

資源組織與檢索規範：蝴蝶詮釋資料格式的發展
Resources Organization and Searching Specification:
The “Butterflies of Taiwan” Project

羅思嘉

Szu-Chia Lo

國立成功大學圖書館講師

Instructor, National Cheng-Kung University Library

陳雪華

Hsueh-Hua Chen

國立台灣大學圖書資訊系教授兼系主任

Professor and Chairman, Department of Library and Information Science

National Taiwan University

【摘要】

為配合國科會數位博物館專案計畫中「蝴蝶生態面面觀」之資料的著錄，「資源組織與檢索規範」研究群自 1998 年 8 月開始著手進行對蝴蝶資料的瞭解以及資料著錄項目的研究。「蝴蝶生態面面觀」是由國立暨南大學、國立自然科學博物館共同提出的先導計畫，主要目的是以台灣蝴蝶為內容主題，建立一個可以支援科普教育的數位博物館。「資源組織與檢索規範」研究群則是配合計畫內容主題及系統設計，以資訊儲存及檢索需求為基礎，發展蝴蝶資料詮釋格式。本文將針對計畫在資訊組織與檢索規劃過程中對資料欄位訂定考量及結果做分析與說明；並就訂定資料之結果欄位與都柏林核心集欄位做比較，以促進資料的交換。

Abstract

“Butterflies of Taiwan” is a sub-project under Taiwan Digital Museum Project (TDMP), sponsored by the National Science Council of Taiwan. “Butterflies of Taiwan”, a cooperative project, was proposed by National Chi-Nan University and National Museum of Natural Science; its metadata was developed by Resources Organization Searching Specification (ROSS, also a sub-project under TDMP) Research Team. In order to design the appropriate elements and create the butterfly metadata, ROSS started to gather relevant information on butterfly and information cataloging in August 1998. The main purpose of this project is to establish a digital museum to support and promote science education. Task of ROSS is the following: with respect to information storage and retrieval demand, to develop butterfly metadata format and design system specification based on the project content. This article presents the metadata format created for butterfly project and discusses issues related with its implementation. In order to promote information exchange, mapping of butterfly metadata to Dublin Core will also be presented.

【關鍵字】

資訊組織、資訊檢索、詮釋資料、都柏林核心集、蝴蝶

Information Organization、Information Retrieval、Metadata、Dublin Core、Butterfly

壹、發展背景與進行模式

國科會數位博物館專案計畫主題之一的「蝴蝶生態面面觀」，是以建立一個可透過網路檢索、使用的蝴蝶數位博物館為目的，希望藉由博物館所提供的資訊讓檢索者能夠對台灣的蝴蝶有更深的認識。這個計畫是由國立暨南大學（以下簡稱暨大）及國立自然科學博物館（以下簡稱科博館）共同提出的。

1998年圖書資訊學界的數位教授執行一項名為「資源組織與檢索之規範」（Resources Organization and Searching Specification, 以下簡稱ROSS）專案計畫（註1），作為國科會「數位博物館」專案計畫之下的一個分項計劃，其研究目標涵蓋在中文環境數位圖書館／博物館中，有關於資訊組織與檢索各個層面的重要議題；包括資料的儲存與管理系統的設計、使用者資訊需求及查詢資料的方式、各系統間整合等。ROSS在「蝴蝶生態面面觀」計畫中所扮演的角色是就資訊特色及運用層面，提供訂定資訊組織格式的意見，以確定資訊組織與整理的有效性，訂定蝴蝶詮釋資料格式。

一、「蝴蝶生態面面觀」計畫目的及計畫進行架構

（一）計畫的發展

台灣長久以來有「蝴蝶王國」的稱號，「蝴蝶生態面面觀」主要是建立一個數位化的物種生態博物館，讓使用者可以透過網路獲取相關資料，進而認識、瞭解蝴蝶。透過網路多媒體技術，除了檢索、瀏覽網頁相關資料外，並可參加虛擬科學講座及自學學程等不同方式以獲取相關的自然科學知識。蝴蝶相關資訊不單可就其傳統名稱、物種類別等方式檢索，同時也可以外形、顏色等形態機制查詢。

「蝴蝶生態面面觀」發展目的是以一個完整的蝴蝶資料庫提供研究與教學相關素材，作為大眾自然科學輔助教材的一部份（註2）。

（二）計畫的發展任務架構

「蝴蝶生態面面觀」這個計畫是由暨大數位博物館研究群與科博館共同合作執行，進行台灣特有蝴蝶典藏數位化的工作，建構一個虛擬的蝴蝶生態區。這個生態區除了內部資訊的豐富，有效提供資訊檢索與利用是影響系統運作的主要因素，而資訊的檢索與利用必須以良好的資訊組織及整理為基礎。ROSS研究群在「蝴蝶生態面面觀」計畫中所扮演的角色就是探討研究各現有數位博物館／圖書館系統所處理的資料特色、資訊組織與檢索方式、檔案架構、使用的相關標準以及系統整體結構，吸取各個不同計畫在建立過程中所獲得的各項經驗，整合使用

