

3.2 SOM 沿革與近期發展

本章節除將 SOM 歷史發展脈絡作一介紹外，另將 SOM 相關期刊與國際研討會進行整理歸納。另外，創始人 Teuvo Kohonen 教授所屬芬蘭赫爾辛基大學(Helsinki University of Technology)的類神經網路研究中心(Neural Networks Research Center)之近期發展亦為本章節之主要內容。

3.2.1 SOM 發展介紹

自從 SOM 於 1982 提出以來已有數千篇相關研究論文發表於國際期刊與著作中。該網路模式擁有簡單卻功能強大的自組織機制與特性，因而迄今在類神經網路研究領域中仍為相當受重視的課題。關於 SOM 創新性的延伸發展，首推自 1997 開始隔年舉行的 WSOM (Workshop on Self-Organizing Maps)，前兩屆於芬蘭赫爾辛基科技大學舉行(Helsinki University of Technology)，對於整合 SOM 相關研究理論與實務技術相當有助益。根據 2001/6 於英國 Lincoln 的 University of Lincolnshire and Humberside 大學所舉辦的 WSOM 2001，其主要議題如表 3.2 所示。

表 3.2 WSOM 研討會議題列表

	WSOM 2001	WSOM 2003
地點	Lincoln, UK	Kyushu, JP
議題類別	<ul style="list-style-type: none">● 樣式辨識與資料分群(Pattern recognition and data clustering)● 非監督式學習(Unsupervised learning using SOMs and other techniques)● 資料視覺化(Data visualization)● 商業與財務應用(Commercial and financial applications)● 多維度信號處理 (Multi-dimensional signal processing)組件及系統設計與狀態監控 (Component and system design and condition monitoring)	<ul style="list-style-type: none">● SOM 理論與在神經科學中所扮演的角色(Theory and role in neuroscience)● 網路智慧應用(Applications in WEB intelligence)● 文字與文件分析(Text and document analysis)● 機器人學(Robotics)● 影像處理(Image processing)● 視覺(Vision)● 樣式辨識(Pattern recognition)● 硬體製作(Hardware realizations)

由上述研討議題類別不僅顯示 SOM 於下列領域有不斷創新的應用發展，更使得 SOM 的基礎模式準則與演算法特性更加釐清：

- 資料分析 (Data analysis)
- 資料探勘與視覺化(Data mining and visualization)
- 序列性與非向量處理(Sequential and non-vectorial processing)
- 收斂狀態(Convergence conditions)
- 非線性之成分分析(Non-linear component analysis)
- 資訊與強化學習(Information and reinforcement learning)
- 統計學習概念(Statistical learning ideas)

儘管 IJCNN 2003 的專刊「Advances in Neural Networks Research: IJCNN '03」對於 SOM 鮮少著墨，但今年 Neural Networks 計劃推出的兩本特刊中，除了探討「視覺與腦」(vision and brain)外，便是「自組織系統的最新發展」(new developments in self-organizing systems)，內容將以 2003/9 於日本九州工業大學(Kyushu Institute of Technology, HIT)舉行的 WSOM 2003 論文為主(預計 2004 年底發行)，所涵蓋的主題將如表 3.2 所示。

1980 年代中期，因自組織映射圖網路、霍普菲爾網路(Hopfield Neural Network, HNN)與倒傳遞網路(Back-Propagation Network, BPN)的提出，使得自 1960 年代中期開始沒落的類神經網路研究帶來的重大突破，在加上電子光電技術的進展，提供「神經電腦」的可能性，以及現代生理學、認知學與神經科學進步亦提供豐富啟示。1987 年類神經網路的研究與應用進入了新紀元，各式期刊、研討會、學會與研究計畫及論文大量出現，顯示研究風氣盛極一時。在 SOM 相關之期刊中，仍以 INNS 籌辦的 Neural Networks 與 IEEE 籌設的類神經網路專刊 IEEE Transactions on Neural Networks 為兩大主要學術期刊，其中包括 SOM 及其他類神經網路最新研究與發展之論文，表 3.3 是 SOM 相關的期刊彙整列表。

