

# 主題地圖及其在索引典之應用

林信成

淡江大學資訊與圖書館學系副教授

E-mail: sclin@mail.tku.edu.tw

歐陽慧

淡江大學資訊與圖書館學系研究生

E-mail: 691070014@s91.tku.edu.tw

歐陽崇榮

淡江大學資訊與圖書館學系助理教授

E-mail: cjouyang@mail.tku.edu.tw

## 摘要

本文首先從主題地圖之發展切入，介紹相關標準及其內涵；接著簡略介紹知識組織、書後索引、術語彙編與索引典等相關概念；然後提出以主題地圖建構索引典之語意網路模型、方法與步驟，實際從教育論文線上資料庫的中文教育類詞庫中，選擇十九個主題詞及其替代詞作為實驗用素材，歸納出其詞間關係並建構語意網路圖；最後，將此語意網路依據國際標準組織的主題地圖規範實作，並在伺服器端建置主題地圖處理器，成功的完成一個實驗性質的線上索引典。

關鍵詞：主題地圖、索引典、語意網路、知識組織、人工智慧

## 一、前言

在資訊氾濫的今日，無序的資訊資源所造成的不便反而比資訊的缺乏來得更嚴重。由於資源媒體愈趨多元化，使得知識組織（Knowledge Organization，簡稱 KO）成為一個極為重要的研究領域。如何將資訊或知識經過適當的分析、拆解、封裝與組織，是使用者能否快速、正確的找尋到資訊，進而選擇、利用、吸收，內化成自己的知識，以滿足資訊需求的關鍵點。

圖書館一向對於知識組織的研究不遺餘力，在圖書館學中，索引（Indexes）術語彙編（Glossaries）與索引典（Thesaurus）是建置圖書文獻及其他資源知識架構的方法之一。而在人工智慧領域中，也有一種知識表示法（Knowledge Presentation，簡稱 KR）的技術稱為語意網路（Semantic Network），是由概念（Concept）與概念關係（Conceptual Relations）組成概念圖（Conceptual Graphs），用以作為人與機器的溝通管道。語意網路可有效表示人類專家的知識，在人工智慧應用中常被用來建構專家系統（Expert System）中的知識庫（Knowledge Base），以支援推論引擎（Inference Engine）的推理與運作。<sup>1</sup>主題地圖（Topic Maps）則是一種類似語意網路的知識表示模式，結合了傳統索引、圖書館學與人工智慧等領域的優點，在資訊世界中，有如 GPS 般提供定址與連結的功能<sup>2</sup>，可以有效

<sup>1</sup> Giarratano and Riley, *Expert Systems: Principle and Programming*, PWS-KENT, 1989.

<sup>2</sup> Pepper, S. (n.d.). "The TAO of topic maps: Finding the way in the age of infoglut." *XML Europe*

的組織知識以利於探索、推理，解決大量無序資訊所帶來的問題。在傳統紙本出版的世界裡，有許多方法可用來組織或指示書本或文獻中的資訊，如書後索引讓讀者可以直接找到他們所感興趣的部分；主題地圖可視為紙本索引的線上版，也十分適用於管理資訊連結，如術語彙編、交互參照、索引典以及目錄等。

由於主題地圖之知識表達是結合語意網路的基本模型與書後索引，因此，本研究嘗試以主題地圖來表現索引典的詞彙結構，利用主題地圖提供使用者自行定義結構的方式，訂定各主題詞的詞間關係，提供使用者以瀏覽的方式來瞭解語彙間的關連性。

本文首先從主題地圖之發展切入，介紹相關標準及其內涵；接著簡略介紹知識組織、書後索引、術語彙編與索引典等相關概念；然後提出以主題地圖建構索引典之模型、方法與步驟；再者，實際以「教育論文線上資料庫」(Educational Documents Online, 簡稱 EdD-Online) 中的中文教育類詞庫，選擇十九個主題詞為例，利用其詞間關係建構語意網路圖，並據以構成主題地圖；最後，我們在 Server 端以 Ontopia 公司的 Omnigator 軟體<sup>3</sup>作為主題地圖的處理器，以提供導覽式線上索引典。

## 二、主題地圖

### (一) 主題地圖的發展

主題地圖的發展與 SGML 和 XML 息息相關，我們可從圖 1 看到其的演進歷程。一方面 ISO Topic Maps 採用 HyTime 來定義其語法，而 HyTime 即是承襲 SGML 架構而來；另一方面，由於 SGML 太複雜導致 XML 的興起，而 XTM 便是主題地圖的 XML 版本，旨在簡化 ISO Topic Maps 並適用於網際網路。

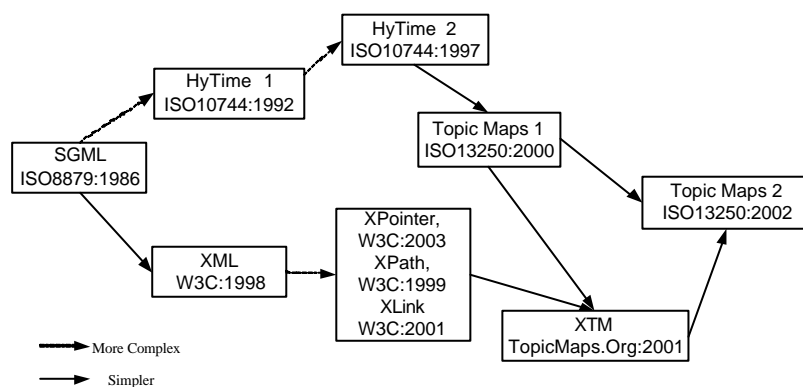


圖 1 主題地圖的演進歷程<sup>4</sup>

#### 1. SGML/HyTime 與 ISO Topic Maps

SGML ( Standard Generalized Markup Language ) 是由國際標準組織 ( International Organization for Standardization, 簡稱 ISO ) 於 1986 年頒佈的國際標準，正式名稱為 ISO 8879 Information processing -- Text and office systems --

2000, Paris, 12-16 June 2000. <<http://www.gca.org/papers/xml/europe2000/papers/s11-01.html>> (Mar. 22, 2003).