者需求研究的成果，應用於研究計畫的各環節中，設定一個適用於計畫同時又具有共通性的資料著錄格式。

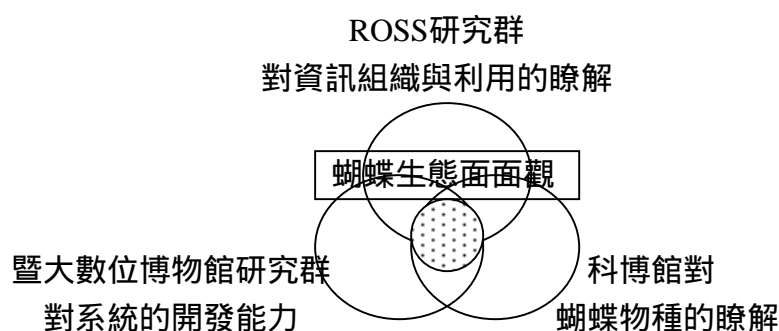


圖1：「蝴蝶生態面面觀」計畫的任務架構

二、著錄格式發展步驟

資訊組織與整理格式的訂定分為幾個部份進行，包括透過文獻分析、實地觀察、訪問座談等不同方法收集分析各相關資料、擬定中文資源著錄格式以及測試中文資源著錄格式等。

(一)文獻分析

ROSS研究群蒐集數位博物館/圖書館資訊組織之相關文獻，研究探討不同的詮釋資料，並分析探討資訊組織方法及檔案記錄結構。其中與「蝴蝶生態面面觀」計畫直接相關的文獻蒐集與研究包括參與博物館資訊交換協會(The Consortium for the Computer Interchange of Museum Information, 簡稱CIMI)的倫敦自然歷史博物館(The Natural History Museum)及北美草原野生生物研究中心(Northern Prairie Wildlife Research Center)的相關計畫。ROSS研究群透過相關資料的蒐集、研究與實際參與協會活動，瞭解各系統之設計特色、系統結構及資料處理方式。

(二) 觀察分析

除文獻研究外，ROSS研究群拜訪科博館，探討其描述藏品的方式，以及實際觀察科博館現有標本藏品及處理相關記錄之特性。透過調查，瞭解科博館的蝴蝶藏品及使用特質，掌握藏品的特性，作為基本資料格式訂定的基礎。

(三) 訪問座談

拜訪以蝴蝶為研究主題或對其有研究經驗的相關學者專家、解說員及對蝴蝶有興趣的資訊使用者，以訪談或者是以座談會方式瞭解各種可能的使用者資訊需求。

(四) 擬定基本資料著錄格式

根據以上文獻分析、觀察分析、訪問座談以及整合「使用者資訊尋求行為分析與使用環境之建立」(註3)部份之結果，擬定基本資料著錄欄位。

(五) 討論基本資料著錄格式並予以修正

就所設計的基本資料著錄格式，與暨大及科博館討論資料的適用性及資訊取得與計畫的配合度，並加以修正。

貳、蝴蝶詮釋資訊格式的訂定

一、科博館既有資料的架構與分析

在國科會數位博物館專案計畫開始前，科博館即已有其管理藏品的系統，主要是供館方在管理及應用藏品上的需求。科博館系統的資料主要是以標本為主，供博物館工作人員記錄各種標本資料的著錄與管理，其中不單包括蝴蝶標本的相關記錄資料，也包括其他昆蟲物種的資訊。由於在「蝴蝶生態面面觀」計畫進行之初，即是希望能夠以科博館既有系統中已儲存的資料為基礎，依照需求，由現有系統轉出至數位博物館中，以將重輸資料的需求降至最低。此外，理想狀況是系統除了滿足外在使用者資訊蒐尋的需求外，也可用於控管內部資料，如同一般實體博物館同時兼具管理及服務的機制。基於上述考量，瞭解科博館現有系統資料架構以及與數位博物館資料需求相比對，成為資料格式訂定初期階段首要工作。

(一) 資料檔案架構

科博館蝴蝶資料管理系統是其昆蟲管理系統的一部份，結構設計為關連式的資料庫，主要處理的資料分為資料收集、資料管理、生物分類、物種生態資料、參考文獻、資料使用記錄以及標本資訊等七個檔案。其中以資料收集、資料管理、生物分類這三個檔案為核心。參考文獻視實際收集狀況處理。物種生態資料、參考文獻、資料使用記錄以及標本資訊等，在討論之初，因尚未實際運用，因此暫不討論；但系統若以管理導向為優先考慮，則將更進一步討論相關細節。

資料收集主要是記錄標本資料的採集資訊，包括物種是在什麼地區收集到的，由誰所收集到的，什麼時間收集的等資訊。資料管理方面則是處理物種標本入館後的資訊，包括流水號、物種分類、標本狀況以及管理人等記錄。生物分類則是將物種分類系統數位化，並與其他資料相連結，記錄物種名、所屬的屬科種、物種發現者及時間等。

除了以標本資訊為主軸的資料外，科博館還收集了物種的參考文獻，並記錄文獻出處、相關主題等資訊；提供檢索者以作者、篇名及關鍵字等不同欄位檢索科博館所收錄的相關文獻資料。

(二) 資料內容的分析

科博館既有的資料記錄除了參考文獻，其餘資料均以標本藏品為中心，透過內部收藏流水號連結不同的資料。科博館目前所著錄的資訊可約略分為標本物種相關資訊、標本收集資訊及標本管理資訊三種類型。物種資訊提供標本名稱、分類及生態等基本資訊；標本收集資訊則記錄標本採集地區、時間、採集者及數量等；管理資訊則是針對資料入館後的狀況予以記錄。

