了什麼事情，就發現這些熱點區啊！附近阿！那些燈啊！街道啊！這街道凌亂就不用講啊！燈幾乎都破掉就是沒有燈，就是他認為這個已經造成一個可以犯罪的一個環境去了！所以他那時候就花了，好像也沒有非常多的，好像六百萬美金去整頓這些熱點區的環境讓他變得哇非常明亮，那個街道很整齊什麼，然後發現他其實大大的降低那個犯罪率，他發現他花得這個錢比他要這個罪犯的成本其實真得太少，那他這個就要結合了一些地政的資料，燈、街道的那些資料，那些其實這也是大數據另外的問題，就是他必須做不同來源資料的整合，不僅僅只有警政署的資料或者是公務局他那些資料，那些都要有辦法整合起來…對所以在犯防治罪這一塊大數據是可以應用得非常的多，而且這將來等於說將來的，抓罪犯的方法跟以往可能會有大大的不一樣，那這個是第一點犯罪防治…

6. 對於公共議題政策方面應做網路輿情分析

…還有第二個就是那個親密關係暴力防治的公共議題，他們這種親密關係得這種他們都有一些固定的那種社群媒體，他都會有一些那個類似粉絲團還是什麼那一種，他們固定會在哪邊 po 一些消息，有些人看了之後，很多人就會有一些意見，有些回覆以往他們就是看一看，因為有時候那些東西太多，沒有特別去分析，那時候我是建議他們說像這一種輿情啦！輿情分析，尤其是一個公共政策，像柯 p 這次會贏他就是靠網軍嘛！網軍其實他就是在做輿情分析，他很多的政策要推出來之前，他其實是先推到網路上、社群媒體上一 po 出來之後，然後很多人就會有一些看法有些意見，然後他也有一個團隊專門是在做這個輿情的收集，但是他現在目前我知道他的做法就是都由這些人來看，看完之後他收集起來就是說你這個政策有多少人持有持哪些反對意見，那反對意見的比例大概有多少？那內容大概是什麼？然後他回來綜合一下，他發現這樣子做可能反彈會太大他就會立即修正，可能一天之內就收集到這樣的資訊，但是如果這樣的東西長期要靠培養一個團隊每天在看這個，其實當他要處理的事情越來越多的時候，其實這個已經是不太可行負荷量太大，所以現在就很多那種網路，網路輿情分析的那種機制、系統

7. 網路爬文技術所蒐集之資料應先分群並找出關鍵字

這個親密暴力關我有一個政策要推出，一回來大家都在討論大概有幾種議題、幾種聲音，因為不是大家講的都一樣可能有人是講某一個方面，有的是講另外一個方面，所以我爬文回來，第一步我就是可能先做一下議題的算分群，譬如說爬了一萬個文回來，這一萬個人裡面其中有三千個是再講某個議題，另外三千個是再講另外一個議題，再來三千個又是講另外一個議題然後接下來的議題爬完回來之後，接下來我們就要自動分析，這議題裡面大概在講什麼內容，都提到哪些關鍵字，其實關鍵字大概就可以代表他的內容，有時候我們看這三千個都是就在講這幾個，大概就知道他在講什麼，所以用關鍵字的方式，然後在透過什麼文字雲啦！什麼那種方式去呈現什麼，決策者其實只要這樣稍微瞄一下，他大概就可以知道人家反應是什麼，然後再來在搭配一些正負評的統計…他們應該要隨時去做這種，輿情監控尤其是一些公共議題的部份，一定要做這個輿情監控，然後很快讓上面的人知道…所以這個親密暴力防治的這個公共議題，我覺得從我的角度啦！我覺得可以做一些這種輿情方面的分析而且那我自己之前是幫財政部國稅局，做那個電子發票的輿情監控做了一兩年，那個時候一開始是想說那個電子發票有什麼好輿情監控，後來發現其實這麼簡單的東西，他的背後還是隱含很多問題，因為這隱含很多利益上的問題…所以所有東西都電子化、透明化之後，你要逃的機率其實非常低…

8. 發票電子化後有助於減低盜領發票之犯罪問題

…他們才發現每年被假的中獎發票被領兩億多為什麼？那些人很聰明就是，他製造那個假發票，假的中獎發票，他不是做那種金額大的，他都做那種三百塊的，三百塊不用證明，你只要到超商說我，中獎他就立刻會給你，他們就是做這種三百塊，然後弄的非常的多，就一直去領，然後一年損失兩億多，那他們現在電子發票一弄上來啊！你只要那個發票，只要去兌獎他立刻上網登記這張發票已

經被領過了!第二次你再去兌他就說你這個已經領過!你怎麼又來領,所以現在這個也是大數據,因為他現在有一個發票雲這也是一個大數據,有了這個發票雲,之後所有電子化之後,他馬上就又抓到那個電子發票…

(二)政府不同機關間的資料庫連結(或資料釋出),您認為要注意哪些重點?可能會面臨哪些困難?該如何解決?

1. 資料庫連結應集中式管理且要有整體的規劃

政府單位各部門的資料,絕對不是任意兩兩單位互相做連結,他應該是一個政府要整體的規劃,然後要集中式的管理,那我記得政府現在已經有朝這個方向在做,他們現在就是找那個國家高速網路中心,因為國家高速網路中心弄了很大一批雲端設備在那邊,可是他弄了之後呢?其實他基本上使用的效率非常差,沒幾個人用,因為那是國家的那到底誰要來用你這個雲端設備,所以他們現在初步就是說,一樣讓這個雲端高速網路中心,要讓他有 function,要讓他能夠起來,那剛好我們政府要 open data,所以 data open,最好就全部到這個雲端,因為雲端這個國家高速網路中心,因為那邊他有很好的雲端設備,因為你資料量一大的時候,你需要處理的那個量,還有那個計算能力要很好,像我說那個發票雲光是建制發票雲的設備,花了兩三個億啊!在買那個設備所以就知道要,要來弄這個東西不是這麼簡單…

2. 機關之間的資料釋出要有統一的交換格式

那剩下的一些政府單位比較小的,我覺得要集中在國家高速網路中心,而且要統一的交換格式,他們現在都有統一的交換格式,就是說你資料傳不是說,我的 sql server,我就用這個 sql server,我設 oroco 我就傳 oroco 那是誰有辦法跟你連,所以他們現在傳上來,我知道他們規劃就是有一個標準的交換格式,那些沒有傳上來的那邊,他的格式也要完全改成是這個格式……其實還是同一個概念就是說,他原來的資料還是可以用他原來資料的方式儲存,但是你要送到國家高速網路中心的時候,你必須做一個格式的轉換,就像以前公文不能交換,現在都可以用電子公文交換,各部會他就是大家講好一個格式…

3. 資料量尚未足夠且未全面電子化

那現在比較需要可能要克服的就是,那他到底擁有哪些資料,我知道很多地方的資料都還沒有 key in 進去,大部分有些有 key in,有些用書面的,有些他要將這些資料要把他全面數位化,還有以前的資料要不要數位化,因為你不能說我只有最近兩個月的啦!沒有以前的資料…沒有就是根據你以往的資料有就有,沒有就沒有就老老實實的把他呈上去,把他記錄下來…但是就是說你以往那些東西,很多都還不是電子化的!

4. 資料庫連結重點是權責應劃分清楚並且要有獎勵措施

…其實各部會我覺得問題是那個主事的老闆!那個長官啦!他到底有沒有要做啊?很多人都說大數據很重要,然後真的要叫他做不編啦!他認為那個不重要,其實我覺得最大困難是人,然後你把那個權責劃分清楚,他到底下面做事的人到底有沒有權,出事要不要負責,把這個劃分清楚,有沒有想要做這個事情,因為他們認為做這個事情也沒有 credit 啦或是什麼啊去推這個,因為現在其實也不是說他們怎樣,所有老師也都這樣,看 credit 做事啊!做這個沒加分我就不做

啦!要有，但是我是覺得政府規劃鼓勵的措施要出來，要有鼓勵措施，獎勵措施沒有出來做完什麼，有什麼獎勵，沒做有什麼罰責，這個沒去訂定，跟他用口頭講沒有用，我是覺得，應該是從人那邊開始措施那邊先開始，這些都搞定了馬上動起來，因為太多人都看錢做事嘛!對越來越多，所以你只要很多東西，你有獎勵制度出來馬上前仆後繼啊!困難都不是困難了!

(三)您認為政府或是社會上有哪些資料(或資訊、資料庫)是與犯罪或親密關係暴力有關的?或是對犯罪預防或親密關係暴力防治有用的?有哪些是基礎的條件(例如資料的建置程度、資料內容、個資保護…)?又需要哪些配套措施?

1. 網路意見領袖對公共政策議題具影響力

其實這個跟以往的一些觀念又有點不太一樣，以往一個人就是一張票，可是在網路裡，網路的世界裡面你會發現，很多人他有影響力，他就是意見領袖，意見領袖講出來的，他下面的人基本上都會遵守，遵守的比例很高啦!所以你說一個人他在那邊發言不代表大家對不對?可是那一個人可能抵一百個、一千個!所以在那裡常常有很多人發言，有一些粉絲團，所以像柯 P 可能也還沒做到，就說我們現在做的分析，分析這些人的互動的關係，找出意見領袖，只要把這幾個意見領袖搞定，以後選戰我就太好弄了!因為如果這些人都支持我，他下面的粉絲基本上就會被他影響，所以這個跟以往的那個方向又不太一樣，雖然說一個人可能沒有辦法代表全部，可是 he 會影響下面的人，下面的人可能影響他家裡的爸爸媽媽，他可能不上網，可是會被他的子女所影響，所以這個沒辦法估量啦…

2. 網路輿情分析並不會有涉及個資法的疑慮

…那當然這種在分析的時候，當然是說個資的保護!那基本上現在只要牽涉到個資法，在台灣其實是都不能做啊!那但是我們現在能收集到的資料，例如說網路上像公共議題的這個部分，這個部分沒有所為個資法保護，因為那個都是 open 都可以上網去分析，而且我沒有分析到到詳細到每個人，我們只是把他發表的議題，拿出來看看說過去有多少的輿情產生，那沒有牽涉到說，針對某個人怎麼樣，如果這樣就可能，比較會有個資法的問題，所以我們現在能抓的到的資料，其實台灣現在在個資法上算做的非常的好，所以我們能抓到這樣，事實上都是已經是個資法在規範下的資料…

3. 資料的建置應分層級閱讀

不同層級的人他能夠 touch 到資料的程度，這個是要分級，像你如果開放到、open 到給一般民眾在用的時候，可能他就是一些比較無關緊要的資料，例如說這個地方有幾個停車場，或是說這個地方的一些道路資訊，那個跟個人資料是沒有關係的，那就可以開放給民眾，如果要做到另外一個部分的犯罪防治，我可能需要一些專家學者進來，那可能就是要跟我們有簽約的，那些專案他可看得到哪，到哪一個層級的資料，那如果已經要到查案了!那我必須要知道他，他們家的爸爸媽媽是誰各是住在哪裡，那個又是另外一個層次，所以資料不是 open 就沒有 open 沒有這樣，應該是說 open，但是是有分層級的，那每一個層次下有哪些規範，誰可以看誰不能看，那這個其實要有嚴格的規範，那在這個規範之下，大家就可以按自己的職責或者權力，可以去行使資料得這個分析，那這樣又可以沒有牽涉到個資法…

4. 使用者應負資料保密責任

根據那個個資法還有資訊公開保護法，那現在就是這兩個在把關！一個是個人資料保護法，還有一個是資訊公開保護法，那他要去分析那個資料的時候，就是說那資料我可不可以帶回家來分析，如果不能帶回家分析，那他是不是要提供場所給我分析，他提供場所他的資料提供機器，要提供設備對他來講又更麻煩，所以他現在那邊在解釋說，個人資料保護法是不是我們只要有簽約他就可以帶回來，洩密就是你自己的問題，那還是什麼現在其實都還在解釋討論，那現在傾向是說你如果簽完約你可以帶回來…所以他說如果要 open，請先解決權責的問題是誰該負這個責任，長官沒有一個要負責的啊！每個都說是不是我做的，不是我做的那我來做，那我對不對我扛那麼大的責任，所以變成說現在那時候 open data 一直推不起來就是，就是就推得很慢就是這樣，因為下面的人他只是領死薪水，那 open open 是給人家用，對我一點幫助都沒有甚至沒有績效，所以剛剛講的就是說，現在大家不過現在就是漸漸有在聚焦，傾向是使用者要這個負責任，那這樣長官沒有責任，他也沒責任他就 open 了，他就他就願意去 open 這個事情，所以這個前提就是說，剛剛講的那個個資法保護部分，到底是誰該負這個責任，這個權責要先釐清，不然你根本就推不動，然後簽到什麼層度你才能夠看，資料的建制層度，就是說逐步比較重要的議題的資料，可以先建制，那逐步應該是全面要要建制起來…就是議題現在比較熱門，這個比較重要的要先建制，其實他們都知道哪些比較重要，我上次去他們都知道，那建制完之後，其實大部分我看到的現象就是，這些政府單位建制完之後，都不知道要怎麼分析，他只有 data 所以你一定要就是我剛剛講的，資料內容要公開，要的哪個層級要趕快訂出來，然後這時候他就可以發包給各個研究單位，大家一起來幫他做分析…

(四)近年來恐怖情人之類的親密關係暴力遽增現象，您認為運用大數據分析可行性如何？可能受到的限制或面臨的困難為何？如要運用大數據分析，進行的步驟應先從哪些方面著手？

1. 先將資料電子化在分析出家暴加害者的型態特徵

其實我是覺得將他，以往處理的那些例如說那個暴力，親密關係的那些家暴，那些資料是不是能夠先把他電子化，然後把那些關係人的那些資料都去記錄下來，那我們大概就可以，就這邊其實就可以去分析說，例如說以家暴來講，哪一種型態的人或哪一種特徵的人，他其實比較容易是家暴的高危險群，其實光是這邊就可以做分析…

2. 可分為內部及外部雙向的資料收集

但是資料如果都能 key 進去，我就可以大概可以篩選出有一批是還沒有發生，但是我覺得是高危險群的……電子化然後可以開始去收集一些網路上面的一些有關這種暴力防治的那些，因為你暴力防治一些法條出來，你都會公佈在很多的的地方，那些公佈的之後，你有沒有去處理那些人家的回應，那些資訊那些輿情資訊…可是以往是都沒有去收集，因為他光自己的內部就沒有電子化了！那怎麼去收集這個網路的，所以內部跟外部，雙向的這個資料的收集，其實我是覺得，外部主要是看公共政策，內部就是我要分析這些這些關係人，其實從這個分析的裡面你也會發現其實，以往有被家暴者，他常常會變成是家暴人家的，從你如果電子化之後，你可以找到這個原因，因為他小時後被人家家暴，他長大他就不知不覺，他也會去家暴人家，可是你以往都沒有這個數據，可以來佐證這個…

二、大數據分析

請問欲分析大數據，於資料庫的建立是以統合成一個資料庫來做分析較佳？還是分成多個資料庫形成資料倉儲來做分析較佳？

請問以大數據儲存的資料庫類型，哪方面的資料庫較合適？(EX. 關聯式資料庫、物件資料庫等)

請問您有無推薦分析大數據較適合的方式？國內是否已有類似成功的實作案例？

請問現行資料庫資料表的模擬分類是否滿足大數據分析的要件？

請問欲分析的資料量若過多(例如有警察資料庫、司法資料庫、社會福利資料庫、醫療資料庫等)，是否影響分析結果的正確性？

經過多次分析所得結果，所占百分比高低或是獲得的高密度關係與低密度關係資料，是否均可以被運用？

1. 應先統一格式統合在同一個地方形成一個資料場所

所謂統合是說統合在同一個地方，但是你不同政府的單位，當然還是要自己設一個資料庫，你不可能把我的跟你的合在一起，合不起來，那就是說放的地方要集中，因為你到時候集中之後大家都放在這邊，你有要做訊息的交換，你就統一在同一個地方就可以，你如果說放在我這邊的話，你來什麼你公文給我，我也不知道要傳去哪裡給你，所以應該就是要統合在同一個地方這樣，像國家高速網路中心這個就是統合在同一個地方，但是不同政府單位的資料，當然還是自己是各自是一個資料庫，變多資料庫形成一個資料場所，但是這個資料庫的這個形式，還有那個儲存格式，要依照之前規劃好的那種方式，一定的規格，統一的規格，其實政府都已經有在做…

2. 適合用 NO SQL 的資料庫在 HADOOP 架構下進行分散式計算

其實傳統都是關聯式資料庫系統，但是以大數據現在比較…強調要高速運算，高速運算的話就要透過那種…雲端的那個 HADOOP 不知道你有沒有聽過，這個 HADOOP 就是一個軟體架構，就是說他可以在這個 HADOOP 架構下，你底下可以有很多台電腦可以做分散式計算，你如果計算能力不夠的時候你只要加電腦進來，他馬上可以再分給那個電腦一起幫忙做運算，就有點類似 GOOGLE 的搜尋，GOOGLE 的搜尋你給他一打，答案馬上就出來，其實他背後是好幾萬部電腦在運作，那好幾萬部電腦都有一樣的功能，那這時候你只要下一個查詢，他看哪一台電腦比較有空他就叫那台電腦去查，就等於有點是說，我底下可能有一萬個人，你今天有一個人有問問題，一看十號那個人現在有空你去幫忙他解答…所以跳號啊，那 HADOOP 的架構就有點類似這樣，就是說我在這 HADOOP 下，我可以有很多部台電腦，電腦只要計算能力不夠，我只要一加進去他馬上就有計算能力，你不用買一台很強的伺服器，他說你這伺服器可以處理幾萬個人，總有一天你還是爆掉，他那個可以加平價的電腦一直加進來，他就會分給他們，所以這個是在一個 HADOOP 的架構下去做這種分散式的運算處理，但是在 HADOOP 下面他比較 PREFER 的那種

資料儲存方式就不是關聯式資料庫，他比較是大數據的，因為關聯式資料庫他以前都是用那個 SQL，SQL 的指令在做數據的查詢，那 HADOOP 下面他比較強調用 NO SQL，NO SQL 就是一般就是類似 GOOGLE 的那種檔案處理方式，這種檔案格式就是比較適合快速的資料處理，那這個可以跟 HADOOP 搭配，因為他如果他是 要做分散去處理，那一般傳統的關連式資料庫可能比較沒有那麼適合…(物件資料庫是)傳統，那些都是傳統，比較適合用 NO SQL 的那種資料庫。

3. Dexter Mining 是作為公共政策的分析，家暴人的特徵適合用 Data Mining 資料的分析就是有分兩種，一種就是 Dexter Mining 文本的分析，一個就是 Data Mining 資料庫的分析，那像輿情分析這個就是文本的分析，因為他這是人家他要發表的一些言論，那這個就牽涉到文本，那我說這個家庭暴力防治，那他以往的那些資料，那像家庭暴治防力他原來的那些資料，他裡面就會記錄他是什麼地方的人、他幾歲、他家庭的關係成員是怎麼樣，這個是原來的 Database 裡面的資料，這就屬於 Data Mining 的分析，我就可以分析說通常有家庭暴力傾向的人，他家庭的關係是怎樣，他的上一代家庭是不是也有家暴的情形還是怎麼樣，我就可以把這些，特徵找出來…(輿情)是屬於 Dexter Mining，那另外一個就是你原來 Database 的 Mining 的就是 Data Mining，例如說家暴的人的一個特徵，什麼樣子的人比較容易有家暴傾向，那這個就用 Data Mining 就可以做，那 Dexter Mining 一般是作為公共政策的一個分析，就是說這個政策出去有沒有有些人有些意見，有一些這個看法，那可以提供做參考，這個就是 Dexter Mining。

4. 現行的資料是不完整，尚未能滿足大數據分析要件，更無法蒐集輿情資料

…應該是說現行資料庫裡面的資料表的資料應該是不夠的，因為我剛不是一直強調資料表要電子化，他所謂的模擬分類可能是說以他目前有紀錄的那個資料，他可能有把有人是說這屬於有家暴的，這個人是屬於什麼，他可能有一些資料的紀錄，但是我剛剛一直強調的就是說現行的資料其實是不完整的，所以他的模擬分類我覺得應該指他現行的資料，那你要做大數據分析，那你當然在不完整的資料上比較分析不出，尤其是政府的資料我覺得更是不完整…政府單位他那邊據我所知那邊的資料其實是相當不完整的，尤其是過去的那些資料，他根本就沒有空去建錄，也沒有編預算去建，因為建的時候沒有用，所以我是覺得目前尚還不能滿足大數據分析的要件，資料的完整性就都不夠，更不要說去蒐集輿情，輿情更是沒有。

5. 專家應控管數據的重要性，避免影響分析結果的正確性

…會影響正確性就是你如果將一些不重要的資料，都加入的話那就像我剛剛講的分析小孩子淹死，誰叫你把霜淇淋的銷售也把它納入，所以你一開始你應該就要慎選，就是從你要分析的這個目的來看，它應該是跟哪些因素，可能有關係那有沒有關係不知道，但是你不能把一些因為當初我現在收集到這麼多數據，全部都丟下去你分析看看，就會跑出霜淇淋的這種結果，然後你就會發現這怎麼會是這個結果，那這當然你就會覺得不正確…所以有時候人家在做大數據分析是說，我們不要管全部都把它丟下去，其實我覺得這都是錯的，人、專家還是要控管那個數據的，哪些數據要進去，哪些數據不應該進去擾亂。人還是扮演一個很重要的地位…所以我這邊強調的是說大數據分析其實人還是最重要的，因為數據量太

多，到底哪些數據是重要哪些數據是不重要，哪些是雜訊哪些是真正的訊息，人還是要佔有蠻重要的一個地位，人要介入，所以跟他無關的就不要下去了，這樣就不會影響分析結果的正確性。

6. 將有家暴傾向者分成五個等級，最危險者應立即處理

當然可以被運用，就是我們把例如有家暴傾向的人排名，那因為我們能力有限，當然是先從機率高的最危險的，例如說你把潛在可能會家暴的人分成五個等級，第一個等級極為重要就要立即處理，第二個等級就是相當危險，那如果人力也是要立即做處理，如果沒人力就是說根據那個等級處理下來，那下面那個機率稍微低點，他可能也不會那麼快爆發，那這樣就是有效運用你的那個人力，不要亂弄…極度危險極度關係…

三、其他有關大數據分析運用在親密關係暴力防治的政策建議？

- (一). 如果將來想整合政府各部門的資料庫，做為親密關係暴力防治的大數據分析，您認為該由政府那一個機關來主責？為什麼？
- (二). 請問是否可以從網路取得本研究欲分析的資料來源？是否需用建立網路爬文技術？
- (三). 其他補充？

1. 由國家高速網路中心主責

……如果是這樣子的話，軟硬體比較不會有問題，因為以往你要做這個，你必須要有雲端設備嘛！雲端設備就是一筆費用，他們可能沒這個預算，沒有這個人來 maintain 這樣的東西，你現在都丟到國家高速網路中心，第一個他已經快倒了！這個好不容易只好就可以救他，然後第二個那些 maintain 的那些東西都由國家高速網路中心來負責，問題就會比較小……

