MAAT

後設資料發展趨勢與數位典藏的應用

陳亞寧
數位典藏國家型科技計畫 後設資料工作組
計畫主持人

2004年12月20日

「後設資料在數位典藏之研究發展：回顧與前瞻」研討會　2004

MAAT

大綱

- 全球發展現況
  - 文獻分析
  - 其他
- 全球發展趨勢之分析
- 數位典藏國家型科技計畫(National Digital Archives Program in Taiwan, NDAP)後設資料應用現況及其分析
- 研究發現
- 結語與建議

全球發展現況－文獻分析[1]

- Dempsey, L. & Heery, R. (1998). Metadata: a current view of practice and issues. Journal of Documentation, 54(2), 145-172.
- 陳雪華. (民86). 網路資源組織與Metadata之發展. 圖書館學刊, 12, 頁19-37.
  http://www.lis.ntu.edu.tw/main/journal/o12-2.pdf
- Gilliland-Swetland, A. (2000). Introduction to metadata: Setting the stage.
  http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf/swetland.pdf
- Turner, J.M., Moal, V., & Desnoyers, J. (2004). MetaMap: Acronym index (Ver. 1.2).
  http://mapageweb.umontreal.ca/turner/meta/english/index acroeng.html

全球發展現況－文獻分析[2]

- **Dempsey and Heery (1998) 依據下列要點進行Metadata分析**
  - 使用環境(environment of use)
  - 功能(function)
  - 產生(creation)
  - 結構(designation)
  - 相關的查詢協定(associated search protocols)
  - 況態(status)

	Band 1	Band 2	Band 3
使用環境	Global Internet search services	Selective Internet search services; directory service	Descriptions of scholarly collections; other importance repositories
功能	Location	Discovery, location, selection	Location, selection, evaluation, analysis, documentation

	Band 1	Band 2	Band 3
產生	Robot generated	Robot plus manual input	Intellectual expertise required, often involving dedicated 'information' staff
結構	Unstructured	Attribute value pairs; limited structure	Subfields; qualifiers, structured mark up

全球發展現況－文獻分析[5]

	Band 1	Band 2	Band 3
相關查詢協定	http with CGI form interface	http with CGI form interface; directory service protocols (WHOIS++, LDAP) with query routing (CIP)	Z39.50 SGML browsers and querying
狀態	Proprietary	Emerging Internet Standards	Domain specific standardisation

全球發展現況－文獻分析[6]

- 陳雪華教授(民86)將Metadata區分為
  - 早已普遍使用－MARC, PICA
  - 描述科技文獻－BibTex, EELS, EEVL, RFC 1807
  - 描述人文及社會科學資源－ ICPSR SGML Codebook Initiative, TEI Headers
  - 描述政府資訊－ GILS
  - 描述地理空間性資源－ CSDGM/FGDC
  - 描述博物館藏品與檔案特藏－ CDWA, CIMI, EAD
  - 描述大量網路資源－ Dublin Core(DC), IAFA/WHOIS++ Templates, LDIF, SOIF, URCs
  - 其他－ Warwick Framework

