

# 網路搜尋活出美好真實的人生

## 「數據謊言與真相」讀後心得

### 一、前言-戴著面具的人生

從出生每個人都充滿好奇心，一直探索這世界的人事物，然而在成長及教育過程中，我們被要求或模塑成社會期待的樣子，因此或多或少我們都戴起了面具，這面具不但保護自己的隱私，也讓我們能融入社會及群體中。

因我們都帶著面具生活，所以在群體生活裡，並不會表現出真實的自我，別人也不知道我們最真實的需求與想法；久而久之，因長期隱藏真我，使我們過著虛偽的人生，最後連自己都迷失了，不知哪個面向才是內心想要的生活。

因此迷失是許多人常掛在嘴邊的鑰字（keyword），筆者在109/3/29上網查，出現約4.63億個搜尋結果；而英文lost更是超過60億個搜尋結果，相近詞失落也有約0.5億個搜尋結果；從這些數據可明白，這是個「失落的年代」（約1.47億個搜尋結果）或是「失落的自己」（約0.37億個搜尋結果）。

這種迷失的感覺如影隨形，因我們總是為社會或親友的期待而生活，所以我們不會表現出真實的自己，也不認識他人的真實。我們常以為自己應該跟別人一樣，甚至要學習成功勝利組的模樣；因

此我們崇拜偶像、明星、師長，想要高富帥、白富美，但當這些偶像的形象破滅時，我們就落入更迷失的困境中。

## 二、 本書核心觀念—透過網路搜尋數據追求真相

本書作者提出川普的勝選，就是虛偽人生的最佳寫照，因為川普挑起種族主義的大旗，犯了美國選舉的大忌，所以在選前的民調，沒人看好川普，但川普卻一路過關斬將，不但贏得共和黨初選，最後還當選總統，令許多人跌破眼鏡。

作者透過數據科學分析，從網路搜尋的統計發現，川普當選才是真實反應美國人內心的想法，把他們自以為沒有種族歧視的問題，當頭打了一個棒喝。

當時許多人以為美國已沒有種族歧視，更何況上一任總統是黑人，但作者研究網路數據後，發現人們在網路搜尋黑鬼頻率較高的地區，跟川普勝選的地區，竟呈現很大的重疊性；也就是說川普勝選的地區，出現種族歧視相關鑰字的搜尋，比其他議題如經濟、就業等還多，換言之，從網路搜尋數據的統計，就顯露出美國人最真實的想法了。

除了美國選舉議題外，作者還利用網路搜尋數據，挖掘出許多真相與事實，有很多結果跟我們以為的大相逕庭。雖然這網路搜尋數據，跟以前社會科學家或民調專家所用的數據型式及來源不同，

卻反而更能找出人們真實的想法。作者最核心的研究心得：人們對民意調查會說謊，但在網路搜尋時並不說謊，因為網路搜尋是因有需要才做的行為。

舉例來說：要怎樣才能找出真實的失業率，以前我們常用民調方式調查，但結果常偏離事實；作者研究發現，失業者在網路最多的搜尋，是跟消遣相關的詞彙，因此將這些搜尋詞彙統計成數據，就能準確預測失業率的比例。

另外許多人以為 NBA 球員，大多來自貧窮黑人家庭的小孩，但深入研究網路數據後，才發現出生在富裕的家庭更有機會成為 NBA 球員。還有我們習以為常的知識，在網路數據上也得到印證，例如：氣溫較低或陰雨的地方，比氣溫高的地方更容易令人憂鬱；因為在天氣好的地區，搜尋有關憂慮資訊的次數，比天氣糟的地方少了 40%。

### 三、 網路搜尋數據—新型數據有新啟發

本書點出網路數據跟傳統數據，有 4 個最大的不同點，那就是：

1. 提供新類型數據、2. 提供誠實的數據、3. 可以聚焦在小子集、4. 可以因果關係實驗。詳述如下：

#### (一) 提供新類型數據

在 Google 搜尋時，會發現 Google 常自動帶入相關搜尋的

鑰字，而從這些相關搜尋中，可延伸探討而找到彼此關聯的研究議題，這是網路大數據所提供的新類型數據。在以前以為沒相關性的資料，經過Google搜尋，可找出具啟發性的關聯數據；筆者也利用這相關搜尋應用在房價議題上，找到一些有意義的發現，這將在心得第伍部分詳述。