物種相關資訊以最基本的物種資料為中心，乃以描述物種為主，包括物種的名稱、物種發現時間及發現者、物種的分類類別、以及其成長與生態環境等。標本採集資訊主要是提供標本取得的背景資料，先由蒐藏批號判斷標本的取得方式：館員採集、與其他單位交換、個人/單位捐贈、或是採購；並進一步在相關資料上記錄物種標本的原始採集地、採集時間、數量、負責人等館藏記錄。標本管理資訊主要包括標本入館後的標本狀況以及管理的相關資訊及典藏資訊，如同一般圖書館流通記錄。

從所括的資料類型而言，科博館資訊系統主要是以內部工作控管及資訊檢索為主，並非以提供一般外部使用者利用為目標。雖部份資料可運用於先導系統中，供外在使用者使用，但仍需增添生態環境、成長週期、各成長階段等方面更詳盡的資料；同時也需要進一步考慮到資訊檢索及呈現上的需求。

二、數位博物館/資料單位在物種資料上的處理

除了瞭解科博館資料架構、分析其可用性及應用模式，ROSS研究群進而參考其他數位博物館或類似計畫資料著錄方式，以提高規畫過程在考量上的完整度。在「蝴蝶生態面面觀」中所參考的網站中，以倫敦自然歷史博物館及美國地質調查局(U.S. Geological Survey)的北美草原野生生物研究中心兩個網站，最有參考價值。

(一) 倫敦自然歷史博物館

倫敦自然歷史博物館為CIMI會員。自1997起，倫敦自然歷史博物館參與由CIMI規劃的「都柏林核心集詮釋資料實驗計畫」(CIMI Dublin Core Metadata Testbed Project)，計畫目的是測試及瞭解日受重視的都柏林核心集，其15個欄位與著錄規範在處理各類型數位藏品的可行性。倫敦自然歷史博物館在計畫中以寄生蟲標本為測試資料，其資料本質與「蝴蝶生態面面觀」計畫處理蝴蝶物種資訊有共通之處。因此在規畫蝴蝶詮釋資料格式的過程中，研究群參考、分析該博物館所訂定的寄生蟲詮釋資料格式。

表1：倫敦自然歷史博物館都柏林核心集詮釋資料著錄格式（註4）

欄位名稱	欄位名稱(中)	欄位說明
DC.Title	名稱	屬種名
DC.Creator	創作者	鑑定者
DC.Subject	主題和關鍵詞	分類 1. 種 2. 屬 3. 標本狀態 4. 寄生蟲 (固定值)
DC.Publisher	出版者	固定為The Natural History Museum London
DC.Contributor	其他貢獻者	採集者
DC.Date	日期時間	採集時間
DC.Type	資源類型	型態，固定為physical object
DC.Format	資料格式	格式
DC.Identifier	識別資料	註冊號
DC.Relation	關連	關連資料 1. 科 2. Higher_Tax 3. Expendition 4. 藏品 5. 寄主
DC.Coverage	時空涵蓋範圍	涵蓋範圍 1. 當地地區 2. 其他國家 3. 生長分佈
DC.Right	權限範圍	資料所有權，固定為 http://www.nhm.ac.uk/generic.Copy.html

倫敦自然歷史博物館運用了都柏林核心集15個欄位中的12個欄位，乃以物種資料為主；簡述、來源、語文這三個欄位不在著錄格式欄位中。

(二)北美草原野生生物研究中心（註5）

北美草原野生生物研究中心是由美國地質調查局的生物資源部門所主導的研究中心，主要任務是收集北美草原各種野生生物數量、生長、遷徙環境等相關資訊，主持各項研究計畫、並透過研究中心的網站，提供研究者或對野生生物有興趣的使用者使用。在各種資料中，生物物種資訊佔了相當大的比例，物種包括蝴蝶。系統按照類別、分佈地區，以瀏覽方式提供資訊檢索，並提供物種的圖檔。

表2：北美草原野生生物研究中心蝴蝶物種詮釋資料範例（註6）

原始欄位名稱	資料說明	資料範例
[Name] ¹	名稱	Two-tailed Swallowtail
[Scientific Name] ²	學名	Papilio multicaudatus
Wing span	大小	3 1/2-5 inches (9-12.7 cm)
Identification	特徵	Upper surface of male fore wing with narrow black stripes. ...
Life history	生態習性	Males patrol stream courses or city streets for ...
Flight	發生期	One flight May to mid-August in ...
Caterpillar hosts	寄主植物	Leaves of ash (Fraxinus), ...
Adult food	蜜源植物	Nectar from thistles, ...
Habitat	分布：小區環境	Foothill slopes and canyons, ...
Range	分布：大區環境	Western North America south from ...
Conservation	保育	Not usually required
Management needs	管理	None noted
The Nature Conservancy Global Rank	全球保育等級	G5-Demonstrably secure globally, thought it may be quite ...
Reference	參考資料	Scott, J.A. 1986. The butterflies of ...
Image ³	圖檔	182877.jpg

註：1.網頁資料上未特別給予欄位名稱，直接著錄物種通用名稱。

2.學名著錄於通用名稱後的括弧內，Two-tailed Swallowtail(*Papilio multicaudata*)。

3.系統所提供的物種影像依狀況各有不同，多為幼蟲、雌雄成蟲及背腹等圖型。

北美草原野生生物研究中心蝴蝶物種詮釋資料為ROSS研究群提供了一個組織整理網路資源的範例。除了瞭解資料組織的架構外，研究群並從中分析、評估數位博物館如何能夠以不同的風貌呈現藏品。

