

學術論述**生態資訊學之發展與應用回顧**林朝欽<sup>1)</sup> 鄭美如<sup>2)</sup> 陸聲山<sup>2,3)</sup>**摘 要**

本文回顧生態資訊學技術之發展，兼論台灣在生態資訊管理系統之建置與應用。生態資訊學發展雖較生物資訊學及生物多樣性資訊學稍晚，但有其特殊目的。該科技為處理複雜的生態資料所發展出來的資訊管理系統與工具。由於資訊之整合與分享是未來的趨勢，整合生物資訊學及生物多樣性資訊學始能互相分享資料。台灣長期生態研究網研究人員自2004年起建構生態資訊管理系統，此系統乃根據生態資訊學的技術標準。2007年台灣長期生態研究網正式完成設置此資訊管理系統，且此系統已為東亞太平洋地區長期生態研究網大部分會員網考慮採用。本文有三個應用此系統之例子可供參考。

**關鍵詞：**長期生態研究、資訊學、資料、資訊、知識本體論。

林朝欽、鄭美如、陸聲山。2008。生態資訊學之發展與應用回顧。台灣林業科學23(Supplement): S1-10。

Review article**A Review of the Development and Applications  
of Ecoinformatics in Taiwan**Chau-Chin Lin,<sup>1)</sup> Meei-Ru Jeng,<sup>2)</sup> Sheng-Shan Lu<sup>2,3)</sup>**【 Summary 】**

In this paper, we review the development and applications of ecoinformatics, an ecological information management system, for the Taiwan Ecological Research Network (TERN). Ecoinformatics is the latest development following bioinformatics and biodiversity informatics. It integrates the functions of bioinformatics and biodiversity informatics and deals with an enormous amount of diverse, complex, available eco-data. A team of TERN researchers began developing ecoinformatics for the ecological community in Taiwan in early 2004 and completed and tested the system in late 2007. Increasing numbers of local researchers are adopting this system. In contrast to other

<sup>1)</sup> 林業試驗所森林保護組，10066台北市南海路53號 Division of Forest Protection, Taiwan Forestry Research Institute, 53 Nanhai Rd., Taipei 10066, Taiwan.

<sup>2)</sup> 林業試驗所林業推廣組，10066台北市南海路53號 Division of Forestry Extension, Taiwan Forestry Research Institute, 53 Nanhai Rd., Taipei 10066, Taiwan.

<sup>3)</sup> 通訊作者 Corresponding author, e-mail: sslu@tfri.gov.tw

2008年8月送審 2008年9月通過 Received August 2008, Accepted September 2008.

systems developed in the West, this system is unique in that it uses EML (ecological metadata language) documents to resolve language and character-set issues. Many workshops introducing the use of EML documents have been held in Taiwan and other countries in the East Asian-Pacific region. Breakthroughs have gained wide popularity among researchers from other member networks in East Asian-Pacific regional ecological communities. Three examples of applications using this system are given.

**Key words:** LTER, informatics, data, information, ontology.

**Lin CC, Jeng MR, Lu SS. 2008.** A review of the development and applications of ecoinformatics in Taiwan. *Taiwan J For Sci* 23(Supplement):S1-10.

## 前言

生物學知識可概略取自分子、個體以及生態三個不同層次的科學資訊(Dial et al. 2003)。生物資訊學(bioinformatics)是發展較早的一個處理生物分子資訊的新學科，它對分子生物數據的管理與應用有相當顯著的成果，美國國家生物科技資訊中心(National Center for Biotechnology Information, NCBI)的entrez系統(Benson et al. 2005)便是一個例子。

生物多樣性資訊學(biodiversity informatics)則是針對個體物種數據的處理進行系統開發與工具設計，尤其是分散在全世界的博物館、標本館的物種標本的數位化與資訊管理的建構，可說是生物多樣性資訊學的重心(Bisby 2000)。全球生物多樣性資訊機構(Global Biodiversity Information Facility, GBIF)的DiGIR系統，目前已有87個國家共同使用(Canhos et al. 2004, Soberon and Peterson 2004)。

本文將論述的生態資訊學(ecoinformatics)發展雖較晚，但卻為整合前述兩個資訊學門所發展的新領域，以處理複雜的生態資料為主要考量，所發展的資訊管理系統與工具，亦朝向整合生物資訊學與生物多樣性資訊學，使此生態資訊學成為可分享資料的大系統(Fegrans et al. 2005, Jones et al. 2006)。美國國家生態學分析與整合中心(National Center for Ecological Analysis and Synthesis, NCEAS)與國際長期生態研究網的生態資訊管理委員會(International Long Term Ecological Research IM Committee)建構的生態元資料語言(Ecological Metadata

Language, EML)與科學工作流程系統Kepler可以說是此學科的代表，且已初步完成將基因整合到生物圈的資訊系統架構(Michener et al. 2007)。