- **Gilliland-Swetland (2000)將Metadata區分為**
  - 管理**(administrative)**
  - 描述**(descriptive)**
  - 長期保存**(preservation)**
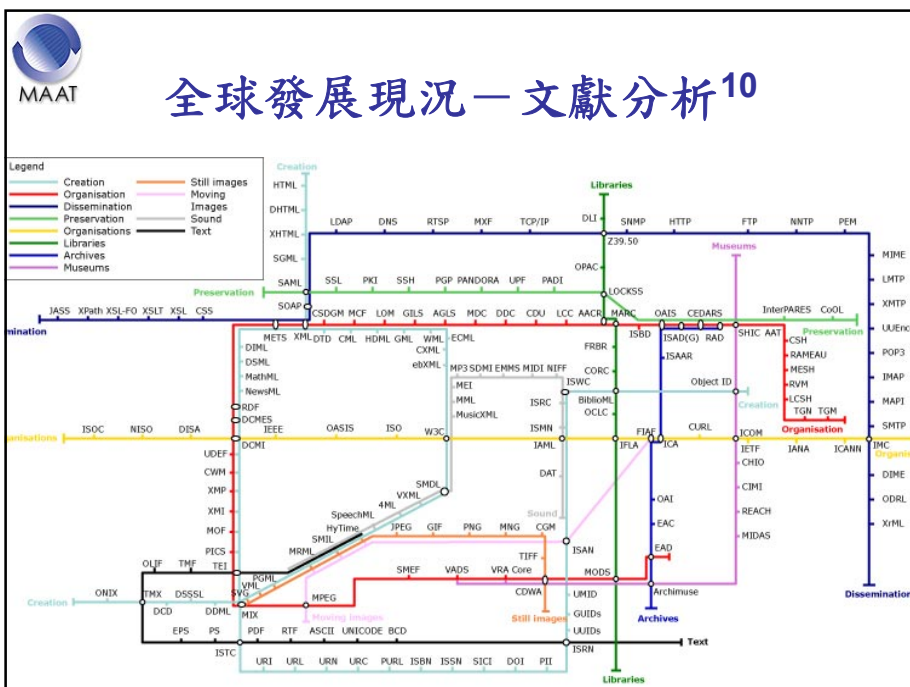 - 技術**(technical)**
  - 應用**(use)**

- **Gilliland-Swetland (2000)將Metadata屬性/特質區分為**
  - **Metadata來源－internal (如file names and header information) vs. external (如cataloging recs, legal information)**
  - 產生方式**－automatic vs. manual**
  - 性質**－creator vs. expert**
  - 狀態**－static vs. dynamic, long-term vs. short-term**
  - 結構**－structured vs. unstructured**
  - 語意**－controlled vs. uncontrolled**
  - 層次**－collection vs. item**

- **MetaMap將Metadata區分為**
  - 功能－產生**(creation)**, 組織**(organisation)**, 散播 **(dissemination)**與長期保存**(preservation)**
  - 社群－圖書館, 博物館, 檔案館等
  - 檔案類型－圖片 **(still images)**, 動畫**(moving images)**, 聲音**(sound)**, 文字**(text)**
  - 機構組織－**ISOC, NISO, DISA, DCMI, IEEE, OASIS, ISO, W3C, IAML, IFLA, FIAF, ICA, CURL, ICOM, IANA, ICANN, IMC**
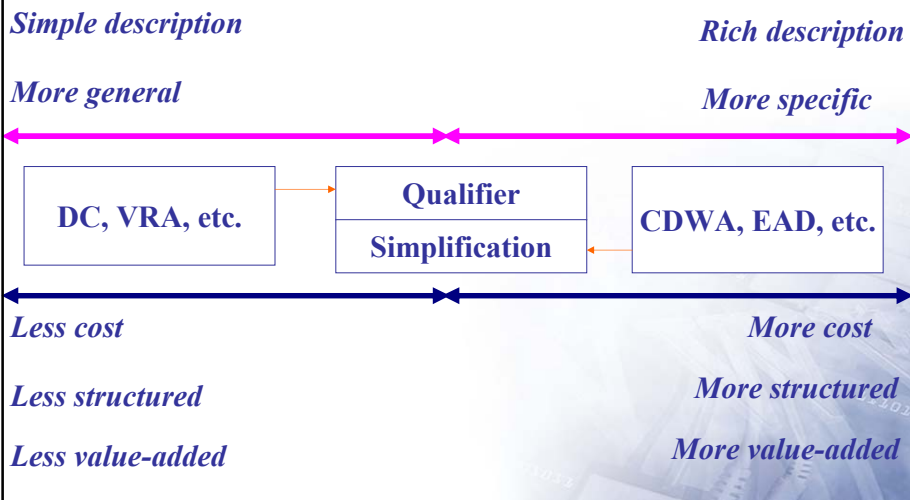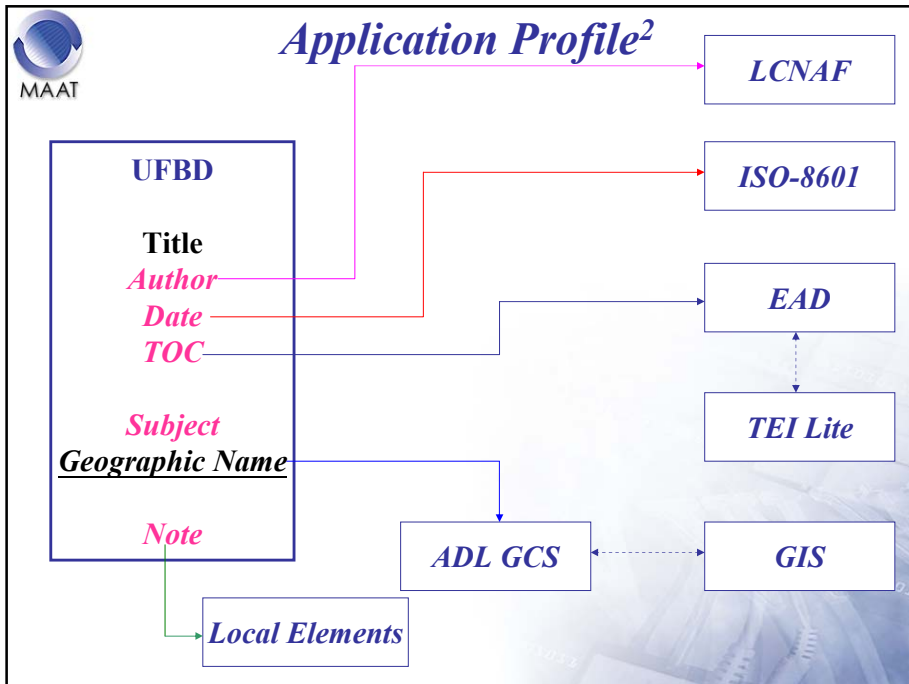
Legend
Creation
Organisation
Dissemination
Preservation
Organisations
Libraries
Archives
Museums
Still images
Moving Images
Sound
Text