2. 爬文技術要建立，資料蒐集後要做分析

這就我剛剛講的輿情分析，那輿情分析你要先知道的就是說，例如說跟我這個科室有關的或是說跟我做的這個主題有關的，通常在網路上有哪個地方是有在討論這些東西，要先去做這個了解，那其實我覺得大部分他們都知道，大部分他們都知道說，網路上大概有哪個地方在討論，那我們政府有公布的那些 FACEBOOK 的那些討論區是哪幾個，所以他大概都知道喔，但就是說有沒有其他地方也有在討論，那有的時候當然就是要透過網路爬文技術，這個定時的將這些輿情抓回來，可是抓回來之後你還要有分析喔，不是說抓回來就就不要分析了，這個只是說在做作答數據的部分而已，（爬文技術）要建立、分析，Dexter Mining…

3. 數據先應電子化，並培養分析的專業人才

建議就是先從像我剛講的大數據你要先從有數據開始，有數據就是你要將它電子化、數位化，甚至去收集一些網路上的數據，那有了數據之後呢我們要做就是開始想培養人，你現在有了數據之後，各科室的人有能力去分析嗎？因為要做你要有能力去做分析嘛，那分析的人有沒有，他能力有沒有培養起來，是不是要開始規劃、培訓，要將這些將來要做大數據的人的能力培養起來最後要用，要怎麼運用，那這要再跟各個科室，如果我把這個分析出來，假設我把家庭暴力分成

五個等級，這五個等級要怎麼運用，那這個人要以及想就是你可以從大數據的從有、做，再用，這方面我們都可以給他建議。

4. 大數據運用在親密關係暴力犯罪防治訪談紀錄(蔡俊輝副主任)

訪談對象：國家實驗研究院國家高速網路與計算中心副主任蔡俊輝副主任

訪談時間：104年4月30日16時00分至18時05分

訪談地點：財團法人國家實驗研究院

訪談者：王伯頌、韋愛梅

一、犯罪防治

- (一). 面對政府要求資料開放、大數據等政策及國際發展趨勢之下，您認為大數據運用在犯罪防治的公共議題上可行性如何？發展的方向？
- (二). 政府不同機關間的資料庫連結(或資料釋出)，您認為要注意哪些重點？可能會面臨哪些困難？該如何解決？
- (三). 您認為政府或是社會上有哪些資料(或資訊、資料庫)是與犯罪或親密關係暴力有關的？或是對犯罪預防或親密關係暴力防治有用的？有哪些是基礎的條件(例如資料的建置程度、資料內容、個資保護…)？又需要哪些配套措施？
- (四). 近年來恐怖情人之類的親密關係暴力遽增現象，您認為運用大數據分析可行性如何？可能受到的限制或面臨的困難為何？如要運用大數據分析，進行的步驟應先從哪些方面著手？

1. 大數據是趨勢，可行性沒問題

可行性當然是可行…這個當然沒問題，這當然是一個趨勢，可是就是說我們資料的蒐整，這件事情應該是只要有資料那資料蒐整清楚之後，後面大數據我覺得那都是應該是有現存的，較不是什麼太大的問題。那可是就是說要有耐心的去把資料彙整。

2. 現階段只能用常用的 case 去摸索

即使現在我們目前研究學者這邊可能還沒有那麼多數字，很多東西只能猜的，可能你還是多想，因為你那個領域越清楚，就會越知道該怎麼樣去整理或者是蒐集你要的資料，那這些東西越清楚那你後面的那個分析，該有的電腦該有的軟體工具…那現階段當然也是只能還是用一些常用的 case 去摸索去了解。

3. 資料庫連結應注意個資的問題

目前資料庫的連結最重要的還是個資的問題…這個因為牽涉到法律的問題，所以這個還是要照法律來解釋，那站在我們的立場當然就是看是怎麼樣

4. 格式的建置重要，不宜過於簡單

就是會有一些衛福部的內政部的一些…至於說資料內容跟資料建置程度還是回到這個格式的建置，剛剛格式的內容我沒有意見，因為我也不太好發表意見，不過我剛剛提到就是說，有些地方盡量用勾選式然後內容文字盡量要有用，就是說你如果勾選太簡單，那你到最後會後悔，就是說你這些資料還是沒用，那個地方可能要想清楚一點，反正就把他印出來密密麻麻也沒有關係，可是裡面的文字如果說是具體有用那對於後續的分析的連接會有比較好的幫助。

5. 大數據分析可行性高

到目前為止好像沒有什麼不能用的大數據，什麼都可以。…因為這牽涉到個人很多的隱私…有一點個人權益的保護…

6. 恐怖情人沒有具體的背景資料，增加辨識之困難

可能受到的限制與困難就是這些恐怖情人可能並沒有一些具體的背景資料去 identify，所以我覺得就必須要靠比較親密的關係的一個周遭 close 的朋友從那邊來獲得這些數據。

二、大數據分析

- (一). 請問欲分析大數據，於資料庫的建立是以統合成一個資料庫來做分析較佳？還是分成多個資料庫形成資料倉儲來做分析較佳？
- (二). 請問以大數據儲存的資料庫類型，哪方面的資料庫較合適？(EX. 關聯式資料庫、物件資料庫等)
- (三). 請問您有無推薦分析大數據較適合的方式？國內是否已有類似成功的實作案例？
- (四). 請問現行資料庫資料表的模擬分類是否滿足大數據分析的要件？
- (五). 請問欲分析的資料量若過多(例如有警察資料庫、司法資料庫、社會福利資料庫、醫療資料庫等)，是否影響分析結果的正確性？
- (六). 經過多次分析所得結果，所占百分比高低或是獲得的高密度關係與低密度關係資料，是否均可以被運用？

1. 除了考量資安問題，資料庫之統合或分類均可

我個人是覺得統合成一個或者是分成很多個其實都不是太大的問題，因為現在網路都很方便，而且頻寬都越來越夠…所以當通訊不是問題的時候，我覺得這個一個或多個來處理都不是太大的問題，除非是有考慮資安，那你大概只要保護那一塊就好，那多個當然就會有很多點去保護，這個當然是他一個的好處，會有另一個壞處就是說缺乏備援，萬一你一個 crush 掉的時候，你就沒有很多個它可以互相備份，那他有使用上的一個方便，所以我覺得各有各的優缺點，可是並不見得這個對分析本身我覺得倒沒有太大的影響。

2. 無法精準對資料庫類型下定義

因為我其實也不是資訊班出生喔，所以我沒有辦法做很精準的定義說關聯性資料庫…對於在做大資料分析者的角度來講，資料越多越好…至於說哪一種類型是不是決定說你要怎麼樣儲存你的 data 的方式，可能是這樣子那我覺得這個還是會 depend on 你的題目本身，我覺得直覺的感覺不見得正確，比較社會科學的問題關聯式應該是比較會有…那物件式通常感覺上比較靜態式，所以譬如說像有些工廠管理或生產管理或什麼喔那大家每天都在做一樣的事情，只是在 tune 一些參數的內容也許是比較屬於這一類的。

3. 依照問題找出適合分析大數據的方式

一開始提到的還是要看你的問題然後怎麼去解，我覺得每一個問題本身就現在來講，都沒有一個現成的套裝工具可以…還是要很多客製化的東西，當然你資料儲存的方式有些軟體什麼 HADOOP 或 SPARK 什麼東西可以去幫您做分散式處理，

可是處理完了以後你還是需要有一個分析工具，分析工具有些套裝像 SARS，可是那一些工具當然你還是要去定義你要的一些東西。

4. 合適的分析方式包括 HADOOP 與 SPARK

所以基本上我覺得合適的方式不外乎還是在處理這一塊有 HADOOP 有類似 SPARK 這種，那我的瞭解因為說實在的 HADOOP、SPARK 我也聽沒有用過因為對我們來講我們那時候只用過 TRAIN UC，這種程式也也是像 HADOOP、SPARK 這種套件，他應該是一個類似 MIDDLE WARE 這樣的統計工具，幫助你做分散式處理，那 HADOOP 是處理比較靜態的資料就是說他不會隨時間一直在長，譬如說我每天每一分鐘要記錄如果可以記錄我的血壓，那我要做大數據，我就每分鐘給他一次記錄，那我就有很多 data，我一年下來就可以很多。可是而且這些數字會隨著時間一直在長，所以這是比較動態的，這種可能就用 spark 條件可能會比較好。那 HADOOP 是一個比較靜態的。

5. 國內成功的實作案例以產業界佔多數

目前用比較多比較成功的大概都是在產業界…那我的瞭解剛剛除了講到那個台積電以外…我知道在台中台中彰化那邊有很多菁英集結，那一塊的產值我加起來其實不輸台積電

6. 資料庫的模擬分類例如發票及收據能滿足大數據分析的要件

這邊提到資料庫的資料表，這種大資料分析做的最好我剛剛講的 google，其實他也沒有什麼資料表…譬如說像我們的發票好了收據，收據裡面其實記載很多資料，時間、地點、物品、價格不知道還有什麼，還有一些隱藏的 code 我不知道，那些都是資料，所以現在大家都會從很多地方去擷取很多的資料，那像這種特殊的事情，你就要設計特殊為了這個事情設計了這個那這樣你的資料是最乾淨的。像發票我剛剛講的也許很多訊息在裡面，可是你可能也只需要時間跟價格，然後你就可以知道什麼時間…那你就有一個消費者那個消費者是誰不重要，可是消費者在什麼時間的消費的數據可以判斷決定他的商業模式，諸如此類

7. 資料量越多，正確性越高

過多絕對不是問題，而且越多應該正確性越高，這個多的意思就有點像三個點要決定這個曲線，可是如果你有一百個點，線就出來了形狀就出來了所以應該是越多會有越正確。

8. 高密度關係與低密度關係資料，均可被運用

我覺得應該都是可以，那當然高是代表很明確的 signal 嘛，然後低可能就比較不一定，可是他也是有機會是一個 signal。

三、其他有關大數據分析運用在親密關係暴力防治的政策建議？

- (一). 如果將來想整合政府各部門的資料庫，做為親密關係暴力防治的大數據分析，您認為該由政府那一個機關來主責？為什麼？
- (二). 請問是否可以從網路取得本研究欲分析的資料來源？是否需用建立網路爬文技術？
- (三). 其他補充？

1. 由衛福部主責以整合政府各部門的資料庫

我覺得如果以主責的概念來應該還是法務…不，真正執行…警察是法務部的話它是內政…就是在司法的那個單位吧，執法的那個單位……那如果配合我們的政府的組織，那這一題就是衛福部。

2. 資料庫才是取得資料來源的重點

網路不是重點，那個資料庫才是重點。那只要有資料庫有足夠好的資料網路都不是問題。

3. 網路頻寬須足夠才能建立網路爬文技術

如果資料庫 access 到，那我網路一定都連的到。透過網路頻寬如果有夠的話，你透過網路來做這些不管爬文分析才有道理，不然速度很慢一下斷線一下在那邊一直轉你當然不可能有那個耐心在那邊做這個事情，基本上我想如果以政府來講的話，若這個資料庫都很 ready 那透過整個政府體系的這些網路圖或資網中心這大概都不是太大問題。

5. 大數據運用在親密關係暴力犯罪防治訪談紀錄(陳泉錫處長)

訪談對象：法務部資訊處陳泉錫處長

訪談時間：104 年 5 月 8 日 15 時至 16 時

訪談地點：法務部資訊室會議室

訪談者：王伯頌、韋愛梅

一、面對政府資料開放及大數據的政策要求下，貴機關有無相應的規劃重點?或是配套措施?

(一). 先找出急迫解決的問題，在評估是否適合大數據之研究：

目前的這樣大力的推動，但 open data 是政府該做的!做到這樣程度我覺得是適度，那在大數據這一塊，我覺得現在重點應該做到說，到底我們現在這個社會、這個國家，急迫要解決是什麼，再來再評估這要解決問題是不是適合以大數據的型態，或者是大數據能不能有足夠的這樣能夠 support 我們來做這個研究

(二). 找到對的題目使用傳統數據也可以做分析，運用大數據分析是其次：

那不然以我們目前政府裡面要用的資料，以傳統的數據的這種處理方式，不一定要那像 Hadoop 不一定要那什麼工具，其實也可以做分析，問題是你們有沒有找到對的題目跟正確的方法，其實我覺得那個正式要用大數據這種技術，我覺得是其次

(三). 尚未有相對的規範重點配套措施：

那至於就這個議題我們現在沒有相對的規範重點配套措施，我來這個回應就是說，其實上次因為再政策要求之下，其餘單位坦白講還想不出什麼議題，那就我們資訊單位來想議題，那這議題變的是代表我們法務部的資料，提到行政院去…因為資料做久了，我發覺我們去觀察什麼政府最急迫要解決問題，什麼問題是比較澎出的所以我們大數據裡面我們希望能夠做一些研究

(四). 先使用傳統資料庫分析研究，尚有欠缺再去找外面資料來搭配：

所以我基本上想法是說我們要丟出議題，那這是第一個認為是值得，相當值得研究的一個議題，大概類似這樣我們是有提一些研究，那目前至於說配套措施，其實對大數據我們目前並沒有非常具體的配套措施，因為你要先知道說你要研究什麼議題，我要真要用到什麼工具對我是有幫助的，這時候才去看看你需要什麼額外的資源，那以目前來看我，我是覺得說我有，我們傳統數據庫的資料，就已經有相當多的資料在裡面，那當然如果從外面再搭一些網路資料來，也許可能是 positive 但是那個多少效果不知道，要不要投資那麼多力量去做這個，他的 in phase 真的有那麼大，我到覺得可以先用傳統資料研究完了之後，要用哪資料，用這樣那如果還有欠缺，在用那時候就資源都會比較相對的適度

二、資料的運用(政府各體系間之聯繫機制與資料庫建置情形)

(一). 近年來恐怖情人之類的親密關係暴力遽增現象，您認為運用大數據分析做為防治策略的可行性如何?可能受到的限制或面臨的困難為何?分析獲得高密度關係與低密度關係資料，是否均可以被運用?

(二). 您認為政府或是社會上有哪些資料(或資訊、資料庫)是與犯罪或親密關係暴力有關的?或是對犯罪預防或親密關係暴力防治有用的?有哪些是基礎的條件(例如資料的建置程度、資料內容、個資保護…)?又需要那些配套措施?

(三). 政府不同機關間的資料庫連結, 您認為可行性如何?面臨哪些困難?資訊要合作分享的前提要件為何?

1. 保護令無法確實保護家暴被害人:

我如果就親密關係這一點, 我們先談到家暴, 家暴目前方法是什麼?發一個保護令給被害人, 有用嗎?沒有用啦!(台語)早晚去殺早晚去砍, 沒有用, 所以社會公益是不能被凸顯得, 可是我一直看不慣說, 為什麼我們整個社會對這個問題一點方法都沒有, 一點做為都沒有, 我覺得是不對的

2. 親密關係施暴者可使用電子監控, 被害人可攜帶警報器:

譬如說電子監控, 我覺得應該可以有效化, 他沒辦法百分之百的解決, 但是我可以解決六成七成 maybe, 方法是說我讓你那個, 暴力的這個親密關係, 就是加害家庭的暴力的, 你施暴者帶電子監控, 被害人帶一個警報器類似這樣, 當他接近一定的距離譬如說五百公尺, 開始 warming, 不是 warming 這個人而以, warming 警察, 那給這個加害者有禁壓力, 當你要去的時候, 警察順便也趕到了, 所以你要不要去, 那這樣當有一些人還是衝過去, 有一些人就已經失去理性還是衝過去, 他至少這個人來的及把門關者嘛!那萬一真的衝進去被殺了!會但是我覺得很多人是比較聰明的人, 他覺得算了!風險大我就不刺, 我覺得因此可以保護可能六成七成的人, 雖然有些人還是被殺, 但是有就會保護更多人, 這個是有效方法, 所以我到覺得說這個也許, 現有的工具會比用數據去分析找方法更快, 這個會覺得說 maybe 這個是比較有效的方法

3. 對於較輕微的毒品犯及酒駕犯, 可使用電子監控以紓解監獄擁擠之情形:

那更進步的提到像我們高雄的, 大察我們高雄監獄很多事件, 問題我寫了一篇文章, 我不曉得各位有沒有看到, 我寫了一篇文章, 寫提出了看法就是說, 這事情大家譴責監獄的時候, 有沒有人看到這背後的一個主要的因素, 人事行政局不用負責任嗎?因為現在監獄各位知道, 監獄的人力我們是一比十四, 外國是一比十, 我們超過人家, 我們一個人要人家三倍武功才能去管受刑人, 我們那麼厲害嗎?沒有嘛!那代價是什麼?風險, 什麼時候出事不知道, 所以這個就是人力不足, 不過人力不足你現在跟人事局叫半天沒用(台語), 什麼方法想一些比較科技的方法, 你讓那些毒品犯, 傷害自己的只要有刑事案你就不放他出來, 傷害自己只是吸毒傷愛自己的, 而且監理面表現良好幫, 讓它變成一個除役的手段, 表現讓你在監獄裡面乖一點, 我就讓你出來帶電子監控早一點出來, 我監獄能不能減少一些人可以嘛!可以減少蠻多的, 而且帶電子監控之後你接受醫療, 因為以前你出監之後就出監了啊!你就很難管他, 依法你也不能管他, 警察頂多兩年之內可以跟他驗尿, 那而且我們現在資源不夠, 驗尿也要偶爾才做一次, 那你帶監控不一樣, 我就請你定期到什麼警察局報到, 可以驗一下尿你不來你就違規你就要進去, 那壓力很大他就會守規矩, 這個是很有效的方法, 你為什麼這樣做, 你把他關在裡面沒用(台語), 出來照樣再犯藥頭一找, 跟你這個有監控之下回歸社會, 我覺得比用個外役監還更有效, 我覺得這個方法恐怕有用很多, 新加坡這麼嚴的機關, 他們用這個對毒品犯, 為什麼不能處理, 這樣我們就可以, 我們監獄裡面兩萬六千個毒品犯, 有一些是販賣運輸的那少數, 可是販賣運輸又兼吸食, 純吸

食大概有一萬三，一萬三裡面把他扣掉一些比較不好管的，放個七八千人是出來有的，那七八千人對我們監獄的，現在超收就很大的解決，接下來談酒駕，酒駕當然肇事的就不要放他，但是酒駕如果是三犯進去，但是他事實上就是愛喝貪杯也沒有犯什麼錯，一定要關他嗎？不見得你讓他戴監控出來，宵禁晚上不能要出去，白天你繼續投入社會的工作，不影響社會生產力，維持家庭，那晚上有管他不要出去，為什麼不？我們監獄又少掉三千人，我一少少一萬一千人，我的管理人力就可以縮減嘛！為什麼不做？所以這個都可以做的，所以我只是回應說，有些方法是在的，不見得一定大家想的方法都大數據，可是我覺得大數據不見得能解決很多問題

4. 必須有一個單位出來統籌大數據研究並且串連各部會資料：

依個資法是要除去你資料提供協助研究，必需要去識別化是依法，每個部會其實每個做法就連不起來了！所以你要做跨部會連結就做不到了！所以你必須有一個單位出來統籌，譬如說現在要做，這是要主責大數據研究的組織單位…那你主單位你就要把各部會資料拿過來做串聯，串聯完了之後，其次規劃這個就是單位，這樣就能解決，阿如果你要各部會都自己去識別化丟不出來，那這樣的話你沒辦法做串聯，當然你這樣符合法的規定不能做串聯，那目前不能做串聯

5. 拿到資料要負起保護資料的責任，並不能是原來提供資料單位負責：

舉個例子說所有政府機關我都要回來了，那我們就負責整合，你這樣給我，我就說我負責，我負責就沒有這個問題，我只要說你的資料給我，我就說我負責，我就可以串，串完之後我其實我要再給你研究就可以了，所以這個問題可以解決，不是不能解決…現在大家都擔心個資法問題，其實以我的觀點是說，如果你被指定要做這個研究，你就要負起資料的保護的責任，你要把他想辦法傳借，像如果說依規定，我們把該把關的資料給人家，我就給阿但是資料保密責任在我家當然我負責，到你家當然是你負責怎麼是我負責呢？因為現在很多協調就是說資料提供出去之後要原來單位負責，我覺得你這樣誰給啊！我當然不給…所以這是觀念的問題，所以現在很多單位都你資料提供給我，還要你自己負責，這我當然不會給你完整的資料…所以這問題其實觀念，不是不能解決的問題。

三、請問貴機關的資料庫，

- (一). 運用在親密關係暴力防治的可行性如何？建議未來發展的方向？貴機關現有的資料庫及其內容，未來運用大數據上有無使用上的不足或受到限制之處？
- (二). 如有不足，由無其他途徑？該如何取得欲研究的資料？是否可從網路取得本研究欲分析的資料來源？
- (三). 欲分析大數據，於資料庫的建立是以統合成一個資料庫來做分析較佳？還是分成多個資料庫形成資料倉儲來做分析較佳？

1. 要依照研究議題才能判斷資料是否足夠的，不排除未來去撈網路上的資料：

所以我們資料庫到底有沒有足或不足的受限，其實還 depends 你要做什麼研究，maybe 有些資料或是不足的，但是我是覺得說要看你研究範圍，那目前因為就目前我提出來的研究議題，我是覺得其實已經有相當的資料可以用，當然你要用的變數未來用的太多變數也許我不見得有，但是變數多少呢？是你研究強度的，或是研究可信賴度的問題，所以我這個問題我就比較難回答說足或不足，可

是我們目前以我們目前我覺得想要做的研究來講，我覺得應該暫時是可以供應的啦！那未來要看什麼地方不足，研究之後再看，這個大概我的觀點是這樣，啊當然我講的是說，已經把各部會資料整進來了！整進來之後...我講的是政府之間的資料整進來，那我這邊所比較可能有不足要民間，包括到網路上去撈那個那一塊，但不排除未來如果有需要再去往這樣做，但政府裡面像我剛提的那議題，其實就跨好幾個部會，跨我們跨衛福部，甚至跨教育部等等都有關連

2. 對於毒品驗尿未來應該配戴電子監控：

其實我的了解他事先通知你來驗尿，事先就告訴你時間，啊我可以做防備啊！這個其實應該是通知，未來應該帶電子監控，通知你隔天就要去，你就要來不及處理也不行...所以剛開始一個禮拜驗尿兩次，我們台灣現在恐怕三個月才驗個一次，所以效果是因為事先通知他過來，他們是剛開始一個禮拜就驗兩次到三次，所以你根本無所遁逃，不過這要資源的投入，那我們現在台灣號稱說對毒品的問題很嚴重很重視，開了很多會不過資源都沒有投入，該有的資源你不投入那你一直在叫有什麼用

