

知識庫在數位化圖書館中的應用

國立臺灣大學資訊工程研究所 林光龍；葉建華；歐陽彥正

【摘要】：本論文主要是討論如何將知識庫系統導入數位化圖書館的建構中。隨著資訊科技的進步與通訊網路的普及，數位化的資訊內容將成為人類擷取資訊的主要來源，而數位化圖書館所典藏的龐大數位化資料，將會是提供數位化資訊的最主要來源。因此，如何有效地組織數位化資料，使其具備提供整合性資訊與知識的能力，正受到各方從事數位化圖書館研究單位的注意。「知識管理」的問題在早期的數位化圖書館計劃中鮮少被提出來；但隨著資訊檢索與資料採礦技術的成熟，如何將數位化圖書館的角色從「資料儲存者」向上提升為「知識提供者」，將會是未來數位化圖書館繼續發展下去最重要的研究課題之一。本論文主要針對「台大數位化圖書館與博物館」典藏的許多數位化史料進行資訊擷取與知識推論的研究，我們計劃藉由物件的「時間知識」來推導出隱含於不同物件之間的关系。

關鍵詞：數位化圖書館；詮釋性資料；知識庫；時間知識；時間邏輯；知識推論；演繹推論

一、序論

拜數位化與資料儲存技術快速進步之所賜，數位化圖書館（digital library）的開發在硬體支援上蒙受到相當大的利益。在此同時，針對數位化圖書館相關技術與服務的研究，諸如有關資料庫建構、資訊檢索與利用、數位化資料保存收藏等各方面的議題，就像雨後春筍般的紛紛被提出來討論（見參考書目 2、16、17、18、21）。

隨著資訊科技的進步與通訊網路的普及，數

位化的資訊內容將成為人類擷取資訊的主要來源，而數位化圖書館所典藏的龐大數位化資料，將會是提供數位化資訊的最主要來源。因此，如何有效地組織數位化資料，使其具備提供整合性資訊的能力，便開始受到各方從事數位化圖書館研究單位的注意。

「知識管理（knowledge management）」的問題在早期的數位化圖書館計劃中鮮少被提出來（見參考書目 4、5、29）；但隨著資訊檢索（information retrieval）與資料採礦（data mining）



技術的成熟，如何將數位化圖書館的角色從「資料儲存者」向上提升為「知識提供者」，將會是未來數位化圖書館繼續發展下來最重要的研究課題之一。

因為目前數位化圖書館中典藏的眾多電子化文件，祇不過是將資料從傳統的媒體轉存至電子化設備，方便後續使用電腦來從事快速的資料索引、查詢與瀏覽（見參考書目 22、23）。對於文件本身所隱含的內容，甚至文件當中所提及的人、事、時、地、物等特定主題之間的關係，並未做任何進一步的分析、組織與管理。

本篇論文將提出一個物件導向（object-oriented）的資料結構來建構知識庫（knowledge base）。這個資料結構主要是用來描述某一物件所擁有的屬性（attribute），及它和其他物件之間所存在的關係（relation）。目的是要重新組織資料庫成為知識庫，之後應用資訊擷取（information extraction）與知識推論（knowledge inference）的技術，針對數位化圖書館中所典藏的資訊內容來進行推演，以獲取內隱的資訊與知識。

舉例來說：如果在數位化圖書館的知識庫中儲存有以下片段的資訊「某甲曾經在光緒年間內任職於浙江海鹽縣擔任縣丞；某乙曾經在光緒年間內任職於浙江海鹽縣擔任知縣。」因此，我們可以從分散在知識庫中的這些片段資訊，透過資訊擷取與知識推論的技術，推論出某甲與某乙可能曾經有過同僚關係。

在數位化圖書館裡所典藏的資料與資訊是相當豐盛而複雜的，如果想要將上述所提及的工作交由專業人員來處理，將會是相當費時且費力，而且無法保證挂一漏萬的可能性不會發生。但

是，如果將這樣的工作交由專門的知識管理工具來處理，不僅可以省時又可以省力，還可以將寶貴的專業人力資源，使用在另一層次的哲學問題分析上，以獲取更深一層的資訊與知識。

近年來已經有許多針對「使用物件導向結構來呈現時間資訊」的文章被發表（見參考書目 13、14）。然而，這些文章所論及的內容都僅止於討論如何紀錄與查詢時間資料。本篇論文想要討論的是，如何利用時間資訊的推導，以獲取其他隱含性知識。在進行隱含性知識的推演過程中，我們採用由 James F. Allen 於 1991 年時所提出的「時間邏輯系統（temporal logic system）」觀念。

