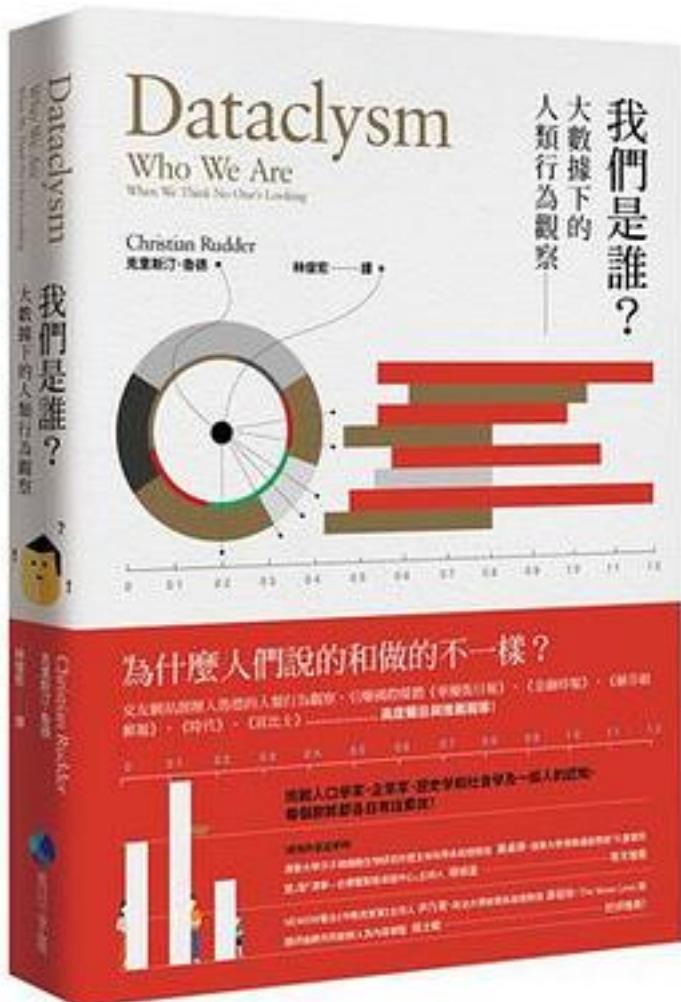


我們是誰？：大數據下的人類行為觀察



[我們是誰？：大數據下的人類行為觀察 下载链接1](#)

著者:克里斯汀·魯德

出版者:馬可孛羅文化

出版时间:2016-5

装帧:平装

isbn:9789865722906

你沒見過的大數據另一面，

美國知名社交網站創辦人用數學證實：

我們其實是這樣那樣的人。

臉書、推特、Reddit、Tumblr、Instagram……等社交網站

一躍為新一代的人口學家。

大膽創新關於人類行為的調查，一場醞釀中的革命正在發生！

一般談到大數據只談商機和操控。

其實，巨量數據正透露出人類以為四下無人時，我們是什麼樣子，

我們想的、說的、做的，其實都不一樣。

這個傳統問卷調查問不出的真相，來自大數據中，人們最真實、最不設防的心聲。

當我們以為每天只是過著自己的小日子，

每一個行為，最後都匯集到數據裡，成為大數據裡的一部分。

美國著名網站OkCupid創辦人魯德透過經營交友網站，

得到一般研究無法取得的巨量數據，而且真實，

因為在交友網站上，人們只依自己的想法行事，誠實透露好惡，

藉由統計分析，解讀用戶創造出來的資料，

用廣大的樣本數來證實一般人會有的猜想，或者破解常見的迷思。

本書觀察到在網路時代，數據所帶出的真相：

1. 男人偏愛二十歲出頭的嫩妹，這是胡亂栽贓還是信口開河？數據證實，這是真的！女人過了21歲以後，在擇偶市場就開始走下坡。
2. 缺陷是一股強大力量！與其當個芸芸眾生，還不如勇敢做自己。

