

The Big Data Revolution

大數據

數據革命如何改變政府、商業與我們的生活

涂子沛 著

T h e B i g D a t a R e v o l u t i o n

大數據

數據革命如何改變政府、商業與我們的生活

涂子沛 (Zipei Tu) 著



香港中和出版有限公司
www.hkopenpage.com

責任編輯	丁志紅
版式設計	彭若東
封面設計	閔 鵬
責任校對	江蓉甬
排 版	潘斯麗
印 務	馮政光

書 名	大數據：數據革命如何改變政府、商業與我們的生活
作 者	涂子沛 (Zipei Tu)
出 版	香港中和出版有限公司 Hong Kong Open Page Publishing Co., Ltd. 香港北角英皇道 499 號北角工業大廈 18 樓 http://www.hkopenpage.com http://weibo.com/hkopenpage http://www.facebook.com/hkopenpage
香港發行	香港聯合書刊物流有限公司 香港新界大埔汀麗路 36 號 3 字樓
印 刷	陽光印刷製本廠有限公司 香港柴灣安業街 3 號新藝工業大廈 6 字樓
版 次	2013 年 3 月香港第一版第一次印刷
規 格	16 開 (168mm×230mm) 352 面
國際書號	ISBN 978-988-8200-34-4 © 2013 Hong Kong Open Page Publishing Co., Ltd. Published in Hong Kong

目 錄

序言一	大數據：為華文世界提出一個重要話題 / 許倬雲	11
序言二	中國的雄心應該拓展到大數據領域 / 托馬斯·H·達文波特	15
序 幕	新總統的第一天	19
	一人一票：把“黑”人送進“白”宮	20
	大國新政：陽光是最好的防腐劑	25
上篇 帝國風雲 得數據者得天下——美國的成功經驗		
第一章	歷史爭戰《信息自由法》	33
	第四股力量：知情權的起點	34
	國會議員：孤獨的戰爭	37
	白宮當家人：一個妥協者和機動者	41
	政府 VS. 社會：舊劇情重現新時代	46
第二章	數據帝國的興起	51
	摩爾定律：全世界半個世紀的發展規律	52
	最小數據集：上升到立法高度的開路先鋒	59
	民意幾時有：選票催生的創新	62
	普適計算：計算機本身將從人們的視線中消失	68
	“大數據”戰略：爭奪全世界的下一個前沿	72

第三章 數據治國	79
循“數”管理：平安大道怎樣鋪	81
數據“驗”平權：民權史上的碑石	89
數據“打”假：最大的爭議就是福利濫用	92
CompStat：街頭警察的創新傳奇	95
第四章 商務智能的前世今生	103
起源：從數據到知識的挑戰和跨越	104
結蛹：數據倉庫之厚積薄發	108
蠶動：聯機分析之驚艷	111
破繭：數據挖掘之智能生命的產生	115
化蝶：數據可視化的華麗上演	118

中篇 法則博弈 大數據時代的發展與危機

第五章 帝國的法則	133
收集法則：減負，為人民減負	134
使用法則：隱私，文明社會的共識	139
發佈法則：免費，人民已經交稅	145
管理法則：質量，互聯網時代的根本	149
第六章 《數據質量法》的困局	153
產業界“俘虜”政府：數據背後的政經戰爭	154
美式“旋轉門”：權、名、利大串場	157
“摻沙子”法案：國會對付總統的獨門秘器	160
環保“風險門”：公共利益常常無人代表	164
集體行動的邏輯：人人都想“搭便車”	167
三權之歧：甚麼是真正的“和諧”	170

第七章 全國隱私風波	175
《一九八四》：零隱私的恐懼	175
大數據就是“老大哥”：中央數據銀行之爭	177
百年糾結：統一身份證	182
“9·11”大拐點：以反恐的名義向左轉	186
萬維信息觸角計劃：追蹤恐怖分子的“數據腳印”	189
6種改變政府的力量：山姆大叔大退讓	193

下篇 公民故事 公民社會與數據互動之雙贏時代

第八章 數據開放運動	201
一個新的世界：從軟件開源到數據開放	201
總統的雄心：公共財政支出透明	209
數據民主：印裔首席信息官的崛起	215
Data.Gov：從旗艦初航到保“數”運動	220
大眾創新：航班延誤之候機經濟學	225
第九章 試金石：白宮訪客記錄	235
總統在見誰：大醫改中的“小”插曲	236
全體美國人的房子：白宮	243
步步妥協：總統與草根的對決	247
從白宮安保到政治監督：執著的公民改變世界	253
第十章 礦難中的歌聲和數據	259
集體行動的號角：你站在哪一邊	262
可以避免的悲劇：數據揭示的全景式真相	268
默認公開推定：和礦難賽跑的原則	272
唯一的道路：民主時時都要“爭”	277

