

第二章 文獻探討

第一節 群體智慧

群體智慧的研究，源自於對社會性動物之群體行為的觀察。例如，螞蟻群找尋食物並尋求最短路徑、魚群和鳥群的集體遷徙行動、蜜蜂找尋築巢地點等等。其中又以螞蟻的行為最具代表性：螞蟻行進的路徑會留下費洛蒙，但是費洛蒙會隨著時間而揮發，所以距離食物來源較短的路徑，費洛蒙的殘留量會比較多，揮發的少。大部分的螞蟻會循著費洛蒙濃度較高的路徑走，為這條路徑產生正增強；經過一段時間，蟻群們會突現 (Emerge)、找出一條較佳、較短的，通往食物的路徑 (Fleischer, 2005)。這樣的群體智慧運作，特點在於自我組織性 (Self-Organized)。群體之中並沒有一個強制性的，由上而下的控制力量；但當每一個個體都遵循著某些簡單、重覆性的行為與互動準則時，卻可以造就群體的複雜行為，達到特定的目標 (Sumpter, 2006)。

(Sumpter, 2006) (2006) 指出，群體智慧行為的準則有下列九項：

1. **完整性與多樣性 (Integrity and Variability)**：如果蟻群中的每隻螞蟻都只沿著相同的路徑走，那麼群體中並沒有辦法得到多樣的意見，自然也不會有群體智慧的產生。因此，個體需要具有高度的多樣性。
2. **正回饋 (Positive Feedback)**：群體透過模仿或追隨他人的行為，將由個體們所貢獻的意見中，較佳者予以增強、散佈開來，使較佳的解決方案能被突現出來。例如，費洛蒙濃度較強的路徑代表的是一個暫時的最佳解，而螞蟻依循著這個路徑行走，又會在再次在該路徑上遺留下費洛蒙，使得其濃度更加的強烈，這就是正回饋。但是在多樣性及正回饋之間必需要取得平衡，才能夠有效的解決問題。舉例來說，著名生物現象——螞蟻的「循環磨 (circular mill) 悲劇」，即是僅有正回饋，而缺乏多樣

性與獨立判斷的例子。因為每一隻螞蟻都只跟著前面的走，而沒有試圖找新的路，最後型成一個循環的路徑，反覆繞行直到力竭而死 (Surowiecki, 2005)。

3. **負回饋(Negative Feedback)**：正回饋可以建立一個群體的行為模式，而負回饋則是使它更穩定。例如，當蟻群找到一個食物據點後，無限的正增強會使得晚到的螞蟻面臨食物已被搬運一空的窘境。因此，負回饋可以驅使後到的螞蟻轉移陣地，尋覓其他的食物來源，維持群體智慧運作的穩定性與永續性。
4. **反應臨界值 (Response Threshold)**：動物的群體行為，會在突破某個臨界值後改變行為屬性。以人類來說，如果街頭上只有一個人突然抬頭凝視天空中某個方向時，旁觀者只有 40% 的機會跟著往那個方向看去。但這個機率會以指數的方式增加，當有五個人都同時仰望天空時，突破了臨界值後，旁人好奇抬頭的機率會迅速增加到 80%，如圖 2 - 1 所示。

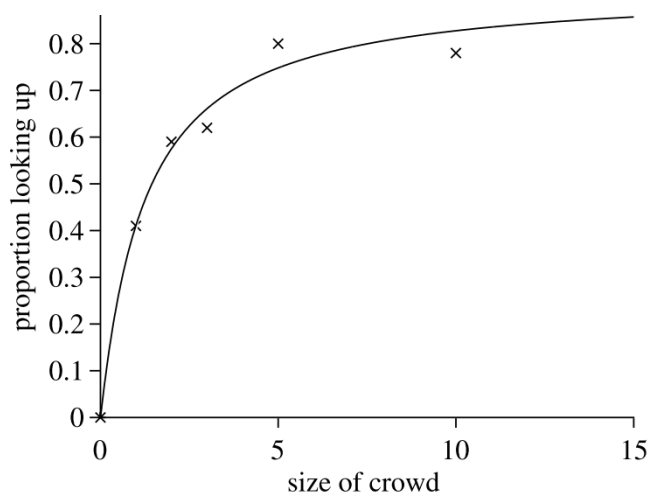


圖 2 - 1 群眾之群體行為反應臨界值
資料來源：(Sumpter, 2006)