表 3.3 SOM 相關論文發表的期刊

Journal	ISSN	Issued Year
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
Neural Networks	0893-6080	1988
Neural Computation	0899-7667	1989
Neurocomputing	0925-2312	1989
Pattern Recognition	0031-3203	1968

1988 年舉行的 ICNN'88 正式揭開類神經網路一系列國際學術研討會之序幕，研討範疇相當廣泛，從生物科學的細胞分子與基因科學開展到工程科學之控制學，機器人學，海洋工程與電力系統等，並且隨著資訊時代來臨，朝向以人工智慧為基礎的自然語言，訊號/影像處理與互動式多媒體系統等應用領域發展，使得電腦科學之智慧型運算，特別是模糊理論(fuzzy sets theory)、類神經網路(neural networks)與遺傳演算法(genetic algorithm)為主的演化式計算(evolutionary computation)三者所構成軟性計算科學(softcomputing)在跨學門之間皆有非常蓬勃之學術發展。

也就因為相關學術研討會如雨後春筍般密集出現，且研討主題會因應當時環境需要而變化，在此僅針對 IEEE 所屬與本研究直接相關且定期舉行之研討會整理如表 3.4 所列。另外，依據與研究主題相關之非常態性研討會亦列舉其中。

表 3.5 針對 SOM 的開發工具加以整理歸納，其中有星號標註者為免費軟體，而 SOM_PAK 與 SOM toolbox of Matlab 5 為赫爾辛基大學資訊與電腦科學實驗室研發團隊所開發，JOONE 則為 Paolo Marrone 的 JOONE 團隊研發。至於其他 eSOM 及 Neuro Solutions / Neuro Solutions for Matlab 則為商業評估軟體。關於詳細資料請參見相關參考網址。

表 3.4 SOM 相關國際研討會列表

國際研討會	簡稱	創始時間	舉辦頻率(year)
Workshop on Self-Organizing Maps	WSOM	1997	2
International Joint Conference on Neural Networks	IJCNN	1989	1
International Conference on Artificial Neural Networks	ICANN	1989	2
IEEE International Conference on Neural Networks	ICNN	1988	variant
Neural Networks for Signal Processing	NNSP	1991	1
其他非常態性研討會			
Combinations of Evolutionary Computation and Neural Networks	COECNN	2000	NA
Combinations of Genetic Algorithms and Neural Networks	COGANN	1992	NA

最後根據赫爾辛基科技大學(Helsinki University of Technology, HUT)的類神經網路研究中心(Neural Networks Research Center, NNRC)針對 SOM 整理出詳盡的參考目錄，分別針對 1981-1997 收錄 3,343 篇相關文獻，以及 1998-2002 的 2,092 篇文獻，共計 5,435 篇相關文獻，建立 SOM 完整且具有索引參考之研究資料庫，並投稿於 Neural Computing Survey 期刊【Kaski, Kangas & Kohonen, 1998】【Oja, Kaski & Kohonen, 2002】。而本研究另外整理出 SOM 相關書目如表 3.6 所列。主要是依據書名含有 Self-Organizing 或作者為 Kohonen 者，其餘則是內容與本研究主題相關者。