三、「蝴蝶生態面面觀」詮釋資料欄位的訂定

除了分析科博館現有資料以及瞭解其他相關數位資訊系統之外，ROSS研究群在訂定系統資料著錄欄位過程中，還必須考量資訊的辨識、檢索、擴充度及分享等方面。首先考慮的是哪些資訊及其呈現方式能夠輔助使用者瞭解物種、運用資料，因此每一筆記錄所需要包括的物種資料及資料單元是研究群的第一考量點。

其次要考慮的是在所著錄的資料中，有哪些欄位是使用者可能用來檢索資訊的檢索值，包括基本的資訊，如物種的各種名稱、學名、中文譯名，以及描述性資料的檢索，如物種的外形、顏色、發現的地區等。直接與檢索點相關的是檢索延伸的必要性，例如系統是否要提供找尋相同科屬種、相同蜜源或者是出現在同一地區、時間範圍物種的檢索功能。對於所查到的資訊，使用者會產生哪些在資料整理上的需求，例如排整、比對、分析及輸出欄位與模式。在ROSS研究群討論過程中，另個受到相當重視的問題是：系統未來在運作上的擴充性及調整度。

有鑑於「蝴蝶生態面面觀」是先導型計畫，許多議題目前無法充分掌握，包括系統需求、使用者層級與使用方式等，同時也要考慮未來在不同物種上的運用，因此在欄位的架構上必須要有某種程度的擴充性與調整度，以容納其他資料的處理。

另一個考量點則是訂定的資料記錄欄與國際標準的相容性。數位博物館/圖書館計畫的主要動機是希望能夠有效地檢索及分享網路資源；資料能否滿足網路跨平台檢索及資訊交換，乃取決於所訂定的資料組織檢索規範與相關標準間的共通性。

(一) 系統需求

1. 系統使用預定的使用對象 – 使用者研究

雖然「蝴蝶生態面面觀」計畫是以科普教育為目標，但在討論初期，參與者對系統的期望是：除了提供對蝴蝶有興趣的一般使用者或從事解說工作的解說員一些可利用的資訊，也希望能夠整合相關的研究資源提供專家學者一個蒐集資訊的管道。在這樣的考慮下，從對蝴蝶認知層面而言，可能的使用者包括初學者、對蝴蝶有基本認識的、及專家學者；從資訊利用目的而言，可分為收集瀏覽認識蝴蝶的資訊、搜尋教材、蒐集研究資料、或者是館方所需的典藏、管理功能。若考慮所有變數，系統則需依照不同族群的需求設計不同的使用介面、檢索方式、資料欄位及資訊呈現方式。

2. 資訊檢索需求

詮釋資料除了提供使用者認識物種過程中所需要的資訊外，另一個重要目的就是做為使用者檢索資訊的依據。「蝴蝶生態面面觀」系統除了提供類別、名稱瀏覽檢索外，也希望設計不同資訊蒐尋模式的檢索功能，包括科別查詢、關鍵字檢索、外形比對、物種特徵及影像內容檢索。

科別檢索，可以是直接鍵入科別詞彙，或者是提供一個科別架構表，以瀏覽方式檢索資訊；關鍵字查詢則是就所輸入的詞彙與系統記錄內容文字進行比對。類別、科別、物種名稱、關鍵字等類型的檢索都是一般資訊系統常見的檢索功能，在欄位訂定與詮釋資料著錄上多無爭議。引發較多討論的部份，是在外形、特徵及影像比對檢索。為了更詳細描述物種，除了物種影像外，各個特徵及外形的描述是不可或缺的--但是，這樣的文字描述是否適用於資訊檢索？以顏色為例，當使用者描述所看到的蝴蝶顏色為白底黑點，而試圖以此檢索條件查詢相關資料，系統如何確認所看到的是背面或腹面的色彩，其中顏色的比例為何？斑點形狀又是如何？更重要的是資訊描述者對物件的描述是否是檢索者看到現象的認知與描述？若只是單純的文字描述比對，其間的檢索效率令人質疑；若要提高檢索效率，勢必加入參數、套版的控制與資訊輸入。

綜觀在系統需求部份，ROSS研究群需要考慮的是，博物館主要服務的對象、

預定的主要對象如何檢索資訊、主要對象需要什麼樣的資訊、使用者希望瞭解哪些與蝴蝶有關的資訊、系統提供的資訊將如何被使用；在考慮主要服務對象的同時，潛在的使用者需求也是不容忽略的。

(二) 「標本導向」還是「物種導向」？

計畫討論初期，參與人員希望系統兼具雙重功能，包括標本管理以及科普資訊的建立；但在記錄單元上，必須在標本與物種為單位考量上擇其一。若以標本為依據，在資料記錄中就需標註標本狀態的相關資訊，而標本相關的物種資訊勢必出現重覆著錄的現象；若以物種為主，則每一記錄在對應的標本項目上就需考量欄位的可重覆性及標本細節的著錄；若兩者均考慮，則在欄位設計上需多面向考慮，同時顧及兩種類型資料的連結。

(三) 資訊著錄的目的

在發展每筆記錄所需包括的物種資料及資料單元過程中，ROSS研究群與計畫各個部門反覆討論每一資料單元、內容著錄的目的及功用，包括可能會被用來檢索資訊的檢索點，如物種的各種名稱，學名、中文譯名，以及描述性資料的檢索，如物種的外形、顏色、發現的地區等。其次也考慮到，有些資訊不僅用於直接檢索，在資訊蒐尋過程中，需要具有檢索延伸的必要性，例如找尋相同科屬種、相同蜜源或者是出現在同一地區、時間範圍物種的檢索功能，或者是具有相同或類似外型特徵的物種。

