

應用主題地圖於知識管理系統

林光龍

歐陽彥正

台灣大學資訊工程研究所

台灣大學資訊工程研究所

lkl@mars.csie.ntu.edu.tw

yjoyang@csie.ntu.edu.tw

摘要

本論文討論如何將知識管理的概念導入數位圖書館/博物館的建構中。綜觀數位圖書館/博物館的發展歷程，早期的研究主要是著重於數位典藏之基礎建設，包括有典藏內容數位化、資料倉儲、使用者介面及檢索功能等各種技術性問題的研究。隨著資料量的與日俱增，造成資料分散各處、資料內容不一致、以及檢索結果呈現高回收率低精確率等現象，致使用者未能有效地擷取到真正所需的資訊。所以，為了能有效管理資料內容，以及提供更精確的檢索結果，第二個世代的數位圖書館/博物館則著重於後設資料（metadata）、索引典（thesaurus）等規範的研究，促使資料庫本身的量與質皆可同時提昇。雖然在後設資料中包含有完整資料特徵的描述與詮釋，進而幫助人類閱讀資料；但卻無法讓電腦發揮有效組織與快速運算的能力，以改善人類瀏覽知識與學習知識的環境。也就是說，在後設資料中儘管涵蓋有人、事、時、地、物等各層面之資料特性的描述與詮釋，我們仍然無法得知他們本身在語意層次上所代表的意義，以及他們彼此之間的關聯與事件等訊息。因此，為了讓電腦能幫助人類學習知識，在下一個世代的數位圖書館/博物館的建構中，引進知識

管理的機制與功能，便成為必然的研究發展方向之一。本論文將依據主題地圖（Topic Map）理論為基礎架構，採用關聯式資料庫管理系統為實作方法，實作一個以全球資訊網為平台的數位圖書館/博物館之知識管理系統雛形。從某種程度上來說，主題地圖主要是用來組織資訊，以期能有最佳化的知識導覽介面。目前主題地圖規範書（ISO/IEC 13250）已通過國際標準組織的採證，在規範書中明白地指出任何事物，不論是具體存在的個體，或是抽象化的概念，皆統稱為主題（topic）。除此之外，主題地圖的文件格式定義（XML Topic Maps 1.0）也已經公佈於網路之上，我們也將依據此份文件格式定義，來制定我們的知識庫結構（schema）。在這個知識庫系統中，除了紀錄與描述每一主題本身所具有的資訊資源外，更將透過主題地圖來建立所有主題之間的語意網路，以及整個知識庫的本體論（ontologies）。經由此一知識庫系統的建置，我們希望在電腦的輔助之下，整理出更多隱含性的知識，並且依據使用者個人對於資訊的需求和知識的瞭解，自動化、動態化地提供知識領域內由淺至深，由粗糙至精細的形式化概念（formal concept），以改善知識管理介面，幫助使用者導覽與認識知

識領域內的知識。

關鍵詞：主題地圖、知識管理、數位圖書館、本體論、Topic Maps、Ontology

一、前言

隨著網際網路不斷地蓬勃發展，全球資訊網 (World Wide Web) 中的網頁數量，每天都以成千上萬頁的速度快速地成長。為了幫助使用者，能快速的在這巨大的資料庫中檢索資料，「網路資源搜尋器」也就是搜尋引擎紛紛的出現[12]。

搜尋引擎依據使用者所下的查詢關鍵詞，進行檢索之後，列出網頁中有出現該關鍵詞的所有網頁連結清單。使用者可依據此清單，進一步地逐一檢視每一筆網頁內容，以確認此網頁是否真的是他所想要的資訊資源。這樣的動作實際上可視為一種篩選的過程。因為其主要過程是透過搜尋引擎對網頁資料庫進行初步的篩選後，再交由使用者自行篩選出真正的答案。雖然這樣可以相當有效率地節省使用者搜尋資料的時間；但面對查詢的結果裡，種類龐雜的網頁資源，或在網頁內容上，彼此互不相關的網頁。使用者真想從中獲得他想要的資訊，或學習到什麼新資訊，透過祇具有篩選功能的搜尋引擎是很難辦到的。舉例來說，當我們透過搜尋引擎，檢索到幾萬筆有出現「莫札特」這個關鍵詞的網頁後（圖一），在它們當中有莫札特「生平事蹟」、「音樂創作」、「評論」等不同主題的網頁。