另外預測未來也是數據科學最常運用的領域，在運動界(如：棒球、籃球)或賽馬界等領域，已使用傳統統計方法來預測選手的表現，當作每一場比賽調度的依據；但除了傳統數據外，藉由相關搜尋技術，還能找到以前未注意到的關鍵因素。書中以賽馬為例，為找尋未來的賽馬明星，結果從相關搜尋發現，馬跑步的速度可以從馬的內臟大小來做推測，不但打破過去以血統為主的預測方法，也帶出新的醫學觀點，給賽馬領域者全新的啟發。

## (二) 提供誠實的數據

自從網路盛行後，人類在網路上所留下的文字紀錄已經遠遠超過書籍出版的數量了，加上 Google book Ngram viewer，把 1800 至 2010 年世界上大部分的書籍數位化了，因此藉由詞彙的搜尋，可看出這些詞彙彼消我長的情況。書中舉例：香腸跟披薩，從文字紀錄就清楚看出，香腸是很早出現的產品，而



了解人們的心思意念，才能真正滿足其所需。

### (三) 可以聚焦在小子集

由於網路數據是具有各種面向的資料，例如：年齡、性別、地理位置、活動內容、商業行為、健康情況等等，所以在網路數據裡可針對某個面向或子集，更直接地做相關性研究。作者藉由網路數據找出棒球迷會瘋狂支持某球隊的原因、也找到成人常有政治傾向的原因、以及發現報稅不實的原因。而特別的是，有些研究主題就如不實報稅的原因，是用民調也無法問出真正的原因。

另外作者提出了「分身搜尋」的技術，藉此可預測人類的行為、健康狀況、運動員表現、消費喜好等。這個技術在許多商業機構已有運用，例如：亞馬遜公司利用分身搜尋技術，對網路讀者提供推薦書單；醫院利用分身搜尋技術，能做出更準確的病情診斷。這些都是藉著網路搜尋數據，經過統計並預測類似條件者的行為模式或健康狀態。

### (四) 可以因果關係實驗

在生活中充滿了兩兩之間的相關性或因果關係，例如：醫學發現適量飲酒有益健康，但酗酒會造成身體傷害，這樣的研究最常使用的方式，就是利用隨機對照實驗。作者提出我們常

上的社群網站（如臉書）或購物平台，每天都在進行隨機對照實驗，雖然很多使用者是渾然不知，但卻已在網站上留下喜好的紀錄，而這些紀錄常被用來改善企業的廣告內容或其產品，或是找到對產品真正有興趣的人。

以前這種隨機對照實驗，需要花大量的時間金錢，但今天藉由網路，這些社群網站、購物平台，每天都可進行隨機對照實驗；而且網路世界是最隨機的狀況，因完全無法限制誰來網站拜訪，也不知這些使用者的特性，所以是最簡單方便的實驗場所。

許多美國明星歌手、政治人物的粉絲頁，也都有進行這種隨機對照實驗，以了解群眾對自己的喜好程度；而工商界更常利用網路隨機對照實驗，以了解消費者對其產品的喜好；據此同理，政府對於政策的形成，也可以利用網路來做隨機對照實驗，如此更能確定何種政策最容易被民眾接受。

#### 四、心得—沉迷網路抑或由網路數據得解脫

網路已是每天生活花最多時間的部分，例如：我們最常用的社群軟體 LINE，就是隨時保持上網的軟體，隨時等待訊息進來，或是即時分享訊息給親友；另外我們有任何問題或需求，就是上網問 Google 而求得解答；還有許多人戴著智慧手環，也是把身體狀況，