上述三類與生物學有關的資訊技術可能整合成為「新生物資訊學」，此構想的需求性可以從2003年《科學》(Science)與《自然》(Nature)兩本學刊同年刊載大猩猩面臨瀕危的報導一窺端倪(Kaiser 2003, Whitfield 2003)。大猩猩並獲選為《自然》該期的封面主角。該等研究報導非洲大猩猩因感染伊波拉病毒與人類狩獵造成族群大量減少的危機，此危機的解決顯然涉及廣泛性的生物資料需求，它涵概了基因學、流行病學、棲地學、族群動態學、人類社會學等知識。要了解此一危機勢需整合多領域的研究資料，否則在管理上很難有一全面的明智判擇標準。

## 生態資訊學的技術發展

長期的研究生態與蒐集資料過程中，累積了數量龐大且差異性極高的原始數據，這些數據大部分是由個別科學家或研究室管理，因而分散各地且使用的倉儲工具不同，使得其他科學家無法取得這類數據，更談不上分享了。1970年以後個人電腦出現，大部分研究數據開始數位化，但所使用的軟體差異很大。1980年之後微軟的辦公室軟體(Office)上市後，很多研究以表格型態(spread-sheet)記錄原始數據。此種記錄方式易學

與易操作，但缺乏一個完整的資料結構與內容描述的管理機制，使得資料的保存與流通產生困難(McCullough and Wilson 1999)。

為解決此一困難，生態學家與資訊工作者合作，引進關連式資料庫(relational database)，原始數據以資料庫方式管理，提供較為具體的管理機能(Porter 2000)。但關連式資料庫需透過資料模式(data model)的建立才能有效的倉儲與使用，因此需以研究計畫為單元建構資料模式，形成了計畫導向的資料模式，因不同研究者建立的資料模式不同，使得相似的資料分享與整合出現另一個困難(Hass et al. 2002)。因此要整合類似計畫的資料需先整合資料模式，才能擴大資料的應用面。

為解決資料模式整合的問題，資訊技術提供資料倉庫(data warehouse)的概念，以垂直的方式將資料庫所建的資料模式加以對應，並建立一個更大的資料模式(Gray et al. 1997)。這樣的解決方式雖然擴大了資料庫的內容，但也因為整合過程中，不同的資料模式無法完全相容而犧牲許多資料內容，不過資料倉庫仍然提供了相當的資料保存與使用價值。

由於仍然有許多資料因缺乏有系統的管理而流失，不但減損了這些資料的價值，同時也相對提高了研究成本，更無法整合既有資料，產生新的研究價值(Michener et al. 1997)。因此如何在龐雜、分散與異質性的資料中確保資料的長久保存與取得，以及如何發揮資料的分享與整合功能，必須要有一個良好的生態資訊管理系統。更甚者，這樣一個系統需能協助研究人員妥善管理自己的資料，從資料開始蒐集前的設計到進入資料庫後的分享機制均需考慮，才能讓資料不再因計畫結束後即流失(Cook et al. 2001)。其次，這樣一個系統也需能實現資料再利用、延續資料生命與提昇資料價值。為了達成這樣的目的，首先需要建立一套共同的標準來描述資料內容，此一標準可讓各種類資料描述有所遵循，才能進行資料的存取、索引、檢索與共享，元資料(metadata)便是扮演這樣的角色與功能，元資料導向即是在一個標準化格式下描述所蒐集資料內容的標準，因此元資料

標準的選取、推動與實施是生態資訊管理系統最基本的工作與新的嘗試(Michener 2006, Tannenbaum 2001)。

生態元資料語言(EML)在1990年代開始發展，它是根據生態學科而發展出的元資料規格，由美國長期生態研究網、美國國家生態學分析與整合中心(NCEAS)、加州大學聖地牙哥分校之高速電腦中心、德州科技大學及美國國家科學基金會等共同贊助與發展的生物複雜性知識網路(Knowledge Network for Biocomplexity, KNB)計畫項下所開發出的，迄今它已成為生態資料庫聯結與整合的共同標準(Fegrans et al. 2005)。建置與產生符合生態元資料語言標準的文件，有利於日後和其他研究資料整合，並使生態研究因資料使用生態元資料語言的共通標準而更易於交流、合作，使生態研究邁向國際化(Lin et al. 2006)。

雖然元資料導向的生態資訊管理解決了生態研究領域的大部分資料管理問題，但仍然有許多分散且不易整合的數據島(data islands)(Jones et al. 2006)。這些數據島因資料格式、資料邏輯模型與資料語意(semantic)的問題，阻礙了整合與分享的進行，加上生態資料時常需結合基因、物種等生物與生物多樣性資料，而這兩種資料涉及生物與生物多樣性資訊學的標準與協定，使得生態學家仍然需要花費許多時間與精力處理原始數據。因此元資料導向的生態資訊管理尚需加以改善，近年來語意導向的新生物資訊學開始出現，此一導向強調應用知識本體論或稱存在論(ontology)作為建構新生物學領域內所有學門數據管理的基礎(Bard and Rhee 2004, Berners-Lee et al. 2001, Borgida 1995, Madin et al. 2008, Williams et al. 2006)。