接下來的幾節我們將進一步地探討幾個重要議題。其中第二節是討論如何呈現與管理一物件的時間知識，以便後續的知識推論能夠順利地進行。在第三節我們主要是說明何謂外顯知識、何謂內隱知識，及它們之間的關係。在第四節當中列出了我們實驗的資料，並說明由 James F. Allen 所發展的「時間邏輯系統」理論（見參考書目 1），以及提出一物件內隱關係的演繹推論程序，用來推導出隱含於物件之間的關係。至於如何推廣演繹推論工具的應用將在第五節提出說明，這當中包括有物件垂直關係與資訊一致性問題的討論。最後一節將對本論文的整個研究下一個結論，並提出進一步研究的方向與數位化圖書館的發展趨勢。

二、時間知識的呈現與管理

本系統主要是配合「台大數位化圖書館與博物館」計劃所要典藏的內容來進行規劃與設計。由於「台大數位化圖書館與博物館」的特質是它珍藏了許多具歷史性的題材，因此「時間知識



(temporal knowledge) 在本應用系統中是最重要的知識型態。所謂「時間知識」我們用下面的一段例子來說明。

「有兩個人他們從前在學校時是同學關係，畢業後在職場上他們也曾經有過同事關係。」

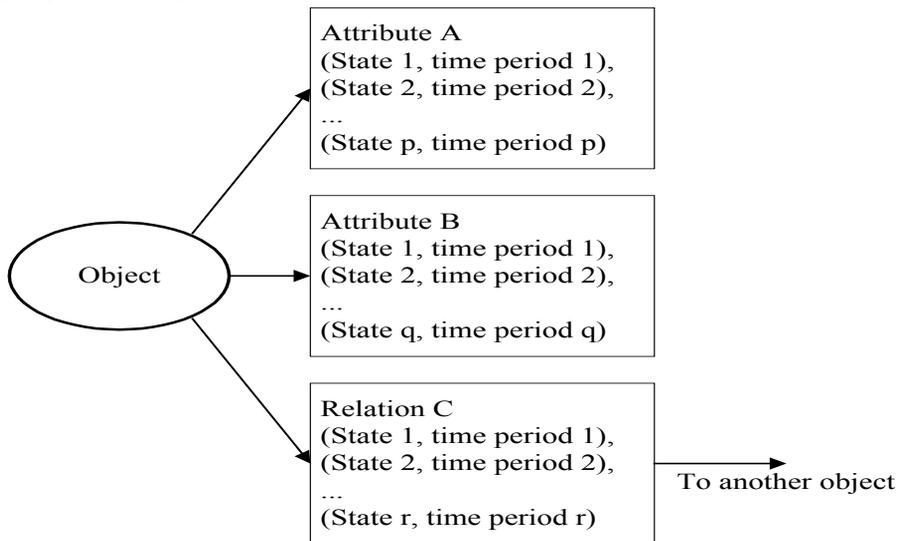
我們將上述的這段話稱為「時間知識」，然而想要獲得上述的這段知識，首先你必須將介於這兩個人之間的這段同學與同事關係，在時間的限定上有明確地描述與詮釋後（見參考書目 1、10、13），才能透過時間推論與時間邏輯運算的機制獲得。

由於我們所欲紀錄的資訊及知識是關於歷史性人物及事件，為了能夠清楚地紀錄一個物件所有有關時間方面的資訊，我們設計「狀態組值 (state tuple)」，如下所述：(狀態名稱, 發生時點, 終止時點)。每一組狀態組值都是專門用來紀錄物件所擁有的屬性, 和關係中具有時間限制的資訊。

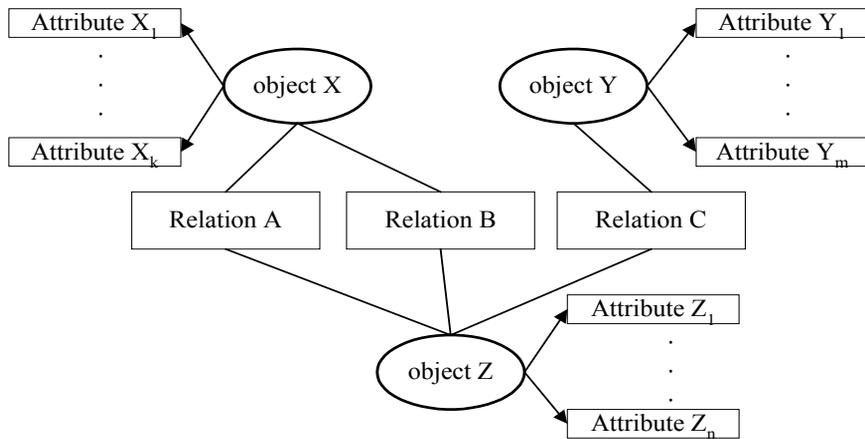
在這個狀態組值中, 「狀態名稱」是用來存放「屬性名稱」或在一個關係底下該物件所扮演的「角色名稱」。「發生時點」、「終止時點」則是指這個狀態的有效期間。舉例來說：紀錄某位官員的經歷 (雲南河陽縣縣丞, 1669/01/01, 1673/01/01)、紀錄同窗關係的角色 (高中同學, 1987/9/01, 1991/06/01)。