3. 可以從網路公共論壇去發掘資料，但基於個資法的保護有些資料不能公開：

所以如果是就這邊剛剛有哪些這樣講，我到覺得說如果就親密關係的這種暴力如果從網路上的語言找到一些蛛絲馬跡，我覺得應該是有機會的因為以這個犯罪的性質來講，我覺得應該有機會...不過我回頭馬上再問就是說，人家的這個交談的原因，他不見得是一個公共論壇，若他上 ptt 去講那 fine，那個公開的可他若是親友之間、他夫妻之間、家裡之間的你去跟人家 ask 這些資料，因為這種親密關係嘛！因為要找到這樣親密關係，他們之間言語之間你去網路上，你有沒有合法去抓這些資料，我還真的是有一點憂心，就是說這樣聽起來也許從他言語，言語的情緒性的言語有一些找到說，也許到已經暴發點了或是已經在醞釀中，是有機會我覺得在網路上，但是我就想說我們有合法的角色去要這些資料嗎？.....就是說原來有譬如說刑案裡面記錄，那未來看要不要開放，不過這個又是很多，法律上的爭議，如果說你想查詢一下，是以目前我們國家法律是不太允許，因為他個人的犯罪資料是不能隨便給他人的，那你也不能說我因為要跟他交朋友...在社會上實在是有這樣的一個需要，不過目前國家法令也不行，我們是刑案裡面是有這些傾向的人，暴力犯這些人是有，但是問題是說我們前科是保護得非常牢，他還是我們個資法裡面的 title 資料，更不能隨便接漏所以這還真的是一個難題，剛剛找了說網路上去發現，網路上就是說公共論壇是 ok 啦！那若不是公共論壇上，又是一個很大的問題，所以這個目前有一點.....看來是說從網路上去歸納，哪些看來是比較 sensitive 的言語的語彙，然後未來在去回頭對照說譬如說輔導這個一班學生，這一班學生在網路上有沒有使用過這幾個比較 sensitive 的語彙，如果 yes 那可能是 something wrong 要多給他關切，這也許是方法之一，我到覺得這個這是比較真的用就用大數據的這個，這個領域我到覺得有可性，不過這個你要去匯整網路上哪些是屬於這種比較 sensitive 的語彙，這倒是真的要有一些很大力氣的一個研究，

4. 可以從立法方面修訂，給被害人關於加害人的相關背景資料：

可是未來就是說我們在家暴的相關法令上，或是親密關係的法令上，認為說已經有攻擊了！我必須對他有一些防衛的一個，因為已經也事實存在了嘛！所以未來立法上，當然可以建議往這個方向思考，我們國家應該給他適度的說，這個親密交往的人，已經有一個家暴的事實發生，我能不能去更進一步瞭解他過往的一

些犯罪的背景，若是我研判完之後，我應該做個檢定說，我應該分手(台語)，所以我到覺得剛剛這樣講這個我們可以從修法上做建議，當然我想這個研究案也許，剛剛這個案例我倒覺得說，應該若有暴力的事情發生，那你去報警，應該可以成立說，已經有構成原因說，我們國家應該給他，相對的資訊讓他能夠防衛自己，我覺得是有正當性的，那目前法是沒有，我覺得可以建議修法來解決這個問題，我覺得應該是合理的一個做法，已經有暴力出現…所以就若剛剛像提的這個領域，坦白講我真的能夠貢獻的比較少，但若說是要類似剛剛英國案例，而當然這有立法問題，那未來如果以這樣的事實出現，我們歸類應該 provide 這樣的資料，我們這邊就有貢獻，我們可以 provide 這樣的資料給他，合理法規，這個我們是可以來提供，以正當性來講應該這樣

5. 把資料整合出來是有困難的，除非是委託研究案才有辦法去各部會拿到資料：

你做要做這議題，可能就要把這議題先想說，你可能需要用到哪幾個因子來做為你分析的一個變數，那所以呢你要把這問題提出來，然後要哪幾個因子在看這幾個因子的資料欄項在哪幾個部會有，也許你不知道他在哪部會不管，那你可以說 ok，橫向跑不能就問說法務部你能不能給我這幾個資料，啊告訴你我有這幾個，但我們一定會問說你為什麼要他，這個依法一定要問，所以你要講出個理由為什麼要，啊如果這種法務部委託研究案當然就沒有話說，這是我們法務部自己委託研究，啊如果說是，不一定法務部研究，譬如說這一次有一個所謂 top-down 的還一個 bottom-up 的一個大數據的一個需求，假設是一個 bottom-up 就是說有學校自己想題目，然後來幫政府做研究，那政府看要不要買這個帳，買這個帳就做研究，啊所以做研究就要提供資料，啊你要看看你為什麼當初設定這個議題，你需要哪些資料做研究，才能夠 fulfill 你研究上的需要，啊是這樣就跨一個部會，就要分別跨去跟跨得部會來協調，目前有一個機制是有張副院長自己主持就說好，如果這個跨部會的議題，如果真要協調，很多單位不願意提供，啊就提到那個張副院長主持那個會議上，當然這層級很高他來協調，所以上次我有聽到這樣一個機制，所以這議題如果現在變成，譬如說 bottom-up 提出來，然後也有部會買帳，應該已經有部會買帳你才會做下去嘛，要不然你也沒有經費來源，所以部會買帳就科技部才會給錢，給錢做好(台語)，變成一個議題之後就有相關的部會要提供資料，還是要 depend 說你要研究，你接著下來寫變數來 support 你的研究，啊每個各部會去要來是這樣，要不然就透過這個機制協調，但現在不做的人還是這樣子，所以還是 event to event，就是說你要做哪個 case 要做哪一個議題，是這樣去找議題，那現在要把各部會直接把全部的資料都去識別化的提供出來，實務上會有蠻大的困難的。

四、其他有關大數據分析運用在親密關係暴力防治的政策建議？

(一). 如果將來想整合政府各部門的資料庫，做為親密關係暴力防治的大數據分析，您認為該由政府哪一個機關來主責？為什麼？

(二). 其他補充？

1. 應該由衛服部主導法務部協辦，其他相關聯單位教育部也應加入：

目前家暴的這一塊是好像歸屬在衛生福利部理面，所以如果以這樣的一個情形，但親密關係是不是就是等於家暴，我不是那麼得確定，不過應該有一定的關連，所以我到覺得說可能衛福部，應該會是，當然法務部坦白講也脫離不了關係啦！衛福部主導，我們必然是一個協辦單位，若法務部還是有一些角色在啦！我個

人是這樣見解，所以可能是，衛福部主辦，那法務部協辦，我的建議是這樣來做... 其實是沒錯，如果說硬要把它放回，變成比較周延一點都有關連，就像我們剛剛講的毒品也是，你教育部當然更重要啊！現在你教育部，你教育沒有弄好，這一塊跟教育坦白還是有關連，所以你要是說，如果就是比較直接相關得應該就是衛福部主宰，我們法務部來協辦，若是要把相關連的部會都找進來都有關連，教育部也跑不掉。

6. 大數據運用在親密關係暴力犯罪防治訪談紀錄(張秀鴛司長)

訪談對象:衛生福利部保護服務司張秀鴛司長

訪談時間:104年5月14日14時10分

訪談地點:衛生福利部保六樓保護服務司會議室

一、親密關係暴力防治現況與困境的探討

(一) 您認為近年來恐怖情人之類的親密關係暴力增加現象的可能原因為何?

媒體會是讓我們覺得親密關係暴力案件增加很多的因素，但另外因交通方便、整個社會開放下，男女朋友交往已是正常事情且有年輕化趨勢，這些社會因素、個人因素，再加上外在因素如媒體推波助瀾，的確讓人覺得親密關係暴力幾乎每天都在上演。

至於增加原因為何，個人認為與少子化有關連，小孩成為是獨有的、特有的，也很少去學習包容、體諒、為對方著想，但在情感上很多時候是不能唯我獨尊，要彼此互相體諒，為對方著想，不是佔有。若二位青少年男女朋友從小都是被養成唯我獨尊的，於交往中天天上演武打行大概都可以被理解的。若再因佔有慾，我得不到的別人也休想得到，而做出殘害性行為，也有可能性。所以這些個人的、家庭的、社會的、外在環境的因素都為親密關係暴力增加現象增添了助燃力。

(二) 您認為哪些因素在親密關係暴力的防治上是重要、關鍵的？在主管機關的立場面臨的防治困境為何？

親密關係暴力案件從受理到社政單位介入，都已是三級了，事件都已發生，要根本解決問題是要更前端，像學校的情感教育、心理輔導應落實做好。例如對目睹父母婚姻暴力的兒少，如在求學階段有老師或生命中貴人幫他做些提點、輔導規畫如何避免父母覆轍，或可避免重演暴力事件。但目前學校老師常基於隱私或因懼怕而不願投入處理，再加上社區的心理健康資源的近便性不足，常導致年輕男女朋友間有衝突，他們能去找誰？除了自己好朋友，也無法去找專業單位或人士諮詢，給比較恰當的、健康的引導。以致到後面要花很多的心力，頂多也只是收場殘局。親密關係暴力防治的重點應放在前端，但在此一、二級過程中，學校、社區是很少參與加入的。

(三) 您認為現行推動的親密關係暴力防治策略中，最有具體成效的策略為何？理由？

親密關係的再犯預防部分以高危機與網絡合作的策略作為來防治，確實做

得比較緊密的與紮實，彼此可以知道你在做甚麼，且能夠針對問題進行解決。

對有家暴但尚無致命危險的處理，是從被害人的需求出發，有需要保護令即協助申請，未申請保護令則由員警加以約制告誡，若要暫時離開家庭，則提供庇護、心理諮商，若也不離開，只需有人陪伴，在二線上也有些短期性資源提供，不會因其無保護令或未被評入高危機，就停止服務。但前提是被害人要有意願，尤其在英美國家對成年人更是強調個案的主體性與主導性。

若再有餘裕預算，想加強推動 1、2 級，讓社區里鄰當支持者，當事件發生時里鄰不再冷漠。在社區部分宣導防暴的觀念，在學校部分對目睹家暴兒童加以輔導；另外是對初次暴力的輔導方案，讓小情侶有情感挫折時，有地方去尋找諮詢、得到正確觀念。不管有無其他單位投入資源，這樣資源是越多越好，因為需求永遠是不夠，當然也會有漏網之魚，但至少將資源配佈好，讓民眾知道資源在哪裡。

二、資料的運用(政府各體系間之聯繫機制與資料庫建置情形)

(一) 您認為有哪些資料庫(或資訊)，是與親密關係暴力有關的，或是對親密關係暴力防治有用的？

我們發現在親密關係內可以介接的資料庫系統有二個，一是自殺防治系統，一是精神照護系統，因為我們看到親密關係暴力有二到三成是與自殺、精神疾病有關係，若彼此系統可以互相勾稽，就可以看到危險因子，自殺與精神系統可先去處理，而我們家暴處理人員也不可將之單純看作家暴事件，而要敏銳察覺到可能有疾病的因素，去訪視時要提醒定期用藥等。

再者是與健保資料庫的介接，例婚暴婦女從家暴問題開始到底使用了多少健保資源，可以最近十年身體受傷、精神耗弱、偏頭痛、憂鬱症…次數、情況來勾稽，因健保內有 ICD 碼對應各疾病類別，不管是自費或健保給付，皆有顯現即可互相勾稽，就可以推算一位婦女受家暴她會花多醫療資源。此外，警察刑事資料、校園的輔導紀錄，社工紀錄，都是可連結的資料庫。但光是我們自己所有資料庫就應要先能掌握，例如經通報後輔導個案有些沒有再受暴、有些再次被通報，那麼是甚麼原因作為導致其被輔導成功，尚未利用家暴資料庫做此類進階加值分析。例如吸毒者回流率高達百分七、八十，那另百分之二十未再入監的，原因到底是甚麼，若拿找出來，就可運用至那些再犯者。但目前要運用數據資料作為政策形成之引導還須一段路。

(二) 貴機關現有的資料庫及其內容，在運用於親密關係暴力防治上感到的不足的地方是什麼？理由？

透過對重大兒虐、性侵、家暴案件進行事後檢討時，發現其中未成年新手

父母對子女照護總是七零八落的，而呈現相關性。另對吸毒的父母，他的朋友都是吸毒，當他們被逮入獄時或畏罪潛逃時，他的子女就會成為包袱而變成人球，甚會被臨時照護者虐待；如果可以透過現有資料庫以此關聯性進行分析而找出潛在受害人，則可及時加以保護；但目前尚未作如此的開發運用。

再者是對現有資料庫未能作再進一步分析或作次級分析，找出為何個案成功輔導與不成功的關鍵因素，或者保護個案可能危險因素、保護因素，以提供實務執行單位參考運用。

(三) 政府不同機關間的資料庫連結，您認為可行性如何？資訊要合作分享的前提要件為何？

在沒有侵犯人權或違法之情事下，開放是趨勢，當年為建置家暴資料庫，也是有婦女團體反彈，但運用至今，已無反對聲音，因為科技是被你所用的，不是增加你麻煩的，而且運用範圍是比您想像中大更大；但前提是要有人去建立資料。

現經法務部解釋個資法第 16 條後，現資料庫釋出已無大問題，如現在疾管局也願意給我們小孩的預防注射資料，對兒虐很有幫助，因從是否固定、按時前往接受預防注射，就可推定父母是盡責，有些吸毒父母常疏忽或只打一、二針預防注射就再也不來了，就可評估為高風險的潛在危因。另外是沒有繳健保費的資料，排除台商、疏忽等因素外，那些經濟不良者可能連帶有家庭問題，這時就可篩選出潛在危險家庭前往訪視。

三、資料庫建置的內容與大數據分析運用在親密暴力防治的可行性

貴機關資料庫的建置，未來的重要規畫與運用，在政府資料開放、大數據等政策要求下，推動的目標？困難所在？推動策略等相關因應作為或配套措施為何？

就從台大高材生案子來討論，我覺得大數據的介入對未來案件影響絕對有，舉例他在高中、大學、研究所等求學階段所接受輔導包括情感，是重要資訊來源。因為形成人的觀念、價值絕非一天所造成，一定是長期累積，所以在求學階段所使用過相關資源，如就醫、輔導等資訊與刑案資料等互為勾稽，或能發現情殺者於過往求學階段已出現徵兆、有脈絡有尋的。更可見要在前端要多做一點，若預防端沒有多做些則會加重後端獄政的使用、家庭的破碎，對二造雙方不好，同樣也對國家不好。

現在部內統計處有健康加值中心，將健保、疾管局 CDC 預防注射、中低收入、社會福利、家暴等資料去識別化後，集中匯入健康加值中心，對外開放給學者申請運用，例如要研究家暴與經濟、疾病的關聯性，則可申請運用這些資料去運算歸類。

四、其他有關大數據分析運用在親密關係暴力防治的政策建議？

(一) 如果將來想整合政府各部門的資料庫，做為親密關係暴力防治的大數據分析，您認為那些部門那些資料庫是可以考慮整合的？可能會面臨那些困難或阻力？如何克服？

除上述自殺防治、精神照護資料庫外，還可以將健保、醫療、警察、學校等資料整合。至於各機關將其資料釋出，在行政院力推大數據、數據公開、並運用至各政策規畫與行業，開放是趨勢。

至於整合運用各部門資料庫時，主導機關應由發起者為中心，因他看見問題、想要解決問題而發起，並因此需要其他機關協助。反則全由一機關單位來做，則限於人力、資源有限，再來是有些部分他並不注重，或不知其中運作關鍵點，反而難以成事。

(二) 貴部未來是否有進一步進行與本案相關議題之研究規劃？或者對將來後續研究的建議？

本年度規劃進行社區普及預防，但因經費有限，目前先導入家暴熱點預防地圖，打算利用 google map、GIS 與家暴資料找出那些鄉鎮、村里，容易發生家暴、性侵等保護案件，並區分為三級--紅燈、黃燈、藍燈，有點像防災概念，對於嚴重的紅燈密集區，需要特別注意，是不是因此區老人多、外配多..而造成事件率高，將原因找出來然後將預防資源放在這裡。將來即可用此數據、視覺化地圖，去遊說鄉鎮、村里長，共同提出規畫策略及作為，並在未來預算挹注下，補助村里來推動。

對於兒虐的預防目前已辨識出較多的因子，但親密關係暴力部分尚未聚焦，但建議可以從憂鬱症去分析看看，對已婚而經常去看身心症、憂鬱症者去投石問路。另從婚暴婦女的就醫史，探討其進出外傷急診資料，即使年輕男女朋友未結婚族群則再加入學籍資料庫，再與健保資料勾稽篩出，以即時給予關懷。

7. 第一場焦點座談

一、時間：104 年 7 月 8 日下午 3 時

二、地點：台灣防暴聯盟(台北市中正區林森南路 130 號 4 樓)

三、研究主持人：范國勇教授(共同主持人助理教授韋愛梅)

四、參加人員(按筆畫順序)

S1 王珮玲國立暨南國際大學教授

E1 姚淑文東吳大學健康暨諮商中心主任

W1 陳佳琪桃園市家庭暴力暨性侵害防治中心秘書

P1 廖美鈴警政署資訊室主任

五、座談討論

議題一、貴單位與親密關係暴力防治可能有關的資料(庫)有哪些？內容為何？有關的理由為何？這些資料有無開放？

W1:

就實務端我們在看這樣子的議題，我覺得大數據對我們這邊真的有點遙遠，我們能夠思考到的部分其實是有兩塊，就我們實務端在使用資料庫的部分，基本上比較是案件進來之後，已經透過這個個案要使用的有哪些，在社政端我們會使用到哪些系統，保護系統一定會因為今天個案進來我們就會知道這樣的資料，那用了保護系統就會用戶役政系統，查說他的戶籍資料整個家庭的關係，因為我們要做的是一個家庭的服務，那會用到戶役政系統的時候，我們其實最常使用還有另外一個部分就是社政系統，還有所謂的入出國查詢的部分，我們了解這個人到底是本國或外國，因為現在有很多外配的資訊，這其實在我們的實務端會常用到，在社政端可能會用到這些系統。另外一個部分就是跨到網路，我們其實會用到司法院我們想知道這個刑事犯罪的東西，他到底有沒有什麼前科，其他就是醫療就醫的部分。其實在使用這些系統時我們直接服務工作者，才會有直接使用到這些東西，我們分成家暴跟性侵，家暴類我們知道部分就是加害人之前到底有沒有其他的家暴通報事件或後續的筆錄是不是有實用。戶役政很好使用的方式就是，我發現我們的加害人常改名，今天要用這個名字的時候，可能我在資料庫查不到加害人，可是我們去查戶役政發現他曾經改過名字，我們用其他的名字去查發現他以前有很多的家暴通報紀錄，這就變成我們實務端在使用系統上，就是你有沒有工作的經驗或是什麼，其實當你在查詢就會發現他有多奧妙的一個地方，那我想講的部分是我實務端會用到這些。(W1)

在使用上我覺得經過這幾年我們所謂的高危機網絡平台，至少可以把社政衛政還有我們的學校，這幾端的資料都可以要求在平台上使用，可是現在還有一塊要跨出去的部分是，今天如果是要跟法務部司法，變成我只能要求你的地檢署檢察官

過來提供資料，可是我們跟法務部要資料其實沒有那麼好要，可能要你的行政繫屬發到你的法院高等或最高，這個部分如果是兒少法案件，我們用兒少服法 70 條是帝王條款來要求使用，可是今天親密關係暴力的部分，他能不能提供到相對人，這又會牽涉到個人資料保護的議題。所以我剛剛講說分兩端就是對我來講我們實務端對大數據老實說這概念其實是比較模糊的，可是我只能想到實務端的部分是在使用上，如果這些是有一些系統可以去操作，他在維護個人保護資料的前提之下我們都可以使用，那當然前提都要去做一些控管，那這些控管會有很多實務上的漏洞，或者是我們現在看到說可能小時候是家暴被害人或什麼，長大他來當社工，那回來也考慮過我還可以上自己的保護系統，甚至可以查到我以前被通報的紀錄、我的社工怎麼寫我的東西，我覺得這個其實未來都一定會看到。這個東西其實我不知道跟我們現在討論的議題有沒有相關，可是我們實務端會遇到這樣的問題。(W1)

議題二、家暴法已將未同居親密伴侶納入保護對象，預定明年實施，對在校學生、無犯罪前科的年輕人，如何運用大數據分析找到有危險徵候者，及早預防暴力發生？

E1

我覺得我們在校園裡面看到未同居親密關係暴力納入保護的對象，是不是能夠從大數據找到危險的潛在因子這個二級預防的保護隱密性，這是一個很大的問題，因為未同居親密關係暴力跟已同居關係就只是差在他們有沒有同居，所以在校園裡面不分他有沒有同居，本來我就應該要做好，所以我要花那麼大的精神去處理這些數據問題的時候，倒不如就直接做好所謂的預防推動三級預防裡面發展介入性處遇性的預防，就能夠解決這個問題了，我為什麼還要透過數據之後，找到一個潛在危險因子，那像我個人以前會認為初級有使用過暴力有被通報的這群人應該要加強輔導，如果說學校本身在系統裡面在這個工作做得好的話，他就可以解決你第二個問題了，所以並不是說我們今天一定要透過很穩定的數據分析，才能達到這樣的功能，因此我會覺得有點本末倒置。就是我為什麼要做這樣的事情，當然如果我們再去仔細思考這些議題有沒有什麼樣的連結性，我會覺得如果今天以家庭暴力的資料庫親密關係來講的話，今天他的家暴時間超過幾年，他的孩子正值青少年階段，那這群我們所講的目睹兒童在長期目睹暴力的情況之下，孩子的社會學習過程當中有可能使用暴力，再加上其他的危險因子介入，是不是我就能夠預測今天校園內這些學生他們可能會潛在的，可是這樣子的話過於將人性本惡做一個分析，覺得在這裡面有一個過渡詮釋的問題，這樣詮釋其實對人是一種傷害並不是一種幫助，所以我覺得就學校的一個輔導系統裡面他不會因為運用這種的資料之後就去做。(E1)

議題三、親密關係暴力防治要運用大數據分析，需要具備那些條件？如要對不同機關間的資料庫連結，建議進行的步驟？

W1

第二部分就是怎麼去做預防這件事情，我還蠻認同小姚姊這邊的說法，就是我對親密關係暴力的預防，目前本來就很多不管是三級的部分在學校端或是怎麼去看這樣的東西，我們有辦法透過什麼樣的數據知道他可能就是衛福部會是一個危險的加害人，如果他已經是個通報紀錄那還好，可是今天如果他是一個學生，他可能是目睹暴力長大的孩子，那我就可以說因為有目睹暴力他長大就會有暴力的模仿，會成為施暴者，這其實是個標籤的問題。我現在能想到的是以現有的資訊系統，我只能看到犯罪資料或通報資料或是以前的相關資料，這些對以前人部分都是一些負向的東西，那怎麼去做預防，這一塊我實在想不到，可是針對性侵害的部分確實以我們現有所謂的性侵害通報表，或者有些數據，我們也會去分析說，針對全國各縣市會去看說誰的通報數字最多，因為像台中大高雄他的人口數多，如果我們用通報總數的話可以看到，今天通報量最高的是哪個縣市，人口少本來就少，其實現在比較會做精密的分析是因為我們用每萬民的人口數來看他的一個家暴通報或性侵通報；就以我們桃園的例子來說，桃園市人口不是全國第一名第二名，可是每次我們只要用每萬民人口數來分析性侵害通報，就是全國第二，然後這個部分轉過來長官問說為什麼桃園是全國第二，我們每年在做資料常常會被提這樣的問題，第一年我知道如果以每萬民人口數換算下來的話，家暴性侵台中最高。如果用桃園的比例來看的話，每萬民人口數我們家暴大概是第五或第六，可是性侵永遠是第二。那桃園市的這個數字我就很想要去了解為什麼桃園特別多，是因為我的人口年輕嗎還是怎麼，其實每次長官來問我這個問題，我都沒有辦法回答，我覺得數據的東西真的很重要，可是另外他還要回歸的部分，現在目前的數據我們很仰賴保護系統，保護系統看到的就是人數人次因為大家都是用通報表去勾選，那通報表還有一個錯誤通報表的議題，然後你今天通報的來源可能我同樣通報這個A被害人，可是這個A被害人在社政或在警政或在醫療被通報的數據他可能在勾選選項都會有錯，那這樣數據的落差有沒有辦法改善，如果沒有辦法改善，我要再把這些放到未來我們要去做的的大資料庫裡，其實會有一些研究結果出來他就有很奇怪謬誤的東西。那我現在只是針對我看到這些東西做一些分享。