外篇 天下趨勢 大數據社會未來之世界走向

第十一章 大趨勢	285
數據權：大不列顛的碩果	285
大合流：國際開放聯盟	290
雲計算：新的航向	294
再造互聯網：從網頁相連到數據相“聯”	298
第十二章 大挑戰	307
逐鹿政壇：得數據者得天下	308
數據競爭：企業贏之道	315
下一波浪潮：從大數據到大社會	322
尾聲 挑戰中國：摘下“差不多先生”的文化標籤	329
大事記 20世紀大蕭條後美國信息開放、技術創新之路	339
本書涉及的美國政府及民間組織機構譯名表	343
關於本書數據、專有名詞體例及圖片版權的說明	346
後記 搭建“大數據”的世界	347

T h e B i g D a t a R e v o l u t i o n

大數據

數據革命如何改變政府、商業與我們的生活

涂子沛 (Zipei Tu) 著



香港中和出版有限公司
www.hkopenpage.com

一個真正的信息社會，首先是一個公民社會。

——作者題記

薛華成〔中國管理信息系統的主要開創者、澳門科技大學行政與管理學院名譽院長〕

通過歷史故事的講述，用近似新聞評論的筆法，《大數據》這本書以美國為例介紹了由經驗管理到數據管理這場正在發生的革命性變革，通俗形象地講清楚了科學的道理，實屬難得。這對學者們擴大視野，對管理幹部深化科學發展觀的認識，都有十分重要的價值。

施伯樂〔復旦大學計算機研究所所長、上海（國際）數據庫研究中心主任〕

《大數據》這本書揭示了數據在未來社會中的革命性作用。大數據將是下一個社會發展階段的“石油”和“金礦”。無論是個人、企業還是國家，誰能更好地抓住數據、理解數據、分析數據，誰就能在下一波的社會競爭中脫穎而出。關於數據的知識，將成為個人知識結構中的必備要素和基礎。

趙嘉敏〔東西網總裁、譯言網創始人之一〕

子沛兄的《大數據》這本書，幫助我們形象地了解信息自由、技術創新與社會進步的關係。我最感興趣的是美國信息自由制度建立的過程。它讓我看到，一個社會、一種文化，是如何通過自省式的努力，認清比自身存在更本源的客觀規律。這種制度一旦建立，就如同打開了控制某個基因群組的開關，後面所呈現的演化，雖然眼花繚亂，反而顯得平淡無奇、順理成章了。

董潔林〔蘇州大學企業創新和發展研究中心主任、教授〕

18世紀啟動的工業革命為人類創造了巨大的財富，也極大地推動了社會變革。20世紀中啟動的信息革命對人類社會也將產生巨大的衝擊，而這個進程剛剛開始。涂子沛先生所著的《大數據》一書，用有趣的故事帶我們回顧信息時代的里程碑，使我們透過未來的變幻不清的迷霧，展望和暢想信息革命的大趨勢。

汪小帆〔上海交大電子信息與電氣工程學院教授〕

《大數據》值得仔細閱讀與思考。要讓數據更好地說話，既依賴於數據科學與技術的發展，更需要政府、機構和公民的數據意識的不斷提高。

張暉〔卡內基梅隆大學計算機系終身教授〕

數據科學正在成為一個新興的學科和領域。涂子沛先生的《大數據》這本書，為我們揭示了大數據浪潮發生的前因後果，以及這種技術進步對政治、經濟、文化的種種影響。希望讀者從本書中獲益，以積極的態勢應對中國的大數據挑戰。

Ramayya Krishnan〔卡內基梅隆大學海因茲學院及信息管理學院院長、終身教授〕

Big Data has the potential to fundamentally transform Society. However, unlocking this potential will require careful attention to data governance and insightful application of data analytics combined with an environment that spurs managerial innovation. Zipei's book provides a roadmap for the Chinese professionals using the American experience as context. As the dean of the Heinz College and its School of Information Systems at Carnegie Mellon University, I am very proud of what Zipei has accomplished. His book combines government, business and technology concepts in a holistic manner, embodying the unique strengths of programs and faculty of the Heinz College. I believe this book will stimulate new ideas in China to meet the big data challenge.