## 知識本體論導向的挑戰與解決

知識本體論是哲學領域中研究存在(being as being)問題的學科，一般將之譯為「存在論」或「本體論」。它源自古希臘，至西元十七世紀才漸成為哲學上研究與論述的課題之一(Smith 2001)。ontology在希臘文中指

“onta”的學問(logy)，“onta”在英文中意指存在(beings)，因此中文多把知識本體論依英文意思翻譯為「存在論」或「本體論」。本體一詞在中國思想史上可追溯到宋明儒學，中國古儒所稱之本體與知識本體論所意指的內容並不全然相同，古希臘哲學家把知識本體論定義為「對存在事物的客觀描述」，因此被廣泛運用在定義客觀世界中的事務。並進而分類事物實體的討論與研究，它也蘊涵了對事物關係的推理(Smith 2001)。近代資訊學(informatics)引用知識本體論的特性作為組織、管理大量資訊與知識的基礎，尤其在網際網路發展的過程中，巨量的資訊湧現在網頁(web)中(Nagao et al. 2001)。如何提供快速有用的資料搜尋服務，是網路商業競爭成敗的技術。因此知識本體論在1990年代被資訊學引用，而成為目前語意網(semantic web)中不可少的基本理論(Fensel et al. 2007)。特別是搜尋引擎網，以往使用字串比對查尋方式，或稍後發展的元資料方案，使用者雖很容易獲得大量的查尋結果，但卻也獲得大量無關的信息，不但不符快速要求，更難以達到有效目的。改以讓電腦能自動判斷使用者輸入字詞的意義，進而由機器處理、解析，讓搜尋快速得到精確想要的資料，這樣的技術正成為目前網際網路上發展的主流。

就上述的演變過程，我們可以說從哲學範疇引進資訊學的知識本體論已被賦予新義，它在資訊學中的定義已與哲學不同。資訊學在知識本體論當成知識/訊息的概念定義、分類以及概念間關係的推理規則，因此資訊學中的語意指文件中的資訊，這些資訊是機器能自動運用與處理的，與一般只能由人類才能處理的語言上的語意是不相同的。

知識本體論的內容包括概念具體模型、概念明確化、形式化以及能分享四個層次(Pinto and Martins 2004)。概念具體模型通常指將一些抽象的客觀現象相互關連而得到一個具體的模型，在模型中的概念必須有明確的定義，進而把這些概念用某種電腦語言書寫成機器可處理的形式，最後成為可以共享用的集合。上述四個層次的架構，提供一個知識領域內共同理

解的知識體系，在此一體系下，該知識領域具有共同的詞彙，並且可以給出詞彙間之關係。

由於知識本體論的建構是由不同層次的抽象概念組成，不但具有分類學上的層次結構，更因為概念間的邏輯關係具有推理功能，因此在知識檢索中可以用以解決傳統上檢索的缺點，因此已被大量的應用在知識管理與檢索上(Madin et al. 2007)。

知識本體論在生物學及生態學研究上可以有什麼應用？我們可以說這是一個全新的課題，但卻不是生物學及生態學家完全陌生的課題，分類學上所用的分類系統及檢索表即是一個典型的生物知識本體系統(Kennedy et al. 2005)。如果生物學與生態學界有一份共同的或可連結的知識本體系統，各具共同詞彙，並定義出詞彙的關係，那麼電腦可以瞭解要驗證的模式語意，它就可以自動的擷取資料庫中的不同資料，無誤的計算出研究者所要的結果。這樣的敘述只不過是把傳統上認為必須由人到現地蒐集並完成資料整理的流程由機器來作，但目前聽起來仍像天方夜譚，讓人難以置信。

生態上具體的知識本體系統尚未出現，其原因是知識本體系統涉及到的不只是生態學家的專業本身，它還需要資訊學家的技術與邏輯學的觀念關係建立(Keet 2005)。建立了知識本體系統之後尚需將其中觀念與原始數據的內容、單位、型態等進行對位(mapping)，完成這些工作後還要開發解析軟體與使用者界面，因此在短期內要完成如此龐大而複雜的工作相當不易。但此項工作已在美國國家生態學分析與整合中心與國際長期生態研究網開始進行。美國國家生態學分析與整合中心已初步完成生態觀念與量測知識本體系統(measurement ontology)，並開發科學工作流程系統Kepler作為使用工具(Pennington and Michener 2005)。國際長期生態研究網也開始結合各區域網先建立生態元資料語言系統，再進行不同語言知識本體系統(language ontology)的建立(Lin et al. 2006)。因此可以預見未來生態資訊學將朝可以整合由基因到生態系層級資料的「新生物資訊學」方向的發展。

