

# 圖書館顧客關係管理與資料探勘

楊錦潭      陳玟志      邱魏津      朱永方

## 摘要

本文旨在探討透過資料探勘 (Data Mining) 技術來瞭解和影響圖書館讀者的顧客行為, 以獲取新讀者與提高既有讀者忠誠度的一種經營模式。數位化圖書館的管理, 的確帶給現代圖書館的顧客或讀者很大的便利, 然而如何更進一步地利用這些讀者顧客行為的資料來達到更高層次的「顧客關係管理」, 其中包括基於對圖書借閱行為而對讀者的個人化或適性化的服務, 皆可借助於資料探勘的技術。資料探勘技術將是未來數位圖書館提升效率與效益的利器。本文從圖書館的顧客關係管理來談資料探勘技術, 讓圖書館對於書籍、讀者及讀者的借閱行為等方面具更有效的管理和服務。未來資料探勘除了被用來蒐集與儲藏資料之外, 亦可作為推動數位學習網, 成為現代化圖書館必然的典範轉移。

## 一、前言

「顧客關係管理」(Customer Relation Management, CRM) 是在商業組織與顧

---

**關鍵詞 (Keywords) :** 資料探勘; 顧客關係管理; 適性化服務; 數位化圖書館; 合作過濾  
Data Mining; Customer Relation Management; Adaptive Service; Digital Library; Collaborative Filtering

楊錦潭: 國立高雄師範大學通識教育中心教授、南台科技大學數位設計學院院長;  
E-mail: yangdav@nknucc.nknu.edu.tw

陳玟志: 國立高雄師範大學資訊教育研究所碩士班; E-mail: koach@ccsh.kh.edu.tw

邱魏津: 國立屏東科技大學圖書館館長; E-mail: chiuwei@mail.npust.edu.tw

朱永方: 國立高雄師範大學資訊教育研究所碩士班; E-mail: fang@mail.fhjh.tn.edu.tw

客透過不斷地互動及有意義的溝通來瞭解和影響顧客行為，以獲取新顧客、留住既有顧客並提高顧客忠誠度及獲利率的一種經營模式。<sup>[1]</sup>而Tiwana對顧客關係管理的定義是，組織從各種不同的角度來了解、區別顧客，以發展出適合顧客個別需求的產品或服務的一種組織程序和資訊科技的一種組合模式<sup>[2]</sup>，目的是用管理組織與既有顧客的關係，以促使顧客達到最高的忠誠度、留存率與利潤貢獻度，並同時有效率選擇性的吸引好的新顧客。<sup>[3]</sup>將之用於圖書館的管理上是未來圖書館必然要走的路。其理由有二。

首先圖書館的定位問題：圖書館一直以來所扮演的角色，大部份是有關館藏書籍的採購管理、借閱、還書等行為而已；然隨著資訊時代的來臨，圖書館的定位應該隨之有所調整；過去處於被動的行政業務應該要積極地以服務讀者為目標，設法吸引更多的讀者來參與圖書館的活動，如開闢讀者群的交流園地，讓讀書不再是人和書之間的關係而已，更可以是增加人和人互動的一項活動，這不僅是可以吸引更多的人來閱讀的行銷手法，也是拓展閱讀效用的空間。因此，圖書館必要條件是要擁有足夠的資訊流，讓所有的參與者形成一個虛擬社群（Virtual Community），使其成為圖書館的經營策略。如同亞馬遜書店善加運用核心顧客資訊的管理，形成其商業模式（Business Model）。提升核心顧客與價值鏈廠商的滿意度，因而大大地增加它的盈收。

其次，數位化時代的來臨：隨著資訊時代的來臨，愈來愈多傳統資料被數位化，同時也逐漸在改變人類的行為，以往用紙張的記錄方式，已經被鍵盤、滑鼠和硬碟所取代，這些數位化的資料一直持續在增加；以圖書館而言，購置新書後的建檔工作，讀者的借閱記錄（Reader Record）等，都將資料一筆一筆地進入電腦的硬碟裡。然而這些數位化的記錄資料，在21世紀之前，它只能幫助管理圖書館的流程，頂多再加上對一般讀者作推薦之用、亦或借閱排行榜等宣傳而已。接著那些逐年累月的資料（Data-warehouse），就很少再有被利用的價值。然而近年來，隨著資料探勘（Data Mining）的技術被提出來後，這些歷史資料重新被挖出

[1] Ray McKenzie著，張晉綸譯，《究極CRM》（臺北市：麥格羅希爾，2001）。

[2] Tiwana, A., *Essential Guide to Knowledge Management--The e-Business and CRM Application* (Upper Saddle River, NJ: Prentice-Hall, 2001)。

[3] 林東清、劉勇志，企業顧客關係管理(CRM)的一些主要經營架構與流程模式，載於：淡江大學主編，《第十三屆國際資訊管理學術研討會論文》（臺北縣：淡江大學，2002），頁569-578。

