

## 第二章 研究回顧法

本章節的主要目的為回顧先前與 SOM 或分群相關的調查性研究，共分為三大部分整理分析。首先，介紹研究回顧的基本概念，及其各階段研究流程；接著針對以「SOM」、「Clustering」與其他相關主題為研究主題的回顧性研究進行整理分析。最後引用 Cooper 對整合型回顧研究所提出之五階段研究模型【Cooper, 1982】作為本研究的主要架構與流程，並結合「主題層面配對檢索策略」【Harter, 1986】與本研究提出的「檢索策略矩陣」作為資料蒐集與文獻檢索的主要研究方法。

### 2.1 研究回顧概述

#### 2.1.1 研究回顧背景

研究的意義為在一些知識領域中小心、系統、耐心地研究與調查，並試圖發現事實及建立原則【Grinnell, 1993】。研究是一種結構性的探索，其使用科學的方法論來解決問題，產生新的可應用知識。科學方法包括系統性的觀察、分類及資料的解釋。研究也可說是一種系統性的調查以發現問題的答案【Burns, 1994】。總而言之，研究乃是一個搜集、分析及解釋資訊，以回答問題的過程，必須儘可能受到良好的控制、非常嚴謹、系統化、具有效性、可驗證性、實證性與批判性。

以科學的工具來追求知識是一個相互合作與依賴的過程，科學家或研究者曠日費時地進行各種科學研究，基本上對這巨大的科學拼圖僅僅貢獻了其中的一小塊，任何單一研究的價值不只來自於其研究本身的特性，而其如何完成或擴展前人的努力也是具有相同價值的。雖然有一些研究獲得比其他的研究更多的注意，但這通常只是因為他們所解出拼圖的部分（或是所介紹的拼圖）是非常重要的，而不是因為這些研究的本身便是解答或答案。

既然科學的本質是累積而來的，因此對過去研究的信賴與回顧，便成為有順序建構知識的必需條件。但是，當今研究生態很明顯地缺乏注重研究者如何去找尋、評估和整合過去研究。這種輕忽特別在今日知識經濟與資訊爆炸的時代造成研究的主要瓶頸，再加上時間限制的外在因素，使得大部分的研究學者不太可能跟得上所有初探性研究（primary research），只能專注於研究者感興趣的主題上。

另外，近年來獲取各種科學資訊的能力已有了巨大的改變，特別是要取得過去的研究成果，便可藉由電腦輔助的文獻搜尋來協助，以大幅提昇研究者獲取研究實證可行性與精確性。這使得對過去研究的依賴因電腦科學與資訊科技的逐漸專門化而日益殷切。因此，為了因應各學科研究之快速興起與溝通需求，例如電腦科學，很多以理論和研究報告為對象的研究回顧已逐漸萌芽中【高美英，民 88】。

### 2.1.2 文獻回顧定義

文獻回顧通常出現在介紹新初級資料的報告上，或是在一些較詳細的獨立研究中【e.g., Harper, Weins & Matarazzo, 1978; Maccoby & Jacklin, 1974】，在介紹新初級資料中的文獻回顧之範圍通常都很小，被摘錄以做為介紹其他初探性研究的研究將會被限制為與新資料所陳述的特殊議題有關的。當文獻回顧從新資料的介紹獨立出來時，無疑地便有更廣泛的目標。文獻回顧可以有許多不同的焦點、目的、觀點、涵蓋的策略、組織和聽眾【Cooper, 1988】。例如，文獻回顧能將焦點放在研究結果、研究方法、理論與應用上，文獻回顧能夠整合其他研究成果、批評以前的學術研究、建立相關領域之間的橋樑或確認在某一領域的中心議題。

最常見的文獻回顧是組合了兩組焦點和目標，第一種類型的文獻回顧是整合型的研究回顧，乃整合許多不同但具有相同或相關的假設性研究，試圖獲得一個完整的結論來綜合整理過去的研究。整合型的

研究回顧者希望能呈現相關的知識，和強調以往研究所沒有解決的重要議題。從讀者的觀點來看，整合型的研究回顧想要「替換那些從研究前線失去光采的論文」【Price, 1965】，並且指導未來的研究方向，所以將會產生大量的新資訊。