隨時記錄到網路上；換句話說，真實的自我每天都在網路上留下足跡、紀錄、問題及喜好。

再者網路進入 4G 時代後，手機智慧功能也越來越方便，甚至已經像是隨身的小助理，用講的也能通；一拿起手機進行語音搜尋，馬上就能找到答案，而這語音內容跟搜尋紀錄，也變成最真實的數據資料庫。例如：想了解這附近哪裡有好吃的美食，想知道某某名人的背景、年紀，想要去的地點走哪條路最省時間，想了解頭暈的原因等等，都會想利用網路搜尋，同時也在網路留下搜尋紀錄跟我們的語音，而這些紀錄都變成網路大數據了。

網路除了提供線上遊戲、線上交友，也蘊藏著龐大的資料，因此我們每天都花不少時間上網；加上我們的成長過程，常被教導要成為別人期待的樣子，因此我們常把真實的欲望藏在內心裡，外面表現是順著父母、師長、老闆的要求，走在一條社會期待的道路，但久而久之，在現實與期待的差距甚大時，許多人就花更多時間沉迷於網路了。

本書提到在 21 世紀初，人們在網路搜尋的紀錄，平均一天就積累 8 兆 gigabytes (10 億位元組) 的資料，是多麼龐大且難以形容的大數據。這些數據也成人工智慧的基礎，讓機器學習這門技術有長足進步；例如：機器能聽懂人類的話，是因成千上萬的人，在網



路上留下語音以及其所配對的文字；藉由這大量的資料庫加上網路傳輸快速，類似 Siri 或 OK Google 等個人助理軟體，就變成人工智慧機器人了；不但聽懂你所說的話，也能跟你對話，更能提供生活需求的建議，因此當數據科學越來越成熟時，網路大數據加上人工智慧就比你更了解你自己，讓你人生過得更美好。

## 五、學習印證得發現

筆者為學習並印證網路大數據的威力，嘗試挖掘人們真實的想法，109/4/4 在 Google 搜尋輸入「房價」鑰字，結果出現 1,360 萬搜尋結果，確認房價是大家關心程度很高的議題。當天搜尋結果出現房屋仲介的廣告、房屋交易網路平台、實價登錄相關網站、房價相關的新聞；而搜尋結果第 1 頁底下的相關搜尋，則出現 XX 房價 2019、XX 房價走勢 2019、XX 房價退燒、XX 房價崩盤等詞彙。

從上述搜尋結果，可看出關心房價的人，除了注意房子交易價格、也會到房屋仲介網站找尋出售的房子；而在 Google 相關搜尋詞彙部分，大部分是相關房價未來走勢，由此推測不少人也搜尋合適的買房時間，期待可以買到合理房價的房子。

筆者同時間利用 Google trend 查詢「房價」鑰字，在房地產領域中，可看到從 2016 年起的搜尋熱度，在 2017 年底有較高的搜尋熱度，而最近一年的熱度是稍微減少的，趨勢如圖 2 所示：