來，開始受到資訊界的重視了。<sup>[4]</sup>

所謂資料探勘（或資料挖掘）主要來自現今數位化時代裡，「氾濫」而幾可成災的大量資料卻不見得有用。換句話說，它導致「只見資料爆炸，未見知識」的現象，因此資訊學家與專業領域的人士共同設法利用既有的大量數位資料來產生「有用」（Usefulness）或「有意義」（Meaningfulness）的資訊。此亦被稱為知識發現（Knowledge Discover）。舉例來說，當人類對某一件事情的經驗很豐富之後，就能夠較易輕地掌握事情的運作過程，因為人的大腦可以將這些經驗自動地轉換為知識，因此在下次遇到類似的狀況時，便可運用這些知識來解決問題；相較於大量的數位化資料，就如同人類日常生活的經驗一般，然而電腦無法自動幫人類消化這些經驗而產生知識。因而資料探勘功能就扮演著如同大腦的運作一樣，希望能幫人類從大量資料堆裡得到一些規則，發現隱藏於其中的知識。

隨著資料庫技術的日趨成熟，電腦化圖書館已成主流的今日，使得圖書館存在著大量數位化資料，這些資料通常包括書籍資料、讀者資料、讀者借閱記錄等。這些資料隨著新書購置、讀者借閱等動態因素，隨時處於更新狀態。這些數位資料如果朝向個人化或群體化的新書推薦功能，即可以落實圖書館的顧客關係管理。因為資料探勘技術為圖書館的既有大量資料引進了新的契機，這些記錄在經過一翻整理、分析、推演才有可能發現一些隱藏於其中的知識，這些知識一般很難從人工處理的方法得到，只有透過電腦強力的運算處理功能後，才可能挖掘出隱含於其中的寶藏知識。

## 二、顧客關係管理的演進

雖然早在50年代，顧客關係管理就已經存在，但是由於相關科技的不發達，使得在收集、分析顧客的客製化資料時，不是那麼方便，而且成本亦相當高，所以並未造成主流。一直到90年代後，隨著資訊科技的發展，如全球資訊網（Web）、網際網路（Internet）、資料倉儲（Data Warehouse）、資料採礦（Data Mining）等，顧客關係管理的觀念才再逐漸受到重視，因此資訊科技可說是顧客關

---

<sup>[4]</sup> Fayyad, U. M., "Data Mining and Knowledge Discovery: Making Sense out of Data," *IEEE Expert* 11:5(1996), pp.20-25.

係管理的促動與催生者。<sup>[5]</sup> Swift 認為支持顧客關係管理的主要 IT 架構應包含資料收集 (Data Collection) 資料儲存 (Data Storage) 資料分析 (Data Analysis) 資料應用視覺化 (Data Application Visualization), 如圖一所示。<sup>[6]</sup>

資料應用視覺化	主管資訊系統	策略資訊系統	決策支援系統	使用者自建系統	網頁			
	結構化查詢	線上查詢分析程序	資料探勘		統計	機器學習		
資料儲存	關聯式資料庫		資料倉儲		資料超市	知識庫		模式庫
資料收集	銷售點情報系統	電子訂購系統	電子數據交換	企業資源計劃	客戶服務中心		問卷調查資料	

圖一：CRM中的各種不同的資訊科技應用架構

資訊時代的圖書館，必須從以往以圖書館的建設工作導向，轉變為以讀者服務為導向 (Service-oriented)；數位化的管理系統，使得圖書館的管理上減少了很多的人工作業，一切的資料管理及報表工作已都能透過資訊工具來完成，因此目前圖書館的經營重心，將致力於對顧客服務的提升。解決此問題的策略在於完成開闢讀者群的交流園地或互動社群，因而可以讓讀者群可以在資訊流的參與中提升讀者群的滿意度，如此一來，CRM觀念的引進，對於圖書館將是非常的重要且迫切。

CRM的簡單概念，就是在適當的時候，提供適當的服務給適當的人，這種以客戶需求為導向的服務，由圖書館的角度來應用，便是在適當的情境之下，推薦某位讀者他可能需要或感興趣的圖書，如當讀者在借閱小說魔戒的時候，我們推薦給他精靈寶鑽；理由是相信對魔戒小說有興趣的讀者，應該對於作者托爾金 (1892) 也不陌生，然而並非每個人都能清楚了解托爾金的所有著作，因此主動推薦有關托爾金其它的著作給讀者，以激發讀者進一步借閱的慾望。以上的例子僅是簡單地以讀者為考量來提出CRM在圖書館上的應用，然而真正在做為推薦時的

<sup>[5]</sup> 蕭振農，「探討壽險業顧客關係管理的策略與運作流程 以南山人壽為例」(高雄市：中山大學管理學院高階經營碩士學程碩士在職專班碩士論文，2003)。

<sup>[6]</sup> 同註5。

依據，是可以往廣度、深度來發展的，如推薦同為描述中古世紀的奇幻小說（非全為托爾金所著）或是推薦與托爾金同一時期的作家，甚至可以推薦長期借閱托爾金小說的讀者所借閱過的小說等。

CRM在圖書館的應用，並非於數位化的時代才開始的，傳統是以人力的方式來推薦讀者相關的資訊，然而圖書的增加和借閱者的次數，使得以人力來運作CRM的效率備受爭議，除非圖書館管理員對於該領域非常的專精或是對該借閱者非常的熟悉，否則單憑記憶的方式要達到真正廣泛服務的價值，終將不可行；因此透過圖書館數位化的記錄分析，來達到CRM的理念，會是較可行、有效率且有品質保證的方法。