「出醜效應」告訴你，有特色才讓你更出色。

3. 使用推特會讓人類的寫作能力下降？大數據說，反而提升了用戶的寫作程度。

4. 亞洲女性受到所有種族男性的高度青睞；白人和拉美裔男性和女性都比較受歡迎，不過，亞洲男性和黑人，受歡迎的程度偏低。種族偏見一直都在，即使嘴裡不承認，但潛意識則不然。

5. 女性美不美，影響的不只是在交友網站上的人氣度，

從網路數據顯示，外貌就是一切，不僅影響工作面試，連工作表現評等也與外貌有關。

6. 為什麼女人……會變心？為什麼男人……要說謊？

為什麼穆斯林……恨美國？為什麼亞洲人……看起來都很像？

這些搜尋紀錄讓Google聽到了我們所有人的告白、關注與祕密。

萬事問Google，問出的不只是答案，還暴露了你的內心。

7. 按讚洩漏的訊息，遠比你想像得多。

一個人按讚的模式還能拿來預測智商：這個模型能夠精準預測某人的標準智商測驗得分，而且一個問題都不用問。

8.

活人獻祭不是過去野蠻時代才有。一名高中女生納瓦茲發的一則無腦貼文，讓我們見識到在網路上是如何地失控擴散，不到24小時，就在740萬人眼前遭到辱罵，普通鄉民逐漸演變為酸民、暴民，大家爭相上前補刀。

9.

同性戀者是後天養成的嗎？大數據考查出美國各州都有穩定比例的同性戀人口。而同性戀是否願意出櫃，與當地的接受度有關，被迫躲在祕密裡的人，會在Google搜尋表態，因為Google是最沒有社會壓力的地方。

本書透過魯德生動有趣的敘述，每個數據都各自有話要說，為我們展示人類的行為、看到整體人性。現在，數據科學家成為新的人類學家，數據能透露的訊息，遠比我們想像中多，因此許多社會或政治上的爭辯，也可透過數據證實，而非以訛傳訛或自以為是，也因為透過分析數據，我們得以看出人類的心理，建構人類的故事。

《華爾街日報》、《金融時報》、《赫芬頓郵報》、《時代》、《富比士》、《紐約客》……

多家媒體一致推薦報導，引爆國際媒體高度矚目與討論

[推薦序]

最有趣的大數據提問，最誠實的數位田野調查

黃貞祥

清華大學分子與細胞生物研究所暨生命科學系助理教授

「know thyself」 ($\gamma\gamma\theta\iota$
 $\sigma\varepsilon\alpha\tau\nu$ ，認識你自己)，相傳是刻在希臘德爾斐的阿波羅神廟的三句箴言之一，也是其中最有名的一句。

臉書剛風行時，許多朋友除了玩開心農場，另一個上臉書的目的是玩心理測驗，是為了更瞭解自己，還是讓朋友來瞭解呢？然而，因為種種的社會規範，我們即使是在做匿名或自己才看得到答案的心理測驗，可能心中想的和表現出的，是一個自己希望成為的人，而非真實的自己。

萬萬沒想到，在臉書背後玩大數據的資料科學家，可能早就能從我們按讚的模式，判斷出一個人是外向還是內向，是開放還是保守，是否是異性戀或是同性戀等等，彷彿那些

演算法，比我們自己還瞭解自己。

沒外人真正清楚臉書等社群網站背後真正且完整的商業動機和模式，不過收集大數據來瞭解人類的行為，肯定是其中一項，這是司馬昭之心。至少，有個交友網站OkCupid的創立，除了讓無數寂寞難耐的男女能夠找到人生另一半，另一個目的就是為了寫這本書吧，真是一魚二吃啊。

不否認，我也是OkCupid的註冊會員，當時是讀了一本書說，有個交友網站能用各種問題來判斷兩人是否適合交往，裡頭可供會員回答幾百個問題，從嗜愛興趣到生活習慣和宗教政治都有，可是其中最準確的居然是喜不喜歡看恐怖片（我不愛）等等看似無關緊要的小問題。

