

「中草藥用藥」之主題地圖式 數位學習教材建構與應用

施如齡

副教授
淡江大學教育科技學系所
E-mail: juling@mail.tku.edu.tw

施百俊

副教授
大仁科技大學資訊管理系

摘要

本研究之主要目的在於運用資訊科技來建構數位學習教材，規劃並建議學習之應用方式。我們在「神農e經：中草藥用藥典範數位學習內容開發及實地實驗研究」中開發了多層次的主題地圖數位教材模組，能夠自動擷取各種不同來源的中草藥用藥學習素材，由課程引擎使用自我組織映射(SOM)演算法，自動編製成多層次的主題地圖，有效的將原始資料組織成具有專業意義的學習材料。主題地圖之教材應用具有幾項特點：呈現豐富多元且專業的資料與其知識架構、提供客制化與互動化的學習機制以支援系統化與建構式學習、經由自動演算系統提供簡單易懂的操作機制。因此，學習者得以藉此獲得最即時與正確的中草藥用藥安全知識。

關鍵詞：數位學習教材，主題地圖，自我組織映射，中草藥用藥

前 言

數位時代的來臨首先為教育領域創造了無限的教學資源，而數位科技的進步亦賦予市場更大的空間來開發電腦輔助教學系統與軟體。以傳統方式編製教材已無法即時因應快速成長的知識內容。因此，電腦與網路除了建立新的學習環境，也成為有利的教學工具。在資訊過剩而資訊品質無法掌握時，資訊數量經常造成教學者與學習者的負荷；如何有效地運用科技，結合知識管理的技術，來協助學習者處理與管理資訊，已成為當前的問題，亦是本研究的最終目的。

2006/10/21收稿; 2006/11/24修訂; 2006/12/27接受

在傳統的教學型態中，教師須於其專業領域中廣泛閱讀相關書籍與資訊，內化之後再進行整理，規劃主題，研擬大綱，藉以指引學生掌握完整的原始資料。其過程之瑣碎與繁複無以言傳，教師在無窮的知識海裡，耗費相當大的時間不斷反覆其歷程。如何能更有效地整合資訊，將之形成可用的知識，實為一大挑戰。

當時代轉換，學習也從教師導向變成學生中心。現在，電腦輔助教學工具能利用視覺化、情境化、多元化的技術來幫助教師快速、有效、正確地進行例行工作，省下許多時間進行更人性化的教育思考，並將課堂時間保留給師生進行人際互動。同時，透過資訊整合，增進學生進行自我導向的學習。數位教材成為補充與擴充教材，提供學生依照個人的學習風格、先備知識、學習需求與學習路徑，進行客制化的學習。為因應更多元化與個人化的學習，數位教材資源不僅須要深度亦要廣度，才能支援眾多學生的不同需求。這項任務若全然由人工處理，則有如夸父追日，難以達成。

面對無數來源與格式不同的資訊，如何編製系統化的專業領域教材或解決漫無目標的學習，是許多教師與學生的困擾。如何將知識系統化、規則化、有序化？如何抽取冗長資料的關鍵主題？如何整合各種不同資料間的一致性、相似性與差異性？又應如何擬出專業領域的知識結構，有效傳達給具有不同個體特性的學習者？

根據 Fisher, Wandersee 與 Moody (2000)，主題地圖包含了一整群的物件，以知識網絡的型態呈現。這個知識網絡是概念的相互連結而完成的，並以骨架的結構呈現。它並非只是一組組的概念而已，更是這些概念組織與連結的型態。因此，借用數位科技長處，運用知識管理技術，主題地圖似乎是整合、組織、導覽大量且持續成長的資訊之有效解答。它能夠協助搜尋、創造、結構、呈現、使用有系統的數位資源，使之清楚而有效率地整合既有資訊，使資訊從其整合性結構來產生新的涵意和理解。

主題地圖大多是在資訊管理與知識管理的領域中被研究與開發，亦多在企業中被使用來了解大量的市場資訊；在此研究中，希望能將此技術結合進教育領域，用來提升教學效率。

本研究深入探討主題地圖技術如何運用原則來組織資訊，以成為可用知識；如何處理不同的資料來產生關鍵主題；如何呈現主題使之易於搜尋與存取；之後，再陳述主題地圖數位教材的建構於教育應用裡的意涵。

本文將從教材建構與教材應用兩部分著眼，交互穿插說明，內文中詳盡描述教材建構歷程，以教育的觀點來看這樣的科技進展對於教學領域所帶來的意涵與潛力。在文獻探討方面，先從知識管理的角度出發，尋找專業學習領域所擁有的不同的本質，再探討數位時代運用知識管理概念於主題地圖的技術發展

與運用。然後，藉由我們進行數位學習國家型科技計畫：「神農e經：中草藥用藥典範數位學習內容開發及實地實驗研究」的經驗，首先陳述建構中草藥用藥的主題地圖的建構流程與方式，再進一步探討藉由此技術所開發出的數位教材在此專業學習領域裡應用的可能與架構。希望藉由此研究，能提供科技與學習整合的另一想法，亦提供其他具有相似特性的學科參酌運用。

二、文獻探討

(一)知識管理與搜尋

1. 知識管理

數位時代裡，智慧型的資訊擷取 (information retrieval) 和知識發現 (knowledge discovery) 的功能，乃學習材料知識管理重要的支援項目，使能更有效的處理、管理、分享並應用為數可觀的數位資料。相關的文件或文字探勘 (text mining) 技術，諸如文件向量法 (Salton, 1989)、文字採掘 (Sullivan, 2001)、文件分群 (Kohonen et al., 2000) 及知識地圖 (Chen et al., 2001) 等，經過系統化的設計與運用，可以把各行業別的專業知識進一步昇華為數位化教材，促使新進研究人員更有效的入門，也更有利於知識的累積與傳承。