<sup>3</sup> Omnigator 是由 Ontopia 公司所開發出來的軟體，為一伺服器端的主題地圖處理器，全稱為 omnivorous topic map navigator，以結構化方式顯示主題地圖，可於以下網址下載試用版 <<http://www.ontopia.net/download/index.html>> (Mar. 18, 2003).

<sup>4</sup> 此圖原始資料來自：Michel Biezunski, "Introduction to the Topic Maps Paradigm" in *XML Topic Maps*, ed. Jack Park, Boston: Addison-Wesley, 2003, p.24., 經本文作者參酌其他相關資料修訂而成。

Standard Generalized Markup Language (SGML)。<sup>5</sup> 具有跨平台、可攜性、結構化、移植性、重覆使用等特性，提供了一個製作電子文件的國際標準；它設計週詳、功能龐大，以致很快地便成為各界遵循規範。

而 HyTime 即是 SGML 的一個應用，提供 SGML 的多媒體與超連結 (Hyperlinking) 功能<sup>6</sup>。最早是由於 SGML 創始人之一的 Goldfarb 為了能表達時序性資訊，<sup>7</sup>遂與 Steve Newcomb 合作，試圖發展一種不論在任何時間與空間，提供對任何種類的資訊做定址與連結的方法。<sup>8</sup> HyTime 於 1992 年正式成為 ISO 10744 標準，全名為 Hypermedia/Time-based Structuring Language，並於 1997 年修正為第二版。它是一個用來提供呈現靜態與動態資訊的工具，用於超文字 (Hypertext) 與多媒體 (Multimedia) 應用軟體的處理與交換。為了將多媒體資訊的時間與空間列表，HyTime 提供一個標準化的架構，來具體指定在文件內或與其他資訊物件間的交互連結。<sup>9</sup>

但由於 HyTime 太複雜，圖形傳播學會 (Graphic Communication Association, 簡稱 GCA, 現改名為 IDEAlliance) 便贊助一項名為 CapH (Conventions for the Application of HyTime) 的研究計畫，旨在發展一個比 HyTime 更容易的瞭解與應用的子集，由 Newcomb 與 Michel Biezunski 所主持。為了能夠自動整合與處理索引，Newcomb 提出一個可以記錄內含於書後索引的知識架構，並與 Biezunski 共同設計出解決方案，稱之為「主題導航地圖」(Topic Navigation Maps)，便是「主題地圖」的前身。這項工作於 1996 年轉移至 ISO 委員會，並於 2000 年 1 月正式頒訂編號為 ISO 13250 的主題地圖標準<sup>10</sup>。採用 HyTime 的結構化欄位模版使其可以在標準語法上，自由地提供欄位類型的子集合，可以為欄位名稱、屬性名稱等重新命名。因此相當符合出版社及其他高階使用者的需求。主題地圖將各種概念具體化為主題，擁有自己的名字與資源指引，再定義各主題與主題間的關聯<sup>11</sup>。

由於 ISO 13250 採用 ISO 10744 HyTime 標準來定義主題地圖的語法，是一個 SGML DTD，因此又被稱為 HyTM (HyTime Topic Maps 的簡稱)。

## 2. XML 與 XML Topic Maps (XTM)

當 HyTM 完成後，由於不是採用 XML 語法，使其應用上頗受限制。因此乃由 TopicMaps.Org 組織著手制訂一個採用 XML 的新語法，即是 XTM (XML Topic Maps specification)，並為多數人所採納<sup>12</sup>。XML 是由 W3C (World Wide Web Consortium) 於 1998 年 2 月發佈的標準<sup>13</sup>，旨在簡化 SGML 並使其更切合網路應用。由於 XML 的出現與所受到的支持，促使了 XTM 的產生。在 ISO 13250

---

<sup>5</sup> Erik Naggum, "SGML: Erik Naggum's Brief Description" Feb 07, 1995. <<http://xml.coverpages.org/naggumWhat.html>> (Mar. 12, 2003).

<sup>6</sup> Steven R. Newcomb, "A perspective on the quest for global knowledge interchange" in *XML Topic Maps*, ed. Jack Park (Boston: Addison-Wesley, 2003), 38.

<sup>7</sup> 同上註，頁 45.

<sup>8</sup> Steve Pepper and Lars Marius Garshol "The XML Papers: Lessons on Applying Topic Maps", October 2002, <<http://www.ontopia.net/topicmaps/materials/xmlconf.html>> (Mar. 5, 2003).

<sup>9</sup> ISO/IEC 10744 HyTime (second edition), <<http://www.ornl.gov/sgml/wg8/docs/n1920/html/n1920.html>> (Mar. 18, 2003).

<sup>10</sup> 同註 8。

<sup>11</sup> 同註 8。

<sup>12</sup> XML Topic Maps (XTM) 1.0, <<http://www.topicmaps.org/xtm/1.0/index.html>> (Nov. 22, 2002).

<sup>13</sup> W3C, "Extensible Markup Language (XML)", available at <<http://www.w3.org/XML/>> (20 Feb. 2003).