5. **領先者 (Leadership)**: 雖然自我組織性似乎不該有領導者的角色，但是科學家發現，在動物的群體行為中，仍然會有少數個體扮演著領先者。例如，魚群的遷徙中，事實上只有少數能掌握方向的個體，就能帶領整個族群前進，其他的個體僅是靠著正回饋或負回饋，跟隨其他的同伴前進。
6. **抑制 (Inhibition)**: 消極的抑制行為，和負回饋比較相近；而積極的抑制行為，則是群體的成員可以主動的去抑制某一行為的發生，而引進另一解決方案。例如，當螞蟻在選擇適當的築巢地點時，抑制的機制會讓蟻群在多個可能的地點之間進行選擇，最後某個地點會達到臨界數量，而其他地點則被抑制。
7. **重覆性 (Redundancy)**: 昆蟲社會的運作，幾乎不會有毀壞 (Crash) 的情形發生。不像電腦系統，每個各別的元件都負擔有特殊的任務，任何零件的故障都可能導致整個系統的停擺。昆蟲的社會中，成員角色的重覆性相當高，因此部分成員的消失並不影響到整個群體的運作。
8. **同步性 (Synchronization)**: 同步性最具代表性的例子，就是觀眾的鼓掌行為。同步性可以視為是一種正回饋，但是是時間上的回饋，而不是空間上的回饋。
9. **利己性 (Selfishness)**: 什麼樣的動機，會讓群體中的個體願意與其他同伴合作，參與群體行為的運作呢？Sumpter 認為是利己的動機考量，加上天擇的結果 (2006)。個體以自身的存活為前提，在參與群體運作之下，進行各種生存的行為，比起單打獨鬥可以更容易存活下來。透過天擇的結果，使得該物種逐漸趨向群體行動。

(Surowiecki, 2005)則認為，群體智慧的產生，必需具有下列四個要素：Cognitive Diversity (多樣性)、Independence (獨立性)、Decentralization (分權化)和 Easy Aggregation (整合機制)。

1. **多樣性**：如果群體的成員組成同質性很高，可能在決策時會採取類似的思考模式，提供的資訊趨於一致，妨礙群體智慧的產生。相對的，如果成員們能提供比較多元的思考方向，則會增加整個群體的廣度，減少盲點的發生，激發出更有價值的結果。
2. **獨立性**：承上，當成員具有多樣性後，如果個體未能獨立思考，純粹以旁人的行為作為自己決策的依據，則個體無法充分的貢獻其獨特的想法，自己所擁有的資訊將會被隱藏，或者變得只接受來自外界的資訊。這就是缺乏群體智慧獨立性的一種典型，個體依舊無法對群體提供價值，最後使群體智慧的結果產生偏差。
3. **分權化**：簡單來說，是鼓勵竭盡局部知識做到專業化，使個體能依其自身經驗與能力，來處理最擅長的問題。然而，此分權化的概念必須在群體內部具有一種能整合（aggregation）、溝通各個子集資訊的管道或機制，才能做成較好的群體決策。
4. **整合機制**：一種將個別資訊轉換成群體決策的機制，也就是個體透過內在的文化跟習俗、外在的規範法律或者該種群體的共通經驗等等默契，吸收並分析外在的資訊後，做出判斷，自然而然的整合成為群體智慧的運作結果。例如公共場所人群的移動，看似各人按照自己的意志行走、毫無規則可循，但潛在的統合力使得群體能夠自然地找到移動的節奏與方向。排隊、先到先排也同樣概念。

