

魚類標本 METADATA 發展與分析報告 (一)

第二次會議 2001/8/17 17.30 於中研院動物所
中央研究院 後設資料工作組

一、國際生物多樣性 METADATA 之發展

1. 生物多樣性資訊的 Metadata 發展重點：
 - 互通性 (Interoperability)
 - 標準化 (Standardization)
2. 整合生物多樣性資訊的主要挑戰：(METADATA WATCH REPORT #2, 2000)
 - 太多類型的生物資訊，卻沒有統一的標準，甚或還未電子化。
 - ◆ 如：來自 museum, herbaria, botanical gardens, zoological gardens 等不同類的機構各自收藏大量的生物標本及相關的描述性資訊；除此，還有涉及地理分佈 (geographical distribution)，生物命名法 (biological nomenclature) 等其他資訊類型。
3. 資料交換的主要障礙：
 - 調查方法與設計的品質參差不齊
 - 資料來源的問題 (如：重覆，被修改 ...)
 - 著錄的標準詳簡與品質不一
4. 生物資訊基礎建設 / 生物多樣化資訊網路的 Metadata 議題
 - Clearing-House Mechanism (CHM) of the Convention on Biological Diversity 生物多樣性資訊交換機制
 - ◆ The term "clearing-house" originally referred to a financial establishment where checks and bills are exchanged among member banks so that only the net balances need to be settled in cash. Today, its meaning has been extended to include *any agency that brings together seekers and providers of goods, services or information, thus matching demand with supply.* (<http://www.biodiv.org/chm/>)
 - ◆ 主要功能在於倡導保育工作相關科學與技術上的合作，並俾便生物多樣性資訊的檢索與交換。
 - ◆ Metadata 作為 CHM 基本架構，讓相關資訊的載入或轉出皆遵循一定的格式
 - NBII (National Biological Information Infrastructure) 美國國家生物資訊基礎建設 <http://www.nbii.gov/>

- ◆ Metadata Standards: Biological Data Profile、Metadata Clearinghouse
- ◆ 扮演橋樑的角色，讓生物科學的社群與其他相關社群間（如：地理資訊社群）的資訊基礎建設得以連結與互通。
- NBN (National Biodiversity Network) 英國生物多樣性網路
<http://www.nbn.org.uk/>
 - ◆ Data Standard Project: 包括發展與確認 Metadata Standard、Data Models、Data Exchange、Standard Dictionaries、Standard Access terms and conditions 等內容。
- IABIN (Inter-American Biodiversity Information Network) 美洲間生物多樣性資訊網路 <http://www.iabin-us.org/>
 - ◆ 跨國跨政府的活動計畫，旨在倡導泛美洲地區的國家，能更有效地收集、共享與使用生物多樣性相關資訊。
 - ◆ Survey of New World Holdings in Museum and Private Collections in Europe (1999)
 - ◆ Harmonizing Metadata Initiatives Throughout IABIN (1999)
- GBIF (The Global Biodiversity Information Facility) 全球生物多樣性資訊中心 <http://www.gbif.org/>
 - ◆ Memorandum of Understanding for the Global Biodiversity Information Facility : (Dec. 2000, 哥本哈根)
Node – A stable computing gateway that allows real-time inter-operational search of multiple institutional, national, regional and/or subregional databases containing primary or meta-level biodiversity data (such as specimen records, catalogues, bibliographic, sequence, protein and ecosystem data) ...**A node must provide descriptions of an accepted standard of metadata of the contents and quality of each database.** <http://www.gbif.org/statfram.htm>
- CODATA/TDWG Initiative (生物典藏資料檢索工作會議 , June 2001)
 - 旨在為生物典藏資料發展一套共同的資料規格書 (Metadata) , 與共同的軟體架構以支援跨資料庫的分散式檢索。
 - 參與會議的代表包括來自：Australian Virtual Herbarium(AVH), Bio-resource network in Japan(Brnet), European Natural History Specimen Information Network(ENHSIN), Red Mundial de Informacion sobre Biodiversidad(REMIB), The Species Analyst(TSA)
 - 第一次會議於 2001 年 6 月召開，已達數項共識包括：資料規格書將提供以 XML 為基礎的方法、參酌 BoiCISE information model 與 TDWG 為主的 HISPID 為重要依據、以及軟體架構將以 SOAP (Simple Object Access Protocol) 為基礎，發展生物典藏資料的檢索機制。第二次會議預計於 2001 年 11 月於雪梨召開。