身為一個好奇心旺盛的人，我當然立馬去試，一口氣回答了上百個問題，可是送出異性的訊息大多石沉大海……後來就興趣缺缺了。沒想到，在OkCupid成立十年來龐大的資料，卻讓資料科學家發掘出一大堆極為有興趣的現象。這本書就是OkCupid的創始者之一克里斯汀·魯德（Christian Rudder）的現身說法。玩大數據的大企業不少，包括谷歌、蘋果、亞馬遜、臉書等等都是箇中佼佼者，只是大多數資料是商業機密，像克里斯汀·魯德成立一個部落格OkThreads，用一手資料來討論並出書的不多。

從大數據瞭解我們的行為，只是個開始

《大數據下的人類行為觀察學：為什麼人們說的和做的不一樣？》不僅稀有，也還真是本令人愛不釋手的好書，克里斯汀·魯德的寫作功力和他創業及資料分析的功力一樣不凡。我一口氣讀完還感到意猶未盡，書中談到的各種現象，就發生在我們這個有社群媒體的特殊時代，裡頭揭露的各種人類行為，令人感到既熟悉又陌生，熟悉的是你我就是參與者，陌生的是眾人組成的行為模式，往往有些出人意表，意味著你我當常常嘴裏說一套，可是身體卻很老實。

從OkCupid的五百萬會員使用模式中，哈佛數學系畢業的克里斯汀·魯德挖掘出一批又一批資料，在書中畫了一個又一個簡單易懂的圖表，揭示了求偶過程中，人類隱而不宣的各種有趣行為。這樣躲在背後，趁使用者不注意時偷偷觀察各種社會學、心理學、人類學現象，就像是生物學家在野外做田野調查，而非把動物關在實驗室裏頭操作。

克里斯汀·魯德看來泡在大數據中是無時無刻都樂在其中，他旺盛的好奇心，問了一個又一個大家都很感興趣，也隱約以為知道答案，然而卻無從下手的好問題。書中分為三部分探討「我們為何互相吸引」、「我們又為何互相排斥」和「我們為什麼是現在這樣子？」。

書中五花八門的問題包括：男人和女人喜歡哪個年紀的異性呢？缺陷也能為人加分？推特讓寫作能力下降？跨種族的吸引力是如何？外貌究竟有多重要？我們怎麼描述自己？怎麼樣的社交網路有助穩固婚姻關係？還有多少同性戀者未出櫃？同性戀者的比例和居住地區有關嗎？犯眾怒的下場會是？

這些問題不單單是有趣而已，這些問題的答案，也會是重要的公共政策之參考。《大數據下的人類行為觀察學》舉了不少實例，指出好些社會或政治上的爭辯，其實可以透過真實的數據來探究，而非以訛傳訛或自以為是。

例如，準確預測出同性戀者在人口中的比例，是個嚴肅的政治問題，大數據考查出美國各州都有穩定比例的同性戀人口，與各州政治宗教氣氛無關，這也意味著同性戀非後天養成的。很有趣但也可悲的，在美國保守的州，更多已婚婦女上谷歌查詢的是她們老公是否是同性戀，比查詢是否偷吃的還多。可喜的是，從大數據來判斷，美國至少對非白人當總統和接受同性婚姻等等事上，是真的愈來愈開明了。

雖然大數據時代中知識就是力量，可是臉書和谷歌都能夠操縱我們閱聽的資訊，我們也不該完全放心大數據的力量，也不能盡信大數據。即使大數據告訴我們的是準確無誤的，瞭解我們的行為只是個開始而已。科技始終該來自人性，我們在這時候更該問題，是身為人的價值和意義為何？而非隨波逐流地讓科技來決定我們該如何生活和看待世界。

[推薦序]

由量化到質性的大數據行為分析

簡禎富

國立清華大學 清華講座教授

科技部「IC產業同盟」計畫暨「清華-台積電卓越製造中心」主持人

現代生活已離不開網路和資訊科技，但凡走過必留下痕跡，所有使用者在網頁上點擊、停留、交流和社群活動而被自動儲存的巨量資料，創造直接分析大數據以檢視人類行為的契機。本書作者克里斯汀·魯德 (Christian Rudder) 畢業於哈佛大學數學系，為交友網站OkCupid的共同創辦人，他結合大數據