CRM重視與顧客的交流，對於圖書館而言，與讀者最密切的交流紀錄莫過於讀者留下的借閱記錄，這些記錄代表讀者的資訊需求或閱讀習慣，因此圖書館要進行CRM就要由這方面著手；以目前圖書館的資訊化情形看來，要取得這方面的資料不難。主要的核心問題有二：即如何由這些記錄找出適合讀者群的規則，並且進一步推薦適合的資料給個別讀者，以符合其的需求；與讀者之間的屬性關切如何被分析出來，好讓個別讀者此項服務感到滿意是CRM當中相當重要的考量因素。

### 三、資料探勘技術理論

「資料探勘」技術通常會依據所要探勘的內容來劃分，不同的需求使用不同的技術來探勘，比較常用的可分為下列幾種法則：

#### (一) 關聯法則 (Association)

關聯法則通常是在一堆很難透過人或傳統的統計方式直接發現其中規則的資料中，找尋潛在的關聯性；透過關聯法則可以探勘出不同事物之間是否存在著某種程度的關聯；因而，關聯法則的考量範圍包括所有物件，可以決定那些相關物件應該聚集在一起。關聯分組在客戶行銷應用上用來確認交叉銷售 (cross selling) 的機會，可以設計出吸引顧客的產品群組，如超市中將相關的物品 (麵包、果醬) 放在同一展示區之內。在圖書館方面，又如在借閱記錄中，去探勘是否書本A若被借閱，則書本B同時被借閱的機會很高，這可能代表A和B的關聯性很高，若能將這些關係找出，則可用於推薦給借閱者。關聯法則常使用到的方法有Rough Set、A Priori 等

## (二) 序列法則 (Sequence)

這個部份和關聯法則蠻相像的，都是在處理資料關聯性的問題，而序列法則基本上是強調在時間順序的因素上，亦即其所運用的範圍大多屬時間相關的規則；如讀者對於書目的閱讀順序是否存在某種程度的規則<sup>[7]</sup>，從這些規則中將可以對讀者做入門書或進階書的推薦。

## (三) 預測法則 (Prediction)

預測法則通常用來利用某些已知資料去預測未知的結果。如在圖書館的管理上，可以利用已知的書籍資料和大部份讀者的借閱習慣，去預測在新書上架後，其主要的閱讀群會是哪些，進而提供預測所需要購買的數量，以達到管理圖書館的效益和借閱者的方便。換句話說，預測法則是根據對象屬性的過去觀察值來預測該屬性未來值。預測經常使用的技巧包括迴歸分析、類神經網路等方法。

## (四) 分類法則 (Classification)

分類法則是運用對資料已知的屬性，通常是明顯的特徵或是由過去的經驗來做資料的分類，推導出資料中是否存在某些規則，進而可以用來對新的資料做出所屬相似類別的預測。如在「顧客關係管理」是依據分析對象的屬性進行分類，並建立群組，如將顧客忠誠度的屬性區分為高度忠誠度者、中度忠誠度者以及低度忠誠度者。常用到的方法有類神經網路、決策樹等。

決策樹是一棵語意樹 (Semantic Tree)，他與一般的資料結構中的樹一樣有節點與樹葉，每一個節點都被安排一個適當的測試，然後利用該測試結果決定資料並將再利用此一節點的哪一棵子樹作為分類的條件繼續做決策，最後透過節點中的測試達到問題分析的目的。<sup>[8]</sup>

## (五) 群集法則 (Clustering)

群集法則是用來產生聚集分類的效果，其和分類法則有同樣的目的，但是過程是完全相反的；分類是透過已知的屬性分類方法去分類，而群集則是利用於當無法找出明顯的屬性可以用來分類時，則利用物以類似的概念，讓資料產生聚集的效果，把相似的資料聚集再一起；通常是透過電腦做資料分群，再依分群之後

---

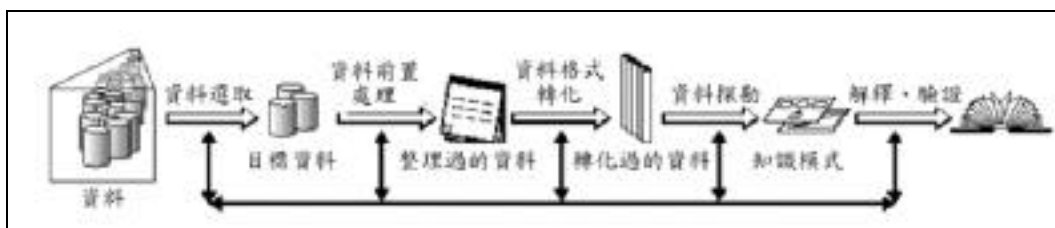
<sup>[7]</sup> 吳安琪，「利用資料探勘技術及統計的方法增強圖書館的經營與服務」(新竹市：國立交通大學資訊科學研究所碩士論文，2001)，頁4，14-20。

<sup>[8]</sup> 邱美珍，「決策樹學習法中連續屬性之分類研究」(中壢市：中原大學資訊工程研究所碩士論文，1995)。

的結果做資料分析、解釋。常用的方法有 K-means、K-medoids、階層式演算法、Agglomeration等。

#### 四、資料探勘的過程

資料探勘是用來發掘出知識的過程，包括一連貫的程序和步驟，如圖二所示，這些過程彼此環環相扣，也可能會出現回饋的現象，即回到前面步驟的可能，Fayyad大致將資料探勘的過程分為九個步驟：<sup>[9]</sup>