## 台灣在生物多樣性與生態資訊環境的建構

掌握知識便能掌控情勢，許多知識力量可來自資訊科技的獲得與應用，資訊科技促成「世界是平的」概念，因為全球資訊的流通已劃平人與人、國與國、洲與洲之間的障礙，讓全球化實質實現。台灣的資訊技術與通訊產業一向發達，尤其硬體產品與「開放原始碼」(open sources)的應用開發至為政府重視，加上網際網路的廣布，使得台灣已成為資訊大國，因此生物、生物多樣性與生態界的研究也受惠於此種發展。

2001年在國際生物多樣性公約之贊助下，國際上正式成立了全球生物多樣性資訊機構，其目的在推動全球各國生物多樣性資訊之整合與資訊分享。台灣亦正式簽署加入該組織，成為副會員(Associate Participant) (Shao et al. 2007)。副會員雖無投票會員(Voting Participant)之投票權，但並不需要繳交年會費，且可享有與投票會員相同之參加會議、發言及交換與分享資料之權利。但需要建立國家入口網站並整合國內生物多樣性資訊，與各國交換資料。因此中研院受我國國科會委託，建置台灣入口網TaiBIF，並成立國家委員會。全球生物多樣性資訊機構雖成立迄今僅七年多，但因在各國政府之大力支助、督導與配合，不但在短期內建構完成全球生物多樣性之資訊交換平台，並已整合全球1.4億筆全球物種名錄及生物標本典藏及分布資料。

1992年台灣建立長期生態學研究網(Taiwan Ecological Research Network, TERN)之初，亦已積極籌思建構生態資料管理問題的方案，想實現資料長期保存、有效管理與共同分享的理念，因而於1995年初選派人員赴美學習資訊管理，並在同年秋TERN開始建立資料管理技能，1998年與美國合作，在台灣主辦「ILTER全球資訊研習會」(International LTER Global Communication Workshop)，進一步的提供亞太地區的生態學者建立資料管理的觀念與技術。不過限於科學家的電腦科技能力、經費短絀與誘

因不足，尚無法落實建立生態資訊管理系統，1995年TERN雖建立了網站與製作了網頁，但對長期生態研究最關鍵的原始資料之保存、管理與分享的技術仍然生疏(TFRI 2008)。

然而在同一時期，在太平洋東岸的美國長期生態學研究網，卻已隨著資訊科技的發展步入「cyber」的另一境界，以生態元資料語言為標準，建立生態資料管理系統。這系統基本上達成了當初長期生態學研究最核心的理想，即依據長期累積的生態資料解釋科學問題(Porter 2000)。

2003年美國國家科學基金會提出「cyber-infrastructure」這個改變科學研究方法的前瞻概念，強調生態研究應該運用尖端科技突破傳統的研究窠臼之限制，例如透過「無線感測網」增加資料收集的頻率與提高資料品質及精準(Atkins et al. 2003)。台灣長期生態研究網的研究人員認識到「進步是靠前瞻的認知與立即的實踐」，於是在林業試驗所內，先更新生態資訊管理內容，邁向建立生態資訊管理系統而努力。當技術成熟後，也協助東亞太平洋地區長期生態研究網發展生態資訊管理，自2004年起開始建構的此系統，經過數年的投入，於2007年正式發佈此一生態資訊管理系統，這個系統的架構以生態元資料語言為標準，而利用網格(grid)技術為核心(Lin et al. 2007, 2008)。

網格技術概念是1985年被提出的，它突破了網際網路的功能限制，強調結合所有電腦成為虛擬的超級電腦，應用網格技術突破生態研究上的困難，如跨地蒐集資料、使用儀器、分析資料變得較容易(Smarr 1998)。21世紀初網格技術備受重視，技術不斷更新，顯然已成為新一代的網際網路。網格技術對生態研究亦造成重大的影響，例如可以執行跨洲際的研究計畫，但研究人員不必長途跋涉到另一大陸洲的現場試驗地，大量的資料運算也不必自己擁有大型電腦，不僅節省時間而且節省許多費用。這項技術目前也被美國長期生態研究網所取用，並建立Ecogrid基本架構來作為生態資訊系統溝通平台，讓生態學家可更進一步藉由Ecogrid上的分享資料分析跨研究地複雜的生態