(W1)

S1：

當時我在做這個危險分析的時候，是用人工的方法找出，我們去找司法院的判決書、警察的移送、法務部的起訴書，然後法院的判決書還要再連結衛福部本身家暴通報的資料，所以你可以想見一個家暴的案子，假設從頭講起的話，他會從警察這邊進來有受理案件的資料之外還有現場資料，到了衛福部這邊就有衛福部的通報資料，然後這個案子如果有走到刑事案件的話，可能涉及到不管是告訴或是非告訴乃論的案件，他可能有殺人、妨害自由、重傷害或傷害，可能就會走到法務部的檢察官這條刑事司法；如果有違反保護令的話一定會走到法務部這邊來；如果申請保護令的話一定會到司法院來，然後到法務部之外到起訴他，可能就有

起訴資料和判刑資料；如果到矯正機構的時候，目前對家暴的加害人我們沒有做處理，所以法務部的矯正系統沒有一個處理的資料。然後我們這些的相對人或是加害人在做高危機案件的時候，我們也去連結衛福部因為這些人很多有毒品他可能是毒危中心的列管的資料庫，有一些人也有自殺的紀錄，我們也連結自殺防治的資料，然後這些人占一定比例有精神疾病的資料，我們有疾病列管的資料在衛福部本身。因此我們看一個人危不危險其實要看的資料是很廣的，除了民事之外，他的一些行為、民事的資料、在司法系統所掌握的資料。而且犯罪前科對於這些人告訴我們在家暴行為是一件很重要的危險因素，他有其他的暴力行為，他是一個低自我控制的人，他對於親密關係是一個危險因素，所以我們現在都是在做這個。當時我們在想這些東西都有關係可是都是用人工 search 的資料。想見今天有一百萬筆的資料，假設以人來串連，以這個加害人來串連，他的資料是散在我們衛福部的家暴資料庫或是散在衛福部的議事處毒危中心毒品危害精神列管自殺防治，然後這些人的前科資料又散在法務部，這些人在監獄的處遇的資料也在法務部矯正署裡面，所以假設我們以這個人為加害人核心，把這些資料都能夠串起來，對於案件的第一步的分析跟掌握，就很重要。那這些東西有什麼好處，譬如說我們從這邊就可以知道，這些再犯的人跟這些有沒有什麼樣的關係，跟哪一些東西是有關係，有進到刑事司法系統的人再犯比較高還是比較低或是怎麼樣，這些初步的資料的分析我們現在基本上都沒有辦法知道，只能憑一個很切割的資料的概念然後去調到，譬如說我們沒有辦法對起來，全部都是切割的資料，法務部給我們一點資料，可是我們不知道這怎麼弄兜不起來，後面我們把那個人兜起來他的背景到底帶來什麼樣的東西我們完全沒有辦法做。這是不是只有大數據才能做我在想應該是。就像有一個東西給我們串，那這些東西可以幫助我們這數十年來再犯比例高的人他具有什麼樣的特色，那這東西沒有串起來這些人有什麼樣的特色，你進了司法系統之後，或者是你申請了保護令之後，這些人再犯年齡到底是不是有降低，這我們現在也完全沒辦法知道。那這些人遊走刑事司法系統有判刑跟沒判刑，到底對他的再犯發生有沒有影響，我們也不知道。(S1)

這些東西我們如果能做一些串連，像學姊講的先做第一初步的分析，把這些資料串起來能夠有一個比較完整的分析，我覺得對於這個領域的掌握很有幫助，比較清楚哪一些的人哪一些的處遇甚至哪一些的情境，比較沒再發生這樣的暴力。所以這個部分是我心裡面在期待說將來我們這個資料庫，如果能夠這樣幫助我們，那這東西真的是有幫助的。因為將來是不是要走更多的司法處遇或是不需要更多的司法處遇，我覺得這對我們政策的一個討論都是有幫助的。如果說這個分析能夠做的好的話，資料如果有達到的話，是不是就可以再去做進一步的預防線，如果我們知道這些人在某些情境下特別明顯，他現在已經放出來了在馬路上、得到社區活動，我們可以怎麼樣去連結他現在可以看得到的資料像網路，我們可以去比對或是說怎麼樣，如果能做到這種預防性的話，我覺得其實更好。或是說他現在的用藥狀況或是精神疾病的狀況，我覺得如果能再做更進一步預防性的話，當然是最好。可是現在這個能夠怎麼樣做，應該是要先從

分析的基礎才能夠有未來運用的可能，這是我初步在想像將來要能夠有用的話，以我們這個角度來看這樣最有用。(S1)

E1：

目前來講的話台灣現有的相關資料庫，第一個我們取得困難，第二個我們根本沒有一個仔細的分析。以目前家暴的資料庫裡面來講的話，他就是一個很大型的文字資料，甚至連所謂的數據上的資料我們叫做 co in 的部分的能力都沒有，他也從來沒有把他引用出來，把這個所謂的區域甚至受暴的歷史方面做過相關的分析，所以現在重新建構這幾十萬筆的資料具有一個分析成效的話，我們都覺得他有一個很大的問題跟困難，第一個部分就是說我們這些資料庫都可以整匯，也可以分析，可是我很納悶就是為什麼會取親密關係，因為親密關係他本來就是一個非常個人、私人去界定他暴力現象的問題然後他到底構不構成所謂犯罪或是雙方兩造互動等等，他其實還有一個解釋上的空間。可是如果說今天我要針對有關於性侵害對於一般人所在的威脅性或是恐懼感較大的話，性犯罪的作為所謂的大數據分析，我覺得可能會更為有用，而且對於所謂的資料跟合法的取得跟未來的接縫或是我們運用的一個概念上來講的話在整個政府的制度上或是民眾的思考，我覺得會比較能夠接受。以親密關係來講的話，會牽涉到所謂的私密性隱私性的問題，所以會覺得說你應用這些資料因為我的家庭裡面有親密關係暴力的存在，其中有一方有精神疾病、一個小孩子有用毒品，那未來就是一個很潛在的關係暴力的問題，我覺得這樣直接的連結其實是一個非常危險，而且會被社會大眾所詬病，我也不覺得這樣的資料出來就是一個實際的。如果以大數據上來講的話要達到一個預測功能，他基本上對基礎的 data 分析就要做可是我們沒有，然後我們還要有一群所謂的專家的團隊去了解這個數據分析的結果，去了解每一個數據他的意涵是什麼，他影響的程度是什麼，你才能夠去都達到一個更進階性的處理，那我會覺得與其如此在談一個很深度的東西，倒不如已經現有的資料分析先做好，我們才能夠看到更清楚的一面，所以我會覺得說這幾個問題，基本的大資料庫先做好再分析這是第一步。(E1)

P1：

我倒是建議未來要去整理可以從兩個方面去著手，一方面是如果從統計數字開始，統計數字本身就是可以做一些很多的比較；另外一方面就是如果說機關跟機關之間他願意有一些資料的內容可以願意去做一些比對的話，再從不同的資料庫研究說到底要萃取哪一些資料內容是有助於我們未來的分析。如我們回到這個主題是親密關係暴力防治，這個親密關係的暴力那個親密關係就有很多的定義跟界定，有婚姻關係、有的是男女朋友的關係或是同居等，這個範圍把它界定之後是一個暴力防治也許可能還沒有到達很多犯罪的階段，他有可能只是打算要分手或可能一些情緒，這些資料要不要進來，如果這個研究可以做到把所需要分析的研究都能夠整理出來，我相信對法務部這是第一階段很大的能夠跨出這一步。

然後接下來才能向實務單位說他其實可以去找哪些機關，可以去把這些資料拿來去做分析，所以我覺得一個階段一個階段，如果像剛法務部說可能有了這個研究案之後接下來可以讓實務機關去定所謂的中長程的計畫。那以政府的部門來說，科技部剛成立，像我們去年在提這個大數據的分析的時候，那時候還沒有科技部，那時候是科技規劃，他們就覺得說我們實在是走得很前面，還沒有一個部會我們都還沒做研究案就已經中程計畫也定出來了。當然因為我是從現有的資料去做，我在想搞不好我們提了四年也不一定做得完，可能未來要做的還很多，因為還有跟很多的部會我們還要跟交通部、法務部都還有很多的，但我們實際上是以犯罪偵查。不過我相信未來這個東西做完之後，或許法務部有需要可以大家彼此因為我已經做了第一步了。如果法務部還要繼續建這樣的資訊系統，因為法務部也有很多封閉式的系統，大家必須要去找一個地方是讓大家可以把這些放上去，譬如說像現在國網中心，行政院那個國家實驗研究院，他有一個全國的一個國網中心，他是有非常好的蒐集群而且他的速度傳輸速率也非常快，也夠安全，那可能找一個地方這個技術面到時候再去研究，最起碼第一步能夠把這些想要分析的資料然後透過大家的一些專業的建議，再看到的問題先從業務的需求面開始，其實這是一個很好的開始。然後未來我們再看看這個大數據這樣的技術面的演算可以幫我們什麼忙。我簡單的建議，也許可能大家有更多的想法，謝謝。(P1)

議題四、其他看法或建議。

W1：

我也想問剛剛講說數字跟文字裡從文字資料來做一些演算，我現在想到的是我們平常在做性侵害被害人我們會提供什麼樣的服務，因為對於性侵害被害人我們要怎麼去協助他去讓他可以心理專案協助或司法協助，可以讓他恢復正常的生活，因為我們社工會打很多的文字紀錄，所以如果對他做了什麼樣的服務，對於這樣的一個性侵害被害人我們要提供什麼樣的服務其實是可以 match 到他，因為我覺得很多的性侵害被害人我們在服務他的時候，其實 lose 掉很多跟他無服務意願，可是這些無服務意願的案件是，在社工來講他沒有服務意願在這個時間點我們去跟他服務的時候他就是不要，可能在我們開接案子就剪掉他了，那他的性侵害創傷 maybe 在結婚的時候或是什麼時候他才會呈現這樣子的一個內容，那我只是忽然曇花一現的想到說，這樣子東西有可能是在我目前的資料庫本來就有在提供這些案件服務，因為我們也會用實務的部分來去做一些研究這些案件可能他需要哪些服務，早日協助他做創傷預防或恢復正常的生活，可是這樣的東西有可能去做到大數據有可能做到這樣的一個協助，就是從這樣的一個數字的東西去抓出，可能這些東西是最可以馬上 match 到他。像這種東西回到實務面的部分剛剛老師也在講說我們怎麼樣讓他資料正確、哪些東西要進去，其實我覺得在實務上確實是很流於形式，就像性侵害驗傷哪些可能是用勾選或是這些東西是可以很明確的，明確的文字後我們可以勾選這個東西我就是要做或者說這個東西我做了就打勾，這樣子的東西其實在這個研究可以去抓出一些我們可能有的東西，可能今天相對

人的什麼東西我一定要去勾選不能去讓他放著，現在我們也在做就高危機的平台裡面要求我的衛政一定要勾你是否為精神病，你不能不空白的鍵因為這個人我已經在服務了，那你就應該要去鍵入這個東西。(W1)

另外的部分就是我覺得對我們來講大數據這個東西並不是資訊那一塊，對我們而言真的是一個比較模糊的概念，這一次的研究跟親密關係暴力防治很相關，他可能會找類似這一塊的人來做，如果下一階段還有焦點團體的話要找其他的一些相關的可能實務界的人，我覺得也可以在事前提供大數據的概念讓我們可以有一些閱讀、事先可以先了解，那其實回來這邊討論的時候可能在要進入會更聚焦。(W1)

P1：

其實大數據是可以去做這樣的東西，因此他的資料要很多，他是來自各個不同的資料。各位不知道有沒有注意到當你在用 google 的時候，他會按照你的習慣，當你經常去查可能是購物網站譬如說你喜歡買皮鞋之類的，他會把你最近去查詢的習慣自動列出來到前面的排行榜，那是一個智慧的分類，所以大數據是在做這種智慧的分類他會幫你判斷你的需要、你的習慣、你比較喜好這個、最近都在查這個。就像剛剛范老師講的那個就是一個價值，他是一個知識的價值，因為你本來已經有點意願想要去查這個，那今天他把你的需要放在最前面，你是不是就很容易進去到他的這個購物網，因為你提供他這個訊息而他已經預判你就是會去買這個東西，所以他把這個東西放出來。一樣的當我們要去判斷說哪一類的人比較容易去做犯罪的時候，你也要把他這種過去的這種 knowhow 這種 learning 放進來，然後讓演算法去幫你跑，所以其實大數據在這裡不是很容易做就是因為大家要集思廣益把所需要的、夠廣的資料放進來，然後再透過技術去把它拿掉。所以大家要集思廣益這個是我們現有的專業。我覺得這樣的研究是很 ok，因為大家可能在不同的領域，大數據是最後的工具，我們可以去搜集這方面的資料，想要達到什麼樣預期的效果，然後再透過這個工具去幫你統計，所以我覺得建資料庫是很後面的事情。(P1)

8.第二場焦點座談

一、時間：104 年 8 月 24 日下午 2 時 30 分

二、地點：銘傳大學基河校區 J310 會議室(臺北市基河路 130 號 3 樓)

三、研究主持人：范國勇教授(共同主持人助理教授韋愛梅)

四、參加人員(按筆畫順序)

S2 王朝煌教授中央警察大學資訊管理學系

J1 朱容辰律師瑜晶法律事務所

J2 高瑞蓮專員教育部學生事務及特殊教育司

W2 張淑慧執行長台灣大學中國信託慈善基金會兒少暨家庭研究中心

P2 莊蕙乃組長桃園縣政府警察局婦幼警察隊

S3 許華孚教授中正大學犯罪防治學系

J3 陳佳秀主任檢察官台灣台北地方法院檢察署

W3 廖書雯秘書長台灣防暴聯盟

H1 盧美凡行動心理師

J1 蘇媛瓊處長法務部統計處

J6 吳永達主任法務部司法官學院

五、座談討論

議題一、大數據概念分析運用在親密關係暴力犯罪防治的可行(能)性，您的看法如何？

S2:

作大數據分析最關鍵之處在取得資料，當今被稱為大數據時代，因為我們很多作業是透過電腦系統處理，而產生許多數據軌跡，也累積出海量資料，若要從中擷取出資訊可謂是千頭萬緒。雖然無法從中作出百分百正確的分析，但無資料就無法分析，所以需有 data for mining。但是，若希望能有完整的資料才能分析，去解決所有問題，那也是緣木求魚。到不如去找個具體的、簡單的、可取得、可掌握的資料庫開始實驗作起，慢慢去解決問題；在過程中就會應需要再去蒐集資料、再去作業，形成一個循環，就會越來越好。所以大數據概念分析運用在親密關係暴力犯罪防治的可行(能)性是可能的，但是要以循序漸進的方式，由小而大、由具體再擴到多樣性。(S2)

W3:

我們非常鼓勵利用大數據，因透過數據分析可以解析出行動形成的原因，就能於行動展開前將原因掌握住，比如說跟蹤是危險情人施行暴力前的徵兆，就可將之列入高危險因子。特別是在台灣，不管數據是否夠不夠完整，但通報網絡在，因

台灣是屬強制通報，有些國家還沒有。所以，我們可以思考台灣可以做出甚麼。我發現美國有些專門公司已從家庭暴力資料中分析出一些資訊，足見此方向是可行的。(W3)

J3：

大數據是一種趨勢，現法務部對毒品案件分析已運用此概念，在親密關係暴力分析上運用大數據是有可行性的；但重點在於數據如何取得。毒品案件偵防大多從行動電話通聯開始，這些資料尚屬容易取得，但現在通聯模式已擴至社群網路，而這樣資料相對是缺乏，雖然社群網路資料是可以拿來分析的，但前提是這些數據如何取得。(J3)

P2：

大數據或是未來的契機，不管是拿資料庫來分析解決問題，或是想解決問題而去檢視重整資料庫，至少讓相關單位——警察、衛福、教育..等等都會去重視資料(庫)的重要，它是可以協助我們去解決問題的，而不是只是單純鍵入填塞資料。這是一個機會，讓前線工作者好不容易收集來的資料，得以運用；也讓辛苦的過程留下軌跡，而不是不見了。(P2)

議題二、家庭暴力防治法已將未同居親密伴侶納入保護對象，預定明年實施，對在校學生、無犯罪前科的年輕人，如何運用大數據分析找到有危險徵候者，及早預防暴力的發生？

E2：

校安通報資料庫內容非常龐雜，以校園性侵害、性騷擾為例，每年約五千多筆，從通報到回報系統加以比對發現屬實案件六成不到。且此資料庫內未含有未同居親密關係欄項，因法源依據須至明年2月才施行。因此，要運用此類別資料還有過渡期。目前我們正建構學生親密關係暴力的量表，預計九月完成，或能於明年實施。(E2)

以預防而言，教育都是要面對的，比如情感教育、性別教育等等。在學生階段是被認為須要接受協助的時期與族群，方法之一是透過量表評估啟動協助機制。至於若已發生偏差性行為則已是屬更嚴重部分，或納入明年施行的家暴法的保護範圍內；屆時資料庫資料可能配合分為量表資料、保護(令)資料。(E2)

H1：

對親密關係中危險情人的定義是很辛苦的，即使以成年人為對象，將之簡單定義為變性人格，也不能算是定義明確，因即使變性人格就一定成為很可怕的危險情人，那也不一定。變性人格可能一開始是跟蹤、騷擾、監視而以，這或許可能是初期徵兆，但也非必然就會演變成危險暴力事件；也有些暴力事件是因被害人所惹起的。所以，要將這些人準確找出來，就是艱困任務。連精障的人雖已被診

斷了，但也未可斷言他必有自殺或殺人的可能。此外，現有學生輔導法，還須經其同意才可進行輔導措施，所以，預防方面還需經過法律鬆綁，不然還有些困難。(H1)

J2：

對於犯罪防治涉及人權思維，在於經評估為有犯罪潛在危險者，他未有犯罪實際行為意即未有著手事實，則警察、檢察官都無法進行刑事偵查作為；若對已有犯罪紀錄者預防則是再犯的預防事項。因此，若透過數據的分析就認定他是犯罪嫌疑者，是值得討論的。如果是在校園、軍營，透過量表分析而發現有可能犯罪傾向而加以輔導，或可理解，但如是一般犯罪預防，則其對象是一般不特定的多數人，就認定你將來會是或生來是犯罪人，這將引起道德的爭論。關鍵是我們使用資料的目的是要從過去的資料去歸納出人格特質，這就值得討論了。尤其是對過去從未有任何犯罪或不良行為案例的人進行預測分析，他並非被列管分子，只因經危險預測因子的分析就推定其有犯罪可能，屆時司法有何強制性作為？是要強迫其上課改造嗎？這樣過程是否發生標籤化結果也需要考量。(J2)

S2

模式建造與預測分析是二件事，但有關聯性。以 Target 公司為例，該公司針對懷孕婦女的行為歸納出模式，後據此做出懷孕預測模型，再運用於商業上行銷。家暴事件不同階段有不同徵候，發生前或有請求警察協調、身體有瘀青等紀錄、將這些被害人的背景、事前跡象作分析建構模式，進行預測分析建立預警系統，即模式建構是對已經發生的為分析母體，再以其模式來偵測未來。

以詐騙事件為例，以被害人角度，經被警察阻擋付錢而避免損失，應該是會感謝而無被標籤化的想法；但從加害人這端，則被預測犯罪潛在危險或許有被標籤化的可能。目前以發展歷程來看，家暴的預警分析重點仍偏重於被害人的保護上，那麼加害人標籤化的機率是相對低的。(S2)

J2：

用在商業的銷售上，他還有權利去決定要或不要，但若放在防治上，是不是為達效果目標就要帶入強制性作為，這是要思考的，因防治是與商業不同的。

從被害者服務的觀點是無標籤化問題，但從加害者預測則有標籤化問題。此外，加害者還要再細分為已經存在加害者與不存在加害者，對已存在加害者可以個案管理追蹤名義加以防治，並從中再分析是否有無奪命行為的傾向；但對不存在加害者部分，要以數據分析結果加以防治，會涉及人權、標籤化等問題。(J2)

W3：

講到資料分析結果運用的人權保護，個人認為保護是很重要但分析也很重要。以少年性侵者的資料為例，為保護他們而要全面保密，以致造成無法取得數據分析，

而無法據之加以輔導防治，導致受害者又變成加害者，是對社會的雙重傷害。所以，這些資料的開放與否要由專業去評估，至於分析出來的資料是否涉及法律關係、人權問題，那是下一步驟要去處理的，但之前是可以利用資料來分析的，透過資料解析建構出行為模式來預測可能性風險。(W3)

J3：

基本上，個人覺得應用大數據找到危險徵候會不會產生標籤化的作用，關鍵在如何運用這些東西，不是說你的運用方法一定會產生標籤化，比如說以找到的危險徵候作為宣導，此時若有此情形者或許能主動求助，就不會產生標籤化。所以應是可以運用的方式為何，例如，將危險情人與校安事件之間找到關聯性，就可以在處理校安事件時順便處理這議題，以危險情人徵候作為防治宣導工具間接阻卻發生，也不會發生標籤化問題。(J3)

J6：

政策工具有些是管制性的、有些是服務性的，所以重點在政策工具的選擇，若是服務性的就無侵害人權問題。(J6)

議提三、您認為(或貴單位)與親密關係暴力防治可能有關的資料(庫)有哪些？內容為何？有關的理由為何？這些資料有無(或規劃)開放？

J6：

目前政府每個單位各有很多資料庫，多屬於基礎建設型的資料庫，宜先將資料庫自行強化，若有不足處再行連結其他單位資料庫加以補強。(J6)

W2：

雖然現在政府有許多資料庫但還須審視其開放程度才可以作連結。比如毒品資料庫、教安資料庫、犯罪資料庫…等都與親密關係暴力防治可能有關，但有許多資料是不開放，甚至有衝突的，或是不能勾稽，比如健保資料庫，為保護病患隱私、為保密而去識別化致無法作勾稽。建議要將可提供資料的資料庫彙列出來，並調查其開放程度、得否連結、能否勾稽。(W2)

J1：

以法務部的資料庫而言，有檢察資料庫及矯正資料庫與親密關係暴力防治可能有關。尤其是檢察資料庫，是統計人員為了分析而去建立資料欄位的，因起訴書是一篇文章而非欄位化的，除自該案分案進入檢察系統所原有標記的分案欄位外，其餘大部分資料非有欄位型設計而難取得資料，是由本部統計人員主動自起訴書類的質性資料中擷取出可增加之欄位資料。在矯正資料分析部分，每半年統計處會作再犯分析，會去串聯警察的資料庫、檢察的資料，看看一年有沒有再犯、五年有沒有再犯，已經有在做大數據的工作，只是侷限在法務部的資料。

此外，建議性別資料及司法院資料也可以運用，但要注意司法院資料庫內判決資料並非全部必然都是確定判決資料。(J1)

S3：

從機關而言，從內政部的警政署、刑事局，教育部的校園(學校)，法務部本身及矯正署、衛福部、醫院，都有相關資料庫，目前已匯入衛福部的平台，以後的應用軸心應在平台建構的良善與否，因是將資料去識別化匯入的，是否可滿足使用者的使用目的，還需有一機關主導相關改進強化事務。(S3)