第二種類型的文獻回顧是理論的回顧（theoretical review），此種理論回顧者希望呈現出能解釋某一特殊現象的理論，並且比較這些理論的範圍、內在一致性和其預測性。理論的回顧一般包含對已經進行或建議的重要實驗之描述、那一種理論是最有力而且和已知的關係是一致的評估、而有時候是將不同理論的抽象概念重組或整合。

通常一個完整的回顧會提出這幾組議題中的幾個，整合型的研究回顧是最普遍的，而理論的回顧則一般會包含一些整合回顧的某些部分，而整合型的研究回顧也常會提出多重相關的假設。例如，一個研究回顧可以檢視幾個不同自變項和單一依變項之間的關係，或者可以試著摘要與一連串相互連結的假設有關係的研究。

本研究主要強調的是整合型的研究回顧，不僅僅因為這是最常見的回顧類型，而且也包含了其他回顧類型所出現的決策點以及一些獨特的觀點。

### 2.1.3 文獻回顧階段

大部分研究方法建議科學的調查需要包括了一組連續的活動，儘管方法學者對研究的階段有其不同之定義，但是階段中最重要的區別在於相當程度的共識（一致性）。方法學者也都同意階段的順序不是固定不變的，實證研究者通常會跳過一個或更多的階段，並且也不只從正向，有時候也會反向進行【Selltiz, Wrightsman & Cook, 1976】。

整合性的研究回顧的過程將會包含五個階段：(1)問題陳述；(2)資料蒐集；(3)資料評估；(4)分析和解釋；以及(5)結果的發表。每一個回顧的階段都有和初探性資料研究類似之功能，例如，初探性研究和研

究回顧兩者在問題陳述階段都包含對變項之定義，而在分析和解釋的階段則都包含了什麼結果是有重要意義的選擇。與初探性資料蒐集者一樣，研究回顧者可以對如何進行其調查有不同的選擇。不同的回顧方法會對回顧結果產生差異，最重要的是，在每一個回顧階段中，方法的選擇可能會損及回顧結論的可信度，而造成對回顧效度的威脅。

和回顧過程的每一個階段相關的研究問題、回顧功能、變異來源以及對效度可能的威脅如圖 2.1 所示，以作為檢視研究活動歷程之重要參考指標。

表 2.1 整合型回顧研究之階段概念 【Cooper, 1982】

階段特質	問題陳述	資料蒐集	資料評估	分析與解釋	公開發表
研究問題	回顧應包括什麼證據？	應使用哪些步驟以找出相關的證據？	回顧應包含哪些取得之證據？	應使用哪些步驟以做出有關整個文獻的推論？	回顧報告中應包含哪些資訊？
回顧基本功能	建構能區分研究相關性之定義	決定要檢視可能相關的研究來源	應用一些標準以區別研究有效性	綜合整理所取得的有效研究	應用編輯標準以區別訊息重要性
造成不同回顧結論的步驟間差異性	<ul style="list-style-type: none"> <li>➢ 所包含操作行定義具有差異性</li> <li>➢ 操作細節差異</li> </ul>	因資訊來源造成之研究差異	<ul style="list-style-type: none"> <li>➢ 品質標準差異</li> <li>➢ 非品質標準影響之差異</li> </ul>	推論規則之差異	對於編輯判斷之指引的差異
可能造成回顧結論無效之文獻來源	<ul style="list-style-type: none"> <li>➢ 範圍較窄的概念會使回顧結論因較不受限而顯得無力</li> <li>➢ 表面操作細節會掩蓋互動變項</li> </ul>	<ul style="list-style-type: none"> <li>➢ 所評估的研究可能在品質上異於研究母體。</li> <li>➢ 在可評估的研究中人的樣本可能會異於人的目標母體。</li> </ul>	<ul style="list-style-type: none"> <li>➢ 非質性因素可能導致研究資訊的不恰當加權</li> <li>➢ 研究報告的漏失會使結論不可信</li> </ul>	<ul style="list-style-type: none"> <li>➢ 區別類型與干擾的規則可能是不適當的</li> <li>➢ 以回顧為基礎的證據可能被用來推論因果</li> </ul>	<ul style="list-style-type: none"> <li>➢ 回顧步驟省略會使結論無法再製</li> <li>➢ 回顧發現及研究步驟省略會使結論報廢</li> </ul>