出版的同時，制訂 XTM 規格的行動也同步開始，由 IDEAlliance 贊助 Newcomb 與 Biezunski 創立的 TopicMaps.Org 所負責，在不到一年的時間內，於 2000 年 12 月 4 日於華盛頓 DC 所舉辦的 XML2000 會議宣布初稿，並在 2001 年 3 月 2 日發佈 XTM 第一版；<sup>14</sup>到了 2001 年 10 月，XTM DTD 被 ISO 13250 所採納，因此在 2002 年修訂的 ISO 13250 第二版同時包含有 HyTM 與 XTM 兩種語法<sup>15</sup>。

在 2001 年 3 月 XTM 發佈後，TopicMaps.Org 與 ISO Topic Maps 委員會便協議採取分工方式，ISO 維持核心標準的發展，而 TopicMaps.Org 則專注於使用者群體議題<sup>16</sup>，並於 2001 年 10 月併入於 OASIS (Organization for the Advancement of Structured Information Standards)。OASIS 成立 XTM 的相關技術委員會 (technical committee)，致力於 XTM 第一版的應用，第一個技術委員會是由 Bernard Vatant 擔任主席，主要訂定 published subject 的發行、管理與使用<sup>17, 18</sup>，並於 2003 年 3 月出版工作草案<sup>19</sup>，其餘兩個則是 XML Vocabulary TC 與 Geography and languages TC<sup>20</sup>。

## (二) 主題地圖基本模型

主題地圖的主要概念是 TAO 三要素<sup>21</sup>：T 是 Topics (主題)，A 是 Associations (關聯)，而 O 則是 Occurrences (資源指引<sup>22</sup>)。TAO 以主題描述知識架構及其關聯性，主題則又可以被對應到其他主題或現實世界中的媒體物件，而資源指引則是指引到一個屬於該主題資源的可識別載體<sup>23</sup>。茲將 TAO 的含意說明如下<sup>24</sup>：

### 1. 主題 (Topics)

在主題地圖中，知識的基本單元稱為「主題」- topic。topic 這個字的來源是希臘字 topos，代表 location 和 subject<sup>25</sup>。在語言學中，主題可以是任何的名詞，一個主題可以有許多連結來指向其所有的資源指引。主題通常都有個名稱，但有時並不盡然，如交互參照中的「見 66 頁」，雖然可以連結到一個主題，但這個名稱並不明確。在主題地圖標準中定義了三種名稱：基本名稱 (base name)、顯示名稱 (display name) 和排序名稱 (sort name)，如「大學圖書館」與「圖書館，大學」，其中基本名稱為必備，而顯示和排序則並不一定要有，但這些名稱都必須是唯一的，這在某些領域會造成困擾，如家譜，或是整合 (merge) 兩個以上的主題地圖時，所以「範圍」(scope) 就是被設計用來處理這個問題的。

<sup>14</sup> 同註 6, 37-40.

<sup>15</sup> M. Biezunski, S. Newcomb, and M. Bryan "ISO/IEC JTC 1/SC34 N323 Guide to the topic map standards" Jun 23, 2002 <<http://www.y12.doe.gov/sgml/sc34/document/0323.htm>> (Mar. 17, 2003).

<sup>16</sup> 同註 8。

<sup>17</sup> Sam Hunting, "The rise and rise of topic maps" in *XML Topic Maps*, ed. Jack Park (Boston: Addison-Wesley, 2003), 55.

<sup>18</sup> The XML Cover Pages "TopicMaps.Org Consortium Continues Development Efforts within OASIS", 2 October 2001 <<http://xml.coverpages.org/ni2001-10-02-a.html>> (Mar. 12, 2003).

<sup>19</sup> The XML Cover Pages "Cover Stories: Draft Requirements Document on Topic Maps Published Subjects.", 10 March 2003 <<http://xml.coverpages.org/ni2003-03-10-b.html>> (Mar. 12, 2003).

<sup>20</sup> 同註 15。

<sup>21</sup> 同註 2。

<sup>22</sup> 目前 Occurrence 一詞的中文譯名並不一致，有：資源指引、資源所在、參考處、出處、關係實例等，本文採用「資源指引」一詞，取其「指引該主題至相關資訊資源所在處」之意。

<sup>23</sup> Alschuler, Liora. "Topic Maps" June 16, 2000.

<<http://www.xml.com/pub/a/2000/06/xmleurope/maps.html>> (Dec. 22, 2002).

<sup>24</sup> 同註 2。

<sup>25</sup> Michel Biezunski, "Introduction to the Topic Maps Paradigm" in *XML Topic Maps*, ed. Jack Park (Boston: Addison-Wesley, 2003), 24.

主題可以被歸類為成群，稱為主題類型 (Topic types)，換言之，主題類型就是主題所歸屬的類別，一個主題可以歸屬一個以上的同類型主題類型，主題類型在主題地圖中也被認定為一個主題。比方說，人、哺乳類、動物分別都是主題，但人同時也分屬於哺乳類、動物這兩個主題類型。

## 2. 資源指引 (Occurrence)

一個主題可同時擁有一個以上的「資源指引」，表示連結一個或多個可定址 (Addressable) 的資訊資源，如某人的出生證明、結婚證書、出版著作、個人網頁或一小段簡介文字 ... 等。資源指引內含在主題地圖內，也可以獨立在主題地圖之外，透過諸如 HyTime addressing<sup>26</sup>或 XPointer<sup>27</sup>等機制來定址。資源指引可以是不同類型的任何成員，在主題地圖標準中，將資源指引類型視為一個角色 (role)，如同主題類型，資源指引角色也被視為主題。

## 3. 關聯 (Association)

主題之間可利用「關聯」來顯示其語意關係，例如「羅貫中」和「三國演義」兩主題之間具有「寫作」關係。不同於資源指引連結到文件來源，關聯表現出一個包含資訊本質、呈現資訊主要價值的知識基礎，一個主題關聯並未限制相關主題的數量。在主題地圖中，關聯同時也被視為一個主題，也有關聯類型 (Association type)，如「寫作」即可視為一種關聯。關聯類型把具有相同關係的主題匯集成群，有助於增加主題地圖的表達能力。

主題地圖就其本質而言是很簡單的，以主題作為基本素材，並利用關聯建立主題間的關係，主題可以有好幾個名稱和資源指引，並利用範圍限制名稱、資源指引和關聯的有效範疇，這就是最基本的主題地圖。<sup>28</sup>

## 三、知識組織

所謂知識組織簡單的說是將資訊或知識精鍊的過程。大陸學者蔣永福和李景正提出七種知識組織方法：知識表示、知識重組、知識聚類、知識存檢、知識編輯、知識佈局以及知識監控<sup>29</sup>。其中「知識表示」是指把知識載體中的知識因子和知識關聯表示出來，以便人們識別和理解知識，是知識組織的基礎，因為任何知識組織的方法都必須建立在知識表示的基礎上。而知識又可細分為主觀知識和客觀知識，後者是存在於各種類型的文獻之中，因此客觀知識的表示就是把文獻中的知識因子和知識關聯用一定的方式表示出來，在圖書資訊學界即是使用分類和主題標目來揭示主題內容。