圖二：資料探勘的程序 (Fayyad, 1996)

1. 學習應用領域的知識 (Learn the Application Domain)：在資料探勘之前，先要對該領域有一定的認識，包括相關的先備知識和資料探勘的目標。
2. 建造目標的資料來源 (Create a Target Dataset)：從相同或異質的資料庫當中，選取出要拿來應用的資料來源。
3. 資料的前置處理工作 (Data Cleaning and Preprocessing)：在選取出來的資料當中，通常會存在著資料不完整、遺漏或雜訊的現象，因此必須先對資料做一翻整理的工作。
4. 資料的簡化和對照 (Data Reduction and Projection)：根據所要探勘的目標，會有很多資料是重覆的或是不具任何代表意義，因此先對資料做簡化的動作，並找出足以代表資料的特性。
5. 選擇資料探勘的方法 (Choose the Function of Data-Mining)：根據目標來決定要採用的方法，常用到的方法可以參照前面所提到的五個資料探勘技術，如關聯法則等。
6. 選擇資料探勘的演算法 (Choose the Data-Mining algorithm)：接續上一個步

<sup>[9]</sup> 同註4。

驟後，決定要採用何種演算法來完成選用的資料探勘方法，如利用 Apriori 來完成關聯法則的資料探勘。

7. 進行資料探勘 (Data Mining)：經過前面的各個步驟，便可開始對感興趣的目標進行探勘；並且用特定的方式來表示出探勘後所找到的規則。
8. 進行結果的解釋 (Interpretation)：上述探勘後所得到的表示法，必須要做一番合理的解釋，並且設法讓使用者能輕易了解發現的結果；若發現並無任何有用的資訊，則必須要回到前面的步驟重新進行資料探勘。
9. 運用發現的知識 (Using Discovered Knowledge)：當透過資料探勘得到知識之後，必須要設法利用這些知識，如用來做為產品行銷的決策參考依據。

針對九個步驟的實際應用，下面我們以國家圖書館碩博士論文系統為例，目標是想找出使用者下載碩博士論文的記錄是否存在某種規則：

1. 必須先了解碩博士論文系統的資料記錄和使用方法，使後續的步驟能順利，這個步驟最好要與該系統的創建者或維護者進行溝通。
2. 在碩博士論文系統中，背後會有一套複雜的資料庫系統，針對我們的目標挑選出適當的資料或是結合其它的資料庫，如各大專院校本身的碩博士論文，來做為後續探勘的資料庫來源。
3. 所挑選出來的資料，檢查資料是否有問題，如使用者的個人基本資料，可能會有許多遺漏甚至隨便亂填的無謂資料，此時必須針對這些有問題的內容先做處理，予以刪除或是以某一特別的記號來代表。
4. 在資料庫的記錄中有些對我們的目標而言是重覆性的，如同一個人對於同一筆資料下載多次，這可能只是個人忘記檔案存放的位置而再次下載，便應歸類為同一筆資料；另外如果我們的目標是去分析在一固定時間內，例如同一天內，使用者的下載記錄是否存在特別的規則，則必須對下載的時間做轉換，考慮同一天的界限為何，將原先的時間記錄轉換為單純的日期記錄，使其容易進行後續的分析動作。
5. 因為我們的目標是在找出使用者下載論文是否具某種關聯規則，因此選擇關聯法則做為資料探勘的方法。
6. 利用 Apriori 演算法來完成關聯法則的探勘。
7. 實際操作軟體，如 MS-SQL、DMAS，來完成資料探勘的過程。
8. 探勘完成後，對於找出來的規則，必須加以詮釋，如 75% 下載過「網頁探勘」



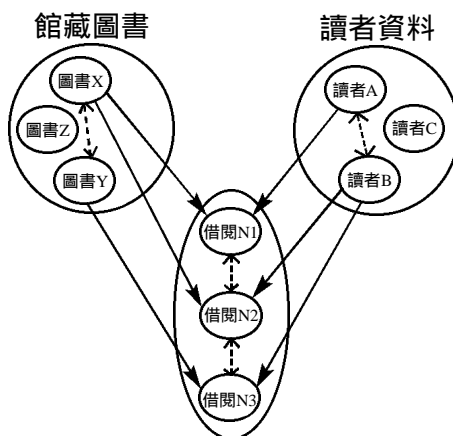
相關文章的人，會下載有關「智慧」代理人的文章，這樣的情形是需要人工去檢視其合理性的。

9. 進一步去應用探勘技術所得到的關聯法則，如上述的關聯性規則可以用來建造相關的「推薦系統」，使下次下載同一類的文章的使用者可以順便接收相關文章；或者是經由大部份關聯性文章的分析後，發現使用者下載文章大都會根據文章內容的引用，這也就等同於建議碩博士論文應該要考慮將相關的引用提供成可以直接下載的連結。

在上面的九個步驟中，值得注意的是在第三、四個步驟，所需的耗掉的精力是佔全部的60%<sup>[10]</sup>，代表在資料探勘過程中的重要部份，事實上是完整資料的準備、淨化和轉換；至於真正探勘的過程（步驟五、六、七）則只大約佔了20%。