#### A. 問題陳述階段

整合型回顧研究的第一個階段是問題陳述，主要是確認研究包括的變項，並皆給予抽象和具體的兩種定義。在本階段，研究者會問：「和

回顧的概念有關的操作有那些？」更廣泛的說，研究者必須決定用什麼來區別有關與無關的材料。

研究回顧者在本階段所碰到的決策點，將可以回答以下幾個問題：(a)會影響研究回顧者特殊研究相關性的決定是什麼？(b)研究回顧者如何掌控包含兩個或更多自變項間互動的假設？以及(c)在陳述問題時，以往的研究回顧應該扮演什麼樣的角色？另外同時也應該釐清研究者應該從已被認定和某問題領域相關的實證研究中蒐集什麼樣資訊。

#### B. 資料蒐集階段

一個研究的資料蒐集階段包括選擇將與母體有關的研究焦點，對研究回顧來說，確認母體是複雜的，因此回顧者應對以下兩個目標加以推論，第一，回顧者應根據所有先前有關某一問題的研究回顧的累積結果。第二，回顧者希望被包括進來的研究會允許其推廣至主要討論領域的母體(或其他單位的母體)。

本階段將針對如何找出適切的研究方法進行探討評估，包括了對研究主題有用的研究來源清單、如何使用最重要的資源以及在每一資源中所包含之資訊會呈現什麼偏誤等。

#### C. 資料評估階段

在資料蒐集之後，調查者需要針對個別資料點的品質進行精確判斷，每一個資料點都被當做周圍的證據來檢視，以判定是否受到與研究主題無關的干擾因子所污染而產生誤差。本階段亦針對如何進行研究品質的評估加以探討，其中包括與信度有關的評估方法與判斷偏誤的來源，同時也必須考慮無法獲取及不完全的研究報告之處置。

#### D. 分析和解釋階段

本階段調查者將所蒐集的不同資料點整合為一有關於研究問題的單一陳述，並從干擾或偶然的變動裡區分出有系統的資料模式並加以解釋。本階段在實際進行時可能會結合一些獨立的研究結果，或採用

一些估計相關性的方法，也同時檢視在不同的研究中所發現用來分析相關性差異的技術。

#### E. 公開發表階段

製作一份描述回顧結果的公開文件是本階段的主要任務，包括引言、方法、結果與討論等部分，並特別將重點放在如何在回顧過程將資料至做成表格，以及能夠實質地解釋回顧的效應。將會提供一些有關於如何發表整合型的研究回顧報告的具體指引。

### 2.2 整合型研究回顧法

本研究屬於 SOM 的整合型研究回顧，並引用 Cooper 的系統化五階段研究模型【Cooper, 1982】，並透過系統化的文獻檢索策略達到了了解「SOM 模式改善的歷史脈絡」與「SOM 與其它分群法之綜合分群應用」之兩大研究目的，同時探討 SOM 演算法在解決分群問題之效能與特色，並試圖建立一綜合性的參考目錄，提供後續研究者參考。

關於 2.1 節所提到研究回顧之五階段研究模型，本屬社會科學中為了解研究母體而進行的整合性研究回顧或初探性研究。本研究有鑒於 SOM 相關研究回顧不足，儘管赫爾辛基大學的 SOM 研究團隊曾先後發表兩篇合計約五千餘篇相關之調查研究，但絕大部分為應用性質，不僅缺乏對改善方法之綜合探討與整理歸納，更無法綜覽 SOM 改善方法的脈絡全貌，因而採用 Cooper 的五階段研究回顧模型，並結合「主題層面配對檢索策略」與本研究提出「檢索策略矩陣」，作為搜尋檢索與評估篩選的準則。最後將所得結果進行構面分析，以下為各階段主要任務與目標：