J3：

目前衛福部有家庭暴力資料庫，是有很多資料，是可以思考去運用的。但此資料庫仍有不足處，如高危機、TIDA 量表、被害人保護管束相關資料並沒有在此資料內，可以考量將這些尚未加入的資料倒入。(J3)

J6：

建議以短中長期策略擬定方式，先臚列出現有可被運用資料庫、開放程度、彼此檢視那些欄位有無欠缺，久而久之就會越完整，機關間的任督二脈就可以被打通。(J6)

議題四、親密關係暴力防治要運用大數據分析，需要具備那些條件？如要對不同機關間的資料庫連結，建議進行的步驟？

S2：

要完全整合各機關資料庫是有現實困難，因此建議由小而大，先掌握自己機關資料庫後，再連結其他機關資料庫。假設還不夠的話，甚至有必要建立機制去蒐集資料。有些所需要的實務界歷史資料是欠缺的，因為那時還未有數據分析的概念，以致缺漏而未建立。另外當初數據資料的建立並非是為了作分析而建立，是因應作業電腦化後而產出的。大數據分析是電子化過程中所附加的價值。

在大數據的摸索階段，還是要從大處著眼小處著手。從目標比較具體可行的從手。先就目前可用的資料分析，再逐步擴張至需要整合機關外資料的階段，最後才考慮建立資料蒐集運用工具。以問題深度來講，先從可以驗證的做起，再延伸至運用數據模型來發現問題。(S2)

S3：

大數據應與小數據作些差別，大數據已巨量至 TB(Tera Bytes)甚至 P(Peta Bytes)，若資料只到幾千筆、幾百筆，頂多用到 data mining。大數據是可以跳脫傳統科學研究方法，進行跨域研究。所以作親密暴力犯罪分析的資料到底夠不夠多，資料是否很零散，要事前研析一下。其中尚有個重點是，有些數據資料是短期的，或是遺失的，甚至有些格式不相容，以致無法鑑別並輸出。(S3)

W2：

數據的遺失或變動是要注意的，以目前衛福部的資料平台，非提供最原始的資料，因衛服部主管健康事項是非常重視個人隱私問題，所以有些資料被去識別化或為保密而遺失。另通報資料庫所依據的報表資料來源，常為新事件發生而修改報表，也造成資料經常變動。(W2)

J1：

從統計人員的角度來看，統計如何產生，就是由業務流程中產出。從警察機關將案件移送進來、檢察起訴、審判終結送到統計、執行端再送到矯正去。整個業務流程中有很多 cut point 產生很多統計。其中結構化資料與非結構化資料，都是大數據可以使用的，若要從小處做起，建議先就政府機關現有結構化資料做起。(J1)

不管是檢察、矯正的統計人員在作分析時，都需要對機關的工作流程非常的清楚，不然統計分析是做不出來的。因對機關工作清楚才可以對個個欄位的定義下得清楚、明確。所以，親密關係暴力防治運用大數據分析，若用到跨部會資料時，那麼它的欄位定義也是要清楚、明確；這一欄位等不等於那機關的欄位，不可看名詞就以為是一樣的。(J1)

另外是個資問題，現在對外提供資料必須是去識別化的，至於機關內部彼此間資料提供是否比照，尚難論斷，但資訊單位仍擔憂過程有洩漏之慮，所以提供給部內運用的資料還是會去識別化，除非有特殊需要才可做回朔，或提供專屬空間在內使用不得外流，如工商普查資料。我們統計單位對外所提供資料必定是無個資的，都是統計數據沒有原始資料(row data)，但大數據分析需要的是原始資料。再者是資料的動態變化問題，如刑案資料今天跑的與明天跑的結果是會不同的。為因應這樣問題，我們運用封存的概念與方法。例如月報表，到月底就封存統計結果，每月有個封存檔，並用統計年月命名建檔，而不以辦結年月建檔。當日所跑的刑案資料，即刻封存其原始資料，之後再要資料，是去找當時封存時的資料來跑的。(J1)

H1：

資料分析若是為了預防犯罪的發生的話，想提供醫療的預防觀念以為參考，醫療係採三級預防的概念。自殺預防亦是用三級預防概念，而最有成效的是在第三級，即對已經自殺過而存活的人加深防治的著力點，而使得自殺率迅速下降，並跌出十大死因之外。其獲得成效的重點就在預防的目標明確地放在第三級，並得到成效。(H1)

親密關係暴力防治是否要分級，將第一級列為的一般民眾宣導、第二級是申請保護令、第三級是高危機或保護令違反者或再犯。第一級又可分出需要通報類的個案，但此類通報不等於就需要保護令或有實際暴力發生。若以此分類方式，則對第一級的預防，難度是非常高的，要如何找出危險徵候是一挑戰。所以，我們若

以預防犯罪為目的去作親密關係暴力防治大數據分析，則是要針對哪一防治第幾級為目標，再就其範圍去取得相關數據、資料。(H1)

另外，很多家暴資料是質性的描述，我這邊會收到警察的移送書、社工資料、高危機…各種通報表，如何透過 data mining 技術去擷取分析是有挑戰性的。一般我們看這些資料時對一些數據及背景資料都很快閃過去，比較著重是看質性描述，看其犯罪意圖、家暴發生犯罪路徑。(H1)

S2：

大數據之資料來源是多樣的，質性資料是其中一部分，這需要技術輔助去擷取，比如以關鍵字、專有名詞等對比分析，找出一些潛在趨勢或徵候；不需要達到百分百的準確，就可提供警察作警力配置的決策，而不用亂槍打鳥式地去防堵。(S2)

P2：

家暴資料庫是屬通報性質，未經精緻篩檢，如二造關係不詳達四分之一，則此資料是不適合用的，但警政資料是經過檢視，從案件進來到成案、結案都在管制流程內。此外，家暴通報資料庫是在家防中心，後端很多資料是上不去的，以致有所缺漏，例如違反保護令罪未聯結。這樣現象致使警察自己另外多做個資料庫去補強，雖然是個小型資料庫，但會依工作需求而慢慢一直增加內容，資料就會愈來愈大，就會漸漸看出趨勢來。(P2)

另一問題是資料庫間沒有串聯，例如通報資料庫與高危機資料庫未連結，造成基層工作者很痛苦，以戀童癖為例，從桃園跑到新北，雖是同一加害人，因管轄權問題，就沒有連結。我們去查看資料庫，因高危機資料庫沒有帶進來，就無高危機的警示，而無法提示員警提高敏感度及警覺度。(P2)

W2：

目前資料庫是很片段的，例如新竹家暴案件，他的相對人是他的媽媽，結果是爸爸殺了小孩，從身分別來分是分屬數個資料庫，但都沒有串起來，同時也為了防弊，沒有社工身分無法取得他類別資料。(W2)

W3：

另對十四歲以下少年事件的資料塗銷，是否有必要全部塗銷、由誰塗銷、塗銷的資料是否仍可運用…等等問題，目前正在做整理，準備向法務部建議應該經過專業評估後再決定塗銷與否，而非全面性塗銷。除了專業評估資料得否用於分析外，要被分析的資料，還需透過第一線的工作人員以其經驗去檢視挑選出評估因子。(W3)

親密關係暴力防治的大數據分析研究才剛初步探索，建議先挑些更有實際作用的、或是極想驗證了解的，先進行探勘，比如說重複性、跟蹤、性主權…等與親密關係暴力間的關係。另也可以看國外所分析出的模式，藉以規劃出策略性規劃，策

劃在何範圍、以何方式去做。(W3)

S3：

目前正進行一研究，以百筆法官一審判決當作資料母體進行 open data 分析，看出其間差異性，這是先期的小範圍研究。但在作此研究時就遇到若干問題，比如這些資料要用何種檔案格式、有無建置搜尋的可能性、資料變相特性、是數字編碼或質性資料、可否轉檔，去識別化的格式數量…等等都需要去構思，最後要交給資訊人員進行跨域協助，去架構與研析出趨勢。這些的過程都需要人才、資金的投入。(S3)

J1：

當資料庫要與其他機關的資料庫串聯時，要同時提供 lay out。另外，還要面對欄位定義不同的問題；建議研究單位可以參考毒品的作法。除欄位定義外，還有原始資料去識別化問題，以前身分證號第二碼為性別碼，現在去掉後，就需要增加一個性別欄位，也就是產生轉換的問題。(J1)

P2：

資料庫的資料若是錯的，是很難以運用的。主要問題還是在定義上，定義沒有共識，這一部會對另一部會所下的定義就如同鴨子聽打雷，就無法登載一致性資料，甚至產出不對的資料。(P2)

E2：

有關校園資料，也有需注意之處，例如學生間關係問題是會被送到性評會去處理的，但是以性騷擾、跟蹤等來定義處理，要將之分類歸入親密關係暴力是無法定義的。(E2)

S2：

親密關係暴力要用大數據分析所面臨資料量問題，雖然少沒有關係，只要注意傳統解決問題方式是用假設去研究探討、用專家學者去分析，但現今利用資料分析、整合技術可以去挖掘出新的資訊，說明有價值的資訊是隱藏在資料內。大數據的重點在搜尋累積資料，分析資料，從資料中去發現問題、透過資料整合、排除迷思或假設。(S2)

J2：

建議從小資料、既有資料先開始作，家暴資料是可以先運用的，因若要跨部門去做，則會遇到欄位定義問題。先從家暴資料去做之後，有加、被害人的，再去比對高危機..等等。於資訊有限下，宜以個案去做。(J2)

H1：

建議可以針對社會既有發生恐怖情人案例進行小規模的分析，找出特徵，比如說自殺是不是危險情人的特質。有些親密關係暴力事件是連通報都沒有的，這就是大數據的好處；透過健保的就醫紀錄，或許設定些狀況，或有機會去挖掘出來。因此，從一些 case study 來瞭解危險情人的徵候或許可能發現跟目前所定義的高風險是不一樣的。(H1)

五、其他看法或建議。

S2：

大數據是蒐集分析找出軌跡，從其軌跡再去洞察行為，看著冰山浮出水面那塊，想辦法去推知沉於水面下那塊，浮於水面上的是受害者，至於加害者部分猶如冰山下，除資料分析推論外亦可透過傳統專家學者分析研究來補強。以現有資料來看，從受害者端入手分析較為容易。尤其所要找的徵兆不只限於一般所認為的酒色財氣這類因子，涵蓋範圍更廣，若從被害高風險著手，則較不陷入個資議題。因大數據資料很多，未來將面對資料淡化問題，在處理過程中如何將不常用資料去除掉，慢慢去淡化它。(S2)

H1：

從防制角度來看，二、三級防治是最有公權力介入與工具的地方，要做第一級防治是資料最少也最難介入之處。家暴致死案件因新聞報導而引起關注，但實際上其所佔的比例是非常少的，而為此而付出很多力量卻沒有效能，是很惜之處。年輕人的觀念有時受娛樂頻道戲劇節目影響，期待對方全心全意投入付出，未達標就想離開，但付出者會心有不甘，如此對立想法下，可能種下殺機。這也是要去宣導矯正想法之處。(H1)

J1：

統計分析就是在解讀資料，如果你業務不熟悉，解讀資料會有問題，此時統計人員會去詢問業務單位意見，統計是就統計數字作基本分析，當然統計人員對機關業務也有一定認識，但不足處會請教業務單位。所以解讀一定要請專家來解讀。此次研究案建議先不要做解讀，先做背景分析 profile，再建議後續研究。法務部對毒品資料庫的運用，第一步想要做的是想獲得吸毒者的背景分析，有此分析結果才能進一步推估出行為徵候。(J1)

J6：

此次研究是杯水車薪資源很少的，但在每個時代都有它問題發生，也都有它解決問題的工具。此次議題在犯罪防治領域還是沙漠狀況，值得探勘與開發，找出著力點是很有作用之處，因此很期待此次研究能借助時代工作來劃出藍圖，非是建立資料庫，是藉此藍圖吸引各領域能注意到這一焦點，來共同討論現有環節還有

那些問題，如資料庫還有哪些缺漏之處。(J6)

跨域管理最重要在資料串連，藉由串聯將任督二脈打通，找出可以遵循的軌跡。

期盼此研究先找出短中長期策略、繪製出研究藍圖，之後再與相關機關單位進行資料串連，逐步作出更精緻的成果。(J6)

9.第三場焦點座談

一、時間：104 年 8 月 27 日下午 2 時 30 分

二、地點：銘傳大學基河校區 J310 會議室(臺北市基河路 130 號 3 樓)

三、研究主持人：韋愛梅教授(共同主持人助理教授王伯頌)

四、參加人員(按筆畫順序)

W5 王惠宜科長台北市政府社會局

H3 李炳樟專員衛福部心理及口腔健康司

W6 林美薰執行長現代婦女基金會

S4 張芳全教授國立台北教育大學教育政策與經營管理學系

W4 郭彩榕視察衛福部保護服務司

J4 黃怡君副司長法務部保護司

S5 黃俊祥科長科技部政風處

H2 黃信得副院長永和康復醫院

E3 詹秀雯老師國立師大附中

J5 詹國裕科長法務部矯正署台北監獄作業科

五、座談討論

議題一、大數據概念分析運用在親密關係暴力犯罪防治的可行(能)性，您的看法如何？

S4:

這個大數據資料我一直在想會不會把它誤用了。什麼叫誤用呢，因為這個大數據就我們理解就是這個必須要有一個資料庫非常非常龐大，不是只有萬筆可能數十萬筆，像我現在在做的研究都是幾十萬筆的。而且不是只有一個橫貫面就是一個年度他是一個縱貫面，就是說每一個年度就是這樣子然後去做統合分析，就是長期的追蹤比較你可以描述的統計和比較簡要的統計或是進階的統計去做比較深入的分析，那找出他的因應喔找出他可能的結果、因素，我想應該是這樣子。那我不曉得說這個委託案所要的那個大數據概念是什麼因為我不太理解啦喔。那還有一個就是說因為這個資料大概有分成有公部門跟私部門，那公部門因為我以前是在行政院經建會，那我當時也建了一些資料，那建資料大概就是人口的成長、教育程度別、教育事業這些比較年度的一個比如說台灣中華民國它的教育程度各階層的程度是多少，只有很簡單的那種描述，它雖然是縱貫性但是它都是很簡要的描述。那樣本也不多。那我現在想要講的就是說，這個大數據應該是它的樣本數應該是很多的，我要講的是說可能是長期追蹤的調查，那像這個我們最近應該說大概有十來年，應該是早期現在就是中央研究院它現在歸併在那個社會研究所，他那邊有一個台灣長期資料追蹤資料庫，裡面有很多的資料就是很多委託案的研究都掛在那邊。我現在想的說因為如果用在親密關係暴力這個其實我不太理解我

們在這個剛剛提到的司法醫學犯罪社工臨床，是不是由我剛剛講的這一些。那之前我看到的初步的一個概念感覺上好像類似我們以前那個教育部統計就是失業人口多少犯罪年齡比率…我的感覺那個應該不叫大數據，這個在我的認定是這樣子，…那它的可行性是如何呢我覺得如果以現在國內的我不敢因為我沒有看到像那次我要來開那個會數據我可不可以先看，因為我沒有看到我不能去做判斷，就像醫生我還不知道診斷我沒有辦法提出那個處方，但是我可以這樣說就是說這個大數據如果做了之後，我是覺得是它是非常科學非常客觀非常具體甚至可以建立學理可以追蹤、可以預測、可以預防，我想這個比以往傳統用個案研究用個案追蹤的方式去理解個案，我想會更具有說服力。同時我想這個要預防嘛預防勝於治療，預防這個很重要。如果我們有這個大數據就是我剛剛所講的那種大型資料庫建置起來的不管是橫貫面或是縱貫面，如果有這樣的資訊呢我想應該是對我們這個議題裡面的親密關係犯罪防治有可以深入地去探討原因或相關的因素可以做一些對策出來。(S4)

S4:

我覺得這個大數據概念運用在親密關係暴力就感覺好像在大海撈針在撈那個親密關係但是現在這個大數據就是有哪些數據你也搞不清楚坦白講我也搞不清楚，現在國內搞不太清楚所以究竟這個親密關係裡面在大數據裡面究竟有多少親密關係的那些資料這個要很明確的界定，簡單講你要從大數據那個概念運用在親密關係轉換成親密關係的大數據，要轉換成變成這樣的定義好像才可以做要不然好像究竟你要大海撈針。(S4)

W4:

其實衛政體系會在談這個大數據的概念，我覺得說跟社政部門的距離有點遙遠，一直想要從這裡面找到一些關聯性喔，一般來講我們做社會科學研究我們會抽樣，但抽樣畢竟是部分樣本他絕對不是母體，那大數據他是數據大到幾乎接近母體因此你所得到的資訊他就是代表某種程度代表真實，但是數據這麼大的情況下你怎麼樣從中找到對你有意義的找到他的因果關係你才會真的說假設有這個因會有這個果我要怎樣的預測出來。然後所以那用大數據的話這個概念親密關係暴力的議題，假設我真的要去找有什麼樣可能性的話我覺得要非常具有想像力才能去想像。那我抽象的思考就是說我覺得大數據對目前來講一個是預測一個是預防，那對預測來講我們本身想要知道說假設你是為了要做資源的配置喔那你可能需要知道說那個趨勢在哪裡我比較能夠知道說我資源配置要配置多少，因此呢有可能說我想要去了解這個親密關係暴力的這個趨勢他可以藉由網路不用特別去調查親密關係暴力現行率的調查可以從現有的既有的一些統計當中，我知道這個東西比較多的時候大概我就知道那個親密關係會比較多的時候這樣我就省去了做調查的成本，但是我們要先去證明說另外這一個趨勢是跟這個有高度相關的舉例來講假如今天全般刑案那全般刑案裡面又分成多種不同的刑案，所以如果說我試

著去跑跑看說能夠去證明全般刑案的那個趨勢的那個線跟我們去做這個親密關係暴力的那個盛行率或者是發生率的那條線是很接近的喔，那或許將來我們就可以用全般刑案就知道說這個我們大概我的資源配置要這麼多，那除了這個刑案趨勢資料之外另外還有一個部分就是我們現在也大概比較清楚知道說原來我們上網常去搜尋一些關鍵字電腦網路就會自動記憶你這個人的偏好而主動傳送一些訊息或廣告來給你。那既然他能夠辨識的出來你的喜好而傳送一些固定資訊那那個資訊有沒有可能置換成我們的政策宣導或者說置換成假設他常常上網尋找一些我們覺得是比較具有危險高度辨識出他是具有危險暴力的，譬如說家庭暴力或者殺人或者說一些我們先辨識出來之後，那將來有沒有可能主動傳送一些有關請檢測你是不是危險情人，主動傳送一些自我檢測量表或者是自我執行的一些東西來給這些人。那另外就醫資料例如有一些人他可能有自殺的傾向，如果能知道這些趨勢的話接下來也許就可以進一步的仿效，那如果說務實面的來看說可以怎麼做的話其實現行已經有一些現在已經一直在做了喔。例如說像我們以往都是能夠掌握到通報出來的案子喔，但是這個案子主要不是關切你知道的部分，那我們過去這麼多年來檢討那些重大的家暴案件發覺說大致上分成兩種，一種是他本來就已經通報進來在 119 個案，所以對於這一類很明確的個案只能去檢討說過程當中我們哪裡做的不夠好，但是對於那一種一開始進來就死掉了他從來前面來講，我們這個部分案件就有一些意義有沒有一些其他的蛛絲馬跡是你可以在那個時候介入而有些提前預防效果他不至於今天在你面前就是一個死亡的案件。那這個目前在做的部份像高風險的概念或者是六歲以下高關懷的概念其實就是這樣的意思。那就例如說他有一些像小孩子出生後沒有去報戶口讓戶政機關幫你做出生的登記，或者是說沒有去按時預防接種那你會覺得說這可能是一個高危險群會做一些介入或者是說他沒有納健保錢繳不出來，這個你可能也會覺得說像類似這樣幾個體系還有那應入學未入學的喔，還有包括說什麼父母吸毒進監獄的，這些假設這些都有各自的系統當這些系統判讀來做一些資料的彙整交叉比對的時候，也許同樣的情況符合越多的體系的這個身分的時候你就是會大家優先出來得到的那一個，那目前就是說有在做的是這個來有另外一個部份就是我們保護資訊系統跟系統司這邊的自殺防治自殺個案的警報器警報系統，這兩個系統他很保護個案，不太容易用到，所以他目前其實大概頂多只有開放給我們去針對我們已經知道是家暴的個案已經知道是兒保的個案，進一步去敲他們的門說請問他在你們家有沒有列管，只是多一道讓我們知道他他是不是有在服務而已，但是對於那些一進入我們的通報案件的那些他們的個案我們是可以知道的。(W4)

S5:

就是大數據他不是一個研究方法，大數據就是巨量資料，他的後端的分析者才是關鍵，提醒我們說現在這樣子的社會有這麼多的訊息在天空像在這裡亂飛，包括我們坐這裡資訊就在這邊跑這些快速的資訊一直在滾動我們一直在丟出去，LINE 也一直 LINE 來 LINE 去，那些資訊都是大量的數據，那現在我們要更快速的現有

的資料譬如說我按一個 ENTER 就有好幾千萬的資料進來了，所以現在是要已經非常及時快速的因為你要馬上地分享出來的。再來我們可能要去注意的是，我們不是要看他的結果我們現在討論到的暴力犯罪我們看到的是個案他已經是結果了，可是我們要往前推的是找關係他到底跟什麼樣有關係，在哪個領域裡面包括社工醫療裡面他已經浮現出那一個數據，還不是數位化是數據，那個數據可能分散在各個地方因為是各個公部門他不會 SHARE 出來，當現在公部門我們叫他 SHARE 出來的時候那個馬上大量的資料浮現出來在某個平台上，就很快速的譬如說剛剛他才去醫院掛號的，回家之後他就打來了，他馬上通報出來的這兩個交叉一馬上他等一下再去做我就可以預測他等一下可能會去跟人家做衝突。對他每做一個場域不同的場域那個場域觀察到的地方馬上就可以做通報馬上就可以按那個鍵，那大量的數據進來之後馬上他的預警系統就會下來就會看了。所以我們可能要對大數據就是第一個首先了解他的特性其實他要有很大的有一個很巨量的資料而且他是隨時隨地都在變化的。那運用在暴力犯罪我們也可能要想到的大家有點疑問就是說個人資料保護法其實我們要去注意的反而是說我們絕對不會去公布個案本身我只要知道什麼你要給我年齡真的我也不知道是誰你只要給我這些在平常的地方我取得到的他竟然說我問他你喜不喜歡看電視喜歡看電視那個就是一個數據，我們把他抓出來那相對在這個數據在別的領域是不是也能夠連結得到他，如果可以的話這就是相關的我們要的就是這關係，其實看似好像是覺得說大數據好像這個方法不能用在這裡，其實我們都把他解開來運用他所需要的一些概念，其實還是可以的我們只要各部門給我們一些這些個案者本身的一些像日常 routine 的生活上的資料就好了，我就可以去做預測，也不會這部分不會，你喜歡吃什麼你喜歡講什麼話或者…通通可以這些資料通通可以只要你講的每個資訊數據接進來，我們後段的分析者才是我們將來面臨到的因為現在各個政府部門很最缺乏的就是統計分析的高手，那據說現在目前有用在大數據學校方面也有人在用了，譬如說吸毒者的那個試劑，很多學校的學生馬上用試劑去測，那個數據馬上就讀到電腦裡面去像可以用試劑去測，那個數據馬上就移到電腦裡面去，其實暴力犯罪他是可以建的，只是要有一點方法而且要持續做，不要只做到前半就停了，其實很多研究案我看都到這邊就停住了，沒有就差一點就貫通了，就差那麼一點，所以我是高度的期待說這樣子的研究能夠繼續維持到這裡來繼續這樣下去而且發展出真的在電腦上以後就有有的東西真的可以用，所以我們建的這些資料進來之後可能要想到的是他的相關的資料也是要從這些相關的再去多問一點多一點訊息進來因為他一定要主控到低量大，第二個這些資訊是可以被後製的有個老師說大數據有個特性就是說當時你建進來的這些訪談資料是個資絕對不能用可是沒有說你建進來之後我再去次級分析好之後我只是把你說講的話再次級分析再做很多資料分析再去寫一個論文然後跟你這個完全沒有關的我只是在做這些現象延伸的探討。這不算是這沒有法律責任也不需要當事人同意所以這些個案進來之後我們運用他相關的資料事實上有沒有違法呢那現在目前是沒有這個法律規範。(S5)