(Heylighen, 1999)認為，群體智慧是群體所擁有的解決問題的能力，並將群體智慧的產生機制，歸納為三種（1999）：

1. **偏好的統合（Averaging Preference）**：達到群體智慧、並解決衝突的最簡單的方法，就是透過投票。個體透過投票來表達其偏好，然後經過統合來決定群體的方向。不過參與的成員數量，在達到一個數值後，對決策的幫助會減少。

2. **回饋 (Feedback)**: 投票機制的缺點在於，每個參與者都必需對問題有全盤的了解，才能夠在所有的選項中挑選出較符合個人偏好者。螞蟻的行為，可以視為是回饋機制的典型代表。每隻螞蟻並不需要走過整個環境來做決定，但是由費洛蒙機制來統合所有的路徑後，較好的路徑可以透過正回饋的循環達到增強，而不好的路徑則會慢慢的衰減。

回饋機制的危險在於，當回饋的強度過於強烈時，暫時性的最佳解會得到過多的重視，使得其他的可能解失去了被檢視的機會，妨礙真正的最佳解的產生。這個問題，可以藉由控制個體對於當下的群體偏好的敏感度來解決。德菲法即是這樣的機制下所設計的群體決策機制。

3. **分工 (Division of Labor)**: 每個個體有著不同的專業性，因此他們於問題的貢獻，會侷限在其所擅長的部分。當個人專注於自己的專長，而能夠統合各個個體的貢獻時，群體智慧便能含蓋較廣的範圍。不過分工時也必需注意到，如果各個專業領域的差距過大，領域和領域之間便很難彼此了解，難以統合出群體的成果。因此，最好在領域與領域之間有著適當的重疊性，讓資訊在各個領域之間更能順暢的流動、交換。

第二節 社會性書籤網站

在 Web 2.0 的風潮下，許多強調使用者參與、分享、互動的網站因運而生，社會性媒體 (Social Media) 或社會性軟體 (Social Software) 即是這個氛圍下的產物，社會性書籤網站也屬於其一。這類型的網站具有下列的特色 (Lerman, 2007a):

1. 使用者可以藉由不同的型式來創造或貢獻內容。
2. 使用者以標籤 (tag) 的方式將內容給予註解。
3. 使用者可以主動或被動的方式給予內容評價，例如投票是一種主動的評價，而系統透過分析使用者的行為，則是一種被動的評價。

4. 使用者可以彼此設定為聯絡人或好友，建立起社交網路。

社會性書籤網站讓使用者分享自己所喜歡的網路連結，並且透過所有的使用者來為這些網路內容分類、評價，藉此共同篩選這些共享內容的品質。歸納來說，社會性書籤網站大概有下列幾項重要的功能模組(Hammond, Hannay, Lund, & Scott, 2005; Lerman, 2007a, 2007b; Lund, Hammond, Flack, & Hannay, 2005)：

1. **文章評選機制與突現的首頁頭條 (Emergent Front Page)：**網路時代所面臨的共通問題，就是資訊的爆炸，帶給使用者的負擔。如何去評定內容的品質，是網路時代資訊處理的一大課題，社會性書籤網站所採取的方法為透過使用者的「投票」推薦。使用者可以將文章連結分享至網站上，然後由所有的使用者對其進行評價，決定是不是值得推薦。首頁的部分，好比報紙的新聞頭條概念。不過相較於傳統媒體是由少數的權威主編決定頭條，書籤網站的頭條不經過人為的選擇，而是根據使用者對各篇文章的推薦情形，透過系統的計算來選出頭條，並放至首頁上。為避免有心人士操弄、鑽規則的漏洞，系統的計算規則通常不會公開，而且會時常調整。首頁中除了最熱門的頭條，還會有最新的文章專區，讓最近被收錄至平台上但尚未得到充份推薦的文章得到曝光的機會。



圖 2-2 HEMIDEMI 共享書籤上的首頁頭條
資料來源：(“HEMIDEMI - 黑米共享書籤”)



圖 2-3 HEMiDEMi 黑米共享書籤(左)及 funP 推推王(右)的「標籤雲」

資料來源：“funP 推推王”；“HEMiDEMi - 黑米共享書籤”