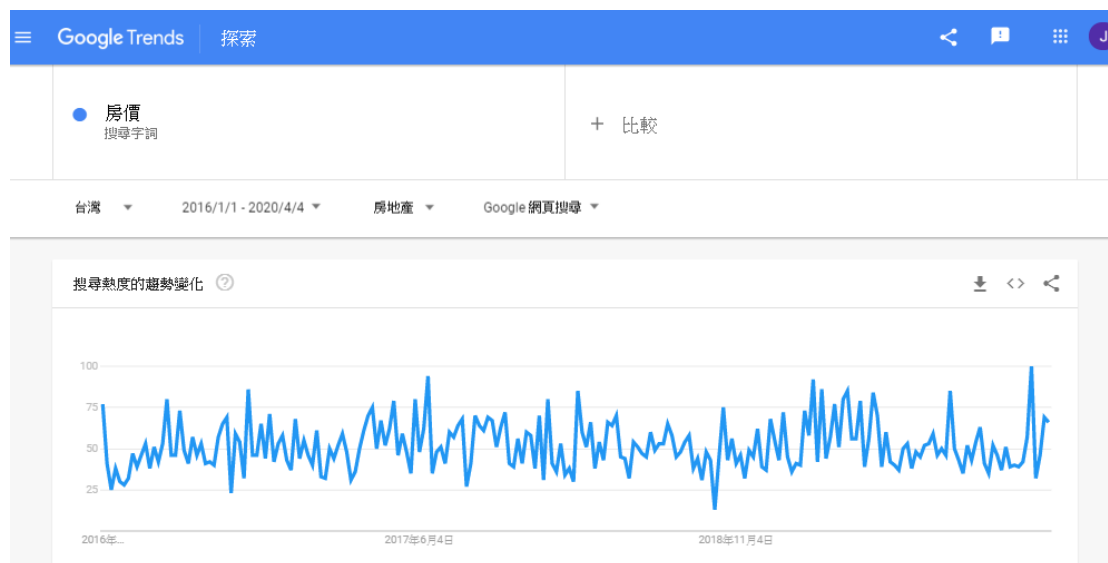


圖 2 Google trend 查詢「房價」搜尋熱度近 3 年圖

(資料來源：筆者 109/4/4 網路查詢結果之截圖)

從圖 3 最近一年的趨勢，2019/6/9 至 2019/8/17 搜尋熱度是 50-75 的稍高熱度，而 2019/12/8 至 2020/2/22 是 25-50 的低熱度時期。

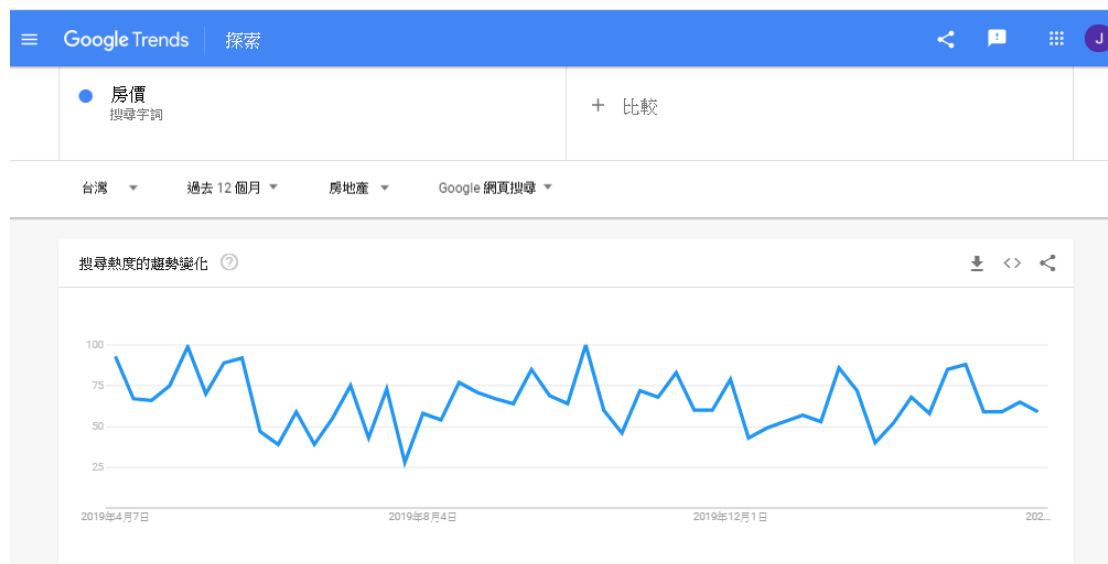


圖 3 Google trend 查詢「房價」搜尋熱度近 1 年圖

(資料來源：筆者 109/4/4 網路查詢結果之截圖)

從圖 4 Google Trend 也顯示過去一年搜尋熱度最高的城市是桃園市，並且從相關搜尋看到桃園青埔熱度增加很多，推測最近許多人對於桃園市的房地產，特別是青埔地區，有相當大的興趣；為了驗證青埔是否為大家關注的重點，在 Google 輸入青埔，就得到 2,230 萬的搜尋結果，印證是大家關注的焦點。