圖一：以「莫札特」為關鍵詞所檢索到的網頁資源。

所以，如何組織這些文字、圖片、聲音、視訊、動畫等多媒體網頁資源，及如何組織網頁之間彼此的關聯，以建構一個最便於使用者利用資訊、瀏覽資訊的環境，將成為知識產業裡大家關心的焦點。1999年12月通過國際標準組織認證的 ISO/IEC 13250 Topic Maps (主題地圖) 標準規範[1]，是一個國際性標準的標誌，用來定義有關資訊資源的主題，以及主題彼此間存在的關聯，其主要動機是提供一個組織、管理和最佳化導覽資訊的資訊技術[10]。

本文首先界定知識的定義和種類，接著介紹知識管理的概念，及相關背景知識與理論。在第四節裡，將簡介我們在實作主題地圖管理系統時，所採用的系統架構與實作方法，之後試著探討此系統架構對知識管理的影響，最後針對本論文的研究過程做一個結論，並提出未來研究的方向。

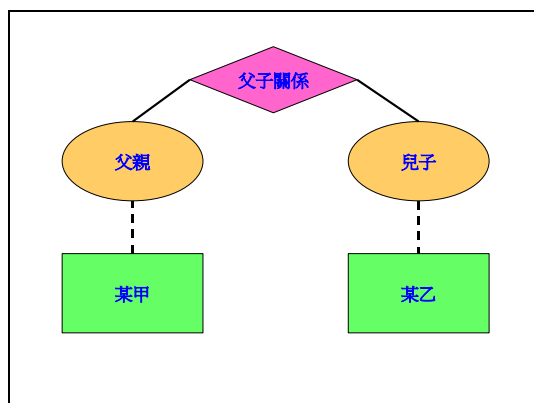
二、知識的定義和種類

資訊是創造知識所必要的元素 [15]，然而「資訊」究竟是什麼呢？最常被使用的意思則是「經過處理後有意義的資料」。所以「資料」經過「處

理」就會變成「資訊」。「資料處理」是將資料予以意義化的過程，這個過程依處理的層次而言，有幾個意思。包括將收集來的資料，按照一定的順序排列，之後編成某種方便利用的目錄或類別，這些目錄或類別就是資訊。另一方面，將收集到的許多資料加以研讀、了解，就成了資訊。譬如我們有一些歷史年表，如果能夠瞭解年代和事件之間的關係，使年代和事件之間產生某種意義，就能使這些歷史事實變成資訊[15]。

知識管理是一種持續改善的過程，包括知識是如何獲取、儲存、使用或創造的過程[7][14]。我們將知識分為顯性知識（express knowledge）及隱性知識（tacit knowledge）兩種[11][13]。

其中顯性知識是指將已知的事實（facts）、關聯（associations）、規則（rules）等資訊，經過組織之後，儲存於知識庫裡的知識。譬如：某甲是父親。某乙是兒子。父子關係是由父親與兒子構成（圖二）。



圖二：從顯性知識推論隱性知識。虛線表示「is-a」。實線用來連結關係裡的成員角色。

隱性知識則是從使用知識庫的過程，經由許多顯性知識的計算或推論所創造出來的新知識。譬如從上述的

知識庫裡，我們可得到某甲與某乙可能具有父子關係的新知識。

三、背景知識

主題地圖就好比是書後面所附的關鍵詞索引[1][9]。一本好書除了內容要豐富之外，更重要的是它必須要有一份很完備的關鍵詞索引。在一份完備的關鍵詞索引裡，必須很完整地整理、組織所有出現在這本書內的重要關鍵詞，以及這些關鍵詞之間的資訊。而這些資訊應該包括有，該關鍵字出現的頁數；以及與該關鍵字相關聯的同義詞或反義詞等。有了這份完備的關鍵詞索引，讀者即可利用這份有如地圖般的指南，間接地反查他想要閱讀的內容。

書本裡的關鍵詞就是構成關鍵詞索引的資訊資源。因此，主題地圖（Topic Maps）可說是一種用來有效組織與管理大量資訊資源的機制。也就是說，主題地圖是一種用來管理知識庫的資訊科技。採用主題地圖做為知識管理的目標是「建立一個個人化與最佳化的知識導覽介面，並提供使用者一個能快速掌握與輕鬆學習知識的環境」。

主題地圖裡所儲存的資訊資源主要可分成三種類別，分別是主題（topics）、關聯（associations）、事件（occurrences）。

「主題」是指使用者感興趣之知識領域裡，各類實際存在的人、事、時、地、物等議題（subjects），譬如：「玄奘」、「瑜伽師地論」、「天竺」等。也可以是存在這知識領域裡，概念上或形式上的本體論（ontology），譬如：「僧侶」、「佛經」、「國家」等。這些