W5：

我對大數據沒有概念，要從那麼大的資料去撈到犯罪確實是不太可能就像隨機殺人事件其實你要找到那種關係是不大可能，你只能講的是說以那種現象啦，但是你要指出說他其實是會一定發生的我不知道。就是說是不是可以從公部門現有的一些資料我們現在就有保護的資訊系統，然後現在也有高風險的層級在中央然後另外還有教育的資訊系統甚至勞動的醫療的資訊系統，那我會覺得這一些系統我們可能可以找出幾個我們不是為了要抓犯人而是找到那個可能的嫌疑人，那種來預測或者是看到它的可能性，就是說這樣的 CASE 裡面他有沒有可能會產生的現象是什麼，這個可能我們在協助要介入的時候比較能夠感覺，例如在教育的領域裡面就可以針對那些 CASE 跟那些孩子他需要提供特別的協助監控他那他要提供關係，那我會覺得這個在做預防是比較有可能性的，因為我也看到是我們如果認定他是一個嫌疑人其實是一個很大的問題。但是如果我們把它轉化變成是說這樣的家庭背景譬如說他有這些的情況其實是我們要特別去關注的提供一些服務的資源給他。那這個服務資源事實上有些方法可以是說我們發現有些家庭他其實是應該提供給十歲以上的家長一些協助因為呢我們現在的孩子他可能未滿十二歲他就進入青春期的，那但是很多青春期的父母親他其實是不知道怎麼面對他青春期的孩子，所以就產生很多的親子之間很大的衝突，那我們就要提早去做這樣的預防，可能要在所有的課程裏面都要增加對於孩子的認識或者是在社區裡面有這些資源，那我會覺得是對這個工作者比較有幫忙，那我會覺得這個價值那當然如果可以從這個大數據是不是可以把他找出這個關聯性出來然後我們可以更聚焦，把把重心可以放在我們主要活動的或主要服務的對象那我們要做一段時間一個長期的觀察之後我們看這個是不是真的有下降，不然的話我們就會覺得是一個方案就是很短暫可能一年兩年然後你也不知道他到底有沒有什麼影響，我看到台灣我們好像沒有一個長期的研究，就很可惜就是說今天你有這些數據但是沒有後面的配套是沒有價值，那你要的一直做到至少十年你去 FOLLOW 到這個孩子長大他可能在婚姻的交流或人際他其實是產生什麼樣的影響那我會覺得這個對我們來講是因為比較有價值的喔。(W5)

J4：

我所理解的現在是因為是從商業界先累積的大量資料然後從這樣去看出一些關聯性得到一些預測但他這樣的預測不是針對個人。可是台灣我覺得我們現在用的大數據我很怕這樣的一個科技國家我的個人因為其實那個毒品的大數據其實我把它定義是叫做個人毒品資料總歸戶，那衛福部那時候的想法其實我們聽起來是毛骨悚然雖然說會去識別化就是用身分證字號去串，我現在只想到要跟毒品資料有關的各個部會的毒品資料用這個身分證去串，也就是我可以把你這個人的身分證這樣去串起來然後醫療要進來愛滋要進來那這樣去串的時候我們就會看到他這些所有的資料然後在去掉那個識別化，可是事實上他沒講到那個識別化他

是有有可能是被復原的，而且誰能夠保證今天科技人員這些東西都保管在他們的資訊或統計單位誰能夠保證今天沒有人會撈出來去回復，然後你的個人資料就是全部在這裡，我們今天是真正的大數據其實是現在公部門資料都是為了行政管理所累積下來的資料，那當然少部分為了通報或某些醫療的通報那這個部分是不是能夠真的找出關聯性那如果要治的話看起來比較是用個人的資料去串譬如說我今天一個人我用身分證字號的時候我去串他的犯罪資料我去串他的醫療資料我去串他的婚姻狀況或者我去串他有沒有自殺的資料或者我去串他種種的資料裡面或許我們去從這個資料裡面去串的時候串很多很多資料的時候，我們可能可以看到說一個有腦傷的病人他其實是家暴的個案，然後他可能是家暴的個案，或者是他自殺一個自殺的人曾經自殺的人他可能是自殺的加害人或是被害人那你有沒有可能從這些資料先去串起來以後先去 RA 起來以後在去串出說告訴我們說有哪一些因素的人他其實容易成為加害人或者成為被害人，但是我是覺得如果你要撈出危險個案的時候我覺得那可能不是大數據的概念，我覺得那比較是個人資料總歸庫的概念。可是對於一個不是加害人的時候你要這樣去弄的時候其實這真的是有人權的議題啦。我真的覺得這人權的議題就是說國家的權力要介入到這麼深如果今天國家為了要防治一個犯罪他可以把你所有的資料去串起來去分析你這個人會不會成為家暴的加害人會不會成為性侵的加害人諸如此類的話那今天國家可不可以把一個人的所有資料都串起來然後預測說你可能是一個，我覺得有一些人權議題是要去討論的，我其實是有點擔心說大家現在在那個大數據這樣的一窩蜂之下你可能會嚴重的侵害到個人隱私。(J4)

W6:

預測一個很廣泛性的社會現象就是像說我今天某一個地方刀子買多了酒買多了棍子買多了這個我就預測整個台北區的這一個的家暴會增加。這麼大範圍還是一個比較小範圍這一個我了解到這個加害人 PROFILE 被害人 PROFILE 就是我要預測這麼小的東西還是要預測一個很大的那聽起來是小的東西我覺得預測他是不是加害人被害人之後哪一類其實不要說到行為模式分析這種東西應該早就有了吧，我們在業界學界應該也都會有一些可能利用以前一些研究然後就可以形成一個系統然後就可以做一個預測，那如果這個預測有跟未來假設你用更多的數據去支持這個那只不過是一個 UPGRADE 的版本，就算沒有到研究其實以實務來講你還是會有 FU 啊，你做久了做到累積一個量你個人其實對於一些判斷其實也是有的只是他沒有到有數據支持的這樣子的一個 AVADANCE 這樣。那所以我現在就是我不知道我們是要把層次拉到哪一個或某一種層次才是我們現在未來可做的，像我覺得你說行為預測的那個部分這一兩年我們也在跟美國的一個學者他在做的是一個「跟蹤行為的預測」的東西，那他那個已經全部電腦化可是我可以想像他們在發展那個行為預測的過程是很複雜的，因為你要把所有加害人可能的這種那個那個就是裡面的變項喝酒犯罪什麼變項通通輸入之後多少幾十萬筆就像剛講的很大的一筆資料然後再分析他會做出來的各種然後於是你請被害人填了這個問

卷之後他就會跑出一個未來你的跟蹤你的跟蹤狂的這個人你要注意哪幾個層面，他沒有分數但是他會跑出整個犯罪的 PROFILE，那因為這兩個我們在跟他合作這個東西我可以想像他們當時一定花了很多心血在做這個可是我不知道這個有沒有接近到這個大數據的概念說針對某個特定的犯罪行為我覺得太廣好像對我們目前來講沒有太大的幫助，要馬就是說你層次很大我可以預測整體台灣未來家暴是更加嚴重的或者是家暴一定會什麼有關可是那個有關又不是我們現在所知道的什麼喝酒這種，因為如果已經已知的我就沒什麼意思了，還要一點未知的像這樣子買棍子的或者是少子化還是什麼之類的會有關連的，有一些新的東西出來，對那要就是說特定一點行為的預測親密關係暴力裡面的某一個行為我不知道這樣會不會比較聚焦，這個問題我也覺得說如果太特定確實我覺得這樣人民的隱私權是受到很大的這個衝擊。(W6)

J5:

大數據的可行性有以下三點要特別注意就是說資料完整性的困難，那運用大數據概念在親密關係暴力犯罪防治上應該是可能的可行的，但是至於說可行性需要四處去蒐集與項目的建置的廣度跟深度例如像生理心理或生活習慣或生活態度等層面還有民眾的接受程度。第二個資料的運用困難方面大數據是一個龐大的資料庫，假設他的資料對於分析者而言是有用原則上是可以幫得上忙但是關鍵是如何分析資料分析者本身的狀況及能力還有人事互動影響的變動的發展性的那研究的結果要如何得來實踐這都是我們要來考量的。再來資料的保密還有人權的質疑就資料的保密性是應該要先納入考量的重點用大數據涵蓋什麼樣的私密性的內容政府是不是會被認為想掌控人民？誰可以從這個來獲得最大的利益。(J5)

議題二、家庭暴力防治法已將未同居親密伴侶納入保護對象，預定明年實施，對在校學生、無犯罪前科的年輕人，如何運用大數據分析找到有危險徵候者，及早預防暴力的發生？

S4:

我覺得把親密關係內涵先界定清楚，他的對象他的適用性他的特質因為我看起來這個還是蠻粗略的，未同居的親密伴侶他上面可能是家屬之類的，量表我是覺得要建立多元多類型而且要有信度效度量表而且要長期的研發，這是在針對什麼如果是按照這個議題的話就是親密關係量表要多元多類型。可能要長期的追蹤我是覺得說諸位剛剛有提到學校階段小學國中高中大學甚至研究所，有沒有想過那個學校輔導中心就是那個階段結束了之後學生到另外一個階段要重新另外一個建置資料，他沒有一個資料的延續性，所以同年代不同階段這個銜接的問題，還有譬如說我已經出社會了那一些社政機構這一些相關的資料的他的一個長期追蹤就是對那一些個案或是這些樣本追蹤我覺得這是很重要的。那還有分析找到那個有危險症候群可能要去歸納這些資料很多過去的這些犯罪像他的特徵像家庭啊家庭特徵是不是低收入戶殘障的或是很多，那個人的特質譬如說比較自閉的酗酒

的這一些那一些封閉的個人特質，交友的狀況教育的狀況生活的習性等，歸納一個基本的參照。那我可能用這個方式去研發我剛剛提到那個量表的那個施測那我們可以從學生開始有去建立這個資料。(S4)

J5:

家庭暴力防治法已將未同居的親密伴侶納入保護對象預計明年實施，還有在對在校的學生還有無犯罪前科的青年如果運用大數據找到有危險的徵候提早預防，那針對在校生可以利用現有的心理量表還有教師的質性訪談還有觀察的資料來下手那找到危險的徵候者其他的一般學生盡量加強心理健康的輔導策略，但是找到了危險徵候者是否反映出教育制度有需要優先檢討或改進或調整因為我們如何定義危險，那定義危險的量表或理論的效度如何判讀有沒有錯誤教育資源有沒有誤用，這個都是我們將來必須要要注意到的。再來針對社會上沒有犯罪前科的年輕人我們是建議說可以在各企業單位或社區活動中心或婚前來提供這個親密關係相關的教育課程，就是我剛剛講說的第一級的這個預防，那必須注意到標籤化這個人性那資料庫的內容普及於每一個人又每一個人的身家背景醫療病例生活習慣都要被拿來攤開來分析，然後被歸在有危險徵候的這個群組裡面的話，被列管或加強輔導恐怕不是任何人都能夠欣然接受的。是不是也要參考先進國家的做法是不是有侵害人權的疑慮。(J5)

議提三、您認為(或貴單位)與親密關係暴力防治可能有關的資料(庫)有哪些？內容為何？有關的理由為何？這些資料有無(或規劃)開放？

S4:

我們在學校裡面沒有這個資料，我建議就是說因為這個大資料庫現在概念模糊但是也不是說不能做，資料庫非常多散在各單位中央地方全部都有，全部要做一個統整跟盤點，不然每個單位都建一個非常雜亂那到底功效如何這個可能有待評估的，建議社政單位醫療衛生如果研究團隊請中央這些單位提供或是你們的人力去蒐集然後來檢討一下哪一些資料庫哪一些資料正在建置的已經過去了可以提供或是不能提供可以整合或是不能整合。(S4)

S5:

其實大數據他有一個迷思和隱憂啦，迷思就是說我們現在對市場是錯的我們以為大數據巨量的資料多就有用這是現在目前官方面臨到的，他以為說現在把大家的東西通通推上來那個就叫做大數據分享，其實我們會不會因為像現在一昧的覺得大數據那麼的有用而一昧的丟數據出去，大家都不是專家了因為他快速的拿走你的資料之後你們不會有專家坐在電腦前面的分析師他才是專家，因為他快速的抓重你的東西，越端越多資料出去事實上你的東西就越來越沒有現在保有的是有價值的數據，他現在告訴你希望你 SHARE 出去 SHARE 出去事實上他就不要你的了，因為你只是他大數據裡面的一小部分，就是這樣 SHARE 出去事實上是不是在為了

跟潮流而潮流，事實上我們自己保有很珍貴本土的犯罪的資料，所以說這就是我們現在可能要考慮到的就是說到什麼樣的程度如果是這些其他相關的因素的這個層次。(S5)

H2:

家庭暴力防治跟精神科的醫療或臨床的醫療在醫療的 DATA 是可以預測的，但是那部分應該不叫大數據我們就今天針對也許做一些問卷然後開始可能這個案的基本的資料，身高體重血壓啦抽血的這些東西我們抽驗他的習慣他的家族史甚至他的社會知識來預測這個人會不會憂鬱會不會得腦癌會不會得什麼病，可以預測他百分甚至預測他的壽命，這是我們運用一些數據來看，那我覺得說防治應該是分兩個部分，親密關係暴力的防治一個是預防就是我看到的那個趨勢好像越來越多了，那治是很簡單一個是很清楚的議題譬如家暴是不是情侶多還是親密關係暴力多還是兒虐多我們可能可以去分類，那我們再來做這樣的議題，所以我們在預防我們常說治療是比較 TOPIC 診斷確定了你就可以去做治療，那我覺得大數據應該比較像是可以做預防的部分。(H2)

W5:

其實是要從日常生活很多的數據而且我們現在看公部門的數據這些是不會存在的，譬如說資訊系統在裡面會去抓到有哪一些是最高的風險，那因為北捷的那個案子所以一直在講怎樣可以預測到這些犯罪人是在哪，但是對我們來講那是困難的，我們也找了一些專家來討論，那甚至就是找了那個心理衡鑑就是針對鄭捷相關差別殺人的老師來談就是說那怎樣可以預測到這些有可能犯罪，那他是說這個確實也很困難，因為他很沒有在一般的系統裡面出現，所以就會變成我們很難抓到但是我們有抓到幾個是比較明顯的，他可能是長期失業然後在經濟上遇到困難或者是那個就業的時間很短他不斷轉換工作的機會，那基本上他在基本就業成就上面跟經濟維持上面就是很大的問題。那第二個就是可能是在學校裡面就是高關懷的孩子，那這些孩子其實應該要建置尤其是家庭的資料，那他後來顯現的可能只是一些行為的資料，那篩出來有一些是來自於精神疾病或者是精神醫療上就醫或者是在醫院裡面的一些相關的資料。那另外我們看到是家暴，有時候也還真的預防不了，有些案子是結案之後才發生，所以就也很難知道在過程到底哪一些是很重要的關鍵，那當然我是覺得如果真的可以去跑到有一些什麼經典的文件，那的確我們就可以把長期一直 follow 但是可能沒有辦法像是社工的 follow，就是應該是要動員有關社區的資源。(W5)

W5:

公部門其實是有非常多的資料庫那在台北市我們是也自己開發了非常多的資料，其實各縣市真的很不一樣，因為在台北市就是會自己開發很多資料然後中央再開發，所以我們常有很大的問題就是可以建議是說在中央的系統要有一個比較完整的

資料庫，我後來發現我們的基礎資料是不足的。(W5)

J4:

我今天真的要用大數據要能強調是，因為我真的覺得政府部門現在所提供的資料是不是能夠有這樣的巨量資料，而且我覺得這個東西他是要去分析出我們現在所沒有注意到的我覺得他才有意義，那我真的覺得大家要非常非常謹慎思考，因為像會議我現在聽起來會比較像是可能是從個人資料的總歸戶，然後去識別化以後然後有把每個人個人資料的總歸戶去連起來，你可能或許可以發現一些這個東西，那這個也是一些大數據的一些運用，只是說你要達到你要蒐集的這樣的數據，然後去看出一一些我們不知道的東西的時候這個過程當中其實是牽涉到你把個人的資料串連，那這個東西我們就會很擔心，你現在來講所謂的去識別化，事實上某個程度我也可以說他是騙人的，因為他們自己也承認說那個時時是很容易回復的東西，其實我很擔心就是說今天政府一窩蜂在用大數據到底那個是不是真的大數據，用的對不對會不會侵害到人民，那我覺得這應該都要考量進去，我不認為他是萬靈丹啦。(J4)

H3:

實務上這些高危機個案很多都已經有精神疾病或自殺，所以其實我們兩邊現在主要面臨資料做串連勾稽，如果說他那邊在案我這邊也在案的，那我們就會特別去提醒訪員公衛部可能去訪視的時候你要特別留意，看看他家人的一些狀況，目前第一階段是做這個去提醒訪員，你的訪視的對象精障自殺又合併有高危機個案那你就是著去留意加強防治，因為實務上發現這一群往往都會最後可能會衍生後續一些刑事案件，那性侵害的部分呢？其實我們也嘗試要做啦，因為我們性侵害有涉及到說他要做危險等級，而且他高治安的危險所以要送到強制治療，那目前這些危險分級都是套一些國外的量表或者是說治療師的治療要你去報到評估小組幫你背書，那我們現在有嘗試就是說利用那個家人的資料庫，我們想透過他的這些初犯的年齡，他的一些犯罪手法態樣場所，我們去給他做一些歸戶，是不是可以找出這群人一些危險因子，做為你那個治療師或評估小組的判斷的一個依據，要不然那個真的是天馬行空，當然如果說他都沒來，警察也說查不到，就說你應該有一個佐證的資料，到時候你要去聲請法院的強制治療，會比較有依據，那強制治療一進去以後幾乎有可能都出來，所以這涉及到很多人權可是我們現在分析，從他們的家人資料庫裡面來做一個歸戶的整理，希望能給治療工作一個參考，目前我們是打算趕做到這種程度…那個矯正機關倒進來的，我們現在只敢看矯正機關倒進來的，其實部裡面也有在做這方面的研究，昨天部裡面還正式有一個巨量的研究計畫，那個研究計畫那是院層級的大計畫，一部份是跟醫療有關的要去做分析就醫民眾的這些電子病例，他的這些比如說X光片這些東西，那另外有一個大區塊就是衛社政的整合，他們也發現高風險族群，所以有設定自殺啦家暴通報兒虐等做一些數據的分析，包括那個後端分析員的評選，所以其實院有交代部裡面

也有積極在做接觸，我們之前也有從那個健保資料庫裡面去分析性侵的被害人，他在後續的就醫行為上面是不是跟一般人不太一樣，或者是說他有特別去看什麼哪一類別科，我們初步分析有一定比例的人去看精神科身心科，所以我們這樣分析以後我們就知道說我們後續治療要放哪裡，我們健保資料庫我們就看被害人去分析他後續的就醫行為，那就跟精神科醫師跟婦產科醫師講，個案剛進來的時候做驗尿採證的時候，你可能就要照會精神科，因為實務上發現有一定的比例會看精神科，所以做驗尿採證要做一個照會的動作，所以我們是從那個基本資料來做分析那這個後續。(H3)

J5:

我們這邊監獄裡面目前大概收容家暴的人數如果是親密關係的暴力防治還有所謂的危險情人，那我們現在個案目前到 8 月 20 號為止有 401 個家暴的個案，那單純的家暴個案大概有 169 名佔 42%，那家暴合併性侵的這個個案有 218 名，然後兩者皆有的話就 112 單純的是 94 個那 18 名是家庭暴力合併性侵，那這個數據是說他的每一個組成類型大都是很多啦喔，那大部分都是說以單純家暴還有家暴還會合併有性侵的，比例也蠻高的，我們裡面台北監獄裡面性侵犯大概有一千兩百多名，那家暴的人數占我們總收容人數三千六百多人幾乎是九分之一，那算是比例也蠻高的。他就是刑事案件，其實會進到我們這邊來應該是有一些保護令已經到監獄應該是狀況都會比較嚴重。然後他這些人的人格特質比較衝動、易怒還有時候很直接個性也很直接、而且會有偏執固執的情形、而且他的文化之間也會比較少認知功能也比較偏低，那這些收容人裡面他也常常伴隨有酒癮、毒癮、精神狀況或生理方面的，那我現在討論這個部分我們講究三級預防，第一級當然就是屬於那種宣導吼對一般性的採取全面性的預防宣導認知的一個宣導，那我想說因為我們這個議題代表我們大數據運用在親密關係暴力防治的探討當然這一定要可行如果不可行那這個議題就沒有討論下去的必要，那我想說我們應該是著定在說致力於這種辨識潛在的這個危險因子，那要可以建立說那個指標一個評估危險的指標，那現在我們可以講到現在到大賣場你木炭的話你隨便都買得到你要去櫃台要跟櫃台老闆娘問報告我要買木炭，那這個就是一個防衛的機制，從社會的一個你只要大數據是一個巨量數據已經接近於事實了那就變成是一種趨勢了，那那你說像危險情人我們其實也一般在做危險情人他個性比較不成熟舉止很衝動然後他比較對事情他有一種很過度很喜歡批判這就是一個指標。然後他比較自我時常會而且遇到事情他比較不容易妥協，根據這些危險情人或親密關係人或他的控制慾比較強那對女性有時候會有貶抑輕視或者用不堪的字言來辱罵，那處事還有跟人互動的時候他可能在一個成長經歷都是用暴力解決的就像小時候青少年的時候像在青少年輔育院像陳進興那個案子他小時候就是它就如果當時能夠在某一個環境就把他矯治好或說給他感受到痛苦他可能就不會延續到後面的這個白曉燕命案挾持武官造成國際轟動的一個情形。那這個可能都還有平常是否患有酗酒吸毒或者還有一個指標虐待小動物，就是說這個東西他是不是有一種他殘

忍的一個性格在裡面或一個殘忍的心態這些都是可以我們可以來預防。那我想說第三級預防應該就是到我們監獄裡面，這等於進到這邊來我們對他的一個狀況做一個治療譬如說給他上課啊對於他的狀況做一個管理處遇還有控制，那將來他出去可以避免再犯。那我想說我們今天討論這個議題應該是在第二級預防裡面建立一個很明確的指標其實我們想一想這個大數據裡面的一個關鍵因子還是怎麼樣我們覺得說可以導向說因為你大數據巨量的數據你沒有做你沒有做篩選做分析甚至於解讀的人他有沒有那個能力還有他能不能擷取要點的話那其實這個大數據其實用在這上面可能就是說大數據是可行的，他是一個參考但是你必須要做篩選出關鍵因子才會比較有用比較有意義。(J5)