圖 4 Google trend 查詢「房價」區域網友感興趣程度圖

(資料來源：筆者 109/4/4 網路查詢結果之截圖)

另外筆者從某房仲網站找到他們所做的民調，從民調結果圖 5 看出新竹以南對房價樂觀程度高於北部，但比較 Google 搜尋熱度來看，台中以南的人對房價並不如北部人熱衷；雖然民調結果是樂觀，但筆者對此結果就會有所保留，不確定這些回答民調者，是真的關心房價還是隨便回答問題？

新竹以南樂觀程度高於北部  
雙北，桃園看跌轉趨持平

房價趨勢	台北市		新北市		桃園市		新竹縣市		台中市		台南市		高雄市	
	2019 Q4	2020 Q1	2019 Q4	2020 Q1	2019 Q4	2020 Q1	2019 Q4	2020 Q1	2019 Q4	2020 Q1	2019 Q4	2020 Q1	2019 Q4	2020 Q1
看漲	18%	19%	23%	23%	30%	29%	42%	43%	34%	31%	23%	30%	17%	25%
持平盤整	42%	48%	33%	40%	34%	42%	36%	38%	43%	43%	46%	47%	37%	36%
看跌	40%	33%	44%	37%	36%	29%	22%	19%	23%	26%	31%	23%	46%	39%

圖 5 某房仲網站對房價趨勢看法統計表

(資料來源：筆者 109/3/29 網路查詢結果之截圖，取自 <https://knowhow.yungching.com.tw/research/trend?nv1=researchR&nv2=trend>)

除了上述發現外，對於有興趣的議題，也可用 Google Trend 做關鍵字相關研究，在輸入兩兩關鍵字後，研究彼此熱搜程度的關係，也可觀察是否跟社會現象吻合？舉例來說，109/3/29 在 Google trend 查詢，看出關心房價的程度，跟關心建案、買房、土地、租屋等，是有不同程度的差異。

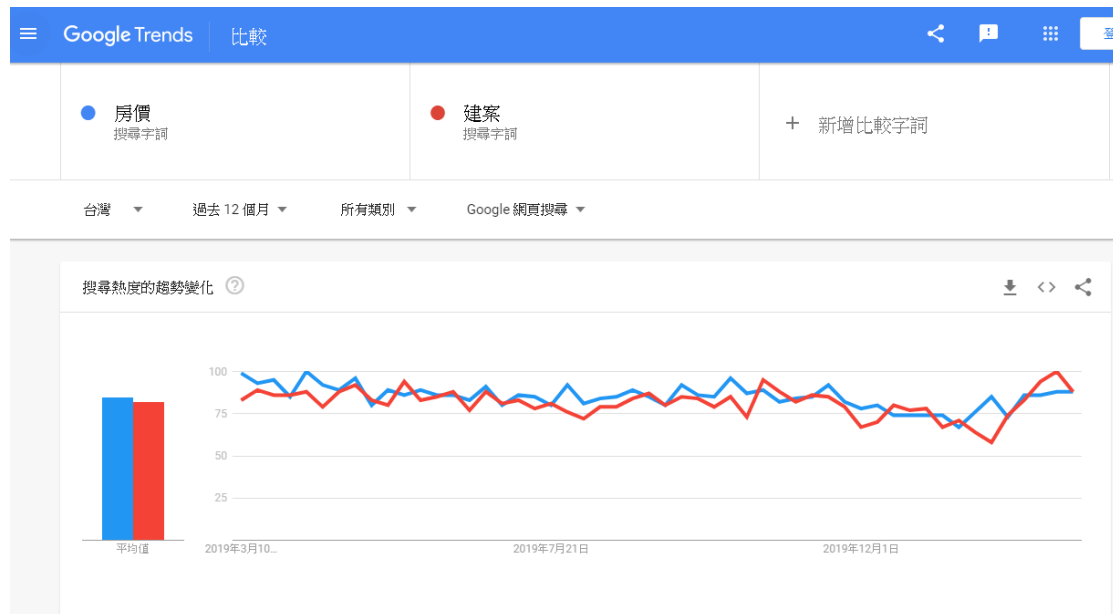


圖 6 Google trend 查詢「房價」、「建案」

(資料來源：筆者 109/3/29 網路查詢結果之截圖)