中文譯文：大數據具有催生社會變革的能量。但釋放這種能量，需要嚴謹的數據治理、富有洞見的數據分析，以及一個激發管理創新的環境。涂子沛的《大數據》這本書總結了美國的經驗和實踐，並為中國提出了一個路線圖。我為他的成就感到驕傲。這本書完整地融合了技術、政治、商業三個角度，體現了我們學院師資教學的獨特優勢。我相信本書能夠在中國激發新的思想和討論，幫助中國迎接大數據的挑戰。

獻給

母親魏木蘭和妻子張煒婷

感恩無私的愛

目 錄

序言一	大數據：為華文世界提出一個重要話題 / 許倬雲	11
序言二	中國的雄心應該拓展到大數據領域 / 托馬斯·H·達文波特	15
序 幕	新總統的第一天	19
	一人一票：把“黑”人送進“白”宮	20
	大國新政：陽光是最好的防腐劑	25
上篇 帝國風雲 得數據者得天下——美國的成功經驗		
第一章	歷史爭戰《信息自由法》	33
	第四股力量：知情權的起點	34
	國會議員：孤獨的戰爭	37
	白宮當家人：一個妥協者和機動者	41
	政府 VS. 社會：舊劇情重現新時代	46
第二章	數據帝國的興起	51
	摩爾定律：全世界半個世紀的發展規律	52
	最小數據集：上升到立法高度的開路先鋒	59
	民意幾時有：選票催生的創新	62
	普適計算：計算機本身將從人們的視線中消失	68
	“大數據”戰略：爭奪全世界的下一個前沿	72

第三章 數據治國	79
循“數”管理：平安大道怎樣鋪	81
數據“驗”平權：民權史上的碑石	89
數據“打”假：最大的爭議就是福利濫用	92
CompStat：街頭警察的創新傳奇	95
第四章 商務智能的前世今生	103
起源：從數據到知識的挑戰和跨越	104
結蛹：數據倉庫之厚積薄發	108
蠶動：聯機分析之驚艷	111
破繭：數據挖掘之智能生命的產生	115
化蝶：數據可視化的華麗上演	118

中篇 法則博弈 大數據時代的發展與危機

第五章 帝國的法則	133
收集法則：減負，為人民減負	134
使用法則：隱私，文明社會的共識	139
發佈法則：免費，人民已經交稅	145
管理法則：質量，互聯網時代的根本	149
第六章 《數據質量法》的困局	153
產業界“俘虜”政府：數據背後的政經戰爭	154
美式“旋轉門”：權、名、利大串場	157
“摻沙子”法案：國會對付總統的獨門秘器	160
環保“風險門”：公共利益常常無人代表	164
集體行動的邏輯：人人都想“搭便車”	167
三權之歧：甚麼是真正的“和諧”	170

第七章 全國隱私風波	175
《一九八四》：零隱私的恐懼	175
大數據就是“老大哥”：中央數據銀行之爭	177
百年糾結：統一身份證	182
“9·11”大拐點：以反恐的名義向左轉	186
萬維信息觸角計劃：追蹤恐怖分子的“數據腳印”	189
6種改變政府的力量：山姆大叔大退讓	193

下篇 公民故事 公民社會與數據互動之雙贏時代

第八章 數據開放運動	201
一個新的世界：從軟件開源到數據開放	201
總統的雄心：公共財政支出透明	209
數據民主：印裔首席信息官的崛起	215
Data.Gov：從旗艦初航到保“數”運動	220
大眾創新：航班延誤之候機經濟學	225
第九章 試金石：白宮訪客記錄	235
總統在見誰：大醫改中的“小”插曲	236
全體美國人的房子：白宮	243
步步妥協：總統與草根的對決	247
從白宮安保到政治監督：執著的公民改變世界	253
第十章 礦難中的歌聲和數據	259
集體行動的號角：你站在哪一邊	262
可以避免的悲劇：數據揭示的全景式真相	268
默認公開推定：和礦難賽跑的原則	272
唯一的道路：民主時時都要“爭”	277

外篇 天下趨勢 大數據社會未來之世界走向

第十一章 大趨勢	285
數據權：大不列顛的碩果	285
大合流：國際開放聯盟	290
雲計算：新的航向	294
再造互聯網：從網頁相連到數據相“聯”	298
第十二章 大挑戰	307
逐鹿政壇：得數據者得天下	308
數據競爭：企業贏之道	315
下一波浪潮：從大數據到大社會	322
尾聲 挑戰中國：摘下“差不多先生”的文化標籤	329
大事記 20世紀大蕭條後美國信息開放、技術創新之路	339
本書涉及的美國政府及民間組織機構譯名表	343
關於本書數據、專有名詞體例及圖片版權的說明	346
後記 搭建“大數據”的世界	347