表 3.5 SOM 相關開發工具整理

開發工具	開發者	時間	產品特色	參考網址
eSOM	Ellipse, Inc. Helsinki	2000	<ul style="list-style-type: none"> ➢ 利用 SOM 分析大型資料中的規則與群集 ➢ Windows 與圖形視窗介面為無經驗者設計 	http://www.ellipse.fi/
* JOONE	Paolo Marrone & the Joone team	2000	<ul style="list-style-type: none"> ➢ 基於 Java 最新技術，提供建立訓練與測試神經網路之免費架構 (Lesser General Public License, LGPL) ➢ 多線程(multithreaded)之核心引擎為基礎，可適用於所有計算資源尺度 ➢ 為專業使用者提供分散式環境同時多機平行處理多個神經網路 	http://www.jooneworld.com/
* SOM toolbox of Matlab 5	Alhoniemi, E., Himberg, J., Parhankangas, J., & Vesanto, J.	2000	<ul style="list-style-type: none"> ➢ 極佳的高維視覺化工具，最適用於 KDD 中資料理解階段，同樣也適用於資料準備、建模與分類等 ➢ 該工具箱模組為免費下載並包含 SOM 的新增、視覺化與分析等功能。 	(Laboratory of Information and Computer Science, Helsinki University of Technology) 【Vesanto, Himberg, Alhoniemi, & Parhankangas, 2000】 http://www.cis.hut.fi/projects/somtoolbox
* SOM_PAK	Kohonen T., Hynninen J., Kangas J., Laaksonen J.	1996	<ul style="list-style-type: none"> ➢ 免費套裝程式，功能健全(ANSI-C) ➢ 但不如 Matlab 環境般具彈性 	【Kohonen, Hynninen, Kangas, & Laaksonen, 1995】 【Kohonen et al., 1996】 http://www.cis.hut.fi/research/som_lvq_pak.shtml
Neuro Solutions Neuro Solutions for Matlab	NeuroDimension, Inc.	1995	<ul style="list-style-type: none"> ➢ 結合模組化圖形介面與先進演算法 ➢ 自動產生 C++原始碼 ➢ DLLs 客製軟體元件 ➢ neuro-fuzzy 架構 ➢ programmatic control from Visual Basic using OLE Automation ➢ Neuro Solutions toolbox 在 Matlab 可支援 15 個神經網路與 5 種演算法，並提供簡易操作之整合介面 	http://www.neurosolutions.com/

表 3.6 SOM 相關書籍列表 (以出版時間排序)

No	Title	Author	ISBN	Year
1	Associative memory: A system-theoretical approach	Teuvo Kohonen	0387080171	1977
2	Self-Organization & Associative Memory	Teuvo Kohonen	0387513876	1989
3	Pattern Recognition by Self-Organizing Neural Networks	Gail A. Carpenter, Stephen Grossberg	0262031760	1991
4	Neural Computation & Self-Organizing Maps: An Introduction	Helge Ritter, Thomas Martinez, Klaus Schulten	0201554437	1992
5	Self-Organizing Maps	Teuvo Kohonen	3540586008	1995
6	Self-Organizing Maps (2 nd)	Teuvo Kohonen	3540620176	1997
7	Self-Organization, Computational Maps, & Motor Control	P. Morasso, Vittorio Sanguineti, Pietro Morasso	0444823239	1997
8	Visual Explorations in Finance: With Self-Organizing Maps	Teuvo Kohonen	3540762663	1998
9	Kohonen Maps	Erkki Oja & Samuel Kaski	044450270X	1999
10	Faithful Representations & Topographic Maps: From Distortion- to Information- Based Self-Organizing	Mart M. Van Hulle	0471345075	2000
11	Advances in Self-Organizing Maps	Nigel Allinson, Hujun Yin, Lesley Allinson, Jon Slack	1852335114	2001
12	Self-Organizing Maps (3 rd)	Teuvo Kohonen	3540679219	2001
13	Self-Organizing Map Formation	Klaus Obermayer, Terrence J. Sejnowski	0262650606	2001
14	Self-Organizing Map Formation: Foundations of Neural Computation	Klaus Obermayer, Terrence J. Sejnowski	0262650606	2001
15	Self-Organizing Neural Networks: Recent Advances and Applications	Udo Seiffert, Lakhmi C. Jain	3790814172	2001
16	Data Mining Introductory and Advanced Topics	Dunham, M. H.	0130888923	2002
17	Advances in Neural Networks Research: IJCNN 2003	D.C. Wunsch II et al.	0080443206	2003

3.2.2 赫爾辛基科技大學於 SOM 之相關研究與發展

赫爾辛基科技大學的類神經網路研究中心由神經計算科學先驅 Teuvo Kohonen 教授於 1994 年成立,該中心不僅主導電腦科學與工程學系(Department of Computer Science and Engineering)中電腦與資訊科學實驗室(Laboratory of Computer and Information Science, CIS)之大部分研究計畫,更於 1995 年迄今均被評選為芬蘭全國優良學術研究室之一,目前由 Erkki Oja 教授主導。表 3.7 為該中心數個重要研究領域及其內容:【<http://www.cis.hut.fi/research/>】