2. **後設資料 (Meta Data) 處理機制：**如何將各式各樣的內容給予後設資料 (Meta Data) 以方便檢索與瀏覽，是資訊處理的重要課題。書籤網站同時採用兩種方法，一個是傳統的資料分類，亦即預先設定好各種文章類型，如新聞、科技、旅行、美食、財經、藝文等等，並且為每篇文章給予一個特定的分類；另一個則是民俗分類法 (folksonomy)。使用者可依照自己對於該內容的解讀，給予數個辭彙當作標籤 (tag)。

相較於傳統分類必需事先完善的定義好所有可能的類別，且一則內容僅能擁有一則分類，民俗分類法擁有相當大的動態與彈性。每個人對同一個標的物的認知、解讀可能有不同的角度，藉由集合所有使用者對同一篇文章所自由給予的標籤，更能為該文章賦予全方位的後設資料 (Meta Data) 標記，且所運用的辭彙可以與時並進。由標籤所組成的超連結，可以方便讀者在閱讀文章時去追蹤所有涵蓋相關主題或概念的內容，而所有標籤所組成的「標籤雲」，則會以顏色及字型的大小，來呈現出最近的文章趨勢。民俗分類法的使用當然也非十全十美，但並不在本研究的討論之內。

3. **社會關係網路：**使用者可以彼此結交為朋友或聯絡人，追蹤聯絡人的動

態，以及聯絡人所推薦、發表或收錄的文章。此外，每篇文章的收錄者、推薦者，以及每個使用者所收錄、推薦的文章皆一目了然。一來，從社交的角度來看，使用者可能想要關心朋友、了解朋友們最近在關心什麼議題；二來，從內容群體過濾的角度來看，朋友之間可能比較會有共通的興趣，朋友的所接觸的文章，可能也是自己所可能感興趣的文章；而同一使用者可能長期對某一相關議題擁有興趣，因此透過使用者之間的社會關係網路，更容易找到自己喜歡的資訊內容。



圖 2-4 HEMiDEMi 黑米共享書籤中，透過文章串聯人際網路
資料來源：(“HEMiDEMi - 黑米共享書籤”)

除此之外，部分社會性書籤網站也有類似群組的設計，集合對某一議題有興趣的使用者，也是一種社交機制。群組的概念可以視為分類及社交的結合，並且可以反應 web 2.0 中長尾效應的概念，具有某一特定興趣的小眾可以自成一箇天地，擁有自己的空間，並能和志同道合的使用者互動。因此，具有頭版頭條設計、以及分眾概念的社會性書籤網站，可以說是結合了傳統媒體和 web 2.0 精神，同時具備大眾傳播及小眾傳播的效果。

綜合上述的幾項功能模組，Klaisubun, Kajondecha, & Ishikawa (2007)，發現使用者通常利用標籤 (tag) 來找尋某一特定主題的文章，然後利用這些文章的推薦人、收錄人來找尋該主題下的相關資訊，如圖 2-5 所示。另一方面，Lerman 則認為有三個因素會影響社會性書籤網站上每篇文章的可見度

(visibility) (2007a, 2007b)，分別是：(1)文章是否顯示於首頁中，首頁中的文章可以獲得最佳的能見度；(2)文章是否顯示於最新文章序列中，新進文章會列在文章序列的前面，故能見度也較高；(3)以及透過朋友的關係而被看到，亦即朋友越多，曝光的機會越大。