Hodge 將知識組織系統分為三大類：(1) 術語集 (Term List)：包含權威檔 (Authority Files)、術語彙編 (Glossaries)、字典 (Dictionaries)、地名詞典 (Gazetteers) 等；(2) 分類系統 (Classifications and Categories)：包含標題表 (Subject Headings)、分類表 (Classification Schemes, Taxonomies, and Categorization Schemes) 等；(3) 關連集 (Relationship Groups)：包含索引典

<sup>26</sup> ISO 10744:HyTime 中的一項功能，主要對於超文件連結與多媒體時間同步性的應用。可參考 <<http://www.ornl.gov/sgml/wg8/docs/n1920/html/n1920.html>> (Mar. 22, 2003).

<sup>27</sup> XPoint，全名為 XML Pointer Language，是 Xlink 的一部份，主要支援 XML 文件內的資源定址，官方網站 URL 為：<<http://www.w3.org/XML/Linking>> (Mar. 22, 2003).

<sup>28</sup> Michel Biezunski, "Introduction to the Topic Maps Paradigm" in *XML Topic Maps*, ed. Jack Park (Boston: Addison-Wesley, 2003), 23.

<sup>29</sup> 蔣永福和李景正，「論組織組織方法」，中工圖書館學報 2001 年第 1 期，頁 3。



( Thesauri) 語義網路 ( Semantic Networks ) 知識本體論 ( Ontologies ) 等。<sup>30</sup>

吳萬鈞則認為，利用各種科學組織系統 ( Scientific Knowledge Organization Systems ) 將科學文獻中經過加工處理後的知識組織起來，是儲存、檢索與利用科學知識不可或缺的關鍵環節。他將科學知識組織系統分為六大類：(1) 科學文獻的知識組織系統：包含等級分類系統 ( Hierarchical Classification System ) 分面分類系統 ( Faceted Classification Systems ) 及索引典系統 ( Thesaurus System )；(2) 科學術語的知識組織系統：如字典、辭典及術語典 ( Terminology )；(3) 學科體系的知識組織系統，如教科書、百科全書等；(4) 綜述 ( Survey ) 述評 ( Review ) 評鑑 ( Assessment ) 及預測 ( Forecast ) 的知識組織系統；(5) 人工智慧的知識組織系統；(6) 模控空間的知識組織系統 ( Cyberspace System )。<sup>31</sup>

主題地圖是一種類似語意網路的知識表示法。在語意網路模型中，主要是利用節點 ( Nodes ) 與連結 ( Links ) 進行知識架構的組成，節點可表示物件、概念或特定領域中的情境，而連結則用來表示與定義節點間的關聯。<sup>32</sup>在圖書館學中，書後索引、術語彙編和索引典是三種常見的知識組織法，以下便從主題地圖的角度切入，概略說明這三種知識組織法及其與主題地圖相關的概念<sup>33</sup>。

### (一) 書後索引

曾有人說過，一本書若是沒有索引，就好比一個國家沒有地圖。一個傳統的索引事實上就是一本書的知識地圖<sup>34</sup>，羅列了這本書所有的主題，並給予相關的指引，如頁碼、同義詞等。以《線上資訊檢索》的中文索引為例<sup>35</sup>：

自然語言 ( natural language ) 4,80,99,108,161-165,169-173

    問題 162-163

    解決方法 163-164

    優缺點 164-165

自然語言索引法 見 索引法

自然語言欄位 171 參見 控制詞彙欄位

其特點包括有：區別不同的主題類型及資源指引；採用「見」( see ) 來處理同義詞；採用參見 ( see also ) 來指引到相關主題；在次要項目提供連結到其他類型的主題 ( 如連結作者與其作品、或上層分類與下層分類間 )；區別同音異義、同形異義、同音同形異義詞；資源指引並不僅止於頁碼，可能還包括章節、附註等，並用不同的字體標示等。總結來說，主要特點有：定義主題、主題間的關係、主題資源指引，並藉此將資訊傳遞給使用者。

因此，一本書的索引即同時兼具有主題、關聯與資源指引三要素，與主題地圖的 TAO 模型不謀而合。

<sup>30</sup> Gail Hodge, "Knowledge Organization Systems: An Overview" in System of Knowledge Organization for Digital Libraries: Beyond Tradition Authority Files, April 2000. <<http://www.clir.org/pubs/reports/pub91/contents.html>> (Feb. 19, 2003).

<sup>31</sup> 吳萬鈞，「科學知識組織系統」，資訊傳播與圖書館學 5:1 ( 民 87 年 9 月 )，頁 19-20。

<sup>32</sup> Freese, Eric. "Using Topic Maps: for the representation, management & discovery of knowledge" XML Europe 2000, Paris, 12-16 June 2000.

<<http://www.gca.org/papers/xml europe2000/papers/s22-01.html>> (Dec. 27, 2002)

<sup>33</sup> 同註 2。

<sup>34</sup> 同註 2。

<sup>35</sup> 蔡明月，線上資訊檢索：理論與應用 ( 台北：臺灣學生，民 80 )，頁 360-1。

## (二) 術語彙編

術語彙編是基本詞彙與定義的列表，如同單一類型的索引，並且只有一種資源指引，也就是該術語的定義。而如同索引，一個術語彙編同時具有「見」與「參見」來連結相關主題，有的還會提供相關資訊，如語言或發音等，但重點在於主題名稱與其定義。因此，術語彙編也非常適合使用主題地圖模型來實現。