全球發展現況－其他

- **Interpretive Rules, Vocabularies and RDF**
- **Preservation & Rights Management**
- **eLearning/LOM, E-Government**
- **CLD Metadata**
- **FRBR**
- **XML DTD and Schema**
- **OAI-PMH**
- **MDRs**
- **Application Profile**

Application Profile[1]

Simple description   *Rich description*

More general   *More specific*

| DC, VRA, etc. | → | Qualifier / Simplification | ← | CDWA, EAD, etc. |

Less cost   *More cost*

Less structured   *More structured*

Less value-added   *More value-added*

Application Profile[2]

LCNAF

ISO-8601

EAD

TEI Lite

GIS

UFBD

Title
Author
Date
TOC

Subject
Geographic Name

Note

ADL GCS

Local Elements

全球發展趨勢之分析

- **Functionality** — **Resources Discovery, Description, Administrative, Preservation, & Rights**
- **Domain** — **eLearning/LOM, E-Government, & Science**
- **Community** — **Library, Museum, Archives, & Internet Community**
- **Level** — **Collection/CLD, & Full-text/TEI Lite**
- **Application** — **Application Profile**
- **Conceptual Model** — **FRBR, CIDOC CRM, Records Continuum Model, OAIS Reference Model, RDF Model**
- **Semantics** — **Rules, Vocabularies and RDF Schema**
- **Interchange** — **XML DTD and Schema**
- **Protocol** — **OAI-PMH, SRW/SRU**
- **Framework** — **Warwick Framework, RDF**
- **Registration** — **MDRs**

NDAP後設資料應用現況及其分析[1]

- **CDWA AAT**
- **CIDOC Ethno Core Data Standards**
- **VRA Core Categories**
- **EAD**
- **MARC21**
- **FRBR-based Metadata (ECHO Project)**
- **HISPID**
- **SPECIES 2000**
- **Darwin Core**
- **CSDGM**
- **LOM**
- **ECML, and XrML**
- **OLAC**
- **DC**

NDAP後設資料應用現況及其分析[2]