## 五、資料探勘在圖書館方面的研究

卜小蝶指出：「圖書借閱行為乃是讀者實際資訊需求的最佳證明，讀者、圖書及借閱行為為三個主體的資料記錄」。三個主體的其關係，詳見圖三所示，這些資料就是用來實作資料探勘的主要內容，利用這些資料來進行下列三種模式的探勘工作：圖書與讀者、讀者與讀者以及圖書與圖書的關係。<sup>[11]</sup>



圖三：圖書、讀者之間的關係

<sup>[10]</sup> Cabena, P, et al, *Discovering Data Mining: From Concept to Implementation* (Englewood Cliffs, NJ: Prentice Hall, 1998)。

<sup>[11]</sup> 卜小蝶，以圖書館借閱記錄探勘加強圖書資利用之探討，《中國圖書館學會會報》，66 (2001)，頁59-72。

針對各項圖書館資料探勘研究的特點，大致可分為三個階段來說明：

### (一) 資料準備階段

在資料準備的階段中，可分為讀者資料和圖書資料兩方面來探討；讀者方面的資料，孫冠華將讀者的身份做概念化階層，推導出一般化 (Generation) 和特殊化 (Specification) <sup>[12]</sup>；所謂概念化階層是將各種屬性以意義的廣狹來定義其階層關係，如大二機械系學生⊂大二工學院學生⊂大學生，當知識是往高層的概念移動，則稱為一般化，反之則為特殊化。而圖書的資料方面，戴玉旻提到可以利用原先圖書館常見的分類方法「中國圖書分類方法」來做分類 <sup>[13]</sup>，同時也可以對圖書做概念化階層的推導。

以上是針對輸入模式的處理，其強調的是個人化資料模組的建立，如何建置、記錄相關紀錄，包含使用者的個人資料和行為記錄。

### (二) 探勘階段

Pu指出以讀者借閱習慣的時間屬性來對讀者做分類，然後找出和圖書的關聯性 (Association)；在找尋關聯性的法則時，常會運用到最小支持度的觀念，最小支持度就是用來定義符合關聯法則的最小門檻，以避免無意義的少量關聯；吳安琪認為除了運用 Apriori做關聯性的探勘之外，同時也採用修改後的 GSP (Generalized Sequential Patterns) 演算法來對圖書借閱的順序做探勘 (Sequential)，找出圖書被借閱的順序關係 <sup>[14]</sup>；而在戴玉旻的研究中，則是運用 H-mine的探勘關聯性和Apriori做比較，同時對不同的讀者族群加入多重最小支持度 (Multiple Minimum Support) 廣義關聯法則，以符合不同族群的需求，找出學校內不同系所的讀者的不同規則；所謂多重最小支持度廣義關聯法則就是對不同階層的分類給予不同的最小支持度，以找出各階層的合適規則。 <sup>[15]</sup>

王毓菁指出：「採用天真貝式分類法 (Naive Bayes Classification) 來對館藏圖

---

<sup>[12]</sup> 孫冠華，「圖書館新書推薦之個人化服務方法」(高雄市：國立中山大學資訊管理研究所碩士論文，2000)，頁10-11。

<sup>[13]</sup> 戴玉旻，「圖書館借閱記錄探勘系統」(新竹市：國立交通大學資訊科學研究所碩士論文，2002)，頁2。

<sup>[14]</sup> 同註7，頁16。

<sup>[15]</sup> 同註13，頁23-27。

書做分類」。<sup>[16]</sup> 當考量圖書和讀者資料漸增的情況下，如何改善資料探勘的效率，因此針對多層更新探勘法 ( Multilayer Update Miner, MUM ) 和Apriori方法做效率的比較，以驗證多層更新探勘法可以得到較好的效率。<sup>[17]</sup>

以上是針對處理 ( Process ) 的部份，其使用到相關的智慧技術，如資料探勘，才能處理複雜資料關係，以得到讀者可以借書行為，詳如表一。

表一：近年來國內對於圖書館資料的探勘研究方法一覽表

研究年份	技術	意義
Pu,1999	Apriori (關聯法則)	在找尋關聯性的法則時，常會運用到最小支持度的觀念以避免無意義的少量關聯
吳安琪，2001	GSP演算法 (序列法則)	對圖書借閱的順序做探勘 ( Sequential )
戴玉旻，2002	H-mine (關聯法則)	應用不同的最小支持度，對不同的族群找出不同的規則
王毓菁，2002	貝式分類法 (分類法則)	對館藏圖書做分類
曹健華，2003	MUM, Apriori (關聯法則)	資料漸增的情況下，改善資料探勘的效率

### (三) 分析及應用階段

在大多數的研究上，是以找出讀者和圖書的關聯性為主，以做為新書或相關圖書的推薦依據。<sup>[18][19]</sup> 吳安琪指出：「探勘圖書借閱的順序規則，使得圖書的推薦更具動態的引導效果，可以運用在引導學習者進入某領域知識的推薦上」。<sup>[20]</sup> 又如針對圖書的自動分類研究的目的是在於可以將未分類的圖書做合理的自動分類，以解決圖書管理人員本身知識不足或主觀意識可能造成的不適當圖書分類。<sup>[21]</sup> 在