## (三) 索引典

至於索引典所強調則是索引的另一層面，它是特定領域中相互關連詞彙的網路，重點在於詞彙間的關連性(relationship or association)，選擇一個特定詞彙後，索引典會提供相同意義、同類型事物的更大類與更小類，以及在其他方面相關的詞彙，更重要的是索引典將詞彙歸類，不僅說明詞彙間具有關連性，更指出如何或為何相關，使得指引更為容易。目前在索引典標準中通用的關連類型有廣義詞、狹義詞、不用、相關詞等<sup>36</sup>。

在下一節中，我們將詳細闡述本研究如何以主題地圖來建構一個實驗性質的線上索引典。

## 四、以主題地圖建構索引典

本節首先闡述索引典的定義、目的、結構與編排，接著界定本研究所採用的詞彙範圍，再據以建構索引典的知識組織圖，最後以 XTM 語法實作成主題地圖，並透過 Ontopia 的 Omnigator 處理器提供線上索引典。

### (一) 索引典概述

#### 1. 索引典的定義

索引典的英文為 “Thesaurus”，其希臘與拉丁的字源本義為「寶典」(A treasury)，自 1950 年代開始發展，廣泛運用於資訊儲存與檢索系統<sup>37</sup>，較常見之定義有：

(1) 索引典是控制索引語言正式組織的詞彙，以便將概念間的重要關係（如廣義和狹義）明確的顯示出來。<sup>38</sup>

(2) 就功能而言，索引典是一種控制詞彙的工具，其用途是將文獻、索引人員或使用者所用的自然語言，轉譯成更規範的「系統語言」(文獻工作語言、資訊語言)；就結構而言，索引典是一部含有特定知識領域的詞彙，詞彙間有語意或從屬上的關係，且詞彙是控制的、動態的。<sup>39</sup>

(3) 就資訊儲存與檢索的範疇而言，索引典乃收集足以表示知識概念的字或詞，並將之以特定的結構加以排列，這些字彙控制了同義字，區別了同形異義字，並顯現各相關詞彙間階層及語意互屬上的各種關係，以做為索引者在分析處理資料及讀者在檢索資料時能選用一致的、經過控制的詞彙。<sup>40</sup>

<sup>36</sup> 索引典標準詳見 Z39.19:1993 [Guidelines for the Construction, Format, and Management of Monolingual Thesauri]、ISO 5964:1985 [Documentation - Guidelines for the establishment and development of multilingual thesauri] 與 ISO 2788:1986 [Documentation - Guidelines for the establishment and development of monolingual thesauri]。

<sup>37</sup> 黃惠株，「淺談索引典」，佛教圖書館館訊 5 期(民 85 年 3 月)，頁 2。

<sup>38</sup> “ISO 2788:1986 Guidelines for the establishment and development of monolingual thesauri” <<http://www.nlc-bnc.ca/iso/tc46sc9/standard/2788e.htm#3>>。

<sup>39</sup> 索引典理論與實務，(主辦單位：美國資訊科學學會台北分會、農業科學資訊服務中心、國立中央圖書館，民 83)，頁 5。

<sup>40</sup> 同註 35，頁 177。

## 2. 索引典的目的

依據 ANSI Z39.19，索引典的目的有以下四點<sup>41</sup>：

(1) 轉化：提供一個讓作者、編制索引者與使用者將自然語言轉換成編索引及檢索用的控制詞彙。

(2) 一致性：促進索引詞彙指派的一致性。

(3) 指示相關性：指出詞彙間的語意關係。

(4) 檢索：作為檢索時的輔助工具。

D. Soergel 認為索引典的概念結構 (Conceptual Structure) 在索引與檢索時能將主題概念表現得恰到好處，而索引典將同義詞聚集亦有詞彙控制 (Terminological Control) 的功能。<sup>42</sup>

## 3. 索引典的結構

一般而言，索引典的詞彙分為標目 (heading) 及參照款目 (cross reference entries) 兩種。通常標目被認為可以使用的詞彙，稱之為敘述語或述語 (descriptors)；參照款目則為不可以使用的詞彙，稱為非敘述語 (non-descriptors) 或被替代語 (use reference)，亦即圖書館書目資料處理時採用的參見 (see) 作法。不被認可的參照款目利用被替代 (use) 與認可述語連結起來，以便指引索引典的使用者參考以選用適切的敘述語。每一個敘述語之下列有各種關係詞彙，這些詞彙以各種記號表明其間的關係，如範圍註 (scope note)、替代 (used for)、被替代 (use)、廣義詞 (broader term)、狹義詞 (narrower term) 及相關詞 (related term) 等。

詞間關係包括等同 (equivalence)、層級 (hierarchical) 和聯想 (associative) 三種關係。<sup>43, 44</sup>

(1) 等同關係：

又稱用代關係，參照符號是 “USE” 及 “UF”，具有這類關係的詞，表示彼此在概念/用法上相同或視為相同。

(2) 層級關係：

又稱屬分、上下或等級關係，參照符號是 “BT” 或 “NT”，屬性相同 (即屬同一範疇) 的詞才能夠成層級關係。具有此關係的詞彙，彼此為上位與下位概念的關係，又包括有屬種 (genus-species)、集元 (set-element) 與整部 (whole-part) 等三種關係。

(3) 聯想關係：

又稱親緣、類緣或相關關係，參照符號是 “RT”，是指兩個描述詞間雖無等同或層級關係，但從索引或檢索效度而言，有相互參照，提醒使用者有另一詞存在的關係。

## 4. 索引典的編排

索引典的基本編排方式有字順、分類與圖形三種：<sup>45</sup>

(1) 字順：拼音語系依字母序或字序排列所有的主題詞；表意文字如漢語，可用「四角號碼」，「注音符號」，「羅馬拼音」或其他拼音等來排列，若首字同音

<sup>41</sup> “ANSI Z39.19-1993(R1998) Guidelines for the Construction, Format, and Management of Monolingual Thesauri”，頁 1。<<http://www.niso.org/standards/resources/Z39-19.pdf>>(Dec. 27, 2002)