J5:

在我們監獄裡面有所謂的獄政系統我想每個單位都有這樣的資訊管理基本資料的我們這在這邊有一個名籍系統可以查到收容人的客觀的一些資料譬如說家庭個人哪還有說他的交友狀況職業等等還有甚至於小孩去見他對他的一個觀察在主要各科室還有衛生醫療還有沒有什麼遺傳疾病什麼的，那在我們的教化的系統裡面對於這個家暴的人他會有一個治療過程或內容但是這個資料是屬於主觀的判斷還有治療的處遇模式這些資料都是很保密的目前是根據這個資料還沒有開放，那未來是不是說基於這個個人資料保護法第 16 條基於公益的這個部分可以或第 15 條可以用這樣的一個解讀。(J5)

E3:

學校可以用的資料庫其實現在新生入學的話我們從新生訓練的時候我們就開始建一些資料庫讓他們可以直接在那一些比較傳統的資料運用，譬如說像校務行政系統像現在所有的輔導老師，輔導紀錄通通進到系統裡面來填報輔導管理系統所有的輔導的紀錄都在裡面。學期成績管理系統社團這裡面包括一些缺曠課還是一些獎懲計畫都在裡面，可是這些是一些比較傳統的一個資料庫的運用，大數據的概念比較像是我們過去可能會假設單親的孩子可能會怎樣，可是大數據他的一個概念分析可假設可能是幾千個幾萬個都有可能就直接從這麼多的母數裡面我們來建構出一個模型而這個模型可能不是我們當初所預想的，他可能某一些相關的概念是超過我們所想的一個因果關係但是他就是出現了而我們不需要去了解為什麼會這樣但是當這一些概念出現的時候他就是要發生了。那這一些概念有可能是什麼那我們剛剛就有模型建立出來之後也許我們就可以有些預防的概念也許說不定是畢業舞會之前還是社團成發可能在學校裡面有這樣的概念，那我們在學校裡面可能我們就可以做一些預防的措施或者是他的搜尋的字也許跟分手劈腿這些搜尋的關鍵字可能有一些關係。那我們目前在做這個階段可能還沒有去考慮到這一些隱私的部分對我們的一個影響，如果我們只是研究出一個趨勢的話，那我們在這個運用上應該是還比較沒有問題的。(E3)

E3:

那我想目前的話比較可以用的資料庫譬如說像教育部有一個校園性侵害性騷擾以及性霸凌事件的統計管理系統，那其他還有校安的通報系統，那這些是共有共通的部分其它我剛剛所提到的校務行政系統輔導管理系統等等這些系統其實是各校各校有各自不同的系統，其實是沒有共通的，那如果我們要運用這樣的資料庫的話是不是可以請那個公部門可以用共同的系統來做建置才會有一個相同的樣貌我們在做資料庫的選用部分會比較容易，那再來就是我們可以建置一些像人格量表等這些部分其實我們在實務上會像興趣量表性向測驗這一些部分那由於我們是升學學校所以學生在這部分真正填答部分還沒有什麼問題，不過像大考中心就給我們看過是有一個是學生對這部分如果比較沒有心的話可能會填寫的像是毛線的編織圖案等那就是會有一個資料真實性的問題，那大數據的概念之所以我們會覺得因為對他來說他覺得這個部分因為去脈絡化的問題的狀況他覺得就這樣搜尋他就是一個安全的情況，所以這個部分會是比较比較真實的，那未來如果我們要把這個部分放進來的話，可能在新生入學的部分有幾題效度他比較可以放下他的戒心而不是一個主要是一個那麼那麼大的部分他覺得他填下去然後你都知道我是誰，他會覺得沒有沒有一個比較安全感。第三個部分是那個轉銜的部分喔，那目前的確是不同的教育階段每一個教育階段都有填了這樣的資料但是都沒有一個轉銜的機制，只有特殊學生的部分，就是從國小一直到到大學他只要同一個 ID 他這些資料會一直轉接上去。那其他的學生其實是沒有的，那這個資料是非常珍貴的如果可以接下去的話資料庫可以做更好的運用。(E3)

議題四、親密關係暴力防治要運用大數據分析，需要具備那些條件？如要對不同機關間的資料庫連結，建議進行的步驟？

S4:

條件上第一個是各單位當然經費的投入，還有研發的委託案、資料的整合、人力的投入還有人力統計，那個法政單位的統計的人非常的欠缺，那我就在想剛剛第一個是從那個單位那個資料庫要那個單位可能要做這些事，那專業能力第一個就是統計分析的能力譬如說層級的設計，統計的素養高階統計像多變樣分析這一些應該要有，我的意思就說要釜底抽薪的去長期去建立去培養這個人力，這個才是對的這是第二個。那第三個那個單位人力的進修很重要，那我是覺得你主管本身你不以身作則，雖然不一定要有這個你一定要採取但是你一定要去尊重因為這是多元的，你要去了解這個是很重要的主管本身這個也很重要這是在第二部份條件上…那第三部分很關鍵的是什麼在資料裡面我分兩部分第一個是盤點，一個是資料的建置，我剛剛講的就是說要各單位願意開誠布公的把它清點一下那些可用哪些不能用有些真的根本不能用了，那這個沒有用那看有缺哪些要來建哪建立，那這個建立呢我這邊提了五點，第一個就是資料結構要建置那個資料結構就是要長期的請這些專家學者看我們需要哪些以這個如果是親密關係，你看看哪些變數哪些變項像，所以在這個資料庫的建置就很重要，就是資料要先夠，那第二個就是我剛剛講的那個變項的界定這個變項的界定很重要，每一個單位項國內就是各吹

一把調甚至我要講的是跨國的資料你的資料庫你的資料的這個大數據是不是跟國際上能夠接軌，你要比較我們常常說犯罪這些要比較那我們比出來的都是雞同鴨腳，思考跨國組織犯罪組織他是怎麼去定義這個。那當然在是那個樣本這個怎麼去追蹤這個就要長期的去追蹤，那再來是要長期去維護這資料庫要常去維護，然後第五個就是要釋出分析，剛剛有提到說有些可能他很保留他不敢釋出去有些有些是可以的。(S4)

S5:

我會建議先從官方資料為主的這個資料庫的分析出再去找數據，這個數據是能夠相關於外面的一些大數據的東西我們要善用相關那兩個字，也就是說我們官方資料現有的資料裡面去找例如說不管是保衛因子還是危險因子而這個因子是跟某些日常生活或者是某些公部門有的資訊是相關的，我只要找出這些相關那我我可以確認的是，我將來的是只要外面有的訊息撈進來我就可以跟我現在有的官方資料裡面我的預測的高風險或者是高保護的一些因子相關的我會建議可以朝向這個方式，這是一種大數據的運用。那現階段我建議第一個官方資料為主或是那我們去分析這些官方資料裡面就有分到其他組例如說衛福部、法務部、教育部內政部、一些社經資料你這些資料要拿出來一個資料庫的平台你要平常就丟給我了，丟給我那我們後端例如說像統資處看是不是有可以做統計分析找出說目前我現在裡面的數據跟什麼有關的跟外面的一些有關的那去做連結，那這樣子我就可以坐在這裡去看到外面的數據在變化的時候，這個數據在變化跟我這個什麼有關，我是覺得如果真的要連結的話兩邊都連結，所以我們可能選擇性的就是說以我們現有的官方資料為主然後再去選擇性的抓取現在在外面私部門的一些現象的東西進來就好了，先選幾個去 run run 看例如說酒、菸、刀、棍棒、就醫紀錄、爭吵次數等等，先 run run 看如果可以再繼續再繞進來，然後要善用電腦數據跟統計。(S5)

H2:

就是說大數據是沒有個別的問題因為他數據太大了大到幾乎掌握方向，但是一旦他有問題的時候其實這時候就必須要各單位的配合，我覺得今天有很多是長官有關政策有一些法令的部分我覺得現在最困難個資，我們今天連一些醫療的 data 根本沒辦法互通，那你沒辦法互通包括家暴這議題，我覺得政府是一體的假如今天我為了整體社會的安全我可不可以說他犯法他所有訊息可以因為犯罪偵防要有一致地那當然還是在保護他個資的部分，不洩漏出去有沒有可能用這樣那這時候已經有點像進入我們醫療體系了，今天你生病了那我今天必須要治療，當他全部的 data 不可以說醫院什麼東西都不給我這樣子不行，那我覺得那是不是有這樣的公權力就是說今天我們以前在開家暴的會我們每次幾乎都要不到啊，尤其是越例如說法務的那個警政他們有的東西資訊不給那社工就很可憐，他報的東西因為不完全那有時候要提追蹤的數據就沒人那其實是很可惜的。我覺得反而是最單

純的但現在最高危險的不是精神病是那個人格壞胚子那些是最困難的，那你只要調他的資料就被告，你為什麼拿我資料你憑什麼反正他也敢噏警察也敢噏法官誰都敢噏，這種反而他最不怕，所以假如說事情上有優先順序我可以先做那一個部分那當然我覺得大數據不可能短期內我們馬上就可以，所以這真的是但越來越聚焦的時候可能就很容易做。那大不了我們做一頁只是看到因為受暴力後會有很多的創傷，那怎麼樣減少創傷，親密關係暴力的高危險個案，只是說我們在他在進入系統前我覺得那已經被標籤把他貼上去，其實我覺得從他周遭的很多數據出來他就可以抓一些東西。(H2)

W6:

大數據他就是要滿足那個大然後又要大到最好不要個人化，然後這個大本身又要帶得出一種預測，那所以我覺得大數據要用我們要用在哪一個層次，那當我們數據再更大的時候，大概就是去佐證我們原來的假設這個 model 對不對再做修正，那我會更好奇的是說我能不能再看到不同的東西，所以如果說我們的作法上能夠在跳脫原來這些模式的話那我會更期待，如果我們初步是用所謂的公部門許多的資料庫去解那麼這樣的資料庫我就會更希望他大到一個不只是我們原來可能是所謂的家暴精神疾患毒藥癮那甚至我覺得是整個所謂的醫療的就醫的紀錄。(W6)

J5:

針對大數據分析需要具備條件就是說建議希望我們是說未來這個大數據的這個部分已經涵蓋三大面向，大數據的內容應該要涵蓋心理生理還有社會的三大面向但入監的收容人他的精神狀況還有成癮狀況人格違常的狀況是比較多，或是這些問題都有的一個多重性的這個案主，這些問題嚴重的個案分析是否適宜遷移到一般社會人士上來來來來評估然後對於不同機關資料庫的連結，必須要思考各個機關資料的真實性不同機關如何連結如何保密開放的程度如何這些資料有沒有開放性的必要這些都是我們將來在做這個研究必須要考慮的。(J5)

五、其他看法或建議。

W4:

我大概歸納成三個部分，我覺得如果在初級預防的部分，大數據他大概做一些趨勢的預測，那所謂的趨勢就是從一些跟這個親密關係暴力有關的一些議題，既有的資料裡面就可以掌握一些資訊，這個部份就是在二級的層面，我覺得也比較務實，他不見得是大數據，但是他是比較類似高風險辨識或者是危險辨識的一些指標，那像例如說那個入學這些學生，他有沒有透過一些特質的檢測，他可以去辨識出來這樣子一個特定對象，那可以多介入去輔導，那第三級的部分照理講已經通報進來就是我們保護的對象，他也不需要再透過大數據在強化我們對這個個案的關心。(W4)

10. 期中報告審查會議紀錄

法務部司法官學院

「大數據運用在親密暴力犯罪防治分析之探討」委託研究案

期中報告審查會議紀錄

時間：104 年 6 月 24 日(星期三)下午 2 時 30 分

地點：司法官學院一樓第一會議室

主持人：周主任秘書穎宏

出席人員：

審查委員：楊委員雲驊、羅委員倩薇、謝委員文彥、委員王雪芳

受委託單位：台灣警察學術研究學會范國勇、韋愛梅、王伯頌

委託單位：司法官學院犯罪防治研究中心中心主任吳永達、組員楊雅芳

紀錄：廖偉廷

會議記錄：

一、主持人：(略)。

二、受委託單位報告：詳見期中報告書面資料及簡報。

三、審查意見(按發言順序)：

(一)王委員雪芳：

- 1、政府資料分散在不同的部門，現行政府單位有哪些屬於家暴或親密關係暴力相關的統計資料，需先進行盤點，才能找出政策缺口；且目前期中報告書所引述主計總處的數據，係屬次級資料，建議調整回歸使用衛福部保護服務司家暴及性侵害相關之原始統計資料，亦可參考引述法務部統計處的檢察統計中家暴、性侵相關案件數據，作為大數據分析的前端基礎資料。
- 2、大數據運用爬文技術要非常謹慎，避免因缺乏精確性而造成錯誤解讀問題。

(二)謝委員文彥：

- 1、研究目的第三點語意不清楚，可參考 p. 36 酌做修改；p12-14 圖表部分年代不統一，建議從衛福部找第一手資料並統一年代，參考文獻資料格式須統一。
- 2、部分引註不清楚或是標題有不明確的部分，和內文關聯性不足，請再逐一檢視調整。
- 3、第二階段焦點座談應列出 3 次座談的時間、地點、主題、參與人及日期，深度訪談與焦點團體的研究工具應分別列出。

- 4、訪談資料分析應以研究者論述為主體進行論述，再以訪談內容作佐證，做一層一層精緻化、概念化，呈現出具體、系統性之分析。
- 5、親密關係暴力相關資料分散在政府不同單位，到底擁有資料是哪些單位？已經各自建置什麼資料？在親密相關的不夠的還有什麼？都需要再梳理澄清！
- 6、在進行焦點座談前，宜先瞭解哪些是解決問題的寶貴資料，可是單位都沒有，再透過焦點座談的討論，決定是否可加進來？未來建立大數據資料不但要弄清楚如何建構豐富的資料，而且要指出主責機關，才能做資料整合，另外，如何在不違反個資法下，建置更細緻的資料，也可以詢問專家學者之意見。

(三)羅委員倩薇：

- 1、有關第一階段深度訪談6位專家學者部分，建議將訪談對象(領域、名字)、時間及訪談紀錄呈現出來，並建議後續的焦點座談也比照相同模式。配合不同領域做不同的題綱設計。
- 2、文獻探討及現況分析、大數據概念發展的內容直接摘錄電子新聞網頁內容，應請團隊進一步驗證內容正確性，及是否有延伸引用的問題。
- 3、不管是文獻的探討或不同意見的陳述，建議以對照表的方式整理呈現，讓研究資料更具可讀性。

(四)楊委員雲驊：

- 1、對未來研究成果，期許能真的落實到政策面，所以，在法律規範層面的思考，希望能做更完整的檢視。例如個資法第6條對個資的例外適用規定，能否作為大數據資料的突破點，即為適例。
- 2、本項研究案，希望把親密關係暴力的預防策略，做為研究重點，舉台大學生張彥文殺女友案為例，大數據分析若能廣泛運用，或許社會就會減少一次悲劇發生。

四、業務單位說明(吳主任永達)：

- (一)研究的目的是在找出解決社會問題的最大可能性，由發現問題到瞭解現有工具的不足，再進一步找到對應的解決之道，在研究報告上都要力求具體明確。
- (二)全書的基礎架構與經費的使用情形，以及爾後的研究需要委託單位做哪些行政協助，亦屬本研究期中報告的程序規範，請於修訂期中報告時，一併補強。

五、主持人結論：

謝謝研究團隊的努力，為使研究成果更臻完善，請研究團隊參酌今天的審查意見，修訂報告內容，並依期限送請複審。

六、散會(下午 4 時 10 分)。

法務部司法官學院

「大數據運用在親密暴力犯罪防治分析之探討」委託研究案

11. 期末報告審查會議紀錄

時間：104 年 10 月 05 日(星期一)下午 2 時 30 分

地點：司法官學院一樓第一會議室

主持人(暨審查委員)：周主任秘書穎宏

出席人員：

審查委員：王委員雪芳、楊委員雲驊、謝委員文彥、羅委員倩薇(按筆劃序)

受委託單位：台灣警察學術研究學會韋愛梅、王伯頌

委託單位：司法官學院犯罪防治研究中心吳中心主任永達、楊組員雅芳

紀錄：廖偉廷

會議記錄：

一、主持人：(略)。

二、受委託單位報告：詳見期末報告書面資料及簡報。

三、審查意見(按發言順序)：

(一) 謝委員文彥：

1. 本案係屬時代性趨勢議題，頗具研究意義。
2. 研究團隊能在短時間內完成不容易的研究議題，並得到豐碩成果，令人感到敬佩。
3. 摘要部分可再精簡。建議第一段大概兩行就可以解釋完，第二段採用研究方法，頂多用兩行到三行敘述即可。研究的結果我建議你們仿造研究結論來寫就可以。建議摘要文字(含關鍵字)控制在一頁到一頁半之間完成，讓人家看了，知道你的研究方法與結論，一目了然。
4. 第一章用兩面篇幅薄弱，第一節可運用第二章的統計資料來補充；第二節何謂親密關係暴力，我建議還是要做個簡單的操作性定義。
5. 將第 17、18 頁及第 31 頁表的統計資料整理出共同的統計項目，譬如性別、年齡等等，共同項目呈現出來後，再以文字描述共同項目及分散項目。
6. 第 57 頁研究架構需要文字說明，第 58 頁焦點團體座談人數應行更正。第 61 頁研究工具五種，請分門別類列出來，不只是目前所列兩類而已。
7. (1)有關 data 命名問題，譬如第 81 頁(一)若改成對於資料流通跟個人資料隱私的價值衝突已減低，一看標題就很清楚。(二)不同來源的資料庫似乎都可以整合及可行性高，這樣的標題命名會較清楚。第 82 頁標題三改成彙整網路上較敏感的語彙是有待研究的，另外(一)改成連結如何克服個資問題是有待商榷的…等等，在資料的連結與解讀上都會比較清楚。

(2)過去資料是否要登入也是一項問題，對沒有明顯犯罪動機又無犯罪前科

者，事實上要建置他們資料是比較困難的。如果你給我這幾個標題出來這樣我就一清二楚。

- (5)第 85 頁大數據應用在親密關係犯罪防治的步驟建議有二種方法，其一將步驟一、二、三 step 弄出來，每一個步驟很清楚了，而哪一些非步驟的也列出來，他應該是相關注意的事項或注意的原則，如果這樣區分，會比較容易聚焦！我甚至建議步驟，可以把它畫成一個圖來顯示，這是一個方式。第二個是結合焦點團體與專家學者的建議步驟，更能顯示出來一起探討。這一部分我們研究團隊再考量如果可以重新思構，我相信呈現出來的貢獻，會讓政府機關更知道如何來處理問題。
8. 結論部分應將專家與焦點團體意見整合，例如：主政單位統合，有需要克服的問題，如何運用步驟，找出權責單位，資料問題範圍的界定，如兼顧資料安全問題，資料如何分析分類處理…等等，結論越具體，政府越清楚該怎麼做。
9. 第 69 頁這邊提到防治親密關係暴力遭遇的困境應回應研究目的二。第一個困境應該要回答第 67 頁，官方資料對親密關係暴力特性所知有限，然後第 68 頁包括為什麼遽增，對暴力的這些特性以及遽增的原因所知有限。第二個就可以回到 70 頁，例如：對年輕老年兩極端了解不夠，年輕或…有一些同居暴力的都出現了，再來外勞問題，也是我們不知道的第三個困境。第三個困境是第 70 頁最底下的如何分類，犯罪類型及加害者的分類還沒有辦法確定。第四個困境是第 69 頁(二)，他有沒有求助、司法對他們的反應是什麼、是否有城鄉差距。第五個困境是如何預防，回到第一個對於一二級預防工作的不足。
10. 第 71 頁(五)標題不清楚，我會建議對親密關係暴力的處理措施是否有效、及保護令的核發是否有效，如果你把這邊變成這 6 點你可大膽的放在 138 頁結論的第二大點。

(二) 楊委員雲驊：

1. 肯定研究團隊，題目新穎，這樣的研究方法與問題解決步驟，將來也可運用在毒品、貪瀆及金融犯罪等其他犯罪問題的研究上。
2. 有些部分較薄弱，例如：研究目的第二點，範圍太廣，每個部分放大開來，都可以成為一個研究案，雖然不可能做到極致，但可以加入不同內容補充(p37-39 稍嫌薄弱)，四個體系所面臨的問題都不一樣，現有的研究成果或參考資料可放入。
3. 恐怖情人法律上沒有定義，各機關名詞、分類有困難，現有的資料系統認定尚有疑義，困境及困難可用實際案例說明。
4. 在國外有沒有做過，具體成效？親密關係國外的例子，也可以引述作為本研究的佐證資料。
5. 資料多但有沒有用，有些資料庫資料是找不到的，大數據分析可以很多資料分析看到不同面向，但結果呈現未必能夠真的呈現，可能現存資料沒辦法

應用判讀，讓資料確實運用及判讀，對犯罪防治上發揮也有一定程度的侷限性。

(三) 王委員雪芳：

1. 第 9 頁研究目的第四點，在第 18 至 24 頁有基本回應，各機關資料應進行盤點，整理出具有指標型的資料，有完備基礎的結構資料，連結到不同資料並進行資料剖析。
2. 個案資料庫是需要串連的，需找到共通的地方。每個資料定義不一樣，如何處理？法務統計的名詞解釋定義方式，可以參考！

(四) 羅委員倩薇：

1. 定義範圍較大，可以透過主管部會衛福部或行政院層級跨部會協調，運用所蒐集的現有數據資料做出模組，分析出來提供相關單位參考，什麼資料是不足的，例如：第 140 頁有甚麼資料庫？有沒有實際上可以建議，到底還缺少哪一塊資料？資料才可以串接。
2. 在建議部分應提出需要一個資料彙集平台，讓主管部會去蒐集各單位的資料。
3. 第 144 頁(二)6. 應改成散落，第 145 頁標題二中長期建議的(三)語意請再補強。

(五) 周主秘穎宏：

為使研究成果更具體，且聚焦在可執行的策略上，認為最末章宜就以下幾點思考，再做調整：

1. 應就目前存在問題與政府已建立資源暨將來解決策略三方面，依各部會逐一盤點，以免責任分散，變成大家都有責任，但誰做什麼，如何檢驗，還是不清楚！
2. 因為我們做的是大數據分析，所以，最主要的解決之道，可聚焦在各單位資料庫之建立上，亦即，各部會資料庫的強項是什麼？不足的地方是什麼（例如：是否有欄位缺損？或格式不相融問題？如何做各自以及彼此串連的補強才能提昇問題解決效益，以及可預期的效益是什麼？
3. 因為這次是先導性研究，革命不可能一次成功，所以，除政策建議，請各單位在政策的興革上，各自努力外，就學術研究的領域而言，下一波，可以做什麼，亦宜有所指引！
4. 綜上，第六章建議調整體例，參考內容如下
第六章 研究結論與建議
第一節 研究發現
壹、 親密關係暴力增加原因的觀察分析
貳、 各政府機關的努力（註：重在資源盤點，請以部會為分析單位）
一、 在資料庫建立方面
二、 在其他防治措施方面
參、 目前尚屬不足之處（註：重在問題盤點，請以部會為分析單位）