#### A. 問題陳述階段

本研究主要的核心主題是「如何改善 SOM 模式以提升資料分群效能」，並進一步分解為兩大子題「SOM 模式改善的相關研究調

查」與「SOM 與分群理論應用」。為回答上述問題，本研究利用「SOM」、「學習演算法」及「分群應用」等三個研究主題層面來定義研究問題，並且規範後續資料蒐集階段文獻檢索的搜尋範圍。

「SOM」是指以 SOM 為主題的研究論文；「學習演算法」則是以演算法改善為主要研究貢獻的論文；而「分群應用」便是探討群集分析與分群理論及其演算法的研究論文。

#### B. 資料蒐集(文獻檢索)階段

採用 Harter(1986)所提出的「主題層面配對檢索策略」，將兩大研究主題分成三大主題層面，並成對地進行檢索後先取交集後聯集之。至於主題層面二與三所產生的交集並不屬於本研究範圍故予以省略不計。本研究並提出一「檢索策略矩陣」，以研究主題與檢索途徑建構資料蒐集空間，並提升文獻檢索的完整性。另外，「引用文獻滾雪球法」將從檢索結果之參考文獻內收錄檢索範圍外確仍滿足本研究問題之重要文獻。

#### C. 資料評估

將檢索策略所得到的文獻進行條件式篩選，以獲得與研究主題直接相關的關鍵文獻。

#### D. 分析和解釋

分別針對不同分區的檢索策略執行並篩選後的目標文獻進行整理歸納及統計分析，以了解與研究主題較為相關的學術期刊及文獻內容。最後將所有目標文獻進行 SOM 演算法的特徵構面分析，以達到 SOM 重點理論與相關文獻探討融合之目的。

#### E. 結果發表

整合分析兩大研究主題所得到的文獻論述，並且尋找兩者之間的關連性，以試圖回答研究問題。

以下分別針對「檢索策略」、「檢索策略矩陣」、「研究主題」、「檢索途徑」與「檢索條件組」加以進一步說明如下：

### 2.2.1 主題層面配對檢索策略

採用 Harter(1986)所提出的「主題層面配對檢索策略」，將兩大研究主題分成三大主題層面，並成對地進行檢索後先取交集後聯集之。至於主題層面二與三所產生的交集並不屬於本研究範圍故省略不計。

簡單的說，本研究擬定的檢索策略就是基於研究目的，針對單一探討主題，於規劃的檢索途徑內進行文獻搜集，至於允收標準便依照定義的篩選條件集合。由下列檢索矩陣中可看出研究主題包括「模式改善」與「分群綜合研究」；檢索途徑亦有兩項：資料庫與期刊；而篩選條件組將會依照兩大研究主題加以定義。以下分別將本研究兩大「檢索策略」加以說明：

#### A. 檢索策略一：「SOM 模式改善」

檢索策略一是將主題層面一「SOM」與主題層面二「Learning Algorithm」予以配對，利用資料庫與期刊作為途徑進行檢索並取其交集，再以所設定的「篩選條件組」進行篩選，最後將其整理分析，以達到了了解「SOM 相關變型模式與改善現況」之研究目的。

#### B. 檢索策略二：「SOM 分群綜合研究」

檢索策略二是以主題層面一「SOM」與主題層面三「Clustering」取其交集，並以資料庫與期刊等作為檢索途徑，再以所設定的「篩選條件組」進行篩選。最後將其整理分析，以達到了了解「調查 SOM 與其他分群法之綜合改善研究及其特色」之研究目的，並嘗試探討在計算科學中 SOM 與其他主要分群法之結合應用與特色。