(四) 交換標準

系統發展最終目標是：建立的資訊，不單用在所設計的檢索系統，還要能夠透過網路提供不同系統或檢索介面使用，因此ROSS研究群在發展格式的標準時，也十分關切資訊互通的議題。需要注意的交換標準包括欄位的共通性，如與都柏林核心集的對照，Z39.50 Profile的檢索欄位、資訊著錄的權威標準，如地名及日期格式、以及資訊組織包裝格式語法的考量，如XML、RDF。

四、詮釋資料的著錄與應用

在資訊組織格式發展初期，其設計原則是系統所提供的資料必須同時滿足資訊搜尋者檢索資料與館方管理方面的需求，也就是說系統設計是以標本導向為起點，以既有資料為基礎加入不足的資料，系統考量甚至包括標本外借條件及管理。因此在初期所規畫的著錄資訊分為系統管理段、著錄段、主題段、敘述段、典藏段、採集段、標本段、重製段、關聯段及附註段。系統管理段所記錄的是該筆記錄的相關資料，多可由系統自動產生；著錄段則記錄該標本的物種名稱、發現者及發現時間等資訊；主題段所記錄的是該標本物種在分類體系中的地位以及物種在外型與特徵上的類別；敘述段則著錄該物種各種生態行為的資料；典藏段則是包括標本在科博館內的典藏狀況，如收藏位置等；採集段則記錄標本採集歷

史資訊；標本段則包括標本製作及保存等資訊；重製段，若標本同時被製作成幻燈片、照片等其他媒體，相關資料則著錄於重製段的相關欄位；關聯段，若標本類別或典藏具關連性，相關記錄在此著錄；附註段則包括其他段落暫時無法處理的資訊。

ROSS研究群就計畫首要目的、系統檢索功能及優先順序考慮、討論草案格式的各資料欄，但受限於計畫主要目標、複雜度及時間考慮，ROSS研究群重新修訂詮釋資料格式並做了大幅度的更動。更改部份包括(1) 詮釋資料將以物種描述為主；(2) 外形、特徵等相關敘述將以描述為主，檢索功能將由檢索程式另外進行內部資料整理；(3) 系統必須能夠支援資料交換及跨平台檢索需求，包括以XML或通用標準格式包裝詮釋資料及依不同對照標準轉換資訊。

詮釋資料格式分為七個部份，包括名稱段、主題段；相關資料段、敘述段、重製段、管理段、附註段等。名稱段為物種的辨識資訊，包括物種的中文名稱、別名及學名等資料；主題段主要是註記物種在分類體系中的所屬地位；相關資料段與敘述段分別記錄物種的生態相關資訊；重製段所著錄的是物種影像檔案的相關資訊；管理段現階段僅記錄最後修改日期，資料著錄格式是依ISO 8601格式著錄；附註段則是記錄其他欄位無法著錄的資料。

表3：資訊組織架構簡表

段落名稱	資料特質
名稱段	中文名、別名、學名、命名者、命名年
主題段	科、屬、種、亞種、中文科名
相關資料段	寄主、蜜源植物、地理環境
敘述段	生態期、外形特徵、分布情況
重製段	檔案格式
管理段	最後修改日期
附註段	備註

七個段落資料可再細分為11個欄位，17 個分欄；生態期則包括12個共用分欄。其中必備欄7個，包括中文名、學名記錄、分類、寄主植物、台灣分佈地區、媒體資訊及最後修改日期等；媒體資訊及各生態期的影像資料欄位是可重覆的。物種型態資訊著錄於各生態期的形狀、顏色、大小、特徵等各欄位，以描述說明功能為主，不用於資訊的檢索。

表4：物種詮釋資料格式

欄位名稱	XML欄名	說明	必備	重覆
中文名	Cname	蝴蝶的中文名稱	M	
別名	Nickname	蝴蝶的其他名稱，包括土名及其他譯名		
學名記錄	Present-SN-record		M	
現有學名	Present-SN	目前學名	M	
學名建立者	Present-SN-author	學名命名者	M	
學名建立年	Present-SN-year	命名年代	M	
分類	Classification		M	
科	Family	以原文著錄	M	
中文科名	Cfamily	中文譯名	M	
屬	Genus	以原文著錄	M	
種	Species	以原文著錄	M	
亞種	Sub-species	以原文著錄		
寄主植物	Hostplant	卵至蛹階段的寄生植物	M	
蜜源植物	Honeyplant	成蟲的蜜源植物		
地理環境	Geographic			
台灣分佈	Taiwan	臺灣本地成長、出現地區	M	
其他地區	Global	其他非台灣地區		
生態期	Life-stage			
卵	Egg	參考生態期的分欄表		
幼蟲	Larva	參考生態期的分欄表		
蛹	Pupa	參考生態期的分欄表		
成蟲	Adult	參考生態期的分欄表		
媒體資訊	Media-information			R
影音格式	Media-type	影像資料格式	M	
描述主題	Media-title		M	
URL	Media-URL	連結位址	M	
最後修改日期	Update	ISO 8601	M	
備註	Footnote			