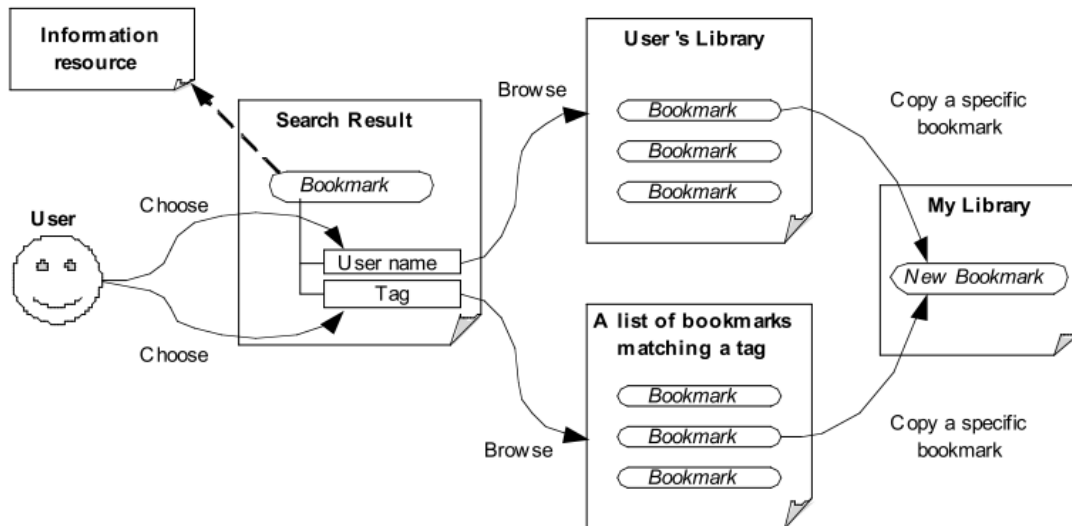


圖 2-5 使用者在社會性書籤網站的資訊搜尋行為
資料來源：Klaisubun, Kajondecha, & Ishikawa, 2007

第三節 資訊彙整

不論從資訊管理的觀點或者是媒體的觀點，核心的精神都是將各種資訊加以彙整、處理，為使用者創造價值。

Zhu, Siegel, & Madnick (2001)將資訊彙整解釋為，透過多個資料來源來匯集、整理相關資訊，並且協助個人或企業能有效的掌握資訊。而資訊的彙整可以分為四個流程，分別是資訊擷取 (Retrieval)、資訊過濾(Filtering)、資訊整合(Integration)、資訊評等與排序(Ranking)：

1. **資訊擷取**：從不同的資料來源來匯集相關的資訊。

2. **資訊過濾**：將匯集的資訊因應不同的需求來做適當的過濾。
3. **資訊整合**：將過濾後的資訊整合。
4. **資訊評等與排序**：將整合後的資訊透過不同的指標、權重的運算來達成適當的排序。

Rayport & Sviokla 也提到類似的概念(1995)。他們從企業的價值鏈出發，認為資訊的價值不該只是傳統價值鏈中屬於支援性活動的角色，而是透過資訊的處理，資訊本身可以帶來許多的價值。資訊的處理也有所謂的虛擬價值鏈，資訊經過獲取(Gather)、組織(Organize)、選擇(Select)、綜合(Synthesize)、分送(Distribute)等再製過程後，會再創造新的價值：

1. **獲取(Gather)**：蒐集、並累積資訊
2. **組織(Organize)**：將蒐集到的資訊以適當的方式排列，讓後續的取用及分析更方便、有效。
3. **選擇(Select)**：從資訊庫中辨識、並萃取出所需要的資訊。
4. **綜合(Synthesize)**：將資訊分裝 (Packaging)，以符合具有特定目的、意圖的使用者的需求。
5. **分送(Distribute)**：將分裝好的資訊傳遞給特定的使用者或顧客。

從新聞學的觀點來看，媒體的核心價值之一——守門(gatekeeping)的過程，就是一種資訊的處理。媒體透過團隊組織，在新聞蒐集乃至於新聞處理的過程中，為讀者／觀眾進行訊息的價值選取與判斷。新聞媒體在其中扮演守門的角色，進行訊息的把關與處理。守門行為包括新聞的蒐集、選擇與處理程序(Baran & Davis, 2005; Bass, 1969)，在實務上包括訊息的型塑、凸顯、時段安排、淡化或保留、重複等等(劉從哲, 2004)。以報紙為例，由報社、記者反覆的判斷、篩選新聞的價值後，經過採訪、撰稿，產生新聞資訊；接著，由編輯在稿件中篩選，並將若干稿件依訊息的主題性、關聯性、對比性等等，加上配稿、套稿的處理組合至版面當中，然後以適當的版位安排、標題及照片圖表的格式設計，完成版面的組合，

