

以嶄新角度切入數據科學

蘇嘉欣強調 溝通技巧重要



▲ 蘇嘉欣認為數據科學家的說故事能力十分重要。

以為能當數據科學家，不是電腦奇才就是數學家，原來凡事總有例外。蘇嘉欣（14 新亞生物醫學工程學）修讀的像是與數據毫無關係，但她卻認為數學及工程是互通的，「其他工程讀的，可能是水管結構與流動，生物醫學工程只是由水管換成了血管，箇中數理邏輯並沒有變。」

中大的生物醫學工程於 2010 年開辦，蘇嘉欣（Winnie）是第一屆畢業生，問她為何敢當「開荒牛」，她笑說背後原因竟是不夠溫柔，「中學時已對心電圖、基因、神經學等範疇感興趣，可惜性格不夠溫柔，也沒有太大耐性，早知不能當上醫生或物理治療師，參觀中大開放日認識到將會有生物醫學工程這學科，毫不猶豫就選了。」

雖然讀書內容均是圍繞醫學範疇，但 Winnie 卻認為生物醫學只是場景，後面工程二字才是重點，「我會用『放諸四海皆準』來形容，正如工程有流體力學（fluid mechanics），我們就有生物流體學（bio-fluid），他們有機械工程（robotics），我們也有生物機械工程（bio-robotics），

這正是跨學科（cross-disciplinary）的概念，彼此傳授的數學及工程理論是互通的。」她甚至認為讀非工程科目的，看待工程會更能跳出框框，「讀書時代表中大參加由 MIT 舉辦的國際基因工程大賽（iGEM），竟然見到有藝術系及建築系的學生參賽，前者用舞蹈來演繹蛋白質形狀，後者則用建築模型來建構基因形狀，那刻才體會到原來可用另類

視角來解讀工程，這對日後工作幫助很大。」

建立台灣電召車派單系統

最明顯例子，要數她畢業後加入 GOGO VAN 擔任數據科學家說起，「當時公司已開發台灣市場，系統應用在香港效率很高，但在台灣用起來數據卻總是有點問題。於是我向公司申請飛往台灣了解，最終發現兩地電召車輛方法截然不同：香港模式是 C2C 的，客人直接電召車輛，司機本人接單，但在台灣卻是 B2C 模式，中間有電商營運，經它分單給電單車手，所以整個算法要修動為 B2C，才能有效派單。」



▲ 在 GOGO VAN 工作時，Winnie 曾與大學物流系合作，研究物流工程相關的機器學習（Machine Learning）應用。

找到問題所在，蘇校友回港匯報她的發現時，同事卻不理解為何要改動系統，「花了很多時間解釋，他們也聽不明白台灣市場到底如何運

作，突然記起讀書時的參賽經驗，於是轉個角度解說：碰巧上司是波蘭人，極為喜歡飲酒，我便探勘 (Data Mining) 港九新界的酒吧位置並羅列在地圖上，假設同事分別管理不同地區酒吧，要如何安排路線，才能令上司可以在短時間內試勻所有酒吧。聽完這個說法，大家即時理解台灣的分單機制，最終也推動到系統算法優化研究的工作。」

經此一役，蘇嘉欣明白數據科學家要懂得把故事說好，「很多同事沒有科技背景，他們不知道數據可以怎麼應用，身為數據科學家不能像其他工種般，乖乖地坐在辦公室，等上司給予指引，然後完成工作；大部分時間是要自覺發掘公司數據，從中找到痛點並作出改善，要對建立數據模型 (data model) 充滿熱誠，還要講出動聽的故事，讓同事明白為何要這樣做，當中涉及人工智能、機器學習、程式編碼等技術，數據科學家只是拿到入場券，溝通技巧亦不能忽略。」

不過即使說服了同事支持自己，但他們不肯用來製做新產品也是徒然，Winnie 坦承初入行時曾因此而一度迷惘，直至最後加入現時公司 AXA 安盛保險後，終於明白箇中原因，「不論是 GOGO VAN 還是後來工作的初創，業務範圍不算太廣，來到保險公司，涉及的商業範疇變得很大，終於明白以前構思數據項目時，還未考慮公司業務上的需要，才會令項目完成後無人問津，這是在 AXA 工作最大的學習得着。」

以炒鑊作比喻

來到新公司，不單要繼續開發新項目，也要推廣數據文化，蘇校友再次發揮說故事的本領，「負責保險後勤工作的同事，每日要處理很多數據，於是我們團隊將之作數據清洗並放上雲端，亦整合一堆分析工具，用家毋須寫程式碼就能自取數據使用。開發完成後，我以炒鑊作比喻：食材由農場運到倉庫，仿如將數據放上雲端，如果你是大廚，懂得處理新鮮食材的，當然可以直



▲ 在 AXA 工作，獲派往法國總公司實習，Winnie (右二) 與當地數據科學家交流。

接寫程式擷取雲上的數據，但如果廚藝一般的，我們數據部門已開發好簡單工具，等於將食材處理好並包裝成即食食物，拿回去可以很快煮好開始吃……聽完這個故事，各部門同事都躍躍欲試我們的新產品。數據科學家不可以只注重開發，開發後也要花心思吸引用家使用你的產品，才是成功。」

參加數據比賽做靚履歷才入行

聽起來數據科學家像是新興行業，蘇嘉欣卻說「這個行業其實一直存在，只是以前喚作研究員或分析員，現時有更好聽的名字罷了！事實上，近年數據科學家也真的很搶手，尤其是銀行及保險公司也出動高薪掘角，大家也意識到數據的用途。」她提醒想入行的，除了要掌握行業所需技術外，最重要是多做數據項目，「自發做些項目放到網站上，或是多參加不同的數據比賽，公司請人就是看申請人以往做過的項目，是否切合到公司的數據需求。」👉



▲ 讀書時參加 iGEM 比賽的經驗，Winnie (前排右二) 明白到工程可以不同角度解讀。

蘇嘉欣小檔案

- 2014 年 ● 香港中文大學工程學士
- 2016 年 ● 香港城市大學工程學碩士
GOGO VAN 數據科學家
- 2017 年 ● ANIWARE 數據科學家
- 2018 年 ● AXA 數據科學家