物種各階段生態期欄位具共通性，分為形狀、顏色、大小特徵、生長分佈、天敵 ... 等11個欄位。

表5：生態期詮釋資料格式

分欄名	XML欄名	說明	必備	重覆
形狀	Feature	外觀型狀		
顏色	Color	背、腹面顏色及圖型		
大小	Size	記錄體長或成蟲展翅長度		
特徵	Characteristic	特殊型態		
圖形	Image			R
	圖形 URL	url	M	
	圖說	Description		
生長分佈	Habitate			
天敵	Predator			
成長天數	days-of-growth			
偽裝避敵術	Defense			
季節	Season			
其他生態行為	Behavior			

肆、詮釋資料格式議題分析

目前先導計畫系統的評估尚未完成，但就目前已建立的資料，就資料格式、檢索及資源交換，可從幾方面來探討後續階段需改善的部份，包括欄位資料著錄規範、欄位的訂定、資訊的檢索及可擴充功能等。

一、欄位資料著錄規範

詮釋資料中有不少是開放式描述欄位，在瀏覽原始資料的過程中發現一旦資料增加到一定程度，開放性文字對相同現象描述就容易出現多種不同的描述方式。雖說現階段許多描述性資料是以說明目的為主，但考慮到關鍵字檢索及日後的處理，遵循一定的描述法則、規範是有其必要性的。同時可參照權威控制資訊的欄位，如分佈地區，應參照目前已被通用認定的標準、格式來著錄。

二、欄位的可重覆性及著錄格式

目前詮釋資料格式僅圖型相關欄位是可重覆的，因此也造成單一欄位多重數值的情況，其中以物種分佈地區、物種天敵及避敵術最為明顯。這樣的情況在描述說明上並無太大的影響，但一旦涉及資料的檢索與交換就會造成資訊的不確定性，同時也需要進一步的分析、整理。因此在格式上應考慮欄位的可重覆及著錄資訊的單一化。

三、檢索資訊的交換

由於系統在部份檢索功能上是以內部程式方式控制，如外觀形狀模組、顏色、體型特徵等資訊，這些資訊除了內部檢索系統使用，未來是否可以支援跨平

台檢索仍有待討論。在討論詮釋資料格式時，為了考慮資料著錄的負擔及著錄值的控制，對於與蝴蝶外觀特徵相關的資訊均以圖像檢索技術處理，不單可以達到資料一致性，同時也較具成本效益。但是以這樣技術處理出的資訊，目前僅可用於「蝴蝶生態面面觀」系統上，尚無法與其他系統分享。

四、與都柏林核心集欄位的對照

都柏林核心集是目前處理網路資源頗受矚目的詮釋資料格式，同時也被相關計畫視為網路資源交換的核心欄位。有鑑於此，將蝴蝶詮釋資料欄位與都柏林核心集欄位做一對照，探討詮釋資料在轉換上的潛在議題。

表6：都柏林核心集與蝴蝶詮釋資料欄位對照表

DC Element	中文 DC	蝴蝶詮釋資料格式	說明
Title	名稱	(1) 中文名 (2) 別名 (3) 學名	物種名稱，包括學名及各種名稱。
Creator	創作者	學名建立者	給予現有學名者。
Subject	主題和關鍵詞	分類	科、屬、種、亞種及中文科名。
Description	簡述	(1) 寄主 (2) 蜜源 (3) 生態期 (4) 備註	包括各生長階段的資訊，寄主植物、外形特徵、天敵 ... 等。
Publisher	出版者	N/A	系統預設，科博館、暨大或國科會數位博物館專案計畫。
Contributor	其他貢獻者	N/A	系統預設，科博館、暨大或國科會數位博物館專案計畫；也可著錄資料整理者。
Date	日期時間	學名建立年	遵循 ISO 8601 以 8 位數格式，yyyy-mm-dd，著錄。
Date.Modified	最後修改日期	更新日期	遵循 ISO 8601 以 8 位數格式，yyyy-mm-dd，著錄。
Type	資源類型	N/A	系統預設，Physical Object。
Format	資料格式	媒體資訊	資料格式，有利於應用軟體的判斷及資料的使用。
Identifier	識別資料	URL	也可加註系統辨識號碼。
Source	來源	N/A	系統預設值，科博館。
Language	語文	N/A	系統預設值，中文，遵循 RFC1766 規範。
Relation	關連	科名	著錄科名，以檢索相同科別物種資料。
Coverage	時空涵蓋範圍	地理環境	物種分佈區域。
Rights	權限範圍	N/A	連結到資料使用說明資訊。

蝴蝶詮釋資料中各欄位的資料在都柏林核心集中均可對應到適當的欄位，但部份欄位在處理上有需討論的課題，包括精確性、分欄處理以及一對一的問題。

(一) 精確性

蝴蝶詮釋資料欄位對應至都柏林核心集欄位有部份是多對一的關係，如名稱欄位。蝴蝶詮釋資料中與名稱的相關欄位包括中文名、別名及學名，從資訊的精確性考量，在資訊轉換過程中是否需要標明所轉錄資料的屬性？名稱一欄是較不具爭議性的，可能產生問題的是在生態資訊部份。蝴蝶詮釋資料中與物種生態相關的資訊多對應至都柏林核心集的Description一欄，包括寄主植物、蜜源、各成長階段可能遭遇的天敵...等，如何在都柏林核心集資料中辨識轉錄資料所代表的意義？是否需要某些識別註記？

(二) 分欄集的處理

蝴蝶詮釋資料中有部份資料是非獨立存在的欄位，如分類中的科、屬、種資料。分類資訊目前是對應至都柏林核心集的Subject一欄，但科、屬、種應轉入同一都柏林核心集的Subject欄位，以符號分隔；還是應將科、屬、種資訊分別寫入不同Subject欄位；各資料項是否需要標示資訊屬性？除分類一欄外，生態期資訊也面臨相同問題，各生長階段的外形、成長天數、天敵、避敵術等都需連結至生長階段才具有其意義。