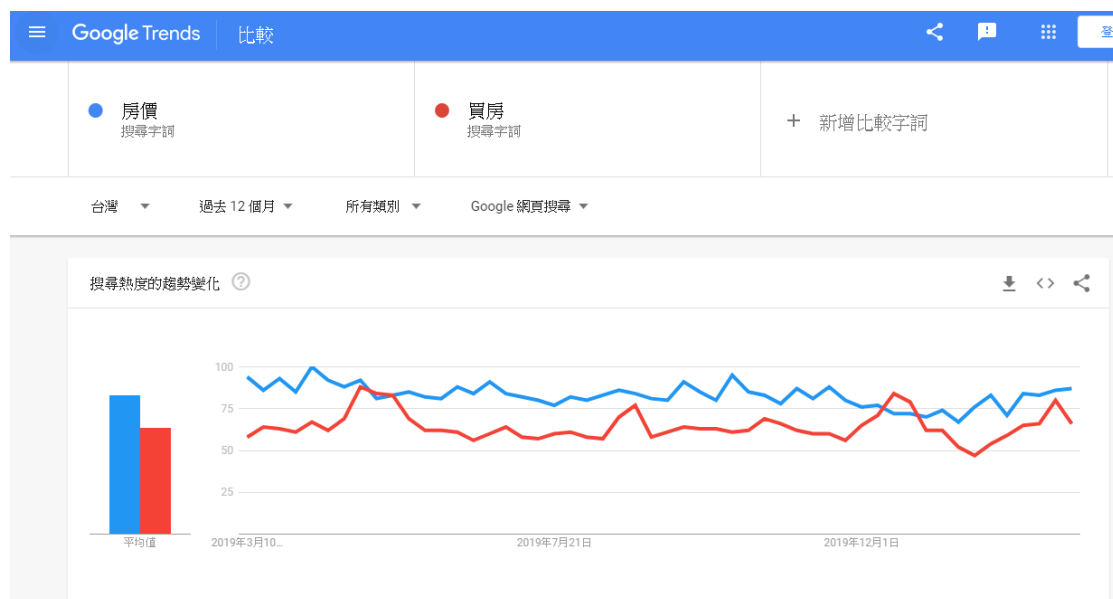


圖 7 Google trend 查詢「房價」、「買房」

(資料來源：筆者 109/3/29 網路查詢結果之截圖)