從學習的角度著眼，知識類別、特性，為資訊處理與管理的基礎。以知識學習的理論來看，認知心理學家大多認為知識具有其構成的基本架構，而其結構 (structure) 來自於概念的相關性與彼此間的連結。Anderson (1980) 將知識分為陳述性知識 (declarative knowledge) 及程序性知識 (procedural knowledge) 兩種，主要在於抽象性概念與實務性步驟的分別，相當於 Krogh (1998) 及 Zack (1999) 所詮釋的顯性知識 (explicit knowledge) 與隱性知識 (implicit knowledge)，以結構化與半結構化的型態呈現；而 Ausubel (1968) 提到漸進分化 (progressive differentiation) 及統整調和 (integrative reconciliation) 兩種策略來進行意義學習 (meaningful learning)，著重於知識學習，由上而下及由下而上的系統性方法；這些理論隱含了兩種意涵：(1) 知識的存在具有其內在的結構性；(2) 以其結構性的模式來進行擷取與吸收，對於學習是最具效果的。

我們觀察在醫藥資訊領域中，病患除了被動接受醫生的建議，他們更可以積極地找尋相關資訊，主動地獲得解答。然而，這些資料往往散落在不同的地方，例如網站、雜誌或書籍，缺乏有效、專業的組織與整理。於是，本研究期望藉由文字探勘技術，協助學習者在大量的數位資料中發現新知識，例如：輸入「人參」這個關鍵字，可以發現哪些藥物有類似的功效；或從相近關鍵字與群集中，了解不同中藥之間的異同。而這些隱性知識，除了可以幫助研究人員進一步探究主題知識之背後原因與內涵，引發新研究議題，並且也是初學者學習的最好材料。

大仁科技大學在行政院衛生署的「中草藥用藥安全資訊中心」計劃中，成立專責小組，建立「中草藥用藥安全」入口網站，利用專家知識以「人工」的方式篩檢整理出中草藥用藥安全相關資訊，統一建立中心資料庫。就知識管理的觀點來看，這是獨立的知識庫，並不具有進階的知識探勘、搜尋與整理的功能。然而，這大量的文件可以運用文字探勘技術，包括：(1)將知識分類群聚 (clustering, categorization)；(2)將資訊摘要；(3)將大量文章以鳥瞰式呈現，或以不同視覺效果 (visualization) 呈現 (Yang et al., 2003)，提供更多知識探索機會，提升其功能。

從數位學習系統的角度來看，利用這些文字探勘技術，可以將原生性的學習材料或參考資料重新「組合整理」，產生有用的知識，變成有意義的教材，讓學習者能夠在處理大量資訊時，也能夠理解、記憶、分析、綜合、運用，得以遂行教學目標。

舉例來說，原本的搜尋引擎資料庫中具有數萬種中草藥材的相關資訊，分別以獨立的文件形式存在。運用自我組織映射 (Self-Organizing Map, 簡稱 SOM) 技術的處理，可以自動整理出最重要的幾個子題，如：炮製、毒性變化……等。更進一步，系統可以再針對個別主題進行深度展開，形成一幅完整的「中草藥用藥」的主題地圖，看到完整的知識架構。無論教師、專家或其他使用者，均能針對資料主題進行個殊化的探索。這也不同于其他資料庫，由資料建置者來訂定主題。

2. 知識搜尋

目前能夠管理資訊、提供搜尋大量數位資料的，乃屬網際網路上之搜尋引擎與專業知識網站。為突顯主題地圖的概念與技術架構，並比較兩者之差異，首先須理解網站之知識管理的概念與技術。

網際網路並未具有組織高重複性資料的特性，使得資料擷取產生困難，亦有資訊負載的問題 (Bowman et al., 1994)。一個專業關鍵詞如：「中藥」，透過一般目的 (general purpose) 的搜尋引擎 (如：Google)，會搜尋到上千筆的資料。對於任何使用者，透過這些網頁的瀏覽以尋找特定資料，事實上是沉重的負擔；而除了普通網頁外，各專業及學術資料庫的網頁 (deep web)，則可能是普通網頁的 400-550 倍的數量 (BrightPlanet.com, 2000)。資訊超載的情況，使得我們有必要利用資訊技術來取得較高品質的資訊。

建構專業知識網站的技術目前大致分為兩種。在一般目的之搜尋引擎與入口網站方面，網路上許多不同的搜尋引擎在進行索引 (indexing)、排序 (ranking) 及網頁文件的呈現 (visualizing web documents) 上，各有其特性，並運用其獨有的演算法。例如，AltaVista (www.altavista.com)、Google (www.google.com)，以及 PageRank (Arasu et al., 2001) 是以排序清單來呈現所擷取的網頁；

而 Yahoo (www.yahoo.com) 則是以類別的方式將相關的網頁放在同一個群集內，以層級的方式將網頁資訊變成許多子集合。這些搜尋引擎多是使用關鍵字查詢，網際網路搜尋器 (Internet Spiders) 是大部分搜尋引擎後端的主要程式，用來收集網頁並且探索網頁中的對外連結，特別是 The Harvest Information Discovery and Access System (Bowman et al., 1994)，具有強大的搜尋技術與速度。但是相對地，一般搜尋引擎所搜尋的結果通常會過多且不夠精確，使用者難以獲得專門及特定領域的資訊 (施百俊等，2005)。

Selberg & Etzioni (1995) 指出，依賴單一的搜尋引擎，使用者可能會遺漏超過 77% 的重要攸關參考資料。NEC 研究機構的研究也指出，網際網路之搜尋引擎無法趕上網路快速變動與成長，每一個搜尋引擎僅能找到所有網站 16% 的資料 (Lawrence & Giles, 1999)。相對的，主題地圖則能將大量資訊整合，將其以眾多關鍵詞類聚，呈現出完整的知識架構，協助使用者進行明確的搜尋。