最後印刷後傳遞到讀者手中。

進入網路新聞的時代，網路新聞的資訊加值鏈可以分為供應端鏈、增值鏈和輸出鏈三個部份(徐世平,2004;轉引自周慶祥, 2004))。新聞前端的供應鏈代表新聞內容的來源，在這個階段完成原始新聞素材之產出；而新聞網站位於中游，做為增值的角色，進行資訊的流程控制，創造資訊的價值；最後的新聞輸出鏈，則是新聞產品真正與使用者的接觸點，使用者透過網站所提供的介面來得到新聞產品，同時網站經營者也企圖在輸出產品的同時，利用資訊科技來蒐集、分析使用者的偏好，以不斷的調整、改進，提供使用者更好的新聞服務產品。

其中新聞增值的方式又包括了篩選、分類、整合和衍生四類：

1. **篩選**：意謂新聞網站對於來自供應鏈的資訊價值進行判斷，為網站端的新聞守門過程，篩選過程就是資訊加值。
2. **分類**：將新聞資訊做有價值分類，將新聞分派至既有的欄目中。
3. **整合**：透過「內容整合」或「形式整合」的方式。內容的整合包括同類事件的整合、某一特定時間內事件的整合、同一事件不同角度之報導整合等；而形式整合則是多方運用文字標題、圖片、多媒體等形式，表現多則內容之間的因果、順序、對比關係。
4. **衍生**：則為新聞網站生產出資訊衍生品，包括簡訊、檢索、電子報訂閱、新聞授權與品牌運作等等。

從資訊系統的角度來看，推薦系統即是希望透過資訊科技及人機互動，替使用者處理、篩選出符合使用者需求的資訊內容。廣義而言，一個推薦系統可以分為三個組成的元件(Swearingen & R Sinha, 2001)：

1. **輸入 (Input)**：使用者必需輸入一些數據或評價資訊，以供系統來進行分析與判斷。這個過程，也就是系統獲取使用者的偏好資訊、評比資訊的過程，又可以分為顯性評比 (Explicit Rating) 及隱性評比 (Implicit

Rating) (Nichols, 1997)。顯性評比是利用事先定義好的模式，讓使用者去輸入評分、排序、評論等評比資料；而隱性評比則是系統暗地裡觀查使用者的行為並加以紀錄，如點擊紀錄、書籤設定、停留時間。這些資訊是由系統主動蒐集的，不需使用者額外輸入。顯性評比所獲得的資料正確性會比顯性評比來得正確；但是相對的，因為增加使用者的負擔，可能也因無法得到使用者充份的參與而導致系統判斷的誤差。

2. **演算法 (Algorithm)**: 這是大部分推薦系統的核心所在。可以透過使用者所閱讀的內容分析，集合眾多使用者的評比資訊，或根據使用者的社交網絡等關係，進行運算、分析，產生推薦的結果。
3. **系統介面 (Interface)**: 推薦系統會配合推薦物、推薦對象及推薦情形，設計出適合的溝通介面。

第四節 小結

社會性書籤網站為利用群體智慧運作的方式，將網路上雜亂的內容加以彙整、篩選。透過使用者的收錄，為書籤網站的平台獲取了資訊的內容；接著，使用者以推薦的方式，做為一種顯性評比，經由系統演算法的運算來定義文章的價值，然後在首頁組織、整合這些篩選後的結果，透過介面呈現給使用者。另外也藉由標籤的運用、分群的設計、社交網路的機制，試圖為文章進行分類、組織、整合，方便使用者獲取相關主題的資訊。

在書籤網站的各個功能模組中，本研究主要針對文章的推薦評選機制來探討。至於分類、社交網路等部分涉及到了語意學 (Semantics)、檔案學 (Archival Science)、社會網路分析 (Social Network Analysis) 等等專門學科，暫不在本研究的範圍之內。