(三) 一對一的問題

目前蝴蝶詮釋資料中有些欄位資訊的單一性還不足，也就是欄位儲存多重數值，這樣的情況造成轉換資料時辨識上的困難，同時較難做到一對一的比對。

五、可擴充功能

(一) 跨平台檢索

「蝴蝶生態面面觀」所提供的資料限於系統本身所整理的資訊，並未對外連結其他相關網站，做一整合性、跨平台的資訊檢索。對於全球蝴蝶物種資訊的完整提供仍有相當的距離。同時前面所討論到的檢索資訊交換議題，在跨平台查詢機制上需要更進一步的討論。

(二) 相關資料的連結

除了科博館所提供的相關訊息，科博館附屬圖書館及其他資訊單位均收藏十分豐富的相關圖書及文獻資料，但目前並未被包括在系統中。這類資訊對使用者進一步的自我教育及資料收集均有相當大的助益，是下一階段系統應考慮整合的資源。

伍、結語

「蝴蝶生態面面觀」計畫到八十八年九月已告一段落，具有初步的成果。根據系統統計有318筆XML資料記錄、208種，分屬11科的台灣蝴蝶物種資料；近兩千個標本及物種圖檔。計畫雖暫時畫下句點，但在讀者需求、使用研究與資訊詮釋資料仍有許多議題需要進一步的討論。除了計畫處理資訊的詮釋資料外，所訂定的詮釋資料格式也需要與國科會數位博物館專案計畫藏品之特性及使用者尋求行為做一分析比較，以期建立各計畫間資料分享及資源檢索共通的管道。

附註

註1： Resources Organization and Searching Specification

<<http://ross.lis.ntu.edu.tw>>

註2： 國科會「數位博物館專案計畫」概況

<<http://www.nsc.gov.tw/y2k/dml/880209DATA2.html>>

註3：「使用者資訊尋求行為分析與使用環境之建立」為ROSS子計畫二

註4：摘錄自http://www.cimi.org/documents/meta_simple_mapping2.html

註5：Northern Prairie Wildlife Research Center

<<http://www.npwrc.usgs.gov>>

註6：Northern Prairie Wildlife Research Center蝴蝶物種詮釋資料範例

<<http://www.npwrc.usgs.gov/resource/distr/lepid/bflyusa/wa/708.htm>>

致謝：

特別感謝暨大數位博物館研究群及科博館共同參與討論並提供相關資料，讓本研究得以順利完成。另外，也非常感謝台大植病所博士班曾兆祥同學在蝴蝶專業知識的提供，以及屈貝琴小姐在資料尋找與整理方面所給予的幫助。

附錄1：青帶鳳蝶詮釋資料

欄位名稱	XML欄名	欄位值
中文名	Cname	青帶鳳蝶
別名	Nickname	N/A
學名記錄	Present-SN-record	
現有學名	Present-SN	Graphium sarpedon connectens
學名建立者	Present-SN-author	Fruhstorfer
學名建立年	Present-SN-year	1906
分類	Classification	
科	Family	Papilionidae
中文科名	Cfamily	鳳蝶科
屬	Genus	Graphium
種	Species	Sarpedon
亞種	Sub-species	Connectens
地理環境	Geographic	
台灣分佈	Taiwan	全臺灣軍有分布、中央山脈四周低平山區街可見到成蟲活動，一般都市近郊亦可見到成蟲活動
其他地區	Global	日本、中國大陸、東南亞國家、新幾內亞、所羅門群島及昆士蘭等地區均有分布
生態期	Life-stage	
卵	Egg	
形狀	Feature	圓球形
顏色	Color	外觀為淡黃色
大小	Size	直徑約為1.3mm，高約1.25mm
特徵	Characteristic	表面光滑，精孔不明顯
天敵	Predator	卵寄生蜂
成長天數	Days-of-growth	卵期約為4-6日
幼蟲	Larva	
形狀	Feature	幼蟲體呈圓筒狀
顏色	Color	一齡幼蟲黑褐色，隨著幼蟲成長，色澤逐漸轉變為綠色，終齡幼蟲頭部綠色，體呈翠綠色，體表密生白色細微絨毛
大小	Size	終齡幼蟲體長約為 35~45mm
特徵	Characteristic	三胸節背方各於左右兩側有一金屬藍色突起，後胸背方有一黃色橫紋，腹節末端有一對黃色叉狀突起
天敵	Predator	小繭蜂、寄生蜂、寄生蠅、椿象、鳥類等捕食性天敵
成長天數	Days-of-growth	幼蟲發育期約需 20~25 日

表6：青帶鳳蝶詮釋資料 (續)