(二) 主題地圖的建構程序

1. 文字探勘

陳文華 (2003) 指出文字探勘技術有別於資料探勘技術 (data mining)。資料探勘技術主要針對結構化的資料，如關聯資料庫中定義明確的表格與欄位；文字探勘則針對非結構化或半結構化的文件資料來進行分析，如新聞文件的本文部分，除內容並無固定格式與架構外，其屬性亦難以直接取得或確認。由於使用自然語言的文件包含豐富的描述性資料，因此，進行文字探勘通常須將文件中的文字轉換成適合後續處理的格式，形成結構性資料—包含斷詞 (word segmentation)、關鍵詞篩選與向量空間展示；再者，進行詞類屬性的判定 (如：感冒為疾病，人參為中藥)。

中文方面的研究尤比英文困難 (許中川、陳景揆，2001)，因其斷詞問題未能以文字間之空白來處理。斷詞方法主要可分為三種：字典法或詞庫式斷詞法 (dictionary approach)、語言學法 (linguistic approach)，以及統計法 (statistical approach) (陳文華，2003)。本研究採取詞庫式斷詞，在演算上比較直覺且容易實作。由於斷詞的品質和詞庫中的詞彙多寡有關，因此本研究根據專家事先建立的詞彙庫，來作為後續文件向量的元素值。

2. 學習材料的組織

(1) 文件分類

在特定的專業領域中，專業知識必須靠學科內容專家 (SME: subject matter expert) 利用知識分類學 (ontology)，以有效分析知識，建構其內容組織，並進行呈現及交換等作業，進而組織成學習材料。一般而言，以人工進行文件分類雖然耗時，但精確度較高。相對地，若用科技進行自動分類，雖較省時，但相對精確度也較低。

儘管如此，要能快速地在眾多文件中定義出其所包含的主題，有兩種方法可以進行自動分類。其一，是自動索引演算法 (automatic indexing algorithms)。它被廣泛的運用在處理文字性資料，用來萃取主要概念，透過某些研究結果發現 (e.g. Salton, 1986)。其效果與人工索引相同。其二，是提供分類工具，讓使用者可以將文件分到不同的類別 (Hearst & Pedersen, 1996; Rasmussen, 1992; Zamir & Etzioni, 1999)。

而文件集群分析通常依據個別文件的屬性或文件的相似性將文件加以歸類。所謂文件屬性，可以是詞彙在文件出現的頻率 (Veerasingam & Belkin, 1996)，或以大小、來源、主題、作者等來組織所擷取到的文件；亦有以其他的屬性歸類者，如 Envision (Fox et al., 1993) 以及 GRIDL (Shneiderman et al., 2000)。以文件的相似性將文件加以歸類的方法，包含一些機器學習的演算法；例如 SOM 是用類神經網路演算法 (Kohonen, 1995)，由電腦自動地「學習」文件主題並予以分類，然後將不同類別的文件分到不同的區域，並呈現區域關係的資料地圖。相似的文件會出現在同一區域裡，而文件內容相近的區域會緊臨在一起，形成依據文件相似性程度而呈現遠近關係的主題地圖。許多使用此技術的系統已建立完成 (e.g. Lin et al., 1991; Chen et al., 1996)。

(2) SOM 演算法

基本上，SOM 包含一組處理單元，每個處理單元被賦予一組權重向量；其權重向量維度與輸入向量維度相同。在初始設定訓練模式時，可能隨機或以某種策略 (例如主成份分析) 來給定權重向量之初值。比對輸入與輸出資料後，除最贏向量 (winner vector) 會進行調整外，其附近之向量亦會隨之調整，以便能讓鄰近集群相似；因此，所呈現的分群結果會讓越相近的分群排得越鄰近 (Kohonen et al., 2000; Merkl, 1998)。

參照陳文華等 (2003)、Kohonen 等 (2000) 以及 Merkl (1998) 等相關研究文獻，其演算法之主要參數有學習速率 (learning rate)、鄰近距離 (neighborhood) 與地圖大小 (map size)。學習速率是用來控制權重調整的參數；鄰近距離指的是最贏向量的影響範圍。參數的調整與初值設定主要透過試誤法 (Try and Error) 處理。它亦能夠將高維度的輸入資料轉換成一個有規則的低維度 (一/二/三維) 矩陣方格。

詳細的演算法請參見圖 1。

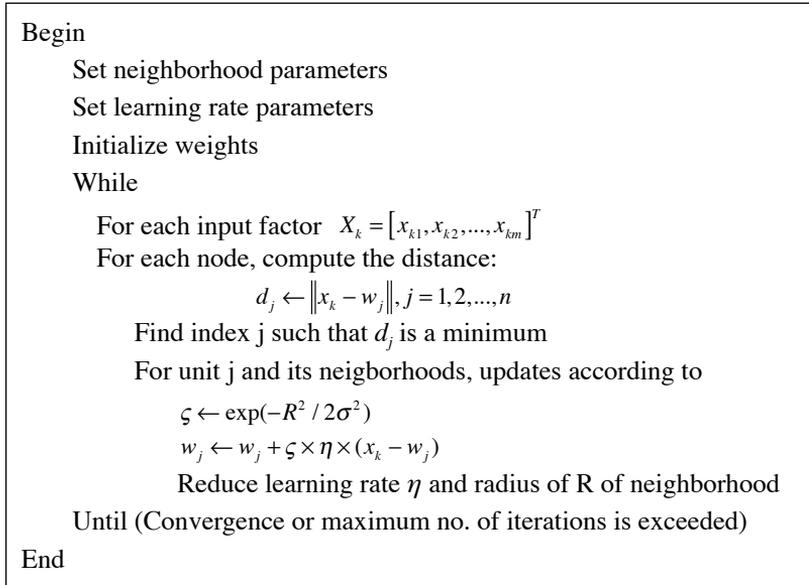


圖1 SOM演算法

資料來源：陳文華等，2003

3. 知識地圖

主題地圖 (topic map) 主要呈現主題 (topic)、關聯 (association) 和連結 (occurrence) 三個項目 (陳文華等，2004)。首先，透過內容分析找出代表各文件之主題，進而將各主題之間或單一主題內的附屬子題，透過「參見」(see) 和「另可參見」(also see) 的機制 (溫達茂、呂春嬌，2001) 加以連結，展示其相互關係，即成關聯。最後，以類似地圖檢索的介面方式表現，提供使用者進行更有效率的瀏覽與檢索。