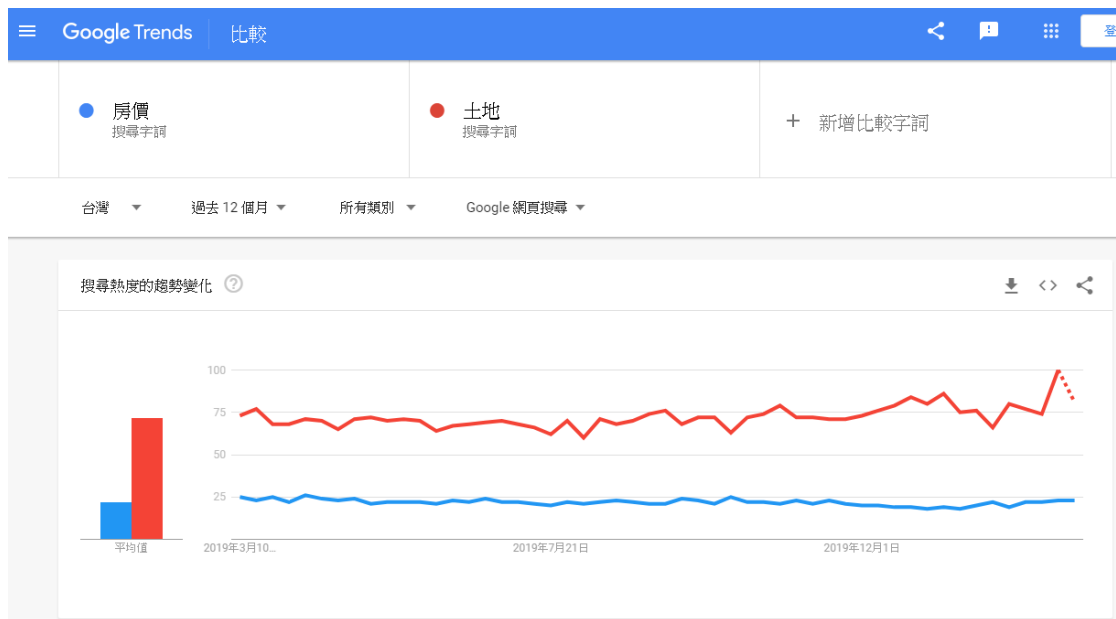


圖 8 Google trend 查詢「房價」、「土地」

(資料來源：筆者 109/3/29 網路查詢結果之截圖)



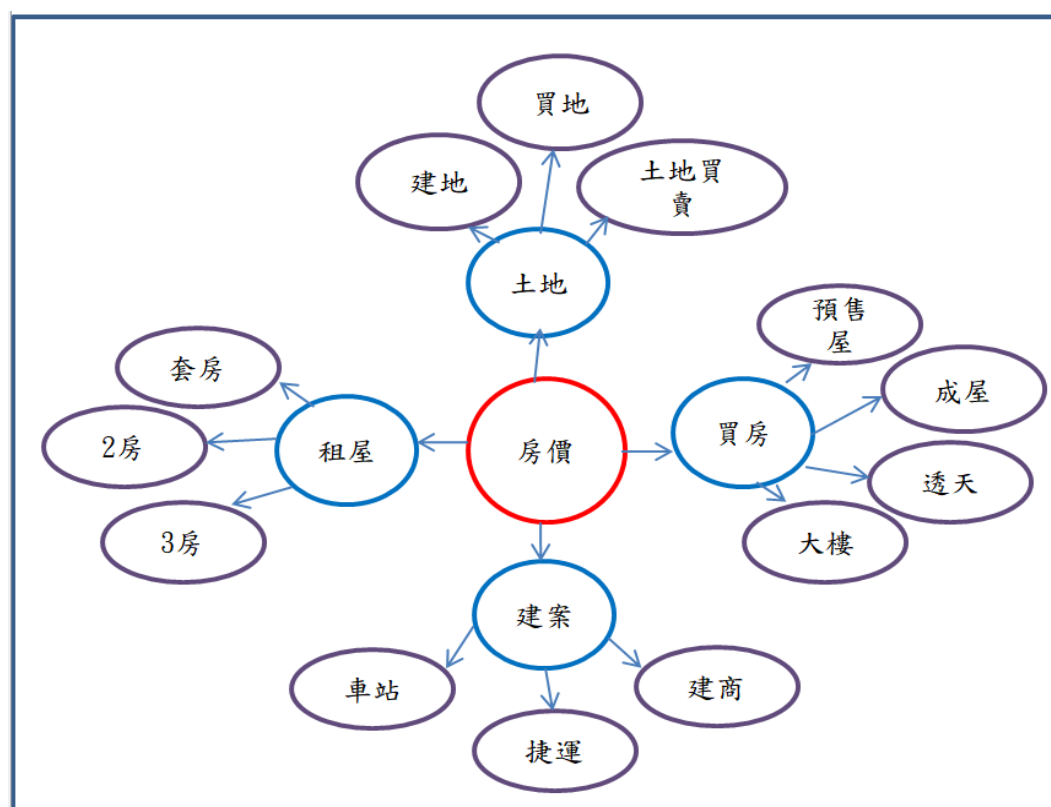
圖 9 Google trend 查詢「房價」、「租屋」

(資料來源：筆者 109/3/29 網路查詢結果之截圖)