每一個物件可以有了一個或多個屬性, 同時也可以與其他物件構成一種或多種關係。每一個屬性及其每一種關係均包含一個或多個狀態組值, 用來細分同一屬性, 於不同時間區段底下的屬性值, 例如: 描述一個人曾經擔任的官職時, 則每一個官職均有明確的時間區段限制。假若一個狀態不受時間區段的限制, 則所屬之時間區段為無限大, 例如: 描述一個人與另外一個人之間的師生關係。圖一與圖二都是一個抽象化的圖示, 圖一是用來表示一個物件的各種狀態組值。圖二是用來表示物件之間的關聯。

圖一、一個典型的物件結構



圖二、不同物件之間的關聯



三、外顯知識與內隱知識

在「台大數位化圖書館與博物館」當中知識的取得主要有兩個來源，第一個是從物件的詮釋性資料 (metadata) 中獲得 (見參考書目 3、28)，我們稱這類的知識為「外顯知識 (express knowledge)」；第二個是使用知識推論的方法獲得，我們將這類的知識稱為「內隱知識 (implied knowledge)」。在「台大數位化圖書館與博物館」中，典藏的物件都有一組相關的詮釋性資料，用來紀錄與描述他們的內容。這些外顯知識都是經由專家以人工方式建立的，這些知識通常是屬於比較直接的，也就是說，可以經由人類的肉眼與認知，直接從物件中觀看出來的資訊。但相反地，內隱知識就不是那麼直接的能從物件的內容觀看出來，而是需要經過一番整理、組織和推敲之後始能獲得。因此，我們想要做的就是建立一個自動推論機制 (見參考書目 19、20)，它可用來吸收外顯知識，經過消化、推論後，產生內隱知識。

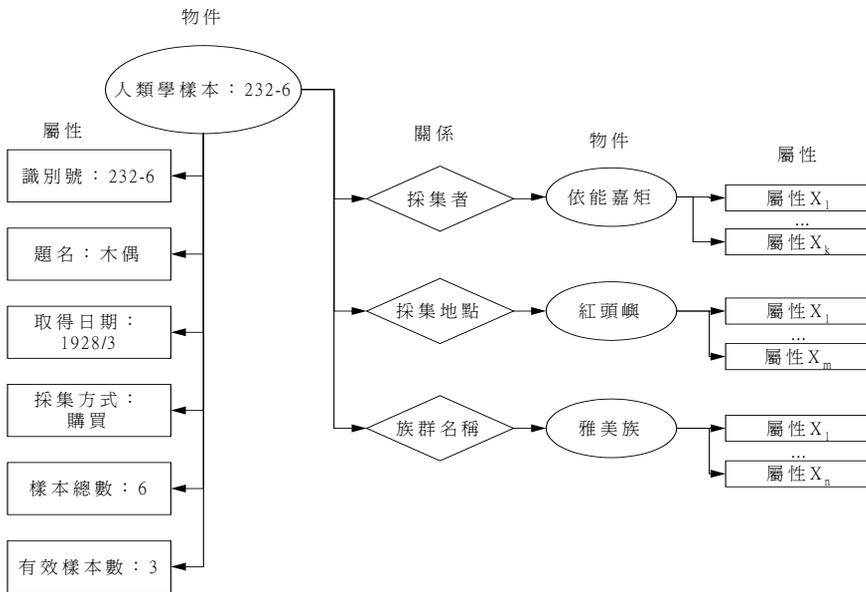
詮釋性資料除了詳細的說明每一物件的屬性之外，也具體的列出每一物件與其他物件之間的關係。表一 是用來說明物件的屬性，如何使用詮釋性資料來表示。圖三 是用來表示物件之間彼此的關係，如何使用詮釋性資料來建構。也就是說，我們使用詮釋性資料來描述一物件所擁有的屬性和關係。

表一、詮釋性資料

紀錄識別號	232-6
題名	木偶
相關族群	雅美族
採集者	依能嘉矩
取得日期	1928年3月
採集地點	約頭嶼
採集方式	購買
樣本總數	6
有效樣本數	3
.....	