目前的主題地圖應用研究包括專利地圖 (patent map) (Yoon et al., 2002)、網路瀏覽地圖 (Yang et al., 2003) 等。通常利用單層平面的表單來顯示分群後的主題，提供使用者一窺領域知識的全貌。針對其中的特定主題，則透過超鏈結 (hyperlink) 的方式去索引資源。其主要缺點有三：(1) 形成的地圖基本上是靜態的時間切面資料 (snapshot)，無法動態的更新；(2) 透過單層地圖上的鏈結所索引到的資料，基本上還須透過人工的整理才能維持其正確性和一致性；(3) 即使使用 Meta search 的方式，動態更新和鏈結所取得的資料，基本上還是缺乏整理的，資訊超載的問題依然存在 (施百俊等，2005)。

所以，在本研究中，我們把每層地圖之間的索引，都變成動態的相關查詢，然後重覆的應用 SOM 去整理查詢結果，編製成多層次多維度的主題地圖學習材料，如圖 2 所示。

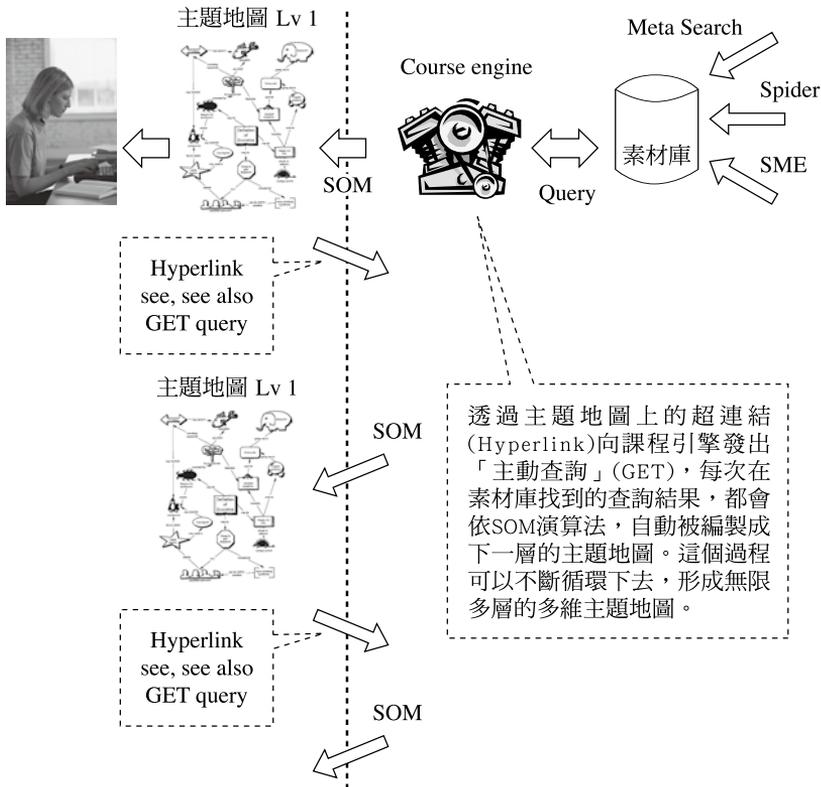


圖2 多層次多維度互動式主題地圖學習材料

三、教材建構與應用

通常所謂的「中草藥」是對中醫傳統藥物的總稱，如果運用得當，可達預防、治療疾病或保健美容的功效，但用之失當則會出現不適的反應，甚或出現對身體的損害反應，即是中草藥中毒（石萍莉，2001）。用藥安全知識的研究、發展及改良，是跨領域、動態的交互學習成長過程。於是，在專業用藥知識的分享與創新上，學習者與教師都必須透過定期的案例分析等途徑來交換經驗和修正處方等。另一方面，面對來自社會不斷要求醫療品質提升，和來自生化科技與醫學工程研究成果導入臨床的壓力，導入數位學習機制是提升用藥安全上必須要兼顧的重要課題。從知識系統產生的主題地圖，將依結構式的學習理論，將大量的數位資訊轉變成為學習材料，所有的內容都將遵循數位學習內容的SCORM規範，以XML的方式編製與呈現，便於學習材料的再利用，也利於學習進度的安排與追蹤。