有意義的議題或本體論統稱為「主題」(topics)，他們構成了主題地圖裡一個點 (node)。

「事件」指的是將主題連接到相關資訊資源的所在之處。這些資訊資源指的是可將主題具體化的屬性 (property)，譬如：玄奘法師的「誕生日」、「圓寂日」；或文字、圖片、聲音、視訊、動畫等，可經由 URI 識別的多媒體網路資源，譬如：有關玄奘法師之生平事蹟的「網頁」。事件的主要功能是提供使用者可從多種面向 (facet) 來認識主題。實際上，事件本質上也是主題。也就是說，「誕生日」、「圓寂日」、「網頁」等事件也皆可當作是主題。

「關聯」是指存在於各個主題之間的關係 (relation)，以及構成此關係的應有的成員角色 (role)。例如圖二所示的父子關係，構成父子關係的兩個成員角色分別是父親與兒子。關聯的主要功能是將整個主題地圖組織成像語意網路 (semantic network) 般地，具有推論能力的知識網路[4][5]。如此，使用者將可從中獲得更多隱性知識。

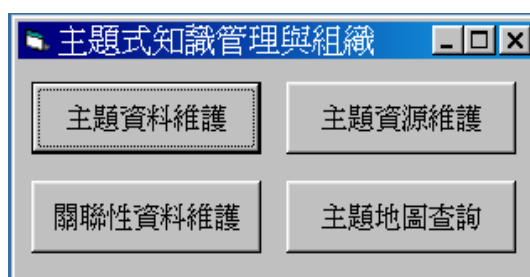
總而言之，「主題」、「事件」及「關聯」就像是三度空間裡的「點」、「線」、「面」，都是用來將一特定知識領域具體化。也就是說，主題地圖是一種知識呈現 (knowledge representation) 的機制，抽象的領域知識經由主題地圖的描述與組織後，將形成一個有如地理空間的地圖。每一位使用者皆可依照其個人對該知識領域的認識與了解程度，從他所熟悉的或有興趣的主題方向出發，在主題地圖的輔助導引之下，將可快速地、清楚地掌握整個龐

大且複雜的知識庫。

四、系統架構與實作方法

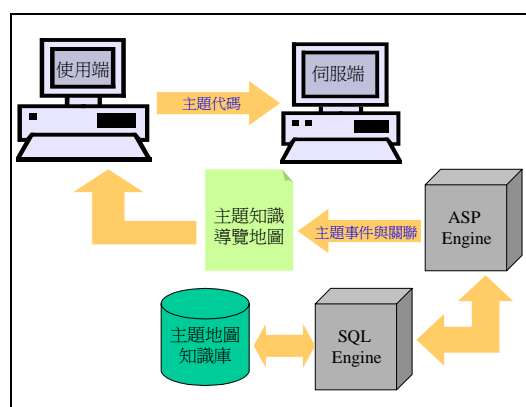
在實作的過程裡，我們將整個知識管理系統分成兩個部份，分別是知識管理與知識利用兩部份。

在知識管理部份，我們設計一視窗介面的管理系統，管理者可透過此介面來維護知識庫 (圖三)。



圖三：知識管理的視窗介面。

在知識利用部份，主要是設計一建構於全球資訊網之上的知識導覽介面[6]。位於客戶端的使用者，可以動態的向伺服端的 Web 伺服器提出瀏覽主題知識的請求，再由 Web 伺服器啟動 ASP (ActiveX Server Page) 引擎，將該主題的所有事件與關聯動態地提供給使用者 (圖四)。

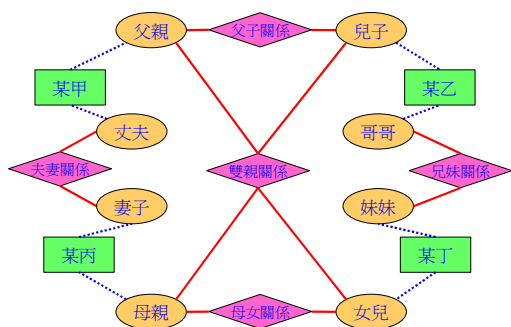


圖四：知識利用的服務流程。

我們依據 ISO/IEC 13250 Topic Maps 及 XML Topic Maps (XTM) 1.0 這兩份標準規格的建議[1][8]，並採用關聯式資料庫模型，制定出主題地圖的

