



## 講「數」助施政——數據科學與政策研究課程

節錄自《CUHK UPDATES 頭條》

根據衛生署的資料，本港 2014 至 2015 學年有 56.8% 的中學生、26% 的高小學生曾經飲酒，反映青少年飲酒情況普遍。食物及衛生局遂於本年 11 月 30 日實施新法例，禁止零售商向 18 歲以下人士販賣酒精飲品。

數字是社會的縮影，可以揭露隱藏的問題、呈現趨勢。我們如何解讀這些數據，將會影響我們的世界觀；為政者如何解讀這些數據，也會決定他們推出甚麼政策、塑造怎樣的社會。

數據科學並非新事物，商界和政府早已廣泛應用，大學也相繼開辦有關課程，但大多數都側重商業應用。然而，新加坡政府已計劃讓約二萬名公僕接受數據科學的正規培訓，中大同樣預見到，把數據科學應用於公共行政之上，勢必成為一門新專業；社會科學院遂於 2019 至 2020 學年開辦新學士課程「數據科學與政策研究」。課程主任黃偉豪教授說：「任何一個範疇的政策，都可應用數據科學。」



### 預測分析助籌謀

新課程是跨學科的課程，學生將會修讀「政策研究方法」、「公共政策的數據分析」和「政策分析及設計思考」等必修科目，以及選修本科、理學院和工程學院的科目，例如「數據科學與監管」、「大數據系統編程」和「電子商務數據採集和分析」等。課程也會涵蓋人工智能，例如怎樣運用電腦模仿人類分析數據等。

數據科學和統計學容易叫人混淆，兩者彷彿都是進行計算，然後從數字得出一些結論。黃教授說明兩者分別：「數據科學是一門跨學科，統計學是其中重要一環。」以公共醫療為例，統計學可以證明某種疾病與地區有相互關係，即是居於某些地區的人口較大機會感染該疾病，但要分析背後原因、制定預防措施，便涉及其他範疇的理論，而整個過程都屬於數據科學。

此外，數據科學是前瞻性的，以預測分析為主。外國的反恐部門會收集社交媒體用戶在網上的行為，如發布偏激言論的頻率，顯示的個人資料如就業狀況等。透過挖掘分析這些數據，評估潛在風險，預測哪些人有可能發動恐襲，採取監視行動。

### 探討法規與道德

課程也會探討數據科學帶來的道德問題，數據科學發展迅速，但相關規管和法例卻見滯後。歐盟剛於 5 月才實施《通用數據保障條例》，當中最為人熟悉的是企業若發現有數據外泄，須於七十二小時內呈報，香港則暫時沒有相關機制，但近期數據外泄事故頻生，促使大眾反思是否要立法堵塞漏洞。黃教授指出：「繼應用、技術開發後，立法規管將會成為第三波發展。」

黃教授舉例說明數據科學發展和道德之間的矛盾：「若將來市民的基因圖譜公開，而某人的基因圖譜顯示他將於三十歲患上慢性疾病，僱主不聘用他，或藉機削減其薪酬福利，這樣是否構成歧視？」

除了課堂理論外，學生也有機會去不同機構進行實習，包括政府部門、非政府組織和私人企業，了解業界情況和數據科學最新發展。政府除了分析數據外，日後還可能開放部分數據，這些工作都要倚靠數據科學的專才。即使是商界，企業持有許多客戶資料，近年也開始投入資源分析數據，期望提供建議回饋社會。因此，學生畢業後，可以投身不同機構，並非局限於公職。

政府坐擁最多數據，日後將會更常運用這些「資源」來協助施政，另一方面，開放公共數據也是大勢所趨，但開放哪些數據、如何開放，開放之餘如何保障市民私隱，這些工作都要用上專業知識。分析數據，制定政策，將會成為廿一世紀社會發展的新方向。



本院成員兼校友黃偉豪教授。