二、魚類標本相關 Metadata 標準 (by approaches)

1. 資訊交換標準/資訊模式

- ASC (Association of Systematics Collections) Reference Model for Biological Collections
<http://palimpsest.stanford.edu/lex/datamodl.html>
http://gizmo.lbl.gov/DM_TOOLS/OPM/BCSL/LIB/ASC.html
- CDEFD Information Model for Biological Collections
<http://userpage.zedat.fu-berlin.de/~wgb/CDEFD/appendix.htm>
- DELTA (Description Language for Taxonomy)
<http://biodiversity.uno.edu/delta/>
- The NBN data model and its implementation in Recorder 2000. Environmental Information Management
http://www.bgbm.fu-berlin.de/biodivinf/docs/archive/Copp_C_2000_-_NBN_Data_Model.pdf
- SDTS Spatial Data Transfer Standard
- ADL (2000): ADL Gazetteer Content Standard. Alexandra Digital Library
http://alexandria.sdc.ucsb.edu/gazetteer//gaz_content_standard.html
- XDF Language for the Definition and Exchange of Biological Data Sets (Taxonomic Database Working Group - TDWG)
- Directory Interchange Format (DIF)

2. Metadata 格式

- Dublin Core Metadata Element Set
<http://dublincore.org/>
http://www.cimi.org/old_site/documents/meta_bestprac_v1_1_210400.pdf
- FGDC Biological Metadata Profile
http://www.fgdc.gov/standards/status/sub5_2.html

目前由美國國家生物資訊基礎建設 (NBII) 指定使用, 係根據 FGDC 標準再延伸關於生物資訊的分類與命名等屬性資訊之 Metadata 格式。

3. 分類資訊標準

- TDWG Taxonomic Database Working Group
- TLR Type and Lectotypification Registers (Taxonomic Database Working Group - TDWG)
- Species 2000 Standard Data
<http://www.sp2000.org/Standarddata.html>