<sup>42</sup> 同註 37，頁 2。

<sup>43</sup> 同註 37，頁 4-5。

<sup>44</sup> 同註 39，頁 44-53。

<sup>45</sup> 同註 37，頁 5。



則可配合筆劃區分排列次序，詞群的基本編排如下：

- 主題詞
- 注釋或定義 (SN)
- 同義詞 (用：USE；代：UF)
- 廣義詞 (BT)
- 狹義詞 (NT)
- 相關詞 (RT)

(2) 分類：除分類表外，還配合引領到分類部分的字順索引。在分類部分，每個描述詞都有一個分類代碼，群詞結構只有主題詞、注釋或定義、同義詞及相關詞，層級關係則由分類法和縮行空隔表示。

(3) 圖形：包括圖形與字順兩部份，圖形只有描述詞，其他詞間關係在字順部分表示，一般分為樹形結構及箭頭關係兩種。

## (二) 研究範圍界定

本研究自教育論文線上資料庫 (Educational Documents Online, 簡稱 EdD-Online)<sup>46</sup> 的中文教育類詞庫中，以主題詞「專家系統」為中心，選取其相關詞彙作為建置實驗用索引典之素材。EdD-Online 是由國立臺灣師範大學圖書館與各師範校院圖書館共同合作建置，內容包含教育論文線上資料庫、教育新聞剪報資料庫、中文教育類詞庫查詢等，收錄國內教育類論文、研討會論文，並提供文獻傳遞服務、教育專題選粹等功能，是一個非營利性資料庫。而中文教育類詞庫係以 ERIC 資料庫 (Educational Resources Information Center, ERIC) 之索引典為基礎，可註記詞彙的適用範圍 (scope)、廣義詞 (broader Terms)、狹義詞 (narrower terms)、相關詞 (related terms)、替代 (used for) 與被替代 (use) 等關係語彙。中文教育類詞庫獨立建置為一線上的全文檢索資料庫，使用者透過詞庫之查檢瀏覽，可直接將單一或多個詞彙轉入教育論文資料庫中檢索，相關詞亦可供一併檢索。<sup>47</sup>

為避免語意關係範圍的無止無盡，我們從 EdD-Online 中選出十九個與「專家系統」相關的主題詞，我們對這十九個主題詞及其替代詞彙進行關聯性分析後，歸納出其詞間關係，如表 1 所示。

表 1 本研究所採用的十九個主題詞與其詞間關係

取材自：教育論文線上資料庫中文教育類詞庫 (EdD-Online <http://140.122.127.251/edd/edd.htm>)

主題詞	替代(UF)	相關詞(RT)	廣義詞(BT)	狹義詞(NT)
專家系統	知識本位系統	電腦、資料處理、知識呈現、人機系統、資訊系統	人工智慧	智慧型教學輔導系統
智慧型教學輔導系統	智慧型電腦輔助教學系統、智慧型電腦輔助教學	電腦管理教學、互動式錄影帶、編序教學輔導、教學輔導方案	電腦輔助教學、專家系統	
人工智慧			電腦、知識呈現	專家系統
知識呈現		人工智慧、電腦、專家系統、自然語言處理		
自然語言處理	自然語言理解系統	人工智慧、電腦、知識呈現、人機系統	資料處理	

<sup>46</sup> EdD-Online 網址為 <<http://140.122.127.251/edd/edd.htm>>.

<sup>47</sup> 教育論文摘要檢索系統(教育論文線上資料庫)：擴建計畫執行報告，國立臺灣師範大學圖書館，<<http://www.lib.ntnu.edu.tw/Education/eplars-per.html>> (Dec. 26, 2002)

主題詞	替代(UF)	相關詞(RT)	廣義詞(BT)	狹義詞(NT)
人機系統	人機對話、人機介面	電腦輔助教學 電腦、專家系統、自然語言處理、線上系統		
電腦管理教學	生涯成熟態度問卷 電腦輔助教學管理、電腦本位教學管理	電腦輔助教學 電腦、智慧型教學輔導系統	資訊系統	
互動式錄影帶	智慧型錄影帶	電腦輔助教學 智慧型教學輔導系統、線上系統	線上系統	
編序教學輔導		電腦輔助教學 智慧型教學輔導系統、編序教材	編序教學、教學輔導	
教學輔導方案	教學輔導計畫 教學輔導服務	智慧型教學輔導系統 教學法、教學輔導		
編序教學	編序學習、編序式自我教學	編序教材	教學法	電腦輔助教學 編序教學輔導
電腦	電腦科技	人工智慧、電腦輔助教學 電腦管理教學 資料處理、專家系統、知識呈現、人機系統、自然語言處理、線上系統、資訊系統		
資料處理	自動化資料處理、資料表格化、電子資料處理	電腦、專家系統、線上系統、資訊系統		自然語言處理
教學法	教學方法、教學實務、教學系統、教學技巧、教學方法	教學輔導方案		編序教學
教學輔導	輔導教學	教學輔導方案		編序教學輔導
電腦輔助教學	電腦輔助學習、電腦本位教學	電腦管理教學 電腦、互動式錄影帶、人機系統、編序教材、編序教學輔導	編序教學	智慧型教學輔導系統
資訊系統		電腦、資料處理、專家系統、線上系統		電腦管理教學
線上系統	互動式系統(線上)	電腦、資料處理、資訊系統、人機系統、資訊系統		互動式錄影帶
編序教材	自我教學教材	電腦輔助教學、編序教學、編序教學輔導		

### (三) 建構語意網路

確定所採用的詞彙及其關聯性後，便可開始著手建構其語意網路。首先，我們依據索引典詞彙關係，提出如下的語意網路模型：以「主題詞」為中心，其等同關係為「同義詞」；層級關係為「廣義詞」、「狹義詞」；聯想關係則為「相關詞」。此模型如圖 2 所示。