一、在資料庫建立方面

二、在其他防治措施方面

第二節 研究建議

壹、政策建議(註：重在問題解決之對應關係，可作短中長期之策略建議)

一、在資料庫建立方面

二、在其他防治措施方面

大數據運用對親密關係暴犯罪防治政策建議一覽表

權責單位	現有機制	不足問題	解決之道	預期效益
行政院				
行政院 主計處				
衛生福利部				
教育部				
內政部				
法務部				

貳、研究限制

參、後續研究建議

四、業務單位說明(吳主任永達)：

希望本研究可以發生幾點效益：

- (一) 資源的盤點：讓各公務機關及學界都能夠完整、有系統性的知道，要解決親密暴力，有那些單位做了些什麼，亦即有那些資源可以互相參考運用！
- (二) 問題的盤點：各公務機關的努力，還有那些問題，有待克服，才能提昇解決問題的效益。
- (三) 方法的盤點：針對不同的問題，可以用那些方法來解決，預期可以產生的效益是什麼。
- (四) 責任的集中：不同的問題的解決，其權責單位是誰，務期定位清楚，問題的解決才會有著力點。
- (五) 方向的指引：告訴不同的公務機關，短中長期的問題解決策略，讓他們可以根據本研究來具體規劃、精進行政作為。
- (六) 後續的研究：本研究係屬前導型研究，如果要接續努力，逐步解決問題，還可以做那些研究，原因是什麼。
- (七) 可以再把以上各點，分門別類，做成表格，以便閱讀及檢核。
- (八) 我的格式調整，只是建議，只要可以把上開需求最大化即可。

五、主持人結論：

謝謝研究團隊的努力，為使研究成果更臻完善，請研究團隊參酌今天的審查意見，修訂報告內容，並依期限送請複審。

六、散會(下午4時10分)。

12. 2015 年犯罪防治研究學術發表會--發現趨勢、創見未來

記錄：

一、時間:104 年 12 月 7 日(星期一)14:00-17:00

二、地點:法務部司法官學院大禮堂

三、開幕典禮、長官暨貴賓致詞:

(一)法務部司法官學院蔡院長清祥

法務部羅部長、內政部警政署陳署長、國立中正大學楊副校長、行政院性平處黃處長、今天論文發表及與談的各位教授、傑出論文獎得獎人，以及與會的來賓，媒體記者、女士、先生們，大家好！

非常榮幸，羅部長、陳署長以及在場的專家學者們，能夠接受司法官學院的邀請，參加這場盛會，司法官學院從 102 年 7 月 1 日改制，成立犯罪防治研究中心以來，無時無刻不在思考，如何運用精簡的人力與經費，協助法務部，擔任國家刑事政策智庫角色，並建立與學術界合作的橋樑，二年多以來，我們除了建置「犯罪防治研究資料庫」網路平台作為發展基地，廣邀學者撰寫刑事政策與犯罪研究專論，編著專書，藉以累積犯罪防治研究量能外，也依照國家犯罪防治研究的現實需要，擬訂研究主題，透過專案委託研究，向學界評選優良研究計畫，以具實證性的深度研究成果，精進政府犯罪防治刑事政策。同時，我們更透過辦理「諮詢會議」、「學術研討」、「學術發表」、「論文評審」、「交流研習」、「發行專刊」等多元多樣型式，促進國內外犯罪問題防治與研究的交流，以期提昇國家推動刑事政策的精準度與執行力。而這些工作，都必須有法務部的鼎力支持，以及社會各界的全力配合，因此，本人首先必須向羅部長，以及所有與會的先進，表示個人最誠摯的謝意。

今天學術發表會，一共有兩個主題，第一個主題是「103 年犯罪狀況及其分析」，為了彰顯這個研究對犯罪防治研究的特殊意義，我們今年特別加上「2014 犯罪趨勢關鍵報告」副標，這份研究係由國立臺北大學許春金教授擔任計畫主持人，中央警察大學陳玉書副教授與蔡田木教授擔任共同主持人，由於研究的型態不同以往，且跨足諸多政府部門資料的統整，辛苦程度，不可言喻；特別是在第一篇從去（102）年的研究開始，引進警政統計資料，不但將犯罪研究的縱深拓展至偵查前的調查階段，並包括與美國、英國及日本等先進國家主要犯罪率與監禁率相比較，都是本項研究的重大創舉，因此，今天也特別邀請內政部警政署陳署長蒞臨發表會，再一次敬致衷心感謝，因為有了貴署的共襄盛舉，才能讓國內整體犯罪狀況的分析，更為豐富、完整。

另一個發表主題是「大數據運用在親密關係暴力犯罪防治分析之探討」，這個研究是由臺灣警察學術研究學會理事長范國勇副教授擔任計畫主持人，警察專科學校韋愛梅助理教授以及銘傳大學王伯頌助理教授擔任共同主持人，研究的發想，在於去（103）年 9 月份發生了台大生張彥文街頭砍殺前女友 47 刀致死事件，原本的親密愛人變成恐怖殺手，其原因何在？有無具體的徵候可尋，從民間到政

府，從警政到醫療到社工再到司法，是否能系統性建立一套有效的防護工程，以減少類此不幸事件發生，並彌平社會傷害，尤其是近年來全球盛行著「大數據」的資訊整合觀念，站在大數據「深度應用、分析」的起跑點上，如何善用政府醫療、社工與司法資料，透過類似個案資料庫分析，運用網站資源，研擬預防與預後對策，以防止悲劇一再發生，也是本研究所欲了解之重要課題。行政院性別平等處黃碧霞處長是性別主流的政府最高主管，也是這方面的專家，相信，透過她的參與，一定能夠讓這份研究成為安定社會的重要力量。

另外，鑑於各大學院校是國家學術發展的搖籃，相關領域碩博士人才的培育，不但攸關未來國家刑事政策發展的興衰，更決定國家犯罪防治研究發展的整體量能，為採取適當的獎勵機制，鼓勵更多優秀人才，投入犯罪防治與刑事政策的研究領域，本學院於 103 年 4 月 18 日研訂「傑出碩博士犯罪防治研究論文獎勵要點」，今年是第二屆傑出論文獎評選，一共有 16 位碩博士生得獎，也利用今天的時間，敬請羅部長公開頒獎表揚，恭禧各位獲獎的研究生，並希望各位能夠繼續努力，在國家刑事政策的興革上，提供更多的研究貢獻。

最後，再次感謝與會所有長官、貴賓的蒞臨，祝福大家身體健康，萬事如意，謝謝大家！

(二)法務部羅部長瑩雪

內政部警政署陳署長、國立中正大學楊副校長、行政院性平處黃處長、今天論文發表及與談的各位教授、傑出論文獎得獎人，以及與會的來賓，媒體記者、女士、先生們，大家好！

非常榮幸，接受司法官學院蔡清祥院長的邀請，參加這場盛會，而且，能夠在發表會的開場，先跟大家講幾句話。

法務部早在民國 53 年即以任務編組方式成立「犯罪研究中心」，雖其間不斷有成立國家正式組織之呼籲與規劃，但礙於國家整體人力員額編制與財政現實，歷經幾十年，一直無法實現，直到 102 年 7 月 1 日才在司法官學院成立正式「犯罪防治研究中心」編制，不過，雖然學院犯罪防治研究中心已經法制化，但仍只能有 2 名常任文官的人力編制，可以想像運作起來，一定十分辛苦，但是看到司法官學院犯罪防治研究中心，從 102 年 10 月 1 日人員才到任運作，短短二年多的時間，就能完成多項創舉，展現豐富的工作成效，謹此特表嘉勉。

今天學術發表會的兩個研究主題，對我們國家社會的安定，都具有非常高的指標意義，第一個由國立臺北大學許春金教授擔任計畫主持人的研究主題——「103 年犯罪狀況及其分析——2014 犯罪趨勢關鍵報告」，我們在這份研究裡，可以看到很多政府刑事政策，所應努力的方向，例如：103 年矯正處遇機構超收問題仍然嚴重（超收 8,859 人，16.2%），檢視 103 年監獄新入監、在監受刑人刑期結構發現，新入監者在一年未滿（含拘役、易服勞役）合計占 80.76%，一年以上者只占 19.24%；反之，年底在監者在一年未滿（含拘役、易服勞役）合計僅占 16.35%，一年以上者共占 83.65%。矯正處遇機構收容之短期人犯高達八成，至年底實際在監者，僅不到二成，顯示其進出頻繁，不但無益

於教化，亦徒增矯正機關之獄政負擔。另外新入監之罪名則以公共危險罪人數最多(其中不能安全駕駛即酒醉駕車更高達 9,631 人)，就刑罰功能，短期自由刑對一般民眾具有威嚇作用；但對人犯而言，進入監所服刑，不僅破壞其社會關係之正常運作，造成日後社會復歸之困難，以及家庭關係之崩解，甚至會產生矯正副文化的負面效應，如何衡平其間利弊，找出對國家社會最有利的對策，正是國內刑事政策發展之重要課題。

又如：對於北捷的隨機殺人案，提出隨機殺人的兇手沒有明確的動機，被害人不特定，與過去殺人案件多半有明顯的恩怨不同。因犯罪人外觀多無特殊跡象，難以立即辨識，因此，警方除應積極開發可以迅速發現網上犯罪預告的技術，並在發現類似模仿言論後立即逮捕處理，以防範未然，並遏制模仿效應的蔓延外，臺灣大眾運輸系統的發展快速，除捷運外，臺鐵與高鐵均隨時有大量的交通人潮往返，一旦發生類此事件，傷亡情形，難以估算。因此，有關大眾運輸體系的安全維護措施，也應全面檢視，並建立可以立即回應暴力事件發生之 SOP 標準處置流程與督導運作機制，以確保社會安全的建議。

對於夜店的襲警案，從第三造警政的觀點，認為維護社會治安的核心力量，除應聚焦於「加害者」與「警察」二造外，更不可忽略「第三造」之社區力量。夜店常是犯罪吸引者，也是犯罪熱點，政府應要求經營者盡力防治店內及周邊的各種犯罪可能。夜店要開店、要牌照、要賺錢，就須提出有效的預防犯罪保全計畫，將保全設備、保全人員、安全訓練、金屬探測器、安檢門、通訊連絡器材、監視器設置、電擊武器、制定保全人員處理顧客糾紛或酒醉的處理標準程序和通報責任等等，均應列入經營成本內，都是非常中肯的政策建議，值得政府參考與省思。

另一個由臺灣警察學術研究學會理事長范國勇副教授擔任計畫主持人的發表主題——「大數據運用在親密關係暴力犯罪防治分析之探討」，在「大數據」時代來臨後，即刻研究蒐集彙整以及全面性檢視與親密關係暴力有關的政府機關各個資料庫，並指出政府機構對於「親密關係暴力」的定義與範圍要有統一的界定，以及須先形成防治網絡的共識，資料庫個案資料建立內容應隨法律及實務需要更新，且應有一致性及延續性等等建議，對政府建立系統性，且有效益的婦幼防護工程，以減少社會悲劇發生，亦有卓越的貢獻，謹此對於參與研究以及今天參加與談的所有學者專家，表達個人最深的敬意。

另外，學術研究發展的傳承，也非常重要，司法官學院非常有遠見設立「傑出碩博士論文獎」，有效吸引國內各大學校院的碩博士生進行刑事政策研究，聽說這次研究的主題，涵蓋人權保障、廉政治理、住宅安全、毒品防治、少年犯罪、性侵犯罪、經濟犯罪、跨國犯罪與鑑識科學等等面向，非常精彩、多元，所以競爭非常激烈，恭禧在場所有勝出的碩博士生，也希望大家繼續努力，讓我們的社會，不斷朝向健康、和諧的方向發展。

最後祝福這次活動圓滿、成功，大家身體健康，萬事如意，謝謝大家！

(三)內政部警政署陳署長國恩

法務部羅部長、司法官學院蔡院長、周主任秘書、吳主任、中正大學楊副校長、行政院性平處黃處長，以及在座的貴賓、專家學者們，午安！

各位都是我們在行政、司法並肩作戰的夥伴，感謝對警察工作的支持。今天非常榮幸 追隨部長及各位先進，參加司法官學院舉辦的「2015 年犯罪防治研究學術發表會」，跟各位專家齊聚一堂學術研討，這些將來都可以做為精進警察工作的重要參考。

談到台灣治安，我們可從幾個數據來看，103 年的刑案發生數是 20 年來的次低，其中影響民眾觀感最深的暴力犯罪、竊盜案件，還有持槍的刑案，更是 20 年來的新低。2015 年遠見雜誌有關縣市長施政滿意度調查中，八個面向中警政治安滿意度高達 71.8%；警政署歷年委外實施的每一季民調，縣市滿意度是 70%、社區鄉鎮滿意度是 80%，與中正大學公布的 104 年上半年治安施政滿意度調查 83%，這個驗證是一致的。

另外，看看外國機構對我們治安評價的幾個數據，第一個是 103 年 7 月美國網站 Lifestyle 公布的治安，我們是僅次於日本、世界排名第 2 的國家；第二個是今年 1 月經濟學者公布的安全城市指數排名，台北市在人身安全部分世界總排名是第 5；第三個是今年 2 月份美國 Lifestyle 網站調查 2014-2015 網路票選全球十大安全國家中，我們排名世界第 2，是僅次於冰島的國家；第四個是今年 10 月份一個研究機構公布的傳統數據裡面我國整體排名第 21，但在安全和治安方面則是排名第 2，也是名列前茅；所以看起來我們的治安排名，至少有二個世界第 2、一個第 5，都是名列前茅的。在外國人眼裡台灣是一個相當安全的國家，這些都是在座大家一起努力的結果，可說在全世界裡面，臺灣治安是給人家一個很信賴、安全、可靠、夜間可以外出的一個國家。

我們雖然是在一個治安穩定的狀況，但還是有努力的空間 我在新北服務的時候，建構了一個 4 年 13 億經費的新北科技防衛城市計畫，做了很多關於交通安全、資訊安全、交通執法、110 指揮管制、天眼監控、刑事鑑識等業務的提昇，這些計畫串連建立起一個科技防衛城。所以個人到警政署後，會朝著科技前進的方向來努力，希望建構全方位的科技防衛，並參照學術界研究的治安對策，讓台灣的治安環境更好。

未來，希望在法務部、內政部以及各位學者專家實用、寶貴的建言引領下，讓我們警察機關跟各級政府部門，可以並肩作戰，把治安做得更好；因此，包括在座的碩博士生的學術論文，都是我們施政的重要參據，同時，也衷心期盼各位專家學者日後都能不斷地提供治安的諮詢與後盾，讓我們將來的治安工作，朝著理論和實務結合的好的方向前進，掌握治安趨勢、建構一個良好的治安環境，最後預祝學術研討會成功，各位身體健康、事事如意，謝謝各位。

四、學術發表會

第一場次發表：103 年犯罪狀況及其分析—2014 犯罪趨勢關鍵報告(略)

主持人：中正大學楊副校長士隆

發表人：許春金教授、蔡田木教授(發表內容如簡報)

第二場次發表:大數據運用在親密關係暴力犯罪防治分析之探討

主持人:黃碧霞處長

發表人: 韋愛梅助理教授、王伯頌助理教授(發表內容如簡報)

與談人:

(一)台灣防暴聯盟廖秘書長書雯

台灣在性別暴力防治的領域當中，不僅是一項學術探討的問題，更是一項需要真實去實踐之問題。對於一個民間團體，我們亦非常在乎該項問題。前二年回台灣時，我想知道究竟我們有多少人死亡導因於家庭暴力，卻發現我們竟然沒有官方正確的統計。因此當時我透過刑事資料當中的殺人罪、傷害罪、重傷罪等等當中，找出其中有親戚的關係作為有多少人死亡的判斷依據。人命的救治是我們最重要的目的。能多救一條人命，即是我們這項工作之重點價值。

非常感謝這項研究報告，使我們得以觀察且密切注意。此外個人非常喜歡今天的題目「發現趨勢，創建未來」，如同研究團隊所言，確實當整個研究完之後，發現我們不瞭解的遠超過於我們所瞭解的，如同有位法官曾經說過，我們不知道的遠比我們知道的來得多。

今年(2015)9月，聯合國做了一項未來下一個15年我們人類永續發展目標的擬定，擬定了從明年開始2016年至2030年17個目標，其中第5個目標為性別平等與賦權女童跟婦女，特別之處在於它中間的具體執行目標，是來自消除性別暴力。換言之，性別暴力的防治在未來當中，將成為人類最關注的問題之一，同時也在未來15年當中，我們該如何面對以解決此項暴力問題，並將其納入具體規範。

資料的本身無法改變生活，但如果要改變我們的生活，一定要有賴於資料，必須透過資料的分析與整理，我們才有辦法做出政策的擬定，在具體推動時才能有心得收穫。對於本領域的發展，中國大陸雖遠落後於我們，但他們現在卻預先洞悉趨勢，在今年已經成立家庭暴力防治研究所，這是當初防暴聯盟想在台灣執行的理念，而我們卻無法自己達成。

研究報告指出了資料庫當中哪一些事實是我們現在急於必須做分析、做歸類與探討，或許接著在大數據的運用當中，我們得以找到一些更具體的方式，不但可以解決我們自己的問題，亦能反饋對這個世界有所貢獻。在亞洲根據一項針對一萬名男性之實證研究，研究結果發現一項很重要的部分，所有的暴力當中最關鍵的原因為性別不平等。從張彥文的案件中我們得以發現性別不平等的歧視其實是存在的。犯罪行為的發生可能早在其青少年階段即有跡可循，因此在預防工作層面上，應更往前一步來進行犯罪預防之工作。

(二)中央警察大學資訊管理系王教授朝煌

1. 大數據的概念

(1)在大數據時代:「凡發生過的，必留下電子軌跡。」

(2)大數據的範圍包含:資料處理系統與決策支援系統資料、資訊科技廣為應用後的各種電子交易軌跡資料、各種感測設備收集的資料以及網友所生

產的資料。

- (3)蒐集分析大數據的主要目的為有系統地蒐集、整合及探勘軌跡資料，發現隱含於大數據的知識，進而增加洞察、掌握先機及提升競爭力，提供適時、適地、適人、及適物的服務

2. 兩個例子的啟示

- (1)一名美國海軍軍官整理航海書、地圖、圖表及航海日誌，繪出一幅全新型式的海圖，加以分析後發現一些經驗傳承的航海模式是非常沒有效率的，為了提高精確度，莫瑞也重新設計航行紀錄表格，並要求所有美國海軍艦艇確實記錄。為了蒐集更多的海流資料，莫瑞要求船長定期將帶有日期、位置、風、及海流資料的瓶子丟往海裡，並且要求船長們如果在海上看到這些瓶子，也務必蒐集起來交給莫瑞分析，藉以找出天然的海上通道。

上述的例子便可得知整合眾人的經驗智慧，可以發現之前未被發現且較有效率的方法，把各種不同的資料整合在一起，可以提高視野增加洞察，蒐集資料的表格如不符實際需要，必須加以修改並確實遵循及在必要時必須建立機制蒐集資料。

- (2)美國明尼蘇達州一家零售商塔吉特的行銷部門寄了嬰兒服及嬰兒床的折價券給當地的一名高中女生，該名學生的父親氣沖沖地跑到門市找經理理論，但該名父親在與女兒深談後，打電話向經理道歉，並且證實預產期在當年八月，為何零售商塔吉特比女兒父親更早知道該名學生的懷孕情形呢？主要是因為零售商為提升競爭優勢，早已進行大數據的蒐集與分析，做好顧客關係管理，零售商的資料分析分為兩個階段，一為準備階段：蒐集顧客相關資料(如辦卡資料)及購物紀錄、依顧客相關資料及購物紀錄將顧客分群及顧客群的購物特徵與購物(順序)型態分析探勘，另一為運用階段：將新顧客依其資料及已知的購物紀錄進行分類及依其所屬顧客群的購物型態進行推薦或行銷。

3. 資料分析與探勘

- (1)橫向：探勘顧客(群)共通的行為模式，例如啤酒與尿布常被一起購買(支持度與信度皆達到一定的水準)，主要技術為關聯規則探勘。
- (2)縱向：個別顧客(或顧客子群)的購物型態分析與預測，例如購買受手機後，會在購買防刮保護貼、及保護套，以及後續的網購方案等等，提高行銷精準度，主要分析技術：分群或微分群、顧客子群特徵化、(時間)順序型態探勘、分類與預測。

4. 大數據在家暴的應用探討

- (1)準備階段：建立監控及資料蒐集系統、社政通報系統、健保就醫資料及報案資料…等，對於已知被害者分析(冰山浮出水面的部分)：蒐集整理已知家暴被害者資料及家暴事件經歷之電子軌跡資料、家暴被害者分群或微分群及家暴被害者子群的特徵或順序型態探勘。

(2)運用階段(發掘冰山沒入水面的部分):建立個案家暴被害風險評估篩選系統、候選個案分類與預測、適時啟動適當的預防策略與作為以及將確定個案資料回饋加入已知被害者資料庫。

- 5.應整合社政、醫療、司法、警政等等相關部門資料，建立家暴資料倉儲，運用大數據分析，便可增加洞察與掌握先機、及時啟動家暴防範措施，達到預防家暴之效果。

五、綜合座談提問與回應(第二場次部分)

(一)法務部周專員提問：

關於資料庫的整合運用，第二個研究案分析得很好，它有提到各個資料庫的資料分析，既然有這麼多資料，為什麼各部會不成立的一個連結性的機制，就是說戶役政、檢察、警察、法院各個封閉資料庫，如能將相關的資料整合、去蕪存菁，那各個部門同樣查詢一筆資料就可以將很多資料顯示出來，不用每筆資料登出、登入、不用再打電話查詢，造成很多文書資料的不便、效率的降低，這是我的一些小小的建議，謝謝。

(二)研究團隊回應：

謝謝您對本研究案的肯定，將親密關係暴力有關的資料庫做跨部會的整合，是好的方式，也是本研究案的建議，希望未來政府部門朝此方向努力。

目前衛福部有幾個主要的資料庫，例如家暴資料庫跟高危機資料庫，其實它已有相當程度的整合，像家暴資料庫裡面有通報資料、檢警處理犯罪資料、法院保護令核發資料，但由於各部會對親密關係暴力的概念定義、資料檔案格式的不一致，未來在資料的整合上就會出現一些困難，另外由於是保護性個案的資料，在資料查詢權限也有其限制，有些資料並未開放，未來各資料庫整合為大資料庫，須在現行基礎下克服一些問題， 要做一些突破，才有可能再進一步談資料的深化與運用。

13. 研究案活動照片

焦點座談會及研究案審查會議照片



2015/7/8 第一場焦點座談會



2015/7/8 第一場焦點座談會



2015/8/4 第二場焦點座談會



2015/8/4 第二場焦點座談會



2015/8/27 第三場焦點座談會



2015/10/5 研究案期末報告審查會議



2015/10/5 研究案期末報告審查會議



2015/10/5 研究案期末報告審查會議

2015 年犯罪防治研究學術發表會照片(2015/12/7)



法務部羅部長致詞



與會貴賓合影留念



韋愛梅助理教授報告



王伯頤助理教授報告



與談人廖書雯秘書長



與談人王朝煌教授



綜合座談法務部周專員提問



司法官學院周主任秘書主持綜合座談