<sup>[16]</sup> 王毓菁，「圖書館閱覽者群組潛在特徵探勘資訊系統」(臺北縣：華梵大學工業管理學系碩士論文，2002)，頁6-10。

<sup>[17]</sup> 曹健華，「應用資料探勘技術於數位圖書館之個人化服務及管理」(嘉義縣：南華大學資訊管理研究所碩士論文，2003)，頁31-40。

<sup>[18]</sup> 同註12，頁5-6。

<sup>[19]</sup> 同註13。

<sup>[20]</sup> 同註7。

<sup>[21]</sup> 同註16，頁25-26。

曹健華的研究中，則是採用了多層更新探勘法來提升效能。以上針對輸出的服務要以提供使用者適合的推薦為主，並且透過適當的方式來傳送給使用者，如 E-mail、個人化網站與行動裝置等。<sup>[22]</sup>

由以上歷年來的研究，不難發現在圖書館的資料探勘工作上，大致是在研究讀者的借閱記錄，並且從中利用各種的資料探勘方法，以求找到真正有代表性的記錄規則，最後可以利用在圖書館館藏的推薦上，增加館藏的利用率；資料探勘可說是在圖書館顧客關係管理應用的基礎階段，了解讀者的行為、讀者可能的需求，才有可能為讀者做出適當的服務。

## 六、圖書館個人化的顧客關係管理推薦

圖書館的CRM必須考慮到讀者個別化的問題，同時了解讀者的需求、讀者的狀態以及讀者對推薦後的滿意度；依照讀者量身訂做的服務，也就是達到個人化 (Personalization) 或所謂客製化的服務。施毓琦指出，大學圖書館的網站，使用者傾向於需要個人化的服務，且重視個人化資訊環境及相關功能，另外也指出了使用者身分別對個人化服務之需求達到顯著性差異，代表了真正因人而異的服務是有實質需求的。<sup>[23]</sup>

「個人化」(Personalization) 的推薦可以分三個類型來探討：

### (一) 非個人化的推薦

完全沒有依使用者的特性過濾訊息，像是傳統的宣傳單推銷方式一樣，直接將所有的訊息送給使用者。

### (二) 短暫的個人化推薦

通常是根據使用者所提供的資料或當時的狀態來提供適合的訊息，像是一般網站的搜尋引擎，是當使用者鍵入關鍵字後，經由搜尋引擎的過濾來提供資訊。目前國內大部份的圖書館所提供給讀者的服務，以此種方式最常見，一般是提供圖書的各種屬性給使用者做為查詢的依據。

### (三) 持續的個人化推薦

考慮到人的因素，根據使用者相關資料或長久的使用記錄來進行資訊過濾，

<sup>[22]</sup> 同註17，頁25。

<sup>[23]</sup> 施毓琦，「大學圖書館網站個人化服務之使用者需求研求」(臺北市：國立臺灣大學圖書資訊研究所碩士論文，2002)，頁13-16。

以達到真正個人化的需求，如在圖書館中，可以利用使用者的身份或長久的借閱記錄來進行個人化的圖書推薦。

目前許多網站均有個人化服務的概念，如國外知名收費期刊文章的收集網站 ACM，如圖四所示。該網站中，每位使用者有自己的個人目錄，可以訂定各種規則來對有興趣的文章做分類，而該網站所收集的文章，若有符合使用者所訂定的規則，會自動落入個人目錄，而且一旦有新文章，也會於下次使用者登錄該網站時告知使用者，如此一來可以讓使用者對感興趣的領域做長期的文獻追縱，減少許多使用者再進行搜尋的工夫。類似個人化服務的功能，在其它網路購物或入口網站也都紛紛加入，如 Yahoo、ebay 等，目的在於提升使用者對該網站的向心力和滿意度。

使用者用自己的帳號登入後，即可看到系統給予的文章推薦目錄，並且自動歸類，讓使用者可以很迅速地查看自己感興趣的文章。



圖四：ACM網站的個人化服務

此型態的推薦是屬於人為導向與內容導向二種。前者是加入合作過濾方法 ( Collaborative Filtering ) 而建立了所謂的個人化館藏推薦系統，做為推薦的依據；所謂合作過濾方法，是使用者利用同儕對於某物品既有的評價來預期自己對該物品可能的評價。<sup>[24]</sup> 後者是運用了內容導向過濾 ( Content-based Filtering ) 的方法，