本研究之目的可分為二：在技術面上，進行文字探勘的工具開發。在教學面上，開發「中草藥用藥」知識領域之教學參考索引。其系統開發流程如圖3。

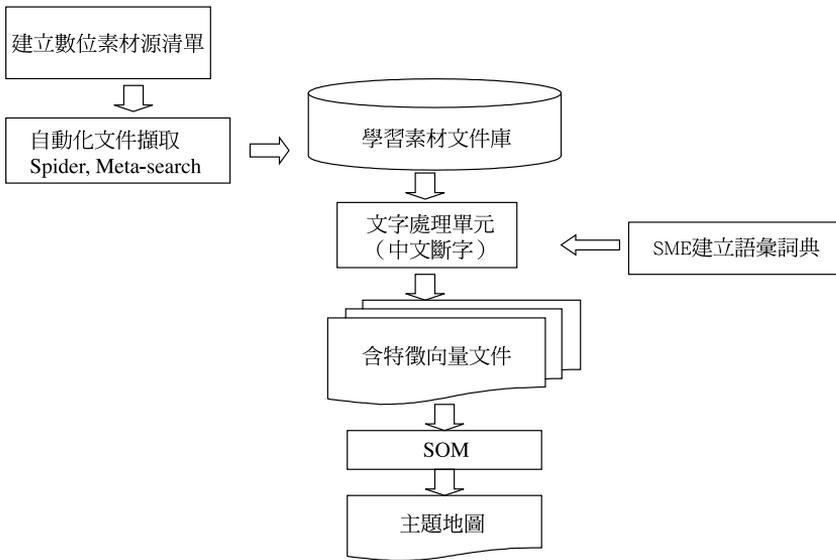


圖3 主題地圖式數位學習教材發展流程

本系統主要由學習素材擷取和文件處理兩個模組整合而成。

(一)學習素材擷取

為確保學習素材的正確性和即時性，在此模組中，利用 meta-search 的智慧型代理人 (smart agent)，多重與同步地在網際網路上的一般性搜尋引擎中，收集有關中草藥用藥安全的資料；另外，我們還製作了 Spider 自動文件獲取程式，邀請行政院中醫藥委員會的專家，羅列出國內外 65 個專業資源網站，定時從這些專業資源網站中，作全自動的文件擷取與更新，獲得最專業、最新的資料，如圖 4 所示。

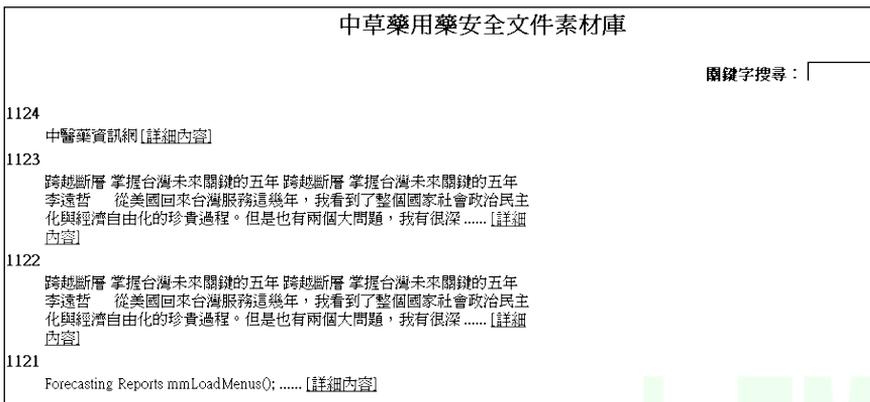


圖4 文件素材庫範例

另外，此模組提供視覺化與摘要功能，讓學習者可以更快地檢視搜尋結果。除了提供具彈性的搜尋方式，也建立一個使用者搜尋概況。學習者可以在資料庫中選擇不同的主題，而代理人則會在學習者從事其他活動時，持續地針對資料庫的資料進行搜尋。此時，學習者即可以針對他們所感興趣的研究，隨時獲得更新後的最新學習材料。

(二)文件處理

1. 特定語彙典與斷詞

各個專業領域都有不同的專業術語。為能斷出好的詞彙，特定應用領域就必須有特定的語彙典。在行政院衛生署的中草藥資訊中心組成了一個專家委員會，責任是審核各種中草藥用藥安全資訊，他們也同時成為本研究第一階段的SME，協助整理專業的語彙庫，作為本研究之文件特徵向量，如圖5所示。

辭彙	
編號	名稱
538	惡寒
537	溫服
3	人參
535	芒硝
35	誤用
38	當歸
39	中醫
44	保管
45	西醫
47	遠志
48	何首烏
49	薑黃
50	決明子
51	冰片
52	延胡索
54	百部
55	地龍
57	白芷
58	石榴皮
59	厚朴

圖5 語彙典範例

本研究根據此語彙典，發展出中文辭典斷詞程式，如Chen & Liu (1992)所建議，以長詞優先擷取的原則，對所有素材文件進行文章斷詞。由於Spider會動態更新，數量並不一定。本期之研究大約網羅十萬份文件。

2. 文件特徵向量

依陳文華(2003)所示之步驟，經過斷詞處理後，有相當多的關鍵詞(數

量級 3-4) 從文章中被萃取出來。接著採用 Salton (1989) 所發展的向量空間展示 (vector space representation)，利用詞彙頻率 (term frequency, $tf_i(d)$) 與文章頻率 (document frequency, df_i) 來進行計算。詞彙頻率是指詞彙 i 在文章 d 中出現的頻率；文章頻率則是指所有文件中有多少文章包含詞彙 i 。通常採用 TF x IDF (Term Frequency- Inverse Document Frequency) 來表示文件向量，如下式：

$$w_i(d) = tf_i(d) * \log\left(\frac{N}{df_i}\right)$$

其中：

$w_i(d)$ 代表詞彙 i 在文件 d 的權值

$tf_i(d)$ 代表詞彙 i 在文件 d 的出現頻率

N 代表文章的總篇數

df_i 代表包含詞彙 i 的文件數

當詞彙出現越多次並出現在較少的文章中 (代表這個詞彙比較特殊)，則獲得較大的權值。此程序用以選出一組最重要的特徵向量來描述每個文件。先將所有文章所斷出的詞彙依據 TF x IDF 值由大到小排序，而且一個詞彙只能被挑選一次，同樣詞彙不重複挑選，選出前 1000 個詞彙，作為特徵向量元素，選出最重要的關鍵詞。因本研究會依不同的詞類加以標記，因此，TF x IDF 的計算還會依使用者需求而給予不同的權重。