問題(Pennington and Michener 2005)。ecogrid系統所具有的平台功能其實是透過目前最先端的網路服務(web services)技術所完成，也是格網運算技術全面發展的先驅，它有別於舊有的網頁瀏覽模式(Pennignton and Michener 2005)。舊有的網頁瀏覽模式來自網際網路的全球資訊網(world wide web, WWW)，人們可以藉由一隻滑鼠，連結到世界上任何一個有全球資訊的網站，瀏覽或下載各種文字、圖片、影音資料。但這種方式屬於人與機器之互動模式，由人解讀機器上有的資訊，這個功能雖已突破WWW初始時期的文字網頁模式，但仍不能滿足當前大量資訊交換、運算甚至快速傳遞的需求。近年來由於電子商業(e-business)的快速發展，機器與機器間之溝通已成為不可或缺的一項技術。ecogrid藉由格網運算技術解決生態研究長期累積了數量龐大，且差異甚大的資料檔案的管理問題，不但用在不同學門上，亦用在不同區域上，突破生態資料之整合的難題(Jones et al. 2006)。藉由ecogrid與網路服務技術，長期生態研究不但已可處理龐大資料的倉儲、整合，更已達到分享與再使用的功能。

其次是兩種可以相容於ecogrid架構下的倉儲系統：Metacat (Metadata catalog)與SRB (Storage Resource Broker)，在ecogrid架構下Metacat與SRB擔任了分散式的資料倉儲與使用功能(Jones et al. 2006)。這兩個伺服器以元資料作為資料交換標準，Metacat採用生態元資料語言作為標準，SRB則由使用者自行決定元資料標準，最重要的是這兩個伺服器屬於跨平台的系統，且使用文字格式作為交換標準，讓機器與機器能互相傳輸資料(Jones et al. 2001)。

最後是系統對資料的品質管制與分析功能，此一功能由資料檢核及科學工作流程(scientific workflow)來完成。科學工作流程其實是把科學家日常的工作過程加以規劃，並以標準的語法建立成數位檔案，讓系統能自動化完成資料分析與整合功能，更且這種功能是可以重覆與分享的(Pennington and Michener 2005)。

以上的系統所使用到的工具全為共享軟體，所有運用在這個系統上的工具，除作業系

統與網際網路系統的軟體外，核心軟體均由美國國家科學基金會所補助的KNB及SEEK計畫下所發展。但資料檢核界面則由林業試驗所的生態資訊管理研究小組與維吉尼亞州海岸保留地研究站(VCR LTER)合作計畫下所開發(Lin et al. 2007)。這個系統已提供完整的生態資訊管理，由最前端的資訊蒐集及到最後端的資訊利用均可充分支援。

## 生態資訊學的應用

台灣長期生態研究網所使用的生態資訊管理系統除提供建立研究資料目錄以外，也可以透過這個系統來結合生物、生態多樣性資訊學所建立的系統，另外也可以利用於解決巨量數據倉儲與模式推導運算的功能，以下三個例子是目前最能代表這個系統的使用。

### 一、基因序列

使用這個系統第三層的資料整合功能，可以透過網路服務與資料轉換工具搜尋儲存在世界各地的基因數據，使用者可以方便得到所需要的數據(Lacroix and Critchlow 2003)。應用該系統的資料整合功能，無需特殊的生物資訊學軟體，僅需建立查詢的工作流程即可。例如使用者想到日本查詢基因資料庫，該資料庫已提供網路服務功能，則使用者僅需透過Kepler並開啟網路服務工作單元(web service actor)及給予要查的基因序列(例如AA045112)，系統將自動與日本之伺服器溝通並取得資料展示出來。Fig. 1為此例子的工作流程及結果。

### 二、生態棲位模式(Ecological Niche Modeling)建構

生態棲位理論是生態研究上相當重要的理論之一，棲位這個名詞的使用在生態學領域由來已久(Grinnell 1917, 1924)。過去100年來生態學家給予不同意涵，1957年Hutchinson把棲位定義為一個n維的超容積(n-dimension hypervolume)，其中n代表某物種主要生存及繁殖的重要環境因子(Hutchinson 1957)。此定義提供

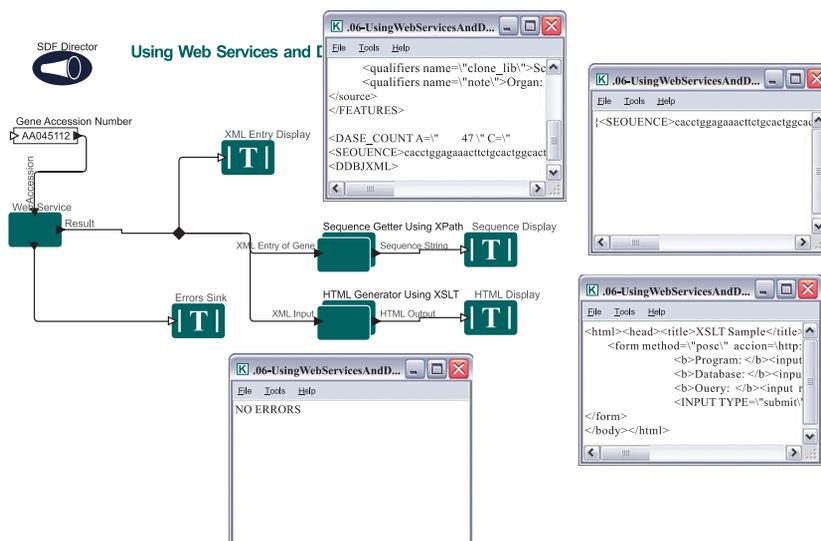


Fig. 1. Using ecoinformatics techniques for a DNA sequence search.