<sup>[24]</sup> 余明哲，「圖書館個人化館藏推薦系統」(新竹市：國立交通大學資訊科學研究所碩士論文，2003)，頁3-4。

根據館藏的各種屬性篩選出相關的書籍，便可以解決合作過濾方式無法有效推薦新書或冷門書的問題。

## 七、亞馬遜書店的成功例子

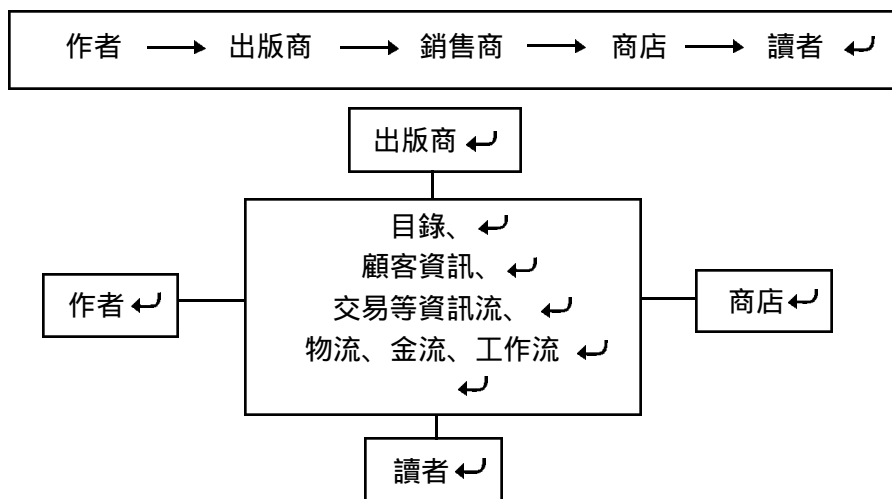
亞馬遜書店 (Amazon.com) 為全球網路書店的領導者，也是結合實質和虛擬價值鏈來增加市場競爭力的典範，Amazon.com 的策略從作者、讀者和書評的回饋得到有用的資訊，並且讓價值鏈中的不同成員都能從中受惠，如其運用「e化的顧客關係管理」(electronic-Customer Relationship Management, e-CRM) 的概念，追蹤潛在顧客的線上行為，然後提供相似顧客有關的建議。此有別於與傳統盲目地透過電子郵件來通知廣大的顧客族群。因為它選擇提供更多個人化服務來造成更多的加值作用。

另外，亞馬遜書店除了提供顧客 B2C 的服務之外，也提供出版社和書店 B2B 的服務，並且讓兩者的服務能相互支援，強調關鍵在於如何加強實值和虛擬價值鏈之間的關係，如顧客資料庫除了可提供顧客更多個人化的資訊，也可以提供書商或出版社相關的訊息。在亞馬遜書店的策略中，有兩項是和 CRM 相關的，其一是社群的經營，讓作者和讀者能快速獲得有關書籍的相關訊息；其二是核心資訊的管理，使系統用來建造和提供書籍目錄和記錄顧客的資料庫及相關的加值資訊。

透過以上的兩項策略，使傳統購書行為由單純且缺乏交流的線性結構，如圖一之上方，轉變為更具有多元網路結構交流的可能性，如圖五之上方。轉變成出版商和書店都能透過亞馬遜書店獲取讀者群的相關資訊，如圖五之下方；再加上分析、電子行銷和個人化的功能吸引更多的顧客與提高顧客的忠誠度。<sup>[25]</sup> 亞馬遜書店結合了資訊流和實值書本，快速適應動態的數位環境，透過線上服務來提供給顧客加值的功能；這之間的關係造就了該公司與顧客之間的良性循環，在顧客參與之後而充實了它的知識庫，而系統便有機會根據知識庫的內容更了解顧客，而提供更好的服務。

---

<sup>[25]</sup> Combe, C., "The Management of E-Commerce Strategies for Sustaining Competitive Advantage in the Online Bookselling Industry: The Case of Amazon.com.," *International Journal of e-Business Strategy Management* (November/December, 2002), pp.153-165.



圖五：傳統購書行為的線性結構和以CRM概念為主的網路結構

由圖五下方數位化進行資料探勘，其中企業成功與否的指標在於企業的資訊流（Flow of Information）、 workflow（Flow of Work）、物流（Flow of Logistics）與金流（Flow of Cash）等四流的優劣。在數位時代中，網路上的資訊流已成為交易降低成本又可讓身處不同時空的人們可以快速聯繫和互動的工具。亞馬遜書店資訊流的主要功能在於：傳遞、轉換及儲存資料，以提高作業效率、提升管理能力、增加競爭能力及掌握顧客需求。

總之，亞馬遜書店有了這些資訊成為對有特殊需求的人提供客制化服務的基石，這些資訊可以由供貨商（上游）、顧客（下游）、與企業之間的溝通來取得有意義的資訊<sup>[26]</sup>，而這些參與者活動加值與深度決定了數位虛擬市場的價值。<sup>[27]</sup>在顧客關係管理價值鏈中，資訊流扮演供應機制的角色，而在虛擬價值鏈中則擔任決策支援的功能；企業目的就是在於將虛擬價值鏈中的資訊價值，轉變為實值價值鏈的加值行為。<sup>[28]</sup>而現代圖書館的經營就可由亞馬遜書店透過資料探勘技術而成功的提高顧客忠誠度的典範，以作為圖書館的顧客關係管理依據。

<sup>[26]</sup> Slater, D., "The Power of Positive Thinking," *The CIO* 15(August, 1998), pp 31-36.

<sup>[27]</sup> Amit, R and Zott, C., "Value Creation in E-Business," *Strategic Management Journal* 22(2001), pp 493-520.

<sup>[28]</sup> Rayport, J and Sviokla, J., "Exploiting the Virtual Value Chain," *Harvard Business Review* 73:6(1995), pp 75-85.

