



## 跨洋合作》文字探勘計畫 嗅出商業情報

【經濟日報/林婉翎】

2008.07.27 02:13 am

財報等報表資訊，是研判投資標的重要依據；然而，報表資訊的內容，往往只呈現「過去的數字」，對於企業的「現況」，陳述有限，投資人若單就報表資訊，難以直接、準確的制訂決策。

有鑒於此，台大管理學院與美國亞利桑那大學資訊管理系暨研究所講座教授陳忻鈞合作，進行由國科會補助為期三年的「文字探勘（Text mining）整合型計畫」，運用資料與文字探勘（Data and Text Mining）技術，分析在網路上「文字類」的商業資訊，盼能結合量化與質化分析，強化商業情報的可信度，協助企業更有效的制訂決策。

### 台大管院 與陳忻鈞合作

「時機已成熟，現在談合作，正是時候。」台大會計系主任暨所長李書行表示，其實，早在2003年，就已埋下雙方的合作種子。陳忻鈞和李書行兩人，是1980代後期，就讀紐約大學研究所博士班的學長、學弟。

1996年，當時擔任亞利桑那大學人工智慧實驗室主任的陳忻鈞，主持的警政計畫，成功開發「COPLINK」系統，幫助警察捕捉嫌犯，並透過即時有效的資訊整合與分析預防犯罪。該系統已被美國警察局廣泛使用。

後續的黑網（Dark Web）計畫，能自動追蹤網路上的恐怖網頁，從分析網路文章的語句、文字風格、相關社交聯網，追蹤恐怖分子在網路上的交流。為美國執法與情報機構執行「防恐」作業，找到了好幫手。

相關的技術，也應用在醫學資訊等領域。李書行回憶，陳忻鈞在研發成果發表後，2003年邀請他前往觀摩。在他親自體驗後，相當驚艷，便和陳忻鈞洽談進一步合作的可能性，但當時，限於經費與相關研究環境未臻成熟，只好暫擱這項計畫。

不過，這段時間，他們一直在找尋適合的時間點。李書行指出，以往，企業進行商業情報分析，為了避免只分析報表等數字資訊，不夠客觀，總會輔以大量的媒體報導等文字資料，但質化的文本分析較主觀，人為判斷有其風險存在，而且耗時、難大量處理。

### 資訊工程 帶動商管成長

李書行說，如今，Web2.0概念風行全球，每天有上萬筆資訊在網路世界流竄，資訊流通更多也更快了，更添企業進行質化分析的難度，當然，有些人會以焦點團體調查，蒐集樣本資料。

「但是，少數樣本分析，資訊豐富度顯得不夠完整，所以判斷未來的風險不小。」李書行說，商業情報分析要有突破，要能在很短的時間內，更有效的蒐集和分析大量的質化資訊。

這便是台大管理學院和陳忻鈞合作的目的。李書行認為，透過這項技術，應用在商業情報分析，不但能在最快的時間，掌握世界各角落的資訊，「也能讓資訊工程帶動商管的成長。」

陳忻鈞進一步說明，這項技術不只能分析，媒體報導的內容，也可以分析散落於網路世界各個部落格的網誌、政府網站等公開資訊內容。「在正式管道還未公開前，很多資訊已經可以在部落格上看到。」他說。

他指出，部落格上的文章，多少帶有部落客的「情緒反應」，透過系統分析，可為文章內的關鍵情緒性字眼，設定不同的「分數」。這些關鍵情緒性字眼的總分分析結果，可做為企業決策判斷的參考。

陳忻鈞坦誠，就現階段的產業環境與技術成熟度，現在開始合作正值「最完美的風暴（perfect storm）」。

### 透過分析 提升決策信度

文字探勘如何提升決策的可信賴度？李書行舉例，負責銀行貸款的授信業務人員，在評估放款風險時，以往只能從財務報表與基本書面資料，執行人為判斷，不免有判斷失誤的風險。但透過系統分析，可釐清該公司的企業關係網絡及經營者的人格特質與社會評價，據此做出更準確的授信風險評估。

另外，這次的合作案，是跨校合作的新突破，不但結合不同學校的會計、資管與保險學者，拉長了雙方研究、交流的時間；另一方面，不同於以往台大與大師級人物合作，多侷限在單次及特定議題，此次合作案「是有深度、長期並跨多項議題的合作交流」。

李書行也透露，台大管院正構思成立「企業風險管理與商業情報分析研究中心」，盼能與陳忻鈞位於美國的研究中心，進行雙邊資源共享交流，也期能借助大師的專業，帶領台灣管學理界跨出步伐，創造世界頂尖的研究領域，最後把研究結果從學界延伸到實務界，帶動台灣產業快速成長。

【2008/07/27 經濟日報】@ <http://udn.com/> 