欄位名稱	XML欄名	欄位值
生態期	Life-stage	
幼蟲	Larva	
偽裝避敵術	Defense	剛孵化之初齡幼蟲首先攝食卵殼，平時多停棲於寄主植物葉表，前胸具黃綠色半透明臭角，平時縮於前胸背板縫內，受到驚嚇或騷擾時，可外翻伸出，並分泌具臭味之黃綠色液體，用以驅敵
蛹	Pupa	
形狀	Feature	蛹為帶蛹，絲帶通過後胸背方懸掛於固定物上
顏色	Color	蛹體呈淡綠色
大小	Size	蛹長約 30~35mm
特徵	Characteristic	中胸背方有一狹長角狀突起，由突起有四條黃色縱脊延伸到尾端
成長天數	Days-of-growth	蛹期約為 30~35 日不定
偽裝避敵術	Defense	老熟幼蟲於寄主植物枝條或附近隱密之植物枝葉間尋找適當地點吐絲化蛹
成蟲	Adult	
形狀	Feature	雄蝶前翅端稍尖，整體接近長三角形。後翅接近三角形，外緣呈明顯波浪狀，不具尾狀突起，後翅內緣反捲，內面密生黃灰色細長鱗毛。雌蝶體型較大，外觀與雄蝶相近，前翅略寬，後翅內緣無明顯性徵
顏色	Color	此蝶前、後翅表底色為黑褐色，前後翅中橫線處有連接的淡藍色塊狀斑紋，後翅亞外緣各翅室有淡藍色弦月紋
大小	Size	本種為大型蝶種，展翅約為 4.5~5.5cm
天敵	Predator	蜘蛛、蜥蜴、青蛙、蜻蜓、螳螂、鳥類等捕食性天敵
偽裝避敵術	Defense	成蟲飛行迅速，捕捉不易
媒體資訊	Media-information	
影音格式	Media-type	N/A
描述主題	Media-title	N/A
Url	Media-URL	N/A
資料更新日期	Update	1999/3/8
備註	Footnote	N/A

附錄2：蝴蝶詮釋資料範例 - XML：青帶鳳蝶

```
<?xml version = "1.0" encoding="Big5"?>
<!DOCTYPE butterfly SYSTEM "butterfly.dtd">
<butterfly>
  <cname>青帶鳳蝶</cname>
  <present_SN_record>
    <present_SN>Graphium sarpedon connectens</present_SN>
    <present_SN_author>Fruhstorfer</present_SN_author>
    <present_SN_year>1906</present_SN_year>
  </present_SN_record>
  <classification>
    <family>Papilionidae</family>
    <cfamily>鳳蝶科</cfamily>
    <genus>Graphium</genus>
    <species>sarpedon</species>
    <sub_species>connectens </sub_species>
  </classification>
  <hostplant>寄主植物</hostplant>
  <geographic>
    <taiwan>全台灣均有分布，中央山脈四周低平山區皆可見到成蟲活動，
    一般都市近郊亦可見到成蟲活動</taiwan>
    <global>日本、中國大陸、東南亞國家、新幾內亞、所羅門群島及昆士
    蘭等地區均有分布。</global>
  </geographic>
  <life_stage>
    <egg>
      <feature>圓球形</feature>
      <color>外觀為淡黃色</color>
      <size>直徑約為 1.3mm，高約 1.25mm</size>
      <characteristic>表面光滑，精孔不明顯</characteristic>
      <predator>卵寄生蜂</predator>
      <days_of_growth>卵期約為 4~6 日</days_of_growth>
    </egg>
    <larva>
      <feature>幼蟲體呈圓筒狀</feature>
      <color>一齡幼蟲黑褐色，隨著幼蟲成長，色澤逐漸轉變為綠色，
      終齡幼蟲頭部綠色，體呈翠綠色，體表密生白色細微絨毛</color>
      <size>終齡幼蟲體長約為 35~45mm</size>
```

<characteristic>三胸節背方各於左右兩側有一金屬藍色突起，後胸背方有一黃色橫紋，腹節末端有一對黃色叉狀突起</characteristic>

<predator>小繭蜂、寄生蜂、寄生蠅、椿象、鳥類等捕食性天敵</predator>

<days_of_growth>幼蟲發育期約需 20~25 日</days_of_growth>

<defense>剛孵化之初齡幼蟲首先攝食卵殼，平時多停棲於寄主植物葉表，前胸具黃綠色半透明臭角，平時縮於前胸背板縫內，受到驚嚇或騷擾時，可外翻伸出，並分泌具臭味之黃綠色液體，用以驅敵</defense>

</larva>

<pupa>

<feature>蛹為帶蛹，絲帶通過後胸背方懸掛於固定物上</feature>

<color>蛹體呈淡綠色</color>

<size>蛹長約 30~35mm</size>

<characteristic>中胸背方有一狹長角狀突起，由突起有四條黃色縱脊延伸到尾端</characteristic>

<days_of_growth>蛹期約為 30~35 日不定</days_of_growth>

<defense>老熟幼蟲於寄主植物枝條或附近隱密之植物枝葉間尋找適當地點吐絲化蛹。</defense>

</pupa>

<adult>

<feature>雄蝶前翅端稍尖，整體接近長三角形。後翅接近三角形，外緣呈明顯波浪狀，不具尾狀突起，後翅內緣反捲，內面密生黃灰色細長鱗毛。雌蝶體型較大，外觀與雄蝶相近，前翅略寬，後翅內緣無明顯性徵</feature>

<color>此蝶前、後翅表底色為黑褐色，前後翅中橫線處有連接的淡藍色塊狀斑紋，後翅亞外緣各翅室有淡藍色弦月紋</color>

<size>本種為大型蝶種，展翅約為 4.5~5.5cm</size>

<predator>蜘蛛、蜥蜴、青蛙、蜻蜓、螳螂、鳥類等捕食性天敵</predator>

<defense>成蟲飛行迅速，捕捉不易</defense>

</adult>

</life_stage>

<update>1999/3/8</update>

</butterfly>