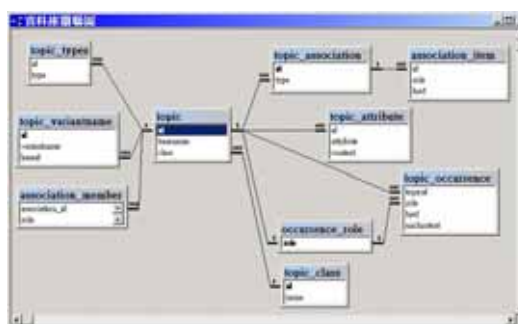
知識庫綱要 (schema) 如圖五所示。

在分析階段裡，我們將龐雜的資訊資源，簡化成一個一個獨立的主題來討論。在討論的過程裡，我們關心有哪些資訊資源 - 也就是主題 - 是重要的。也關心主題之間會有哪幾種關聯性存在，而它們彼此之間分別扮演哪一類的角色。我們使用一個類似分析關聯式資料庫的實體關係模式 (Entity-Relation Model)，將所有主題組織起來，並將此模型具體的透過實體關係圖 (Entity-Relation Diagram) 來呈現 (圖五)



圖五：使用實體關係圖呈現以家庭成員為知識領域的主題地圖。

將主題地圖具體的以實體關係圖呈現出來之後，我們設計出一專門用於儲存主題地圖的知識庫綱要。這份綱要有十份資料表 (table) 構成，它們彼此之間存在有許多必要的關聯性及參考完整性 (圖六)。我們主題地圖裡的顯性知識，也就是已知的主題、事件與關聯分別儲存至主題資料表、事件資料表及關聯資料表。



圖六：主題地圖的知識庫綱要。

因此，整個知識庫裡存放的內容主要有構成主題地圖的所有主題，每一主題的事件，以及主題與主題之間所存在的關聯。這些記錄共同構成一特定知識領域的知識庫。

在建立一個新的主題之時，系統會自動對此主題賦予一編號，以方便於整個主題地圖裡，作為唯一性的識別代號。由於事件與關聯本質上也都是一種主題，再加上主題通常可分成實際存在的個體和抽象化的類別。所以為了方便管理主題，我們大致上將所有主題區分成四種類型來管理，分別為「個體」、「類別」、「關係」及「屬性」。對於主題的別名問題，我們也設計有一別名資料表專門用來存放主題的別名。

關聯式計算 (Relational calculus) 是一種非程序式的查詢語言，比較接近人類所使用的自然語言。使用者不須明白地指出運算的順序，只須說明想要查詢的是「什麼」即可。我們以主題資料表為主，其他資料表諸如事件資料表與關聯資料表等為輔，進行關聯式計算，將可從不同的主題觀點來認識與學習主題地圖裡所儲存的知識。

採用主題地圖做為知識管理的目標之一是建立最佳化的知識導覽介面，並提供使用者一個認識主題知識

的環境。所以我們將知識導覽介面建構在全球資訊網的平台上，由 ASP 引擎動態的依據使用者提出的主題導覽需求，產生一有意義的關聯式計算式，傳送給 SQL 查詢引擎進行查詢。ASP 引擎在獲得查詢結果之後，將依據主題類型整理組織所獲得的查詢結果，提供使用者有關此主題的相關個體、相關類別、相關屬性、相關事件與相關關聯等資訊，如圖七與圖八所示。

五、結論

二十一世紀是知識經濟的時代，知識是個人，也是企業競爭力的關鍵要素，更是企業未來的價值所在。如何將「知識」這項資訊資源有效地管理，成了當前個人與企業界最重要的課題。相同地，引進知識管理的機制與功能於數位圖書館/博物館的建構，將會是數位圖書館/博物館繼續成長的重要關鍵之一。

在知識管理的過程中，第一個問題是「你要如何來組織與儲存在知識領域下的所有資訊資源」。第二個問題是「你要如何呈現知識庫裡的知識」。第三個問題是「你要如何查詢知識庫裡的知識」。

我們應用主題地圖、關聯式資料庫與關聯式計算的技術解決上述幾項問題。由於關聯式資料庫與關聯式計算這兩種技術都是目前市面上被廣泛使用的資料庫技術，所以在理論上與實作上都是相當地成熟。而主題地圖也已通過國際標準組織的認證，並且提出了擴充標誌語言（XML，Extensible Markup Language）版本的文件型態定義（DTD，Documents Type Definition），