了預測物種分布的模式建構基礎，堪薩斯州立大學的自然史博物館發展出棲位模式建構與物種分布預測的方法論與軟體工具(Stockwell and Peters 1999)。因物種分布預測必須使用到物種調查或標本採集的資料及環境的地理資訊圖層，物種資料為生物多樣性資訊學以Darwin code為標準建立，而環境的地理資訊圖層則大部分以FGDC (Federal Geographic Data Committee)為標準建立，因此使用者須切換不同軟體系統來完成物種分布預測。然而若使用生態資訊管理系統的倉儲及整合功能，透過網路服務可以方便的結合生物多樣性資訊學的協定，並整合地理資訊的環境圖層，只要建立整個運算流程，系統可以自動完成使用者所要求的預測結果，Fig. 2是應用此系統結合林試所植物標本館資料庫進行金毛杜鵑(*Rhododendron oldhamii* Maxim.)的分布預測。

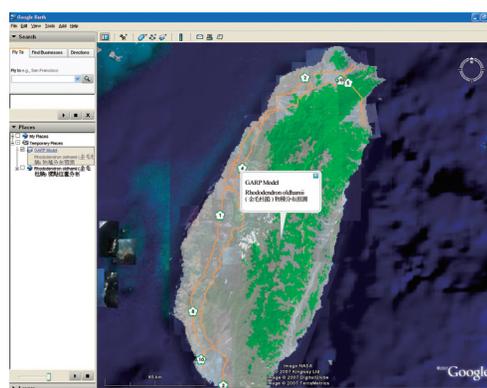


Fig. 2. Niche modeling using ecoinformatics techniques to predict the distribution of *Rhododendron oldhamii* Maxim.

### 三、二氧化碳通量數據管理與計算

大氣中二氧化碳濃度增加導致全球氣候變遷的相互關係已是眾所周知，二氧化碳濃度持續增加與人類活動脫離不了關係，也成為人類社會經濟永續發展的隱憂。當前國際社會許多具體行動也都對這樣環境問題加以關注，

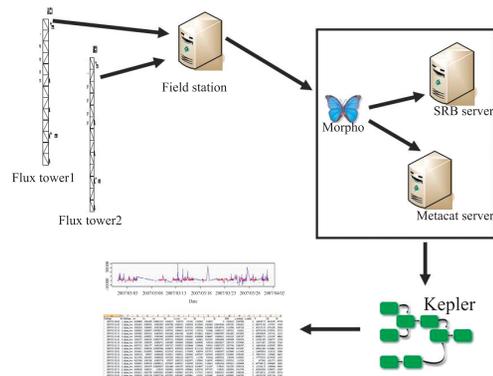
《京都議定書》是國際為保護地球環境，及針對溫室氣體，尤其是二氧化碳減量的具體行動之一，在《京都議定書》裡規範了締約國對二氧化碳的減量指標。因此，各國為爭取最大限度的國家利益，無不全力進行碳源、碳匯與收支平衡的研究，碳循環則是此類研究的主要內容。渦流相關法(eddy covariance)是目前測量二氧化碳通量最普遍的方法，此方法乃透過儀器

直接測量渦流運動的物理性質，以求得生態系與大氣間物質與能量的渦流通量，其原理為利用測定和計算氣體物理量與垂直風速的共變異數(covariance)來求算渦流通量，是目前國際上主要的通量觀測方法(Baldocchi et al. 1988)。然而渦流通量的直接觀測值必須經過一系列的處理，其中包括異常值處理、地形上座標軸的轉換、去除溫度與水氣含量造成的偏差等一系列的數據品質控制過程，另外還包含儀器故障、天候惡劣所造成的缺失值補插運算，方能提供有用的渦流相關通量資料進行生態學意義的解釋。截至目前仍然缺乏一套資訊管理系統，產生完整可靠的數據集，並且可以跨研究社群進行分享。林試所與東華大學合作應用生態資訊管理系統，成功的建構了一個完整的二氧化碳碳數據管理與計算工作流程，此流程使用了所有生態資訊管理系統的所有功能。Fig. 3是這個工作流程的內容與結果。