最後依據 SME 預先定義的語彙庫，找出每份文件對應到這些語彙上的 TF x IDF 值，以產生每份文件的特徵向量，作為 SOM 的輸入值。

(三)產生主題地圖

我們以 SOM 方法產生中草藥用藥安全的主題地圖，每層依主題的重要性展開二維表單，然後採用 LabelSOM (Rauber & Merkl, 1999) 演算法，依詞彙之代表性高低來排序，並設定每個群集產生最具代表性的 3 個標籤詞彙。

在本系統中，主題地圖分別以文字與圖形模式來展現。在文字模式中，首頁呈現一個簡單易懂的目錄畫面，使用者於點選主畫面的主題 (如圖 6)，也可以再點選各群組的主題，進入第二層次 (如圖 7)，繼續探索各文件子集合的分群狀況。此時使用者可以瀏覽各子群組中每份文件的詳細內容，與使用一般主題式瀏覽的搜尋操作方法雷同。

神農 e 經 - 數位學習國

ht

計畫內容 關於我們 相關連結 神農 e 經

多層次主題地圖 Level 1

98345份文件, 1000個詞彙, 執行時間: 15.863055555556小時

No.679群[1]

1. 馬兜鈴酸[1.92432557366265e-33]
2. 山豆根[3.53477316426122e-27]
3. 林克[2.09732374229433e-10]

No.679群[1]

1. 馬兜鈴酸[1.91140171172634e-33]
2. 山豆根[3.51103413152141e-27]
3. 林克[2.09732374229436e-10]

No.680群[1]

1. 馬兜鈴酸[1.95615971353201e-33]
2. 山豆根[3.59324666381668e-27]
3. 林克[2.09732374229433e-10]

圖6 文字模式主題地圖第一層

神農 e 經 - 數位學習國家型科

http://192.192.21

計畫內容 關於我們 相關連結 神農 e 經

多層次主題地圖 Level 1 展開

98345份文件, 1000個詞彙, 執行時間: 15.863055555556小時

本群內容

1. 中醫藥資訊網一年報期刊 [詳細內容]
2. 中醫藥資訊網一年報期刊 [詳細內容]
3. 中醫藥資訊網一年報期刊 [詳細內容]
4. 未命名文件 [詳細內容]
5. 中醫藥資訊網一年報期刊 [詳細內容]

圖7 主題地圖第一層點選後展開畫面

另外，主題地圖亦能夠以圖形方式來呈現，並依顏色來區分各群集的特性，幫助使用者一目了然所有文件的主題分布狀況。下列圖8及圖9以黑白模式呈現（圖9則不另加顏色與說明註記）。

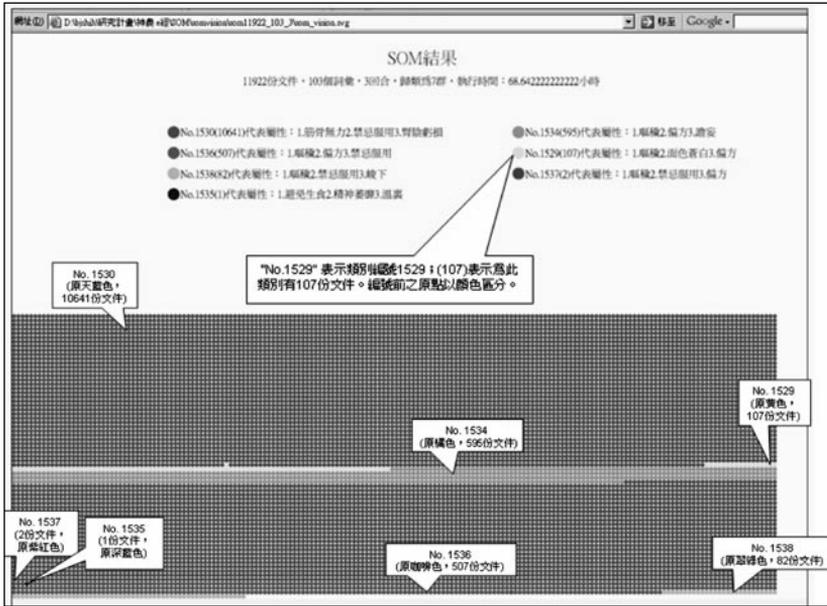


圖8 XML圖形模式主題地圖（11922 x 103）



圖9 XML圖形模式主題地圖（11744 x 84）

圖8中的「11922 x 103」標題代表11922份文件以103個詞彙進行分群的結果，圖9之標記法亦同。在圖中以不同顏色來代表不同的文件群組，如藍色群組的代表主題為「筋骨無力」，次重要的關鍵詞為「禁忌服用」、「腎陰虧損」，即是指在本群集中的文件在「筋骨無力」上的參數值接近。學習者可以一看就知道在文件庫中大部分的文件屬於哪些主題，重要性的分布也很清楚。在圖中的每一個點代表一份文件，學習者可直接點選，進入每一份文件查詢個別文件的內容。

四、結果與討論

本研究的目標在於應用文字探勘技術來產生符合SCORM規範的知識學習材料，最終希望能提供教師結合課程引擎來自動組織出在專業上有效的、完整的學習材料。在此系統開發完成的階段，從中草藥用藥之主題地圖的結果觀察，分群大致上是功能導向的，在中草藥專業上呈現出相當具有意義的學習材料。它具有以下幾個特點。

(一)知識內容與結構

呈現知識架構：對於富含複雜組織系統與內涵的專業知識，知識綱要與導覽有助於使用者掌握其知識要領。它能成為更深入探討的骨幹架構，支援探索。它有助於將隱性知識形式化，使其變為顯性知識，讓使用者更能進行分析處理。