其中在 SOM 資料分析(Data analysis of the SOM)研究計畫中,目前由 Samuel Kaski 擔任計劃主持人,而主要研究內容為發展以 SOM 為基礎的方法來解決資料分析與資訊視覺化問題。除此之外,特別於生物資訊科學(bioinformatics)、學習韻律學(learning metrics)、語音(speech)與自然語言建模(natural language modeling)等方面有跨組性或校際性合作之團隊分組進行研究。目前以「網路自組織」(WEBSOM)【Lagus, 2000】、【Kohonen et al., 2000】,即利用 SOM 探討大型文件集合之組織化問題,其做法是針對文章「內容相似度」來組織大量文件集合(WEBSOM),文件的特徵向量採用字彙的統計方法(直方圖)來呈現,而輸入向量的降維則採用隨機映射法。完成的自組織網路拓樸圖將呈現整體文件集合之綜覽,並可依照使用者指定特定文件進行互動式之摘要瀏覽。WEBSOM 自 1996 年提出經過不斷改良的系統效能與演算法,到 2000 年已將文件規模成功提升至將近七百萬篇內容摘要映射在約一百萬個格點(即輸出單元)的網路拓樸上,並且隨機抽取維度為 500 的文件特徵向量作為起始權重向量,為最新且規模最大的資料探勘應用。至於其他 SOM 的主要應用方向如下所示:【<http://www.cis.hut.fi/research/som-research/>】

1. 自動化語音辨識 (automatic speech recognition)
2. 臨床聲音分析 (clinical voice analysis)
3. 工廠設備與製程之狀態監測 (monitoring of the condition of

industrial plants and processes)

4. 人造衛星影像之雲圖分類(cloud classification from satellite images)
5. 腦部電流信號分析 (analysis of electrical signals from the brain)
6. 大型資料集合之資訊擷取與組織化 (organization of and retrieval from large document collections)(WEBSOM) 【Kaski et al., 1998】,
【Kohonen, et al., 2000】
7. 大型統計資料集合之分析與視覺化 (analysis and visualization of large collections of statistical data)

3.2.3 SOM 近期發展(2004~2005)

SOM 及其相關延伸改善在非監督式學習與資料視覺化的領域使用上為最受歡迎的類神經演算法，目前公開文獻已有超過 5,000 篇論文發表【Kohonen, 2001】，並且許多商業專案計畫更以此解決現實世界多元複雜的問題。2004 年 7 月於匈牙利首都布達佩斯(Budapest)舉行的 IJCNN'04，與 2005/9 預計於巴黎第一大學舉行的 WSOM'05 預期將會產生跨領域的創新概念與軟硬體技術交流，使其普及於工業技術之公眾領域。以下將這兩者研討會 SOM 主要議題類別彙整如表 3.8 所示。

表 3.7 NNRC 主要研究方向現況調查表

研究領域	主持人	研究內容
主要研究領域		
獨立成份分析 (Independent Component Analysis, ICA)	Erkki Oja and Professor Juha Karhunen	<ul style="list-style-type: none"> ➢ signal separation: 利用 ICA 從混合型態的資料如 MEG 腦部訊號、電信訊號或混合語音訊號中找出獨立成份訊號。 ➢ 致力於 ICA 的理論發展與數種應用案例。 ➢ ICA 的理論焦點集中在發展貝氏分布方法與相關快速演算法。
影像適應性分析 (Adaptive analysis of images)	Erkki Oja	<ul style="list-style-type: none"> ➢ 著重發展創新的學習演算法來解決影像處理與分析問題。 ➢ 同時致力於影像處理與分析的理論與應用研究。 ➢ 在 PicSOM 專案中，應用 SOM 進行影像資料庫萃取。 ➢ 在適應性 OCR 研究中發展線上字體辨識之學習法。 ➢ IMPRESS 計畫中成功開發數項工業應用案例。
SOM 資料分析 (Data analysis of the SOM)	Samuel Kaski	<ul style="list-style-type: none"> ➢ 以 SOM 為基礎來解決資料分析與資訊視覺化的問題。 ➢ Bioinformatics, learning metrics, speech, and natural language modeling. ➢ WEBSOM: 處理大型文件集合之組織結構，同時為 SOM 在資料探勘中最大規模之應用實例。 ➢ 睡眠失序研究。
複雜程序及信號之分析與視覺化(Analysis and visualization of complex processes and signals)	Olli Simula	<ul style="list-style-type: none"> ➢ 致力發展具問題獨立性並依據數據處理之智慧型資料分析法。 ➢ 著重類神經網路之應用，特別是 SOM 在複雜系統中的資料分析與資料探勘。 ➢ 已經成功應用在分析、視覺化與流程監控。如鋼鐵業的軋延(rolling)與森林業的製漿流程(pulping processes)等。 ➢ 電信系統應用：利用神經等化器/均衡器結構(neural equalizer structures)來偵測間斷信號。 ➢ 應用以 SOM 為基礎之分析法調查

