

《數據、謊言與真相：Google 資料分析師用大數據揭露人們的真面目》

難易指數：☆

關鍵詞：#大數據 #資料探勘

近年來隨著儲存成本降低與資料獲取量變大，加以物聯網的發展與無所不在的感測器，使得數據的來源與種類更加多樣化，同時數據的收集也更加即時、快速（Veracity），更具真實性（Veracity）。

本書作者有鑑於傳統問卷調查，未必能反應出填答者的真實想法與心聲，反而充斥各式「謊言」。因此，透過研究人們最真實且誠實的新數據－包含網路上的搜尋記錄與行為模式、推特與臉書等社群媒體上的用字與言論、棒球數據，乃至約會網站、購物網站所留下的數位足跡，一窺人們的内心世界，了解他們最真實的想法，以及未經掩飾的内心渴望和價值觀，進而了解真實的社會現況。企圖揭穿意想不到的真相與人們躲在網路背後的真面目－發現他們心裡想的是什麼、想要的是什麼，又或者真正做了什麼。

例如從美國各州人民在網路上對於特定用字的搜尋記錄，可了解各州種族歧視的狀況，進而預測選民的政治傾向與大選結果。然而，作者也在書中提醒讀者，面對數據務必小心，大數據有其局限，帶有偏見的演算法，更是容易衍生公平性與歧視性問題。