圖三、選擇古語釋性資料之物件結構



四、時間邏輯運算與推論

圖四是我們用來作實驗的一組資料，當中列出清朝時期台灣縣衙門官員名冊的片段，我們用它來說資訊演繹推論 (deductive inference) 工具的應用方式。

圖四、台灣縣衙門官員名冊

台灣縣知縣
 沈朝聘
 遼東人。康熙二十三年任（載「名宦」）。丁艱去，今擢大中丞。
 蔣相
 遼東人。康熙二十五年任，有折獄才。二十七年，以清理旗員去官。
 王兆陞
 江南通州人，丁酉舉人。康熙二十七年任；三十年，陞兵部職方司主事。
 錢巍業
 河南彬州人。康熙三十一年任；三十三年以劾去。



縣丞

趙行可

陝西永昌衛人，貢生。康熙二十三年任；二十七年，陞四川成都縣知縣。

陳嘉

浙江仁和縣人，監生。康熙二十八年任；三十年以憂去。

張元英

遼東人。康熙三十年任。陞山西蒲縣知縣。

蔣以選

號千仞，紹興府山陰縣人，由功貢。福州府閩縣調補。康熙三十五年任。

典史

張元初

涿縣人，吏員。康熙二十三年任；二十七年，陞山東濟南府照磨，在臺丁憂回籍。

高娘

山東曹縣人，吏員。康熙二十八年任；三十一年，陞河南懷慶府照磨。

劉蛟祚

江南貴州縣人，吏員。康熙三十二年任。

婁克仁

巡檢

紀文達

陝西涇陽縣人，吏員。康熙二十三年任，二十七年，陞河南府照磨。

常文謨

山東館陶縣人，吏員。康熙二十八年任，三十一年，陞直隸東光縣主簿。

孫禮榮

江南滁州人，吏員。康熙三十二年任。

在這一官員名冊裡主要是列出在清朝時期，任職於台灣縣衙門底下的官員，其個人背景資料與仕途。我們可以觀察到，在這份清單中顯示出某種規則。舉例來說：在每一位官員的描述中都有一段時間區段用來紀錄此官員的任期，而且紀錄的規則是第一段時間表示職日，第二段時間表示卸職日。因此，直覺上我們可針對所有官員的職務與任職時期進行組織，整理出相對應的詮釋性資料後，儲存至物件導向的知識庫。之後，我

們可以自動化地擷取出官員的任期區間，應用時間邏輯運算與知識推論的技術，嘗試從這些外顯知識推論出內隱知識（見參考書目 26、27）。

由 James F. Allen 所發展的「時間邏輯系統」理論是目前被採用於推導隱含性知識最廣泛的基礎理論（見參考書目 1）。當我們在討論時間邏輯時，通常會有一段傳統的述語（predicate）上，加入一時間參數，用來形成一則具有時間涵義的新述語。舉例來說： $P(a_1, a_2, \dots, a_k; t)$ 是一個含有

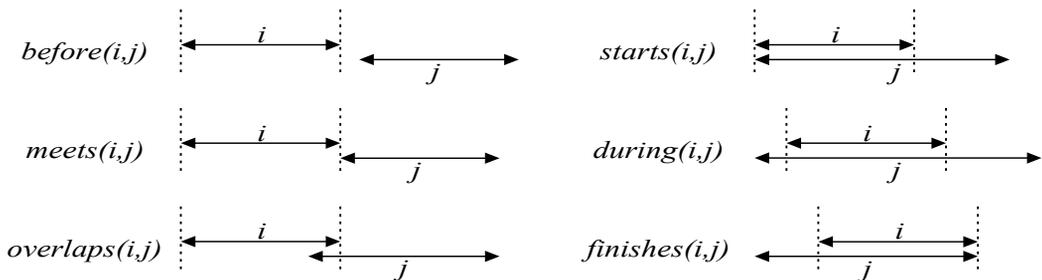


時間的述語，其中 a_1, a_2, \dots, a_k 為非時間性參數； t 是一段與此述語有關的時間區段參數。圖五列出由 James F. Allen 所提出的時間相關性，他指出兩個時段在時間軸上，彼此間可能會發生的六種狀態。圖六則是列出三種基本的時間邏輯運算，包括有時間的交集、聯集、差集。