## 未來展望

過去生態學家囿於傳統觀念，視原始數據之釋出與分享為畏途，加上生態學家多半有升等與發表研究報告之壓力，對複雜且大尺度研究問題的興趣不高，使得生物多樣性與長期生態領域之原始數據難以長期保存與分享，目前已有前瞻性系統與工具，可以說都是未來邁向研究資料完善管理的有力保障。未來網際網路將納入知識內容亦已為趨勢，電腦不但是人類交換文件的媒體，而且將成為傳播知識本體的主要基底。目前生態資訊管理系統，已具備此兩重點之此初步規模，下一步為結合知識本體論，讓生態學研究的知識更能應用於政策研擬上。

推廣生態研究資料的長期保存及建立可以與世界發展接軌的觀念，協助研究人員可以妥善管理自己的資料，從資料開始蒐集前的設計到進入資料庫後的分享機制建立，不再因計畫結束後或其他原因，導致資料流失的情形發生，實現資料再利用、延續資料生命與提昇資料價值，是生態領域研究者未來需更努力的工作。還有如何應用下一代網際網路的尖端技



**Fig. 3. Flux data computing and management using ecoinformatics techniques.**

術，亦需有更多領域的專家投入與結合，人才不足可能是今後面臨的困境，突破這些困境需要政府增加尖端資訊科技的投資與意願，學界改變傳統資料管理與分享的觀念，方能達到資料長期倉儲與應用之長遠目的。

## 致謝

本文感謝國科會與美國國家科學基金會部分經費支持(NSC-95-2911-I-054-003與NSF DEB-06-21014)，並特別致謝促成台灣長期生態研究與資訊管理系統建置的幕後推手金恆鏞博士(前林業試驗所所長)與美國國家科學基金會的Dr. William Chang。此外，特別感謝美國長期生態研究網的Dr. John Porter與其他資訊管理員們及美國國家生態分析與整合中心的資訊人員對此系統建置所付出的貢獻。

## 引用文獻

Atkins D, Droegemeier K, Feldman S, Garcia-Molina H, Klein M, Messerschmitt DG, et al. 2003. Revolutionizing science and engineering through cyberinfrastructure: report of the National Science Foundation Blue-Ribbon Advisory Panel on cyberinfrastructure. Nat Sci

Fund. (National Science Foundation). 84 p.

**Baldocchi DD, Hincks BB, Meyers TP. 1988.** Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. *Ecology* 69(5):1331-40.

**Bard JBL, Rhee SY. 2004.** Ontologies in biology: design, applications and future challenges. *Nat Rev Genet* 5:213-22.

**Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2005.** GenBank. *Nucleic Acids Res* 33:34-8.

**Berners-Lee T, Hendler J, Lassila O. 2001.** The semantic web. *Sci Am* 284(5):34-43.

**Bisby FA. 2000.** The quiet revolution: biodiversity informatics and the Internet. *Science* 289:2309-12.

**Borgida A. 1995.** Description logics in data management. *IEEE Trans Knowl Data Engin* 7(5):671-82.

**Canhos VP, Souza S, Giovanni R, Canhos DAL. 2004.** Global biodiversity informatics: setting the scene for a “new world” of ecological forecasting. *Biodivers Info* 1:1-13.

**Cook RB, Olson RJ, Kanciruk P, Hook LA. 2001.** Best practices for preparing ecological data sets to share and archive. *ESA Bulletin* 82:138-41.

**Dial R, Bond B, Cushing JB, Nadkarni N. 2003.** How trees and forests inform biodiversity and ecosystem informatics. *Comput Sci Engin* 5(3):32-43.

**Fegraus E, Andelman S, Jones MB, Schildhauer MP. 2005.** Maximizing the value of ecological data with structured metadata: an introduction to Ecological Metadata Language (EML) and principles for metadata creation. *ESA Bulletin* 86(3):158-68.

**Fensel D, Lausen H, Polleres A, Bruijn JD, Stollberg M, Roman D, et al. 2007.** Enabling semantic web services: the web service modeling ontology. Berlin, Germany: Springer. 188 p.

**Gray J, Bosworth A, Layman A, Pirahesh H. 1997.** Data cube: a relational aggregation oper-

ator generalizing group-by, cross-tab, and sub-totals. *Data Min Knowl Discov* 1(1):29-53.

**Grinnell J. 1917.** Field tests of theories concerning distributional control. *Am Nat* 51:115-28.

**Grinnell J. 1924.** Geography and evolution. *Ecology* 5:225-9.

**Hass LM, Lin ET, Roth MA. 2002.** Data integration through database federation. *IBM Syst J* 41(4):578-96.

**Hutchinson GE. 1957.** Concluding remarks. *Cold Spring Harbor Symp. Quant Biol* 22:415-27.

**Jones MB, Berkley C, Bojilova J, Schildhauer M. 2001.** Managing scientific metadata. *IEEE Internet Comput* 5(5):59-68.