他屬於一種內容導向之資料交換模型 [2][8]，將使不同的主題地圖，彼此間在資料交換與合併的過程更加地便利，進而擴大知識領域的範疇。

對使用者而言，我們提供他一個可從不同主題的角度，來認識一個知識領域內的所有知識。如此，便可使知識學習的過程變得更生動活潑。但是，目前的知識管理系統裡，對於相同的主題並無法配合使用者對知識瞭解的程度，提供較好的個人化知識管理解決方案。所以，如果能在知識導覽與學習的過程中加上個人化的機制，將可落實個人化知識管理，並提昇個人層面的知識管理應用。除此之外，我們未來也將嘗試利用 SQL 查詢引擎的輔助，來進行隱性知識的推論，以獲得更多隱含於主題地圖裡的知識。甚至是形式化的概念知識，以進一步地改善知識管理介面，以提供使用者一個更有效率的知識導航系統。



圖七：以「父子關係」為主題的主題地圖。



圖八：以「王羲之」為主題的主題地圖。

六、參考文獻

1. Michel Biezunski, Martin Bryan, Steven R. Newcomb, editors. "ISO/IEC 13250:2000 Topic Maps: Information Technology -- Document Description and Markup Languages", 3 Dec 1999. See <http://www.y12.doe.gov/sgml/sc34/document/0129.pdf>
2. Tim Bray, Jean Paoli, and C.M. Sperberg-McQueen, Eve Maler, editors. "Extensible Markup Language (XML) 1.0 (Second Edition)", 6 October 2000. See <http://www.w3.org/TR/REC-xml>
3. B. Le Grand, M. Soto. "XML Topic Maps and Semantic Web Mining", Semantic Web Mining Workshop at ECML/PKDD-2001, September 3, 2001, Freiburg, Germany. See <http://semwebmine2001.aifb.uni-karlsruhe.de/online/semwebmine08.pdf>
4. CMU World Wide Knowledge Base (Web->KB) project. See <http://www-2.cs.cmu.edu/afs/cs.cmu.edu/project/theo-11/www/wwkb/>
5. Eric Freese. "Using Topic Maps for the representation, management and discovery of knowledge". See <http://www.isogen.com>
6. Ontopia: The Topic Map company. See <http://www.ontopia.net/>
7. Erik Jul, "Four Pillars of Knowledge Managemet", Knowledge Management--Methods and System Conference, Taipei, 3 Dec 2000. See http://ceiba.cc.ntu.edu.tw/km/km_conference/session2.htm
8. Steve Pepper, Graham Moore. "XML Topic Maps (XTM) 1.0", TopicMaps.Org, 6 Oct 2001, See <http://www.topicmaps.org/xtm/1.0/>
9. 溫達茂,「主題地圖 (ISO-13250 Topic Map) 的應用」, 國家圖書館館訊, 總號第 91 期 (民 91 年 2 月)。 http://www.ncl.edu.tw/pub/c_news/91/03.html
10. 羅鳳珠,「以XML建置詩的內容標誌在詩學研究上的發展性與侷限性--以蘇軾前期詩為例」, 東吳大學中國文學系, 文獻與資訊學術研討會論文集 (民 90 年 6 月), 頁 339-369。
11. 嚴鼎忠,「新世紀圖書館的文學知識管理--兼論『當代文學史料影像全文系統』」。 中國圖書館學會會報, 第 65 期 (民 89 年 12 月), 頁 89-97。 <http://lac.ncl.edu.tw/會報/lac65/007.doc>
12. 吳育璋,「從YAHOO到Google, 論下一代的搜尋引擎」, 數位觀察者, 第 89 期 (民 90 年 9 月 19 日)。 <http://www.digitalobserver.com/81-90/89/karlon.htm>
13. 邱子恆,「知識管理及其對圖書館的影響」, 中國圖書館學會會報, 第 65 期, (民 89 年 12 月), 頁 99-109。 <http://lac.ncl.edu.tw/會報/lac65/008.doc>
14. 洪筱盈、呂淑惠、胡宇光、范中威、劉曉蓓,「出席「知識管理：方法與系統研討會」報導之二：臺灣地區」, 全國新書資訊月刊, 民 89 年 12 月號, 頁 9-16。 <http://isrc.ncl.edu.tw/isbn/frmContent.asp?Id=266&callerpage=/isbn/frmIndex.asp&Year=89&Month=12>
15. 吳美美、楊曉雯,「圖書館的利用_高中高職篇」, 圖書館與資訊素養叢書Ⅲ, (民 88 年), 頁 16-17。 http://www.ncl.edu.tw/cl_ebook/doc/list_3_4.htm