從圖 6、7 可看出關心房價與關心建案的程度相差無幾，然後比

關心買房的人稍微多一些；而圖 8、9 顯示關心土地與關心租屋的人，比關心房價的人還多 4 至 5 倍，推測關心土地與關心租屋的人並不太考慮房價的變動。另外租屋的搜尋熱度還呈現週期性，暑假期間是高峰期，寒假是低谷期，推測學生租屋族群，是使用網路搜尋的最大族群。

從這些相關性的研究，筆者發現最重要的是問對問題，才能找出真正的關聯所在，這樣的觀念在「大哉問時代」一書中，也有相同的看法。此書提到大家習慣的事，才是最值得探究的，因為在習慣的事物裡，蘊藏許多我們所忽略的關聯性；也就說，這關聯性的探索，是需要一問再問地嘗試。所以相關房價研究議題，可以繼續往外延伸及探討，就如圖 10 所示：從每個鑰字跟房價關心程度的比較，可找出上網民眾對房地產的哪些內容，是最關心的議題。



## 圖 10 「房價」相關鑰字之關聯圖

(資料來源：筆者自繪)

經過筆者實際操作與觀察，再次印證本書作者所說：網路搜尋才是人們最真實的需求，民調數據跟按讚分享的統計，雖然可以推測人們的想法，但那些回答民調跟按讚表現，不一定是內心真實的想法，有很多是戴上面具的回應。而網路搜尋是因需求才做的行為，因此探究網路搜尋行為，就能找到最真實的想法，也不用去猜測面具底下的真意，我們的迷失也能從網路搜尋得到解答。

### 六、 結論—由網路大數據活出美好人生不再迷失

雖然我們是戴著面具生活，但在網路世界已留下真實的足跡，縱然想保護自己的隱私，卻也回不去沒有網路的時代了。現在網路大數據變成社會科學新的研究方向，本書所提的實例在在都顯示，網路搜尋數據把人們真實的想法表露無疑，雖然我們可能表裡不一，但真相已存在網路裡，只待我們去挖掘探究。

書中提到網路大數據的特性與研究方法，亦是我們可以嘗試研究的，當更多網路數據的研究結果呈現，就會開創更多社會科學的新發現；而且藉這網路數據的挖礦，使研究者更了解人們真實的想法與需求，研究者不再投石問路，也不用猜測民調回答者的真實心意，能更直觀地得到真實數據，也更了解人們真實的需求，進而能



幫助人們活出真實的自我，減少陷入迷失的情境中。

隨著網路大數據的資訊累積，使得人工智慧有長足的進步，「在一起孤獨」一書提到，陪伴機器人或機器寵物，已經在美國一些育幼院或老人院裡，扮演陪伴者的重要角色；也就是說，人工智慧的仿真人或仿真寵物，已能陪伴許多孤獨的心靈。

現在不少企業已提供線上客服機器人的服務，而其背後技術是有龐大的語音及問答集資料庫為基礎，能辨識打電話者所說的內容，也能從問答集找到解答來回答客戶。由此推測，網路大數據及人工智慧的持續發展，將來人們可在網路上找到心理諮商師、生涯規劃師或心靈導師，在網路上我們可以坦然地說出自己的問題與疑惑，並能從網路獲得真實的幫助，不再過著迷失的人生了。

參考文獻：

1. Google 搜尋引擎網頁，取自 [https://www.google.com.tw/?gws\\_rd=ssl](https://www.google.com.tw/?gws_rd=ssl)，(109/3/29)
2. Google Trend 網頁，取自 <https://trends.google.com.tw/trends/?geo=TW> (109/4/4)
3. 文字雲網頁，取自 <http://wordcloud.booogle.net/> (109/4/4)
4. 丁惠民 (譯) (2015)。大哉問時代 (原作者：Warren Berger)。

臺北市：大是文化。(原著作出版年：2014)

5. 洪世民(譯)(2017)。在一起孤獨(原作者：Sherry Turkle)。

臺北市：時報文化。(原著作出版年：2011)