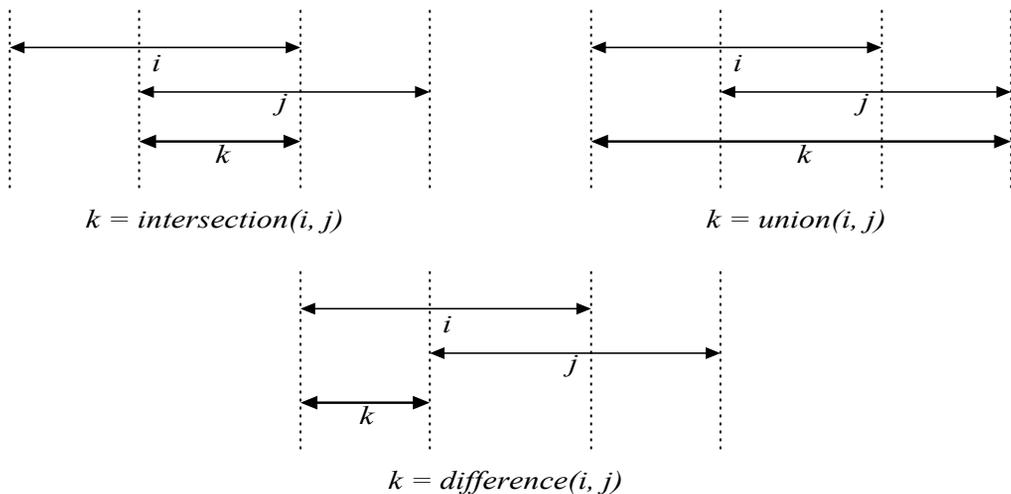
我們將設計出來的演繹推論工具應用在「台大數位化圖書館與博物館」所典藏的物件，嘗試利用儲存於物件之屬性或關係裡的時間資訊進行演繹推論，進而推演出隱含於物件之間的關係。圖七是我們所設計的演繹推論程序。

內隱關係演繹推論程序的流程是從一群包含有 m 個物件的集合，和 n 條時間推論規則的集合開始，使用 $\{O_1, O_2, \dots, O_m\}$ 代表 m 個物件的集合，使用 $\{R_1, R_2, \dots, R_n\}$ 代表 n 條時間推論規則。這 m 個物件是採用上述所提及的物件導向結構所建構而成的，這 n 條時間推論規則，每一條都是用來說明如何從紀錄在物件結構裡的屬性或它所擁有的關係中，推論出一個新的關係。下面這個推論規則表示如何從兩個人的工作經歷，透過演繹推論工具判斷他們是否曾經有過同僚關係（見參考書目 24、25）。

圖五、不同之時段時間區段的相關性



圖六、三種基本的時間邏輯運算



圖七、隕關係演繹推論程序

```

Procedure derive_relations
  for every inference rule  $R_i \equiv P_i \Rightarrow Q_i$ 
  {
     $S_1$  = set of objects that the first argument of  $P_i$  concerns;
     $S_2$  = set of objects that the second argument of  $P_i$  concerns;
    ...
     $S_k$  = set of objects that the  $k$ -th argument of  $P_i$  concerns;
    for every combination of objects  $(O_{i1}, O_{i2}, \dots, O_{ik})$ , where
       $O_{i1}$  in  $S_1, O_{i2}$  in  $S_2, \dots, O_{ik}$  in  $S_k$ 
      {
        for every combination of the states of
           $(O_{i1}, O_{i2}, \dots, O_{ik})$  that  $P_i$  concerns
          {
            evaluate  $P_i$ ;
            if the evaluation results in "True"
            {
              add the newly derived relation into the object structures;
            }
          }
      }
  }
end procedure;

```

$\text{Job}(\text{Person}_A, \text{Company}_x; T_1) \wedge \text{Job}(\text{Person}_B, \text{Company}_x; T_1) \wedge \text{overlap}(T_1, T_2) \Rightarrow \text{Colleague}(\text{Person}_A, \text{Person}_B; \text{intersection}(T_1, T_2))$

3、演繹推論工具的應用

演繹推論工具的應用不祇可以運用在水平關係的推論，在垂直關係上的推論更是能發揮它的優勢。以「禪宗法脈」為例，對於每一位祖師，古當我使用詮釋性資料來描述他的「法脈」資訊時，只需要紀錄基本的「師徒」關係作為推導的

基礎，即可將禪宗的歷代祖師推導出來。此外，演繹推論工具對於解決「資訊一致性（information consistence）」的問題也很有幫助。

表二列出經過整理後的清朝時期官員名冊的部分資料，當中日期部分是指農曆。圖八是清朝時期台灣地方官僚體系的組織架構圖。我們試圖從官員名冊裡推導出「長官/部屬」關係。這裡所指的「長官/部屬」關係必須是滿足下列的條件：乙的位階必須在甲的位階底下，而且甲與乙的任期有重疊，則乙為甲的部屬。圖九是以表二所列