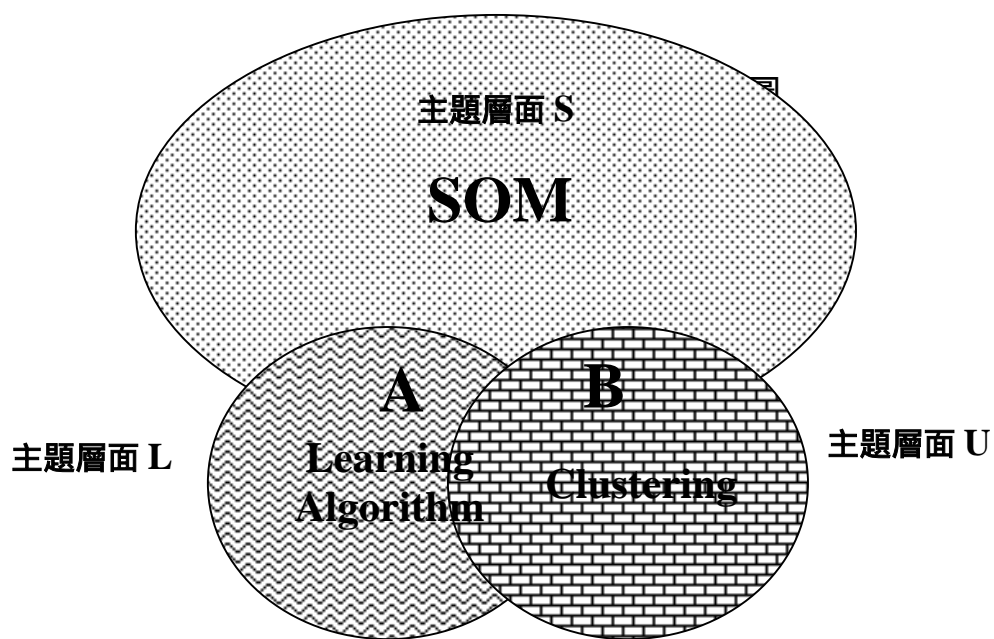
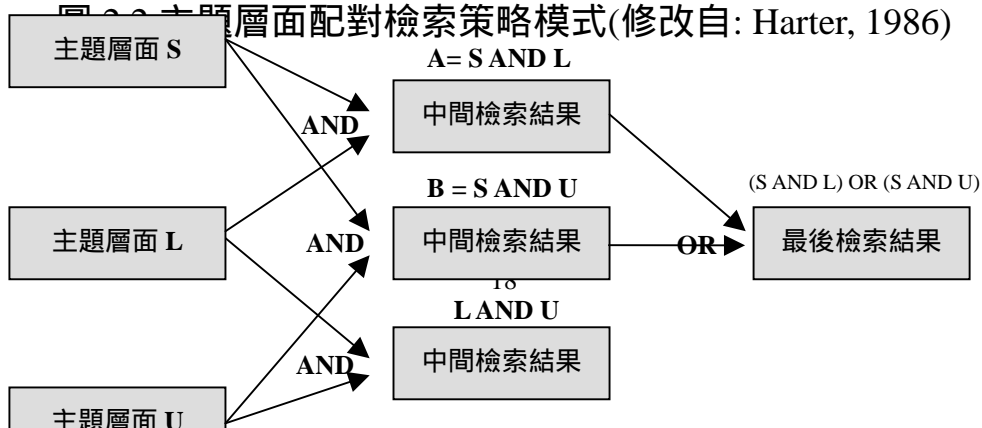


圖 2-2 主題層面配對檢索策略模式(修改自: Harter, 1986)



### 2.2.2 檢索策略矩陣

本論文提出一檢索矩陣概念，擬定以上述兩大研究目的為基礎的兩種文獻檢索策略來進行回顧性調查研究，其中每項檢索策略均包含兩種主題層面、途徑與符合其研究目的之篩選條件組。因此，以此檢索矩陣提供一系統化的文獻搜尋法來進行文獻檢索，並可得到 2\*2 共 4 類的初步檢索結果，再經過所屬策略的篩選條件組進行篩選，最後可獲得 4 類最終的檢索結果，以提供後續分析研究之用。

表 2.2 2\*2 檢索策略矩陣

檢索策略矩陣		檢 索 途 徑		篩選條件組
		D 資料庫(5)	J 期刊(10)	
研究主題	A 模式改善	AD	AJ	SA
	B 分群綜合應用	BD	BJ	SB