**Jones MB, Schildhauer MP, Reichman OJ, Bowers S. 2006.** The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annu Rev Ecol Evol Syst* 37:519-44.

**Kaiser J. 2003.** Conservation biology: Ebola, hunting push ape populations to the brink. *Science* 300:232.

**Keet CM. 2005.** Factors affecting ontology development in ecology. *Lect Notes Comput Sci* 3615:46-62.

**Kennedy JB, Kukla R, Paterson T. 2005.** Scientific names are ambiguous as identifiers for biological taxa: Their context and definition are required for accurate data integration. *Lect Notes Comput Sci* 3615:80-95.

**Lacroix Z, Critchlow T. 2003.** Eds. Bioinformatics: managing scientific data. San Francisco: Morgan Kaufmann. 440 p.

**Lin CC, Porter JH, Hsiao CW. 2007.** Automating analysis of sensor data using Ecological Metadata Language (EML). In Conference proceedings Coastal Environmental Sensing Networks. 2007 Apr 12-13, Univ of Massachusetts, Boston, MA, USA: US LTER.

**Lin CC, Porter JH, Lu SS. 2006.** A metadata-based framework for multilingual ecological

- information management. *Taiwan J For Sci* 21(3):377-82.
- Lin CC, Porter JH, Lu SS, Jeng MR, Hsiao CW. 2008.** Using structured metadata to manage forestry research information: a new approach. *Taiwan J For Sci* 23(2):133-44.
- Madin JS, Bowers S, Schildhauer MP, Jones MB. 2008.** Advancing ecological research with ontologies. *Trends Ecol Evol* 23(3):159-68.
- Madin JS, Bowers S, Schildhauer MP, Krivov S, Pennington DD, Villa F. 2007.** An ontology for describing and synthesizing ecological observation data. *Ecol Info* 2(3):279-96.
- McCullough BD, Wilson B. 1999.** On the accuracy of statistical procedures in Microsoft Excel 97. *Comput Stat Data Anal* 31(1):27-37.
- Michener WK. 2006.** Meta-information concepts for ecological data management. *Ecol Info* 1(1):3-7.
- Michener WK, Beach JH, Jones MB, Ludäscher B, Pennington DD, Pereira RS, et al. 2007.** A knowledge environment for the biodiversity and ecological sciences. *J Intell Info Syst* 29(1):111-26.
- Michener WK, Brunt JW, Helly JJ, Kirchner TB, Stafford SG. 1997.** Nongeospatial metadata for the ecological sciences. *Ecol Appl* 7(1):330-42.
- Nagao K, Shirai Y, Squire K. 2001.** Semantic annotation and transcoding: making web content more accessible. *IEEE Multimedia* 8(2):69-81.
- Pennington DD, Michener WK. 2005.** The ecogrid and the Kepler workflow system: a new platform for conducting ecological analysis. *ESA Bulletin* 86(3):169-76.
- Pinto HS, Martins JP. 2004.** Ontologies: How can they be built? *Knowl Info Syst* 6(4):441-64.
- Porter JH. 2000.** Scientific databases. In: Michener WK, Brunt JW, editors. *Ecological data: design, management and processing*. London: Blackwell Science. p 48-69.
- Shao KT, Peng CI, Yen E, Lai KC, Wang MC, Lin J, et al. 2007.** Integration of biodiversity database in Taiwan and linkage to global databases. *Data Sci J* 6:S2-10.
- Smarr L. 1998.** Grids in context. In: Foster I, Kesselman C, editors. *The Grid: blueprint for a new computing infrastructure*. San Francisco, CA, USA: Morgan Kaufmann. p 1-13.
- Smith B. 2001.** Objects and their environments: from Aristotle to ecological ontology. In: Frank A, Raper J, Cheylan JP, editors. *Life and motion of socio-economic units*. London, UK: Taylor & Francis. p 79-97.
- Soberón J, Peterson AT. 2004.** Biodiversity informatics: managing and applying primary biodiversity data. *Phil Trans R Soc Lond B Biol Sci* 359:689-98.
- Stockwell D, Peters D. 1999.** The GARP modeling system: problems and solutions to automated spatial prediction. *Int J Geogr Info Sci* 13:143-58.
- Taiwan Forestry Research Institute (TFRI). 2008.** Morpho: Ecological Metadata Language editing system. Taipei, Taiwan: TFRI. 152 p.
- Tannenbaum A. 2001.** Metadata solutions: using metamodels, repositories, XML, and enterprise portals to generate information on demand. Boston: Addison-Wesley. 528 p.
- Whitfield J. 2003.** Ape populations decimated by hunting and Ebola virus. *Nature* 422:551.
- Williams RJ, Martinez ND, Golbeck J. 2006.** Ontologies forecoinformatics. *J Web Semant* 4(4):237-42.