官員的職位與任期為測試資料，使用演繹推論工具推導出隱藏於此份列表中，官員彼此之間的「長官/部屬」關係。下面這個推論規則表示如何從兩個人的工作經歷，透過演繹推論工具判斷他們是否曾經有過「長官/部屬」關係。

$Job(Person_A, Position_x, T_1) \wedge Job(Person_B, Position_y, T_2) \wedge Successor(Position_x, Position_y) \wedge overlap(T_1, T_2) \Rightarrow Subordinate(Person_A, Person_B; intersection(T_1, T_2))$

從上面的例子可以知道，在許多場合中，使用資訊科技來針對數位化圖書館所典藏的物件，進行大規模的內隱關係探索，可以用來彌補人類原先對於大量資料處理能力的不足。因此，如果有一個新的物件被加入數位化圖書館的典藏時，可以先從與這個新物件有關係的舊物件開始著手，再間接的擴大探索的範圍，而進一步發現是否有其他舊物件與此新物件有任何間接關係存在。如此將可有效的維護數位化圖書館典藏內容的完整性與正確性。舉例來說：如果這個新物件是早期某原住民部落的頭目，則我們可以從這位頭目居住的村落，及這位頭目所牽涉的相關事件開始向外探索。針對這群大量的新舊外顯知識進行大規模探索內隱知識之後，我們將可以得到許多發生在這個村落裡的重要歷史事件。

另一個應用資訊演繹推論的時機就是用來解決「資訊一致性」的問題，透過資訊一致性的檢查機制，更可以保證數位化圖書館典藏內容的正確性。舉例來說：如果在知識庫中，某乙的「師徒」詮釋性資料欄位裡紀錄某甲為其老師；同時，在某丙的「師徒」詮釋性資料欄位中亦紀錄某甲為其老師。則我們可以將「乙與丙為師兄弟」的關係，分別地描述於某乙與某丙的「師兄弟」詮

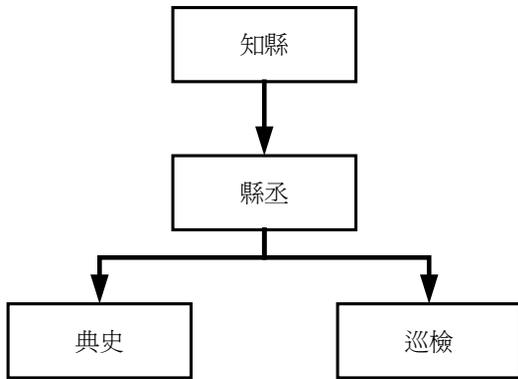
釋性資料欄位裡。若日後發現「甲與乙為師徒」這項資訊有誤時，則之前所得到的「乙與丙為師兄弟」這份關係也就不再成立。因此，我們必須逐一檢視、修正知識庫中，所有倚靠「乙與丙為師兄弟」的這項錯誤資訊，所間接推導出來的其他資訊。在資料的維護與管理過程當中，類似這樣的錯誤常常會引起骨牌效應，如果沒有自動推論機制的輔助，則單靠專家人工的方式執行，則不僅費時費力，還可能由於人為疏失，產生資訊不一致的情形發生。就這個議題而言，自動推論無疑地提供了相當大的好處，因為它可以用最少的人力成本，保證資訊或知識的一致性。

表2：滑州時期官員名冊的部分資料

姓名	職位	述職日(年.月)	卸職日(年.月)
沈朝聘	知縣	1644.01	1644.01
蔣相	知縣	1646.01	1648.01
王兆陞	知縣	1648.01	1651.01
王日俊	知縣	1665.01	1665.01
趙行司	縣丞	1644.01	1648.01
陳嘉	縣丞	1649.01	1651.01
張元英	縣丞	1651.01	1651.01
汪立忠	縣丞	1665.01	1665.01
張元初	典史	1644.01	1648.01
高煥	典史	1649.01	1651.01
李廷貴	典史	1665.01	1665.01
紀心達	巡檢	1644.01	1648.01
常心謨	巡檢	1649.01	1652.01