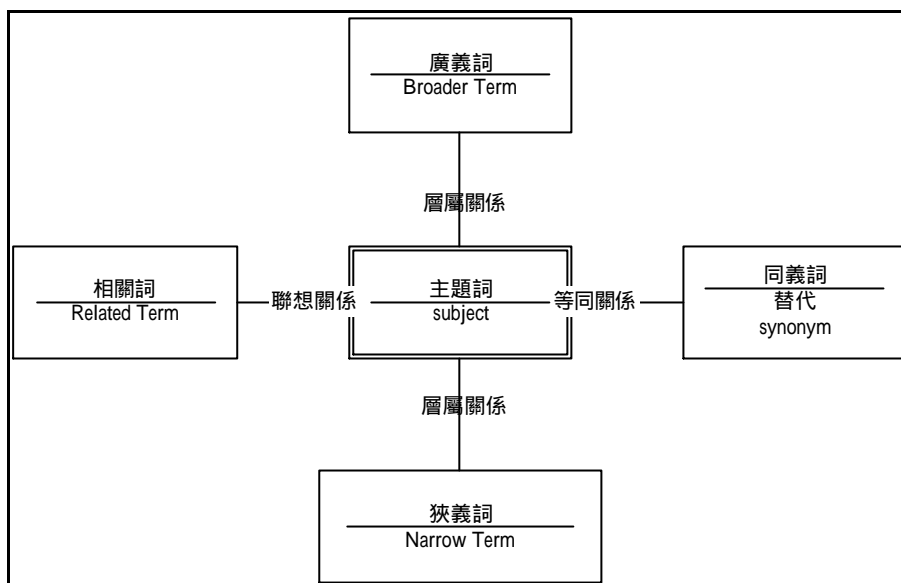


圖 2 索引詞彙關係之語意網路模型

為了更簡化起見，可將同義詞（即替代詞）與主題詞置於同一個方塊內，而兩旁置放相關詞。舉例而言，若以表 1 中的「專家系統」作為主題詞，將專家系統的語彙關係依此方式表示，則結果如圖 3 所示。

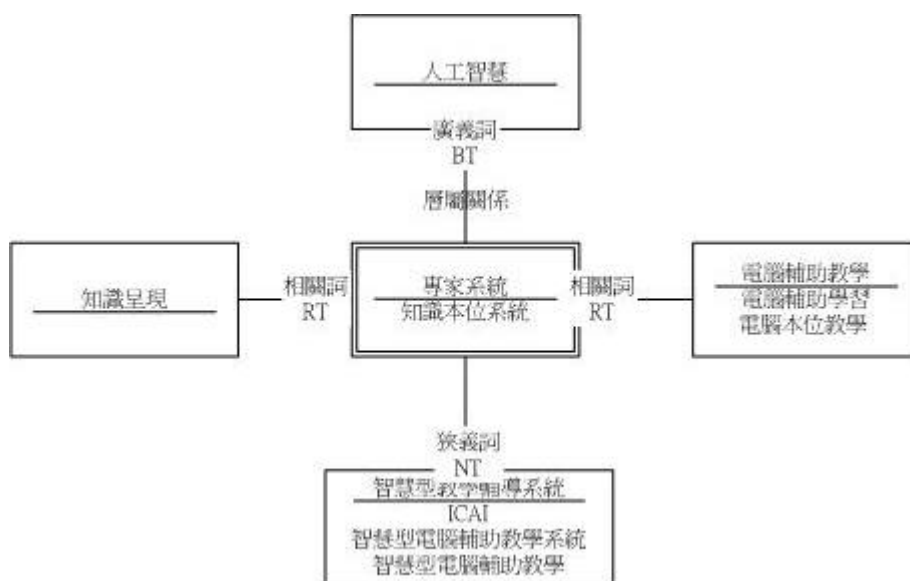


圖 3 以「專家系統」主題詞為例之語意網路圖

根據上述步驟，我們依序將所有主題詞依其關聯性建構成語意網路圖，最後結果如圖 4 所示。其中，每個方塊的上方為主題詞，下方為同義詞，表示同一概念。若主題詞間有層級關係，則以實線箭頭表示，箭頭方向表示上位對下位之關係。若以虛線連接兩個主題詞，則表示彼此間具有聯想關係，且由於此關係並無上下之分，故並無箭頭方向。