#### A. 研究主題與檢索條件組

上述檢索策略矩陣的縱軸構面為研究主題，也就是本研究之研究範圍。在此提出包括研究脈絡與分群綜合研究等兩大研究問題，其主要原因即在於以研究目的為中心並分列子題。而文獻的搜尋便是將主題層面一「SOM」分別與主題層面二「Learning Algorithm」與主題層面三「Clustering」利用各別的檢索關鍵字進行第一階段的檢索；換句話說，不論何種資料庫或期刊的檢索引擎與檢索條件，均一併納入初步的檢索結果。另外，本研究亦針對每個研究主題設計篩選條件組，以作為進一步篩選的依據。其目的一方面為使文獻搜尋能有效地切合研

研究主題需要；另外就是可以避免過於大量的檢索結果，有礙於後續的研究分析。以下分別將本研究「研究主題」及其對應的「檢索關鍵字」、「分析構面」與「篩選條件組」表列如后：

表 2.3 研究主題、篩選條件組及其分析構面

研究主題	研究子題	分析構面	篩選條件組
SOM 模式改善	改善模式 (Enhanced Models)  檢索字串： 「Learning algorithms」	競爭式學習 (competitive learning)	1. 時間範圍(1990~2004) 2. 應提出一嚴謹的研究方法論或建立一套系統，並對於 SOM 演算流程有細部且整體探討者。 3. 應以傳統 SOM 為主體，提出改善原有競爭學習之方法論，並以實例舉證說明。 4. 僅限學術期刊
		資料/集群視覺化 (data/cluster visualization)	
		網路拓樸映射與保存 特性(topology projection and preservation)	
		網路拓樸之動態網格 模式(dynamic grid model)	
SOM 分群 綜合研究	SOM 與其他主 要分群法結合  檢索字串： 「Clustering」	分群理論	1. 時間範圍(1990~2004) 2. 應提出一與 SOM 結合之分群應用，且不出檢索子題範圍。 3. 文獻類別不限，可包括學術期刊、國際研討會與碩博士論文或技術報告等。
		與傳統分群法結合	
		與 Fuzzy Sets 結合	
		與 Genetic Algorithm 結合	
		其他分群法	
		分群法比較	

## B. 檢索途徑

上述的檢索策略矩陣的橫軸構面為檢索途徑，也就是文獻蒐集的對象範圍。在此提出包括資料庫與期刊作為文獻檢索途徑，其主要原因是這些方式至今仍為典型的文獻探討管道且廣被接受。以下進一步針對各檢索目標進一步說明：

(1) 資料庫檢索

本研究選擇五個著名的國際性學術資料庫，分別是 SDOS, IEL, ACM, EBSCO host 與 SCIE，並藉由關鍵字檢索輔以篩選條件進行綜合研判，最後將所獲得的搜尋結果加以整理分析。

(2) 重點期刊檢索

就 SOM 的發展脈絡而言，最初是從國際影像分析研討會提出具完整數理基礎的雛型。到後續生物人工頭腦、認知科學以及神經計算與計算智慧等相關學術期刊，所發表的論文數更是從 1997 年初的 1,500 篇激增至 2001 年約 4,000 篇。有鑑於此，本研究遂以 INNS 的 Neural Networks 與 IEEE transactions on Neural Networks 作為搜尋中心起始點，開始向外向前延伸，將符合目標條件的文獻歸納整理後取前十篇最具代表性者作為 SOM 重點期刊，並將其刊名、國際代碼與發行年份整理如下：

表 2.4 本研究期刊檢索範圍表

Journal	ISSN	Publish Year
Computers and Industrial Engineering	0360-8352	1976
IEEE transactions on Neural Networks	1045-9227	1990
IEEE transactions on Systems, Man and Cybernetics	0018-9472	1971
IEEE transactions on Systems, Man and Cybernetics Part.B	1083-4419	1996
International Journal of Production Research	0020-7543	1961
Neural Networks	0893-6080	1988
Neural Computation	0899-7667	1989
Neurocomputing	0925-2312	1989
Pattern Recognition	0031-3203	1968
Fuzzy Sets and Systems	0165-0114	1975
IEEE Transactions on Evolutionary Computation	1089-778X	1997