圖八、清初期台灣地方官僚體系組織架構圖



圖九、知識推論的結果

關鍵人物：沈朝聘/知縣/1644.01~1644.01
 部屬：趙行可/縣丞/1644.01~1648.01
 部屬：張元初/典史/1644.01~1648.01
 部屬：紀文達/巡檢/1644.01~1648.01
 關鍵人物：蔣相/知縣/1646.01~1648.01
 部屬：趙行可/縣丞/1644.01~1648.01
 部屬：張元初/典史/1644.01~1648.01
 部屬：紀文達/巡檢/1644.01~1648.01
 關鍵人物：王兆陞/知縣/1648.01~1651.01
 部屬：趙行可/縣丞/1644.01~1648.01
 部屬：張元初/典史/1644.01~1648.01
 部屬：紀文達/巡檢/1644.01~1648.01
 部屬：陳嘉/縣丞/1649.01~1651.01
 部屬：高焯/典史/1649.01~1651.01
 部屬：常文謨/巡檢/1649.01~1652.01
 部屬：張元英/縣丞/1651.01~1651.01
 部屬：高焯/典史/1649.01~1651.01
 部屬：常文謨/巡檢/1649.01~1652.01
 關鍵人物：王士俊/知縣/1665.01~1665.01
 部屬：汪立忠/縣丞/1665.01~1665.01
 部屬：李廷貴/典史/1665.01~1665.01



六、結論

這篇論文主要是配合「台大數位化圖書館與博物館」計劃所提出來的。我們嘗試把數位化圖書館的典藏資料庫提升為知識庫，將原資料庫所儲存的數位化物件，使用物件導向結構加以重新組織，經由專家以人工方式建立每一物件的相關詮釋性資料，及對同一物件或不同物件之間在認知上的概念（concept），整理出相關的規則與時間資訊，儲存至物件導向的知識庫，並引進時間邏輯運算與知識推論的理論，設計一具備資訊擷取與知識推論的資訊演繹推論工具，利用物件之間的時間相關性，來挖掘知識庫所隱含的資訊與知識。

數位化圖書館是一個數量豐盛、內容多采多姿的知識寶庫，未來我們將朝此模式，繼續針對其他型態的資訊進行研究，並提出相對應的資訊描述、資訊擷取、知識呈現與知識推論等管理工具。例如：針對「空間邏輯系統（spatial logic system）」基礎理論的研究，提出空間知識（spatial knowledge）的建構、管理與推論等技術（見參考書目 12、15）。除此之外，我們亦將採用 ISO/IEC 13250 所制定的主題地圖（Topic Map），以此來解決大量異質性且沒有組織的資訊資源的問題，提供使用者一個最佳的資訊導航（information navigation）系統，透過此系統我們希望達到知識分享與知識創造的最終目的。

最後引用經濟學家 Lester Thurow 曾經說過的一句話「知識成了人類最大的資產，掌握知識即掌握了財富」。由此可知，在目前以知識經濟掛帥的社會底下，知識管理已成為未來資訊社會中最重要之議題之一。因此，在數位化圖書館的

建構中，融入知識管理的理論與技術，絕對是未來數位化圖書館發展的重要趨勢。

備註：本文係依據下列文章修改而成

林光龍，葉建華，歐陽彥正，〈知識庫在數位化圖書館中的應用〉，《新世紀數位圖書館與數位博物館趨勢研討會論文集》，（新竹市：國立交通大學，民 90 年 11 月 1 日），頁 X-1~18。

【參考書目】

- [1] Allen J, F. (1991). Temporal reasoning and planning. In *Reasoning about plans* (p. 1-68). Morgan Kaufmann Publishers, Inc.
- [2] Arms, W. Y., Blanche, C., & Overly, E. A. (1997). An architecture for information in digital libraries. *D-Lib Magazine*. [Online]. Available: <http://www.dlib.org/dlib/february97/cnri/02arms1.html>.
- [3] Baldonado, M., Chang, C.-C., Gravano, L., & Paepcke, A. (1997). Metadata for digital libraries: Architecture and design rationale. In *Proceedings of ACM digital libraries '97*. Philadelphia, USA.
- [4] Chen, H. (1990). A knowledge-based design for hypertext-based document retrieval system. In *Proceedings of the international conference and workshop on database and expert systems applications (DEXA 1990)*. Vienna, Austria.
- [5] Chen, H., Smith, T. R., Larsgaard, M., Hill, L. L., & Ramsey, M. (1997). A geographic knowledge representation system for multimedia geospatial retrieval and analysis. *International Journal on Digital Libraries*, 1(2), 132-152.
- [6] Chen, H., Yim, T., Fye, D., & Schatz, B. R. (1994). Automatic thesaurus generation for an electronic community system. *Journal of American Society for Information Science (JASIS)*, 46(1), 52-59.
- [7] Cowie, J., & Lehnert, W. (1996). Information extraction. *Communications of the ACM*, 39(1), 80-91.
- [8] Daniel Jr., R., & Lagoze, C. (1997). Distributed active relationships in the Warwick Framework. In *Proceedings of the second IEEE metadata conference*. Maryland, USA.