分析技術與領域知識，分析線上交友網站的第一手資料和深度的文字探勘 (text mining)，提出與傳統愛情專家不一樣的情感分析模式，並以實際案例介紹大數據分析的實質應用，並以簡單明瞭的視覺化方法提出各種洞察與解釋。作者認為，相較於臉書以現實生活人際網絡形成的社群，線上交友網站相對封閉與注重隱私，使用者比較能跳脫社會壓力，而依自己的態度和期望發表意見，不必擔心聊天內容出現在好友動態軸裡，承受他人七嘴八舌的評論。因此，交友網站使用者的表現反而更能反映出真實的個人偏好。例如：本書分析交友網站資料發現種族歧視仍普遍存在社會之中。

隨著軟硬體計算能力和雲網端的發展、處理結構和非結構化巨量資料的能力的進步，以及智能化大數據模式的研究，大數據應用的趨勢已經由量化到質性的分析，例如，本書對於人類行為大數據的觀察和分析應用。我們研究團隊亦透過產學合作計畫，從人力資源的巨量資料分析員工績效表現與其背景、招募管道等因子的關係，改變高科技產業人才遴選決策；此外，在某些職務，經內部推薦管道招募而來的人才表現相對外部招募為佳。根據這些數據分析結果，即可擬定相關決策建議，例如：與相關大學科系建立產學合作研究與實習機制等，以吸引「志同道合」人才；或是設計激勵獎金機制，鼓勵員工進行內部推薦，提升人才招募效率與品質。另外，我們也分析人力銀行的巨量資料，以分析未來類似學經歷的求職新鮮人，卻有不同的待遇，以找出對各種背景、不同領域的有效「加値能力」，避免社會新鮮人自我投資的資源分散和損失。

各國產業結構和社會文化迥異下，所需發展的大數據產業和應用也不盡相同。台灣長期以大量生產、規模經濟的製造模式，在「工業4.0」、物聯網和客製化時代將不再適用，而世界各國都普遍面臨因為自動化和機械人而加劇的貧富差距和年輕人就業供需不等問題。台灣應善用目前在「工業3.0」的供應鏈管理、整合能力和彈性應變的優勢，發展導入具智慧製造和聰明生產服務的「工業3.5」策略，搶先在優勢產業領域與利基市場建立破壞性創新，以人性化的大數據分析和人機協同的智能系統，因應無人化的工業4.0對台灣產業和人類社會的挑戰。然而，在透過大數據發掘潛在未知的人類行為和社群模式時，一方面可以確認一些猜想、破除一些假設；另一方面，網路行為累積的大數據的分析和應用，也可能會創造新的迷思和社會問題，如網路上的造神、霸凌和難以「洗心革面」的難題。

引言

第1部分 我們為何互相吸引

1. 伍德森定律：男人偏愛年輕美眉

- 2. 大眾品味，就敗在太過平凡無聊
- 3. 牆上的字
- 4. 人際關係維繫強度和怎麼連結有關係
- 5. 實際，人不懂自己的心

第2部分 我們又為何互相排斥

- 6. 干擾因子
- 7. 對美貌神話的崇拜
- 8. 讓使用者自己開口說出心裡話
- 9. 憤怒的日子：透過打壓別人抬高自己

第3部分 我們為什麼是現在這樣子？

- 10. 在亞裔裡算是高的
- 11. 資料看到的社會痛苦指數：當有人必須隱藏自己性向時
- 12. 知道你的位置
- 13. 我們的品牌，可能是你的生活
- 14. 數位麵包屑

尾聲

本書資料說明

附註

致謝

- ★《紐約時報》暢銷書
- ★2014年《富比士》推薦書單
- ★2014年全國公共電台最佳圖書
- ★2014年環球郵報最佳圖書
- ★2014年彭博社最佳圖書
- ★入圍洛杉磯時報圖書獎
- ★2014年英國獨立報網路和技術最佳圖書
- ★2015年誠品書店、博客來外文書趨勢類推薦書單