## 八、結 論

現代圖書館已轉變成以讀者服務為導向，數位化的管理系統，使得圖書館的管理上減少了很多的人工作業，一切的資料管理及報表工作大概都能透過資訊工具來完成，因此目前圖書館的經營重心，將致力於對顧客服務的提升，所以「顧客關係管理」觀念的引進，對於圖書館將是非常的重要且迫切。本文提出以讀者為主體的客戶關係管理來服務讀者。

目前由於網際網路的環境的發達，已可提供學習者廣度與深度的不同需求，再加上網路提供了超越時間和地域的學習利基。為了配合新的學習環境，圖書館應如何提升讀者的滿意度就是未來成敗的關鍵地位。數位化圖書館中，透過讀者的借閱資訊即可深入了解讀者的特質與其借閱歷程。同時，資料探勘技術可用於引導讀者進行適性化的書籍推薦。圖書館有了此分析的利器，一方面可獲取新讀者，另一方面亦可提高既有讀者的忠誠度。如此大大的超越了資訊科技只被用來蒐集與儲藏資料的工具，未來將可利用資料探勘技術亦可作為推動數位學習網，與讀者群形成良性互動，為圖書館必然的典範轉移。

## 參考文獻

- 卜小蝶。以圖書館借閱記錄探勘加強圖書資利用之探討，《中國圖書館學會會報》，66 (2001)，頁59-72。
- 王毓菁。「圖書館閱覽者群組潛在特徵探勘資訊系統」(臺北縣：華梵大學工業管理學系碩士論文，2002)。
- 余明哲。「圖書館個人化館藏推薦系統」(新竹市：國立交通大學資訊科學研究所碩士論文，2003)。
- 吳安琪。「利用資料探勘技術及統計的方法增強圖書館的經營與服務」(新竹市：國立交通大學資訊科學研究所碩士論文，2001)。
- 林東清、劉勇志。企業顧客關係管理(CRM)的一些主要經營架構與流程模式，載於：淡江大學主編，《第十三屆國際資訊管理學術研討會論文》，頁569-578。
- 邱美珍。「決策樹學習法中連續屬性之分類研究」(中壢市：中原大學資訊工程研究所碩士論文，1995)。
- 施毓琦。「大學圖書館網站個人化服務之使用者需求研求」(臺北市：國立臺灣大學圖書資訊研究所碩士論文，2002)。



- 孫冠華。「圖書館新書推薦之個人化服務方法」(高雄：國立中山大學資訊管理研究所碩士論文，2000)。
- 張晉綸譯，Ray McKenzie著。《究極CRM》(臺北：麥格羅希爾，2001)。
- 曹健華。「應用資料探勘技術於數位圖書館之個人化服務及管理」(嘉義縣：南華大學資訊管理研究所碩士論文，2003)。
- 蕭振農。「探討壽險業顧客關係管理的策略與運作流程 以南山人壽為例」(高雄：中山大學管理學院高階經營碩士學程碩士在職專班碩士論文，2003)。
- 戴玉旻。「圖書館借閱記錄探勘系統」(新竹市：國立交通大學資訊科學研究所碩士論文，2002)。
- Amit, R and Zott, C., " Value Creation in E-Business," *Strategic Management Journal* 22(2001), pp 493-520.
- Cabena, P, et al, *Discovering Data Mining: From Concept to Implementation* (Englewood Cliffs, NJ: Prentice Hall, 1998).
- Combe, C, " The Management of E-Commerce Strategies for Sustaining Competitive Advantage in the Online Bookselling Industry: The Case of Amazon.com.," *International Journal of e-Business Strategy Management*, (November/December, 2002), pp.153-165.
- Fayyad, U. M., " Data Mining and Knowledge Discovery: Making Sense out of Data," *IEEE Expert* 11:5(1996), pp.20-25.
- Rayport, J and Sviokla, J., " Exploiting the Virtual Value Chain," *Harvard Business Review* 73:6(1995), pp 75-85.
- Slater, D., " The Power of Positive Thinking," *The CIO* 15(August, 1998), pp 31-36.
- Tiwana, A., *Essential Guide to Knowledge Management--The e-Business and CRM Application* (Upper Saddle River, NJ: Prentice-Hall, 2001).

# CRM and Data-Mining in Librarianship

Jin-tan Yang Wen-chih Chen Wei-chin Chiu Yung-fang Chu

## Abstract

The purpose of this article is to explore the behaviors of readers by the way of data mining in order to attain new readers and promote the loyalty of existing readers as new model management in a library. Such kind of digital library really offers efficient and convenient in terms of Customer Relation Management (CRM). Furthermore, there are many important issues on the application of readers' profiles such as personalized recommendations or predictions based on readers' transactions. All of those recommendations can be attained by data mining techniques. To some extent, data mining will be an embedded function for promoting high efficiency and effectiveness in the future digital

---

**Keywords ( 關鍵詞 ) :** Data Mining ; Customer Relation Management ; Adaptive Service ; Digital Library ; Collaborative Filtering

資料探勘；顧客關係管理；適性化服務；數位化圖書館；合作過濾

Jin-tan Yang : Professor and Dr., Center of General Education, National Kaohsiung Normal University ; Professor and Dean of College of Digital Design at Southern Taiwan University of Technology ; E-mail: yangdav@mail.nknu.edu.tw

Wen-chih Chen : Graduate student of Information & Computer Education at National Kaohsiung Normal University ; E-mail:koach@ccsh.kh.edu.tw

Wei-chin Chiu : Director of Library at National Ping-Tung University of Science and Technology ; E-mail: chiuwei@mail.npust.edu.tw

Yung-fang Chu : Graduate student of Information & Computer Education at National Kaohsiung Normal University ; E-mail: fang@mail.fhjh.tn.edu.tw

library. This paper discusses the satisfaction of readers by CRM and the profiles of books and readers ' borrowing behaviors by data mining techniques. Thus, the CRM can provide readers highly efficient management and services. Data mining for readers at library will be both applied to collect and store digital documents, and provide a new paradigm shift as e-Learning web in the days to come.