資料豐富多元且專業：主題地圖來自於語意網的概念，最適用於用在語詞豐富的資料。雖然呈現的介面通常簡單易懂，但底下所蘊涵的是大量來自各處、型態相異、內容複雜，又具專業深度的文件。對於專業學習而言，豐富資源是研究的開始與學習的關鍵。

持續更新資訊：主題地圖的教材「目錄」是動態的，會隨著資料來源不斷更新。在學術與教學領域，新資料與源源不斷的資源具有相當的重要性。

(二)學習型態與優點

支援系統化學習：分層式的主題地圖能夠協助使用者將知識內容分類，以標籤與詞彙的形式，加上詞彙之間於眾多文件裡所發生的次數與彼此間的關聯，讓使用者便於進行系統化的瀏覽與查詢。

適用建構式學習：雖然主題地圖的呈現型態趨於形式化與系統化，它亦適用於任務型的學習型態，在建構主義的觀點裡，學習者需要從多元的來源，獨立地進行資料搜尋。依照個人不同的興趣傾向、學習風格、能力程度等進行探索與發現。重要的是，這個引導者並非教師，而是一群專家，透過專業詞彙庫

的篩選，以自動化的系統整合大量的文件，是個資訊導向的學習。因此，在創造與維持知識架構的同時，此教材幫助使用者「發現」新知識，藉由演算後而呈現的主題，使用者能夠看到前所未有的想法與概念，這於建構學習的基本理念不謀而合。

客製化與互動化：此教材目錄與使用者不斷互動，隨著使用者的選擇，從主層次進而產生次一層次的目錄。由於每位使用者的路徑不同，收穫亦不相同。故而主題地圖能夠適度地支援客制化的學習。

(三)操作機制與功能

操作簡單易懂：一般搜尋引擎、網站、甚至資料庫都具有不同的性質、內容與查詢方式，不熟悉的使用者於查詢相關資訊時須耗時耗力。主題地圖的教材僅僅呈現數個主題目錄，簡單方便，容易掌握關鍵字詞。不須耗費太多時間與心力，即可輕易掌握大量的資料內容。即使不具先備知識的學習者，在缺乏基礎學科概念時，能夠快速上手，從系統提供的關鍵詞開始，進行漸進式的查詢與閱讀。

自動演算與呈現：主題地圖提供教師與學習者一個自動化的摘要環境。即使面對大量的文件資料，電腦系統能夠替代人工，有效率的運算並進行人性化與專業化的資料分析，以最簡易的方式呈現給使用者。未來，教師更能運用於教學網站上，隨時自動產生主題式的教材，省下許多資料整理與複製重整的工作。

五、建議

整合大量資料並編製系統化的專業領域教材已然造成許多教師與學生的困擾。本研究發現，運用知識管理技術，主題地圖是整合、組織、導覽大量且持續成長資訊之有效解答。它所呈現的知識網絡，是概念的相互連結，並以骨架的結構呈現。透過主題地圖，我們能有效地將知識系統化、規則化、有序化，抽取冗長資料的關鍵主題，整合各種不同資料間的差異性，擬出專業領域的知識結構，有效傳達給具有不同個體特性的學習者。

與先前的應用研究（例如：陳文華等，2003、2004）所不同的是，本研究的主题地圖是應用於教育領域，而非資訊管理、知識管理或企業管理上，可用來提升教學與學習效率。另外，本研究的語彙典是由SME所斷出，每一個群集的主題名稱都具有相當的代表意義，並能相當程度清楚地描述該主題群集，如：「毒性」、「採集」……等。

從學習教材組織的觀點來看，到目前為止，系統尚能夠在最後的群聚結果上產生學科專家在專業上可以接受的結果。就開發經驗上，本系統與先前的研

究一樣，必須以試誤法反覆調整各種參數 (learning rate, initial value...)，並沒有一定的規則。未來我們會再加上與專家知識庫交互比對參照，然後再予以微調的功能，經過經驗的累積與系統的訓練，希望能呈現更為完整而可靠的資訊，更精確的表現出知識主題。

最後，本研究在知識管理、資訊科技、教學研究三大理論領域上，創造出新的整合研究方向，展現數位學習研究進入跨領域整合的契機。冀望未來的研究能再以不同的知識領域與內容，繼續探究SOM系統的適用性與完整性，持續推動科技與教學的整合。

誌 謝

本研究為國科會數位學習國家型科技計畫：「神農e經：中草藥用藥典範數位學習內容開發及實地實驗研究」(NSC-93-2524-S-127-001)之成果；並對大仁科技大學「數位內容與學習實驗室」研究同仁之辛勞一併誌謝。

參考文獻

- 石萍莉編(2001)。中草藥中毒急救小百科。台北市：協合文化。
- 施百俊、施如齡、陳瑞龍(2005)。應用多層次主題地圖於數位教材開發：以「中草藥用藥」為例。在陸軍軍官學校主辦，陸軍軍官學校八十一週年校慶暨抗戰勝利六十週年紀念學術研討會，高雄市。
- 許中川、陳景揆(2001)。探勘中文新聞文件。資訊管理學報，7(2)，103-122。
- 陳文華、施人英、吳壽山(2004)。探討文字採掘技術在管理者知識地圖之應用。中山管理評論，12(6)，35-64。
- 陳文華、徐聖訓、施人英、吳壽山(2003)。應用主題地圖於知識整理。圖書資訊學刊，1(1)，37-58。
- 溫達茂、呂春嬌(2001)。知識組織的工具-Topic Maps。在世新大學編，21世紀資訊科學與技術的展望學術研討會論文集(頁389)。台北市：世新大學。
- Anderson, J. R. (1980). *Cognitive psychology and its implications*. San Francisco: W. H. Freeman and Company.
- Arasu, A., Cho, J., Garcia-Molina, H., Paepcke, A., & Raghavan, S. (2001). Searching the web. *ACM Transactions on Internet Technology*, 1(1), 2-43.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Bowman, C. M., Danzig, P. B., Manber, U., & Schwartz, F. (1994). Scalable internet resource discovery: Research problems and approaches. *Communications of the ACM*, 37(8), 98-107.
- Chen, C., Kuljis, J., & Paul, R. J. (2001). Visualizing latent domain knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 31(4), 518-529.
- Chen, H., Schatz, B. R., Ng, T. D., Martinez, J. P., Kirchoff, A. J., & Lin, C. (1996). A paral-

