• 計畫中文名稱	利用擴散理論、語義網絡與時序分析爲基礎之機器學習法建立人類活動與環境變遷間之分析模擬模型(II)		
• 計畫英文名稱	Building an Analytic Enumeration Model to Discover the Association between Human Activities(II)		
• 系統編號	PB9808-2289	• 研究性質	基礎研究
• 計畫編號	NSC98-2221-E038-012	• 研究方式	學術補助
• 主管機關	行政院國家科學委員會	• 研究期間	9808 ~ 9907
• 執行機構	臺北醫學大學醫學資訊研究所		
• 年度	98 年	• 研究經費	655 千元
• 研究領域	資訊科學--軟體		
• 研究人員	蔣以仁		
• 中文關鍵字	--		
• 英文關鍵字	--		
• 中文摘要	當我們試圖分析自然世界中不同或同一族群的影響現象關係時，從小到原子、氣體 分子、生物基因與蛋白質、大到食物鏈、流行病學等的研究上，不得不注意她們彼此 相互間存在的相依關係與交互作用，而這些交互現象所存在的結構，其邏輯上所顯現 的結構爲何？正是我們期望理解與解讀的。 分子（Molecular）間充滿彼此間的交互影響，從最近分子生物學藉由微陣列，研究 大批的基因與某種的生理現象或疾病等的影響揭開序幕，資訊科學的研究，無不在試 圖提供最佳解答的方法途徑上下最大的功夫。 面對這些大量、複雜且無頭緒的資料，進行預測，以爲決策支援，第一步往往是將 資料分群，經由分群延展以了解資料所形成的結構，透過一些正規式方法的描述，是 否能詮釋更深刻的且更形有用的 Ontology 式知識，正是我們所關注的。 近些年來，各項研究，紛紛顯示，越龐雜越鉅量的資料，看似無意義，其實顯現著 無比的規律，而這規律又顯現出相當巨大的能量。而這些能量超過一定的承受程度時， 會造成這些資料所顯現系統的巨變。爲了描述這種隨時間與環境變遷而逐步改變的現 象，或稱之爲滲透（percolation），許多相關的模型陸續提出，這些模型主要在交代系 統的穩態、不平衡態與失控；透過這些模型，除協助我們瞭解資料所呈現的出來的意 義外，並可達成系統模擬，藉以深入剖析在不同狀況下的未來世界；流行病學、生態 學、環境工程、演化等學問，都多少以此爲基底，探求未來。 本研究設計適當演算法，進行高維資料資料之分群，首先將以建立高維資料的同質 性處理，無論是結構式或非結構式的資料爲主之分群爲主，發展文獻與資料自動蒐錄 整理及內容萃取系統，以 Topological Geometry 呈現資料的空間關聯與交錯現象，並 能區分時間，並結合機率規則進行動態變化的判斷推估，本研究中將分別探討隨機、 與特定條件下的變化，當然先行於 Distribution Free 模式，進行推估；隨後進行資 料流動之動態分析，以能自動產生相容且相關之 Taxonomies，並能據以建立內涵且具 決策能力的推論式 Ontology，透過視覺化的人機介面，協助專家瞭解資料中語義內涵 的交叉關聯與隨時間產生的變化，或可稱之爲典範遞移，並利用所隱含知識使可協助 專家修正規則以獲得正確的數		

	位內容之具時間判別之決策 Ontology，最後協助專業人 事進行決策模擬，可利用時間序列動態分群機器學習技術進行鑑別判斷，俾利能修正 決策 Ontology 產生程序，使接近於主題相關之正確 Taxonomic 及 Decision Ontology。 而此一動態模式所顯現的流動性，依資料所顯現的特徵，可輔助我們區隔參數的重 要與否，透過他們彼此間 Likelihood 的比較，產生相對應重要值，也就是所獲得的分 數，此分數結果，可經由 nomogram 之對應建立事前與事後的預測期望，協助專家進行 往後的推理判斷，並做預測。
• 英文摘要	The structure of networks arises in a wide array of different contexts to describe complex relationships occurring from cells to society. The knowledge from scientific researches is getting complex and rapidly expanding nowadays, for example, there are above 15 million literatures in Medline. While considering to analyzing the different targets in a natural world, from the infinitesimal molecular to big food webs, we are not able to neglect the interactions among those species or objects. What is the logical structure among them? To consider their existed interactive structure is our expectation. All information is now behind the large database. Such a high-dimensional data including database contains much more implicit knowledge as we thought. The implicit knowledge is the foundation for us to make decision. However, the complicated of the knowledge behind the data is hard to find out by human. Furthermore, if the time factor has been considered, the problem becomes more complicate and harder than before. In order to provide an optimal decision making in a dramatic changeable area, researchers should get concise and precise knowledge organized from time-variant information. Comparing the time-variant and time-invariant features, we are able to discover significant factors around the data. This project will analyze and design a temporal model for knowledge discovery based on the clustering and ontology construction, which provides the technique for the unstructured data management, especially the huge and divergent literature resources. This research has taken scientific dataset and related articles as subject, implementing “automatic clustering”and “scoring”model to automatically construct the taxonomical and decision supporting ontology. Afterwards, invited experts will evaluate the efficiency of the system with personal diagnosis and feedback the results to the system. With the computer learning technology, the process of constructing the ontology, especially for decision making, will be modified. Finally, the most accurate decision making ontology model will be generated. With the graphic interface of knowledge network, this kind of research process helps professionals understand the correlation between different concepts and further inquire the information which meets their information needs. Thus, the project has based on our topological geometry model and embedded temporal events to a knowledge access and decision-making system with value added, such as comparison, connections, and consequences. The project focuses on designing and developing an ontological model to assist the professionals to organize the structural and unstructured data, and furthermore, extracts the temporal patterns behind those data.