



蕭景燈  
中央研究院資訊科技創新  
研究中心研究技師

## 以結構化的語意資料， 發揮資料應用最大價值

開放資料格式不統一，是目前政府網站的最大問題。如能透過語意上的標註，讓資料可以在不同的應用程式、組織機構與社群之間分享與再利用，才能發揮政府資料創新應用的最大價值。

全球資訊網 (World Wide Web) 的發明人柏納李 (Tim Berners-Lee) 去年12月在雜誌上談到，WWW歷經20年的蓬勃發展，已深入到每個人的生活作息當中；但隨著商業與政治力的介入，資訊網或許還是資訊交流的平台，但「分享」的自由卻逐漸消失了。

根據中央研究院資訊科技創新研究中心研究技師蕭景燈的觀察，「臺灣政府至今仍將網路視為等同於報紙、電視之外的『另一個媒體』，而輕忽了網路所帶來的影響力，往往在不對的地方加以管制。」

蕭景燈以多年前任職於蕃薯藤時遭遇到的阻礙為例，當時中央氣象局網站因為頻寬不足，難以應付颱風、地震時突然增加的網站流量，於是將氣象資料放在蕃薯藤網站上，結果卻遭到部分民意代表的質疑。「這事件反映出行政與立法的步調不一致，以及政府對於開放資料使用的保守態度。」蕭景燈強調。

### 透過語意把網站資料串接在一起

柏納李曾提及，網路的最大特色在於它的通用性原則，不管用什麼硬體、軟體、網路連線或哪國語言，只要運用3個簡單的標準協定：用HTML (超文件標示語言) 格式寫好網頁，用URI (通用資源標識碼) 命名，再用HTTP (超文件傳送協定) 放上網際網路，一個連結就可以連接取得任何資料及內容。



「開放資料格式不統一，或是雖將資料放上網站但限制重重，是目前政府網站的最大問題。」蕭景燈舉例，如果想要評估工業區對環境的影響，就必須取得諸如內政部的地圖資料、環保署的空氣品質監測資料，及中央氣象局的氣象監測資料，這些資料目前雖已可透過網站取得，但使用上仍屬不便。以英國、美國為例，為促進更公開、透明及參與的公共治理環境，成立了專門網站 (data.gov) 將公部門資料以語意網 (Semantic Web) 的格式公開，鼓勵民眾用來創造各種應用與服務。

在數位典藏與學習之學術與社會應用推廣計畫所出版的《Web 3.0 語意網新趨勢》一書中就指出，「產生結構化資料是形成語意網的第一步，再透過標註 (Annotation) 賦予資料屬性與意義，如此就可以達到讓機器直接或間接地自動處理這些資料，最後一步就是將資料發佈到網路上以利他人串接。這些資料可以在不同的應用程式、組織機構與社群之間分享與再利用。」

### 藉助Open Data推動「開放政府」

「語意網主要是技術面的東西，其中包含資料格式與資料發布、查詢的標準，這些標準有RDF、RDFa、URI、SPARQL等；而Open Data談的是如何將這些技術用於實踐資料公開，」蕭景燈進一步指出，像是交通部或主計處已把一些資料公布在網站上，資料量也還算豐富，但缺乏策略性的指導，且因為呈現／提供出來的圖表是固定的，沒辦法做其他運用；如能透過語意把網站資料串接在一起，讓資料可以被運算，就能提供使用者更好的服務。

對民主比較成熟的國家，Open Data更深層的意義是Open Government (開放政府)，「要想讓政府的流程透明、有效率，要促進參與，就必須把政府的門打開，讓更多人可以進來幫你，這恐怕才是推動Open Data需先突破的難關。」蕭景燈強調。