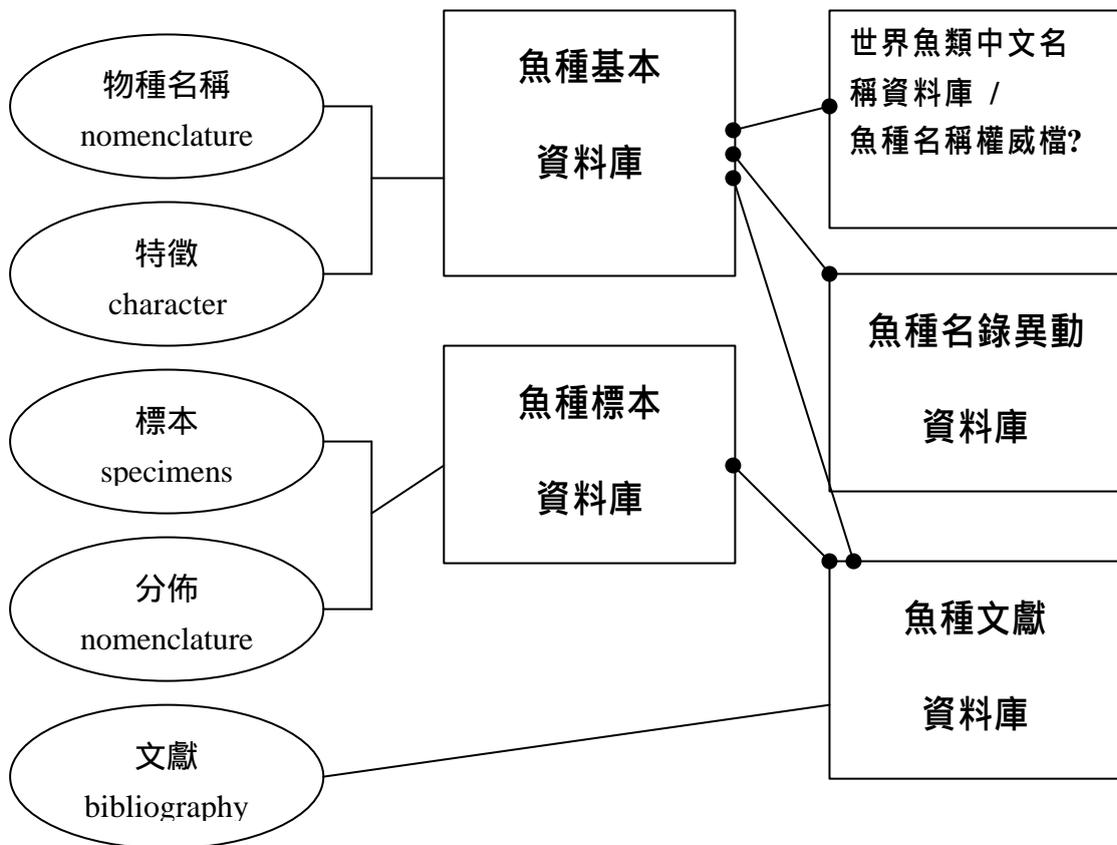
旨在建立全世界已知物種 (species) 的索引, 以作為國際生物多樣性的相關目錄與確認。目前已規範的標準格式包括:

- 統一名稱 (Accepted Scientific Name)【必備】
 - 同義名稱 (Synonyms)【必備】
 - 俗名 (Common names)【非必備】
 - 最新觀察者 (Latest taxonomic scrutiny)【非必備】
 - 資料庫來源 (Source database)【必備】
 - 備註 (Optional / comment field)【非必備】
 - 科名 (Family)【非必備】
 - Index to Organism Names (Maintained by BIOSIS)
 - Integrated Taxonomic Information System (ITIS)
4. 命名資訊標準
- ITRF International Earth Rotation Service Terrestrial Reference Frame
 - MFN Minimal Functional Nomenclator, also known as: DSTI Database Standards for Taxonomic Information (Taxonomic Database Working Group - TDWG)

三、「臺灣動物相典藏之研究：魚類計畫」METADATA 屬性分析

1. 魚類資料 Metadata 欄位清單及結構 (詳附件一 初稿)
- 欄位清單 (Element Set) / 國際標準比對 (Crosswalk for Metadata Sets)
 - Metadata 類別
 - 名稱特性：物種基本資料檔、物種名稱權威檔、物種文獻資料檔、名錄異動資料檔等。
 - 標本特性：標本資料檔。
 - Specimen Management (典藏管理資訊, 如: 典藏單位, 識別碼, 影像檔檔名 ...)
 - Taxonomic Identification (分類識別資訊, 如: 學名, 鑑定者/日期, 選模式標本作者/來源/日期/文獻)
 - Specimen Source (採集來源資訊, 如: 主要採集者, 採集方法, 採集號/日, 計畫/經費名稱 ...)
 - Provenance of Specimen (標本出處資訊, 如: 地點, 經/緯度, 國家公園 ...)
 - Ecological Information (生態環境資訊, 如: 棲所, ...)
 - Fish Description (魚類描述資訊, 如: 性別, 發育 ...)
 - Record Management (記錄管理資訊, 如: 本記錄建檔日期/建檔者/更新日期 ...)

■ 關係結構



四、預計下次的工作項目與成果