- [9] Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. The MIT Press.
- [10] Galton, A. (1987). *Temporal logics and their applications*. Academic Press.
- [11] Gladney, H. M., Fox, E. A., Ahmed, Z., Ashany, R., Belkin, N. J., Lesk, M., Tong, R., & Zemankova, M. (1994). Digital Library: Gross structure and requirements. In *Proceedings of digital libraries '94*. College Station, Texas, USA. [Online]. Available: <http://www.csdl.tamu.edu/DL94/paper/fox.html>.
- [12] Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1996). Statistical inference and data mining. *Special Issue of the Communications of the ACM on Data Mining and Knowledge Discovery*, 39(11), 35-41.
- [13] Goralwalla, I., Leontiev, Y., Özsu, M., & Szafron, D. (1997). Modeling temporal primitives: Back to basics. In *Proceedings of 6th international conference on information and knowledge management (CIKM'97)* (p. 24-31). Las Vegas, Nevada, USA.
- [14] Goralwalla, I., Özsu, M., & Szafron, D. (1997). A framework for temporal data models: Exploiting object-oriented technology. In *Proceedings of 1997 conference on technology of object-oriented languages and systems (TOOLS USA 97)*. Santa Barbara, California, USA.
- [15] Imielinski, T., & Mannila, H. (1996). A database perspective on KDD. *Special Issue of the Communications of the ACM on Data Mining and Knowledge Discovery*, 39(11), 58-64.
- [16] Kahn, R., & Wilensky, R. (1995). *A framework for distributed digital object services* (Tech. Rep.). CNRI. [Online]. Available: <http://WWW.CNRI.Reston.VA.US/home/cstr/arch/k-w.html>.
- [17] Miller, U. (1997). Thesaurus construction: Problems and their roots. *Information Processing & Management*, 33, 481-493.
- [18] Nuernberg, P. J., Furuta, R., Leggett, J. J., Marshall, C. C., & Shipman III, F. M. (1995). Digital libraries: Issues and architectures. In *Proceedings of digital libraries '95*. Austin, Texas, USA. [Online]. Available: <http://csdl.tamu.edu/DL95/papers/nuernberg/nuernberg.html>.
- [19] Park, Y. C., & Choi, K. S. (1997). Automatic thesaurus construction using Bayesian networks. *Information Processing & Management*, 32, 543-553.
- [20] Russel, S. J., & Norvig, P. (1995). *Artificial intelligence: a modern approach* (p. 151-184). Prentice-Hall International, Inc.
- [21] Schatz, B. R., & Chen, H. (1996). Building large-scale digital libraries. *Computer theme issue on US Digital Library Initiative*.
- [22] Schatz, B. R., Johnson, E.H., Cocharane, P., A., & Chen, H. (1996). Interactive term suggestion for users of digital libraries: Using subject thesauri and co-occurrence lists for information retrieval. In *Proceedings of ACM digital libraries '96*. Hyatt Regency Bethesda, Maryland, USA.
- [23] Schutze, H., & Pedersen, J. O. (1997). A cooccurrence-based thesaurus and two applications to information retrieval. *Information Processing & Management*, 33, 307-318.
- [24] Solderland, S., Fisher, D., & Lehnert, W. (1997). *Automatic learned vs. hand-crafted text analysis rules* (Tech. Rep.). CIIR. [Online]. Available: <http://cobar.cs.umass.edu/pubfiles/te-44.ps>.
- [25] Solderland, S. G. (1996). *CRYSTAL: Learning domain-specific text analysis rules* (Tech. Rep.). CIIR. [Online]. Available: <http://cobar.cs.umass.edu/pubfiles/te-43.ps>.
- [26] Tong, W. C., Wüthrich, B., & Sankaran, K. (1995). A temporal and probabilistic, deductive and object-oriented query language. In *Workshop on Temporal reasoning in deductive and object-oriented database* (p. 61-69). Singapore.
- [27] Vila, L. (1994). A survey on temporal reasoning in artificial intelligence. *AI Communications*, 7(1), 4-28.
- [28] Weibel, S. (1995). Metadata: The foundations of resource description. *D-Lib Magazine*. [Online]. Available: <http://www.dlib.org/dlib/July95/07weibel.html>.
- [29] Wong, S. T. C., Hoo, K. S., Knowlton, R. C., Hawkins, R. A., Laxer, K. D., & Tjandra, D. (1997). Issue and applications of digital library technology in biomedical imaging. *International Journal on Digital Libraries*, 1(3), 209-219.