作者介紹:

克里斯汀．魯德Christian Rudder

克里斯汀．魯德是交友網站OkCupid的共同創辦人暨總裁，也是人氣部落格OkTrends的作者。1998年畢業於哈佛大學數學系，曾擔任SparkNotes創意總監，上過國家廣播公司（NBC）的《日線》（Dateline）及美國國家公共廣播電台（NPR）的《面面俱到》（All Things Considered）等節目，作品也得到《紐約時報》、《紐約客》等等媒體熱烈討論。目前他與妻女住在布魯克林。

譯者簡介

林俊宏

師範大學翻譯研究所碩士。

喜好電影、音樂、閱讀、閒晃，

覺得把話講清楚比什麼都重要。

譯有《大數據》、《英語的祕密家譜》、《建築為何重要》、

《大科學》、《群的智慧》、《天空的女兒》等書。

現就讀於師大譯研所博士班。

目錄: 引言

第1部分 我們為何互相吸引

1. 伍德森定律：男人偏愛年輕美眉
2. 大眾品味，就敗在太過平凡無聊
3. 牆上的字
4. 人際關係維繫強度和怎麼連結有關係
5. 實際，人不懂自己的心

第2部分 我們又為何互相排斥

6. 干擾因子
7. 對美貌神話的崇拜
8. 讓使用者自己開口說出心裡話
9. 憤怒的日子：透過打壓別人抬高自己

第3部分 我們為什麼是現在這樣子？

10. 在亞裔裡算是高的
11. 資料看到的社會痛苦指數：當有人必須隱藏自己性向時
12. 知道你的位置
13. 我們的品牌，可能是你的生活
14. 數位麵包屑

尾聲

本書資料說明

附註

致謝

• • • • • (收起)

标签

大数据

人类行为

社会学

台版

繁体

互联网

港台

文化研究

评论

翻譯的質量一般，影響了書的質量。剛開始以為只是一個經營交友網站的老闆想依靠自己網站的一些信息YY一些發現來賺錢的伎倆。直到書末，透過對資訊與人類隱私的深刻思考，透露出一股先人一步而又謙虛謹慎，嚴肅認真的學者態度。

數據祇是一個黑盒子。但這些依舊是不夠透明的，而且僅僅是相關關係，說不出任何因果。我們有意識地在網路上隱藏身份，建構另一個自我。

好吧，就這本書來說，我不想花太多去讀。一天大概從頭到尾翻翻就結束閱讀。但為什麼還是想態讀，因為被伍德森定律吸引到，馬上要21歲的女生當然會很好奇地翻來讀讀看。

愈是做程序、偷窺數據的人，愈在意保護自己的隱私。搜索引擎的搜索軌跡說出我們不願意說的話，但無痕式分頁真的有效嗎？按讚模式、虛擬鍵盤的擊鍵次數、撰寫時間與實際寄出的私訊字元數，這些被數據庫記錄下來的行為，貌似給我們畫出一

個數據肖像，暴露內心的線上行為統統被拿去做X-ray般的散布圖透視。開箱結果蠻有趣，但也挺惊人的。所以搞不好，新的人類學家必須兼能做數據科學家才行。

案例，多数据分析社交行为。

其实全书干货一般。基于单个社交网站的发现可以当做一个有趣的个案来看，得到的结论似乎可以归纳为二：1.大数据可以追踪到比你想象中多得多的信息，勾勒出你和你所属的某些群体的特征

2.不同群体的人在网络上存在差异（关于人类行为的结论，例如性别偏见、种族歧视、同性恋歧视等等，貌似都是前人研究中已经有了定论的。。。好像并没有让人惊讶的发现。。）

整本书对大数据的迷信其实让人有些惶恐：并不是数据多，了解到的东西就多。读之前可能需要对统计有些接触，也需要先认识到人类的复杂性和丰富性吧。

[我們是誰？：大數據下的人類行為觀察](#) [下载链接1](#)

书评

[我們是誰？：大數據下的人類行為觀察](#) [下载链接1](#)