		移動式(mobile)的網路資料。
新興研究領域		
生物資訊科學 (Bioinformatics)	Mannila, H., Hollmen, J., & Kaski, S.	<ul style="list-style-type: none"> ➢ 利用 learning metrics 與其他資料分析法針對功能性基因組資料進行建模。 ➢ 專長發展統計模型與演算法來分析生物過程並探勘生物資訊。 ➢ 兩個密切合作的研究團隊應用現代計算科學與統計方法來解決計算生物學所遭遇的問題： <ul style="list-style-type: none"> ■ Pattern Discovery group, especially the Computational Methods in Medical Genetics and Expression Data Analysis group of the From Data to Knowledge Research Unit (H. Mannila, J. Hollmen) ■ Bioinformatics group of the Neural Networks Research Centre (S. Kaski)
自然語言處理 (Natural language processing)	Bingham, E., Creutz, M., Kaski, S., Kurimo, M., Könönen, V., Lagus, K., & prof. Oja, E., Siivola, V., Somervuo, P., Vuori, V.	<ul style="list-style-type: none"> ➢ 將適應性語言的建模方法應用於資訊擷取、語音辨識、手寫字型辨識與對話建模等方面。 ➢ 可處理自然語言的自動化方法。 ➢ 開發互動式自然語言系統，目標為可使人們與系統透過自然語言完成一項工作，如預定機票。
貝氏方法 (Bayesian method)	Juha Karhunen	<ul style="list-style-type: none"> ➢ 發展變化型貝氏學習法，並應用於解決各種非監督學習問題。 ➢ 研發 Bayesian ensemble learning，並且應用於各種隱性變數之模型結構，相當適合非監督式學習。 ➢ 目前研究重心是以發展 modular building blocks 來建構非線性與非高斯分配之模型，特別是自動產生學習法則與連續變數檢建模。
Data Mining 相關應用		
樣式發現 (Pattern discovery)	Heikki Mannila	<ul style="list-style-type: none"> ➢ 致力發展演算法來尋找複雜資料集合中頻繁出現的樣式。 ➢ 從時間序列資料中發掘規則，從事件次序中找尋關聯性規則與將樣式發現應用於預測及分群。 ➢ 實際應用如：生物資訊學、文件分析、間斷性高維資料分析、基因表示分析與基因映射等。

表 3.8 IJCNN'04 與 WSOM'05 研討會議題列表

	IJCNN'04	WSOM'05
地點	Budapest	Paris
議題類別	<ul style="list-style-type: none"> ● 資訊擷取(Information retrieval) ● 資料視覺化(Data visualization) ● 文字與文件分析(Text and document analysis) ● 動態網格架構(Dynamic grid structure) ● 序列性與非向量處理 (Sequential and non-vectorial processing) ● 樣式辨識(Pattern recognition) ● 資料分群(Data clustering) ● 機器人學(Robotics) ● 方法論改善(Theory and algorithm) 	<ul style="list-style-type: none"> ● 自組織(Self-organization) ● 非監督式學習(Unsupervised learning)：包括 PCA/NLPCA, ICA/NLICA, Principal Curves/Surfaces ● 理論與延伸(Theory and extensions) ● 最佳化(Optimization) ● 硬體與架構(Hardware and architecture) ● 信號處理, 映像處理與視覺(Signal processing, image processing and vision) ● 時間序列分析(Time-series analysis) ● 文字與文件分析(Text and document analysis) ● 財務分析(Financial analysis) ● 資料視覺化與探勘(Data visualization and mining) ● 生物資訊學(Bioinformatics) ● 機器人學(Robotics)