- lel computing approach to creating engineering concept spaces for semantic retrieval: The Illinois digital library initiative project. *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Section on Digital Libraries: Representation and Retrieval*, 18(8), 771-782.
- Chen, K. J., & Liu, S. H. (1992). Word identification for Mandarin Chinese sentences. *Proceedings of the 15th International Conference on Computational Linguistics*. Nantes, France.
- Fisher, K. M., Wandersee, J. H., & Moody, D. (2000). *Mapping biology knowledge*. Boston: Kluwer Academic.
- Fox, E., Hix, D. L., Nowell, T., Brueni, D. J., & Rao, D. (1993). Users, user interfaces, and objects: Envision, a digital library. *Journal of the American Society for Information Science*, 44(8), 480-491.
- Hearst, M., & Pedersen, J. O. (1996). Reexamining the cluster hypothesis: Scatter/gather on retrieval results. *Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Zurich, Switzerland.
- Kohonen, T. (1995). *Self-organizing maps*. Berlin: Springer-Verlag.
- Kohonen, T., Kaski, S., Lagus, K., Salojävi, J., Honkela, J., Paatero, V., & Saarela, A. (2000). Self-organization of a massive document collection. *IEEE Trans. Neural Networks*, 11(3), 574-585.
- Krogh, G. V. (1998). Care in knowledge creation. *California Management Review*, 40(3), 133-153.
- Lawrence, S., & Giles C. L. (1999). Accessibility of information on the web. *Nature*, 400, 107-109.
- Lin, X., Soergel, D., & Marchionini, G. (1991). A self-organizing semantic map for information retrieval. *Proceeding of 14th the ACM SIGIR Conference Research and Development in Information Retrieval*. Chicago, Illinois, USA.
- Merkel, D. (1998). Text classification with self-organizing maps: Some lessons learned. *Neuro-computing*, 21(1-3), 61-77.
- Rasmussen, E. (1992). Clustering algorithms. In *Information retrieval data structures and algorithms* (W. B. Frakes & R. Baeza-Yates, Eds.). New Jersey: Prentice Hall.
- Rauber, A., & Merkl, D. (1999). Automatic labeling of self-organizing maps: Making a treasure-map reveal its secrets. *Proceedings of the 3. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'99)*. Beijing, China.
- Salton, G. (1986). Another look at automatic text-retrieval systems. *Communications of the ACM*, 29(7), 648-656.
- Salton, G. (1989). *Automatic text processing: The transformation, analysis and retrieval of information by computer*. Reading, MA: Addison-Wesley.
- Selberg, E., & Etzioni, O. (1995). Multi-service search and comparison using the MetaCrawler. *Proceedings of the 4th World Wide Web Conference*. Boston, Massachusetts, USA.
- Shneiderman, B., Feldman D., Rose A., & Grau X. F. (2000). Visualizing digital library search results with categorical and hierarchical axes. *Proceedings of 5th ACM Conference on ACM 2000 Digital Libraries*. San Antonio, Texas, USA.
- Sullivan, D. (2001). *Document warehousing and text mining*. New York: Wiley.

- Veerasamy, A., & Belkin, N. J. (1996). Evaluation of a tool for visualization of information retrieval results. *Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Zurich, Switzerland.
- Yang, C. C., Chen, H., & Hong, K. (2003). Visualization of large category map for internet browsing. *Decision Support Systems, Special Issue on Web Retrieval and Mining*, 35(1), 89-102.
- Yoon, B. U., Yoon, C. B., & Park, Y. T. (2002). On the development and application of a self-organizing feature map-based patent map. *R&D Management*, 32(4), 291-300.
- Zack, M. H. (1999). Managing codified knowledge. *Sloan Management Review*, 40(4), 45-58.
- Zamir, O., & Etzioni, O. (1999). Grouper: A dynamic clustering interface to web search results. *Computer Networks*, 31(11-16), 1361-1374.

The Construction and Application of Chinese Herb Medication Digital Learning Material Using Multi-level Topic Map

Ju-Ling Shih

Associate Professor
Dept. of Educational Technology, Tamkang University
Taipei, Taiwan, R.O.C.
E-mail: juling@mail.tku.edu.tw

Bai-Jiun Shih

Associate Professor
Dept. of MIS, Tajen University
Ping-tung, Taiwan, R.O.C.

Abstract

This research aims to use information technology to construct digital learning material, and discusses its design and application possibilities. This paper presents the process and result of our research related to the project "Shannon Sutra: Digital Learning Content and Experimental Research on Chinese Herb Medication Safety" granted by National Science Foundation. In this research, we use multi-level topic map digital learning material model to render tremendous original assets, and then embed text-mining techniques in course engine to produce topical learning materials. The autonomous method can process raw data and original document to generate meaningful search results, and thus develop to be well-structured learning material. Self-Organizing Map system helps to demonstrate rich and professional information and knowledge structure; it provides customized and interactive learning dynamic to support both progressive and constructive learning styles; it has autonomous system to afford easy and convenient operation for information search. With topic map system, learners can interact with the dynamic topic map and receive instantaneous and accurate feedback of most wanted result.

Keywords: Digital learning material; Topic map; Self-organizing map; Chinese herb medication

JoEMLS

<http://research.dils.tku.edu.tw/joemls/>