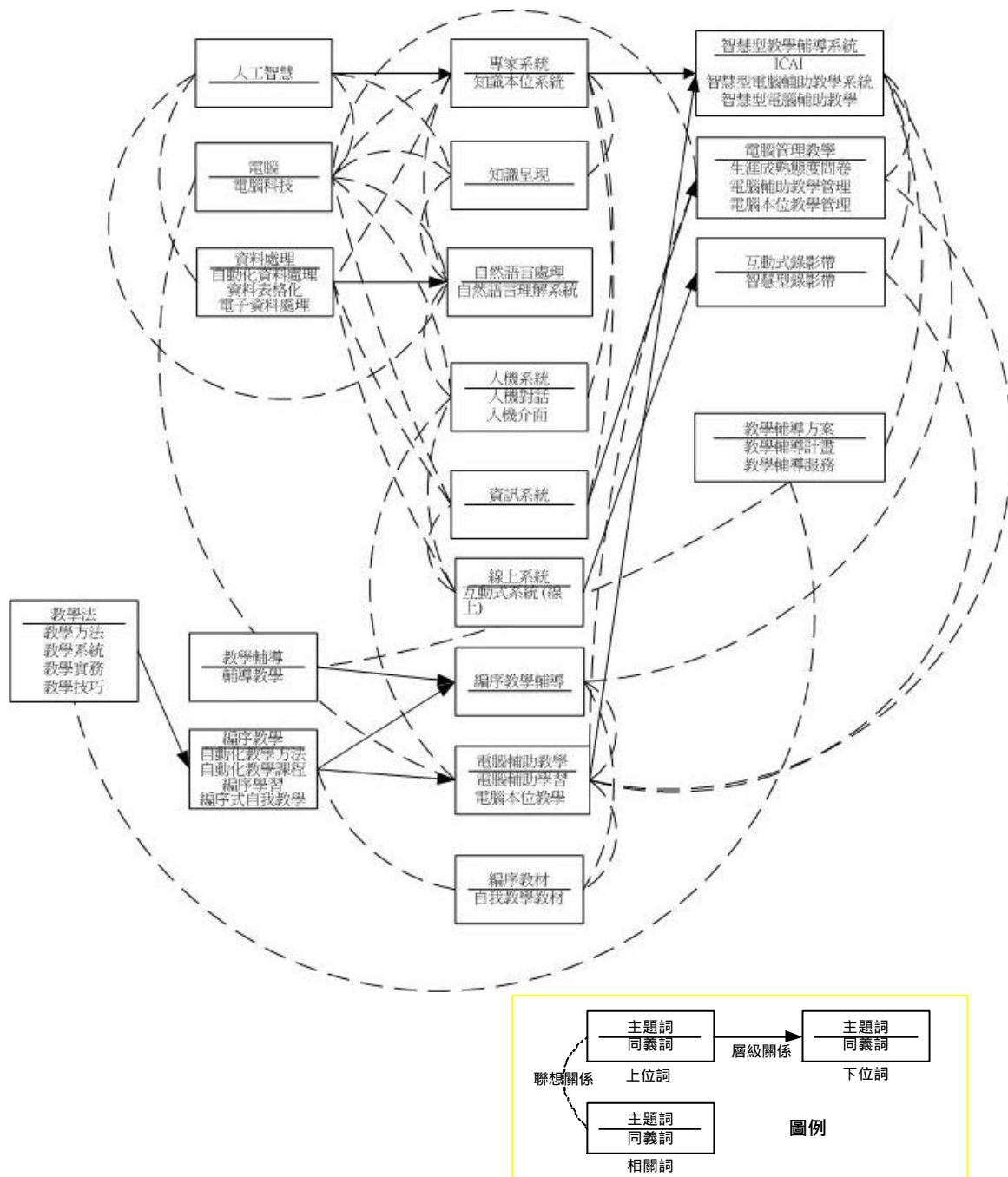


圖 4 以 Ed-D 索引典中的十九個主題詞彙之語意網路圖

#### (四) 實作主題地圖

完成以上十九個詞彙之語意網路後，即可遵循 XTM 標準，建構其主題地圖。首先，將各個詞彙以 <topic> 元素加以標示，並將 <topic> 以 <instanceOf> 加以分門別類，再加上其範圍註 (SN) 作為其資源指引，並以 <occurrence> 加以標示。如下圖所示即為主題詞「專家系統」的 XTM 原始碼片段，至於其他詞彙亦以相同方式建立成不同的 <topic>。



```

<topic id="es">
  <instanceOf>
    <topicRef xlink:href="#subjectTerm"/>
  </instanceOf>
  <baseName>
    <baseNameString>專家系統</baseNameString>
  </baseName>
  <occurrence>
    <instanceOf>
      <topicRef xlink:href="#sn"/>
    </instanceOf>
    <resourceData>
      Computer systems capable of matching a database of factual
      information with a knowledge base of judgmental rules to
      answer questions, make decisions, or teach a skill.
    </resourceData>
  </occurrence>
</topic>

```

圖 5 定義主題詞「專家系統」之 XTM 原始碼片段

接著，我們便可以著手建立各個詞彙之關聯了。不過在建立關聯之前，必須先以 <topic> 將關聯定義成一個主題，並以 <baseNameString> 給予基本名稱。下圖所示即為將「廣義詞」以 <topic> 定義成一個主題之 XTM 原始碼片段。

```

<topic id="bt">
  <baseName>
    <baseNameString>廣義詞</baseNameString>
  </baseName>
</topic>

```

圖 6 將「廣義詞」定義成一個主題之 XTM 原始碼片段

再來，便可將兩個不同的主題詞以 <association> 建立關聯了。例如，下圖是建立「專家系統」與「人工智慧」兩主題關聯的 XTM 原始碼片段。

```

<association id="ha1">
  <instanceOf>
    <topicRef xlink:href="#hierarchical"/>
  </instanceOf>
  <member>
    <roleSpec>
      <topicRef xlink:href="#bt"/>
    </roleSpec>
    <topicRef xlink:href="#ai"/>
  </member>
  <member>
    <roleSpec>
      <topicRef xlink:href="#nt"/>
    </roleSpec>
    <topicRef xlink:href="#es"/>
  </member>
</association>

```

圖 7 建立「專家系統」與「人工智慧」關聯之 XTM 原始碼片段

依據上述步驟，一一將所有的主題詞以此方式進行定義，並將之關聯起來，便完成了一個以主題詞「專家系統」為中心的主題地圖了。

最後，我們在 Web 伺服器上安裝由 Ontopia 公司開發的主題地圖處理器 Omnigator，再將完整的 XTM 文檔發佈至該伺服器上，即完成此線上索引典的製作了。使用者只要用一般 Web 瀏覽器即可使用此線上索引典，圖 8 呈現的是其概觀，分別顯示了主題詞彙、主題詞間的關聯性、在關聯中所扮演的角色、以及資源指引的類別 ... 等。



圖 8 以 Ed-D 索引典中的十九個主題詞彙之主題地圖

若以「專家系統」一詞為例（請參見圖 9），我們可以很清楚的畫面左方看到與其相關之主題及彼此間的關聯性，而右方則是其資源指引的內容。



圖 9 「專家系統」一詞之詞彙關係及資源指引

## 五、結論

無疑的，主題地圖是一項結合了知識工程（Knowledge Engineering）和資訊組織（Information Organization）的新技術。在本研究中，我們成功的提出一個以主題地圖建構索引典之語意網路模型、方法與步驟，並從實際詞庫中選擇十九個主題詞及其替代詞彙作為實驗用素材，歸納其詞間關係並建構語意網路圖，再以 ISO 13250 之 XTM 主題地圖標準語法實作，最後在伺服器端建置主題地圖處理器加以處理，順利的完成一個實驗性質的線上索引典。

展望未來，主題地圖將如同其他新興的標準或技術，一方面在各個領域展露潛力，另一方面則將遭遇各種問題的嚴厲挑戰，是否真能成為知識組織的利器，或是成為知識管理的有效解決方案，實在是一個頗值得探討的議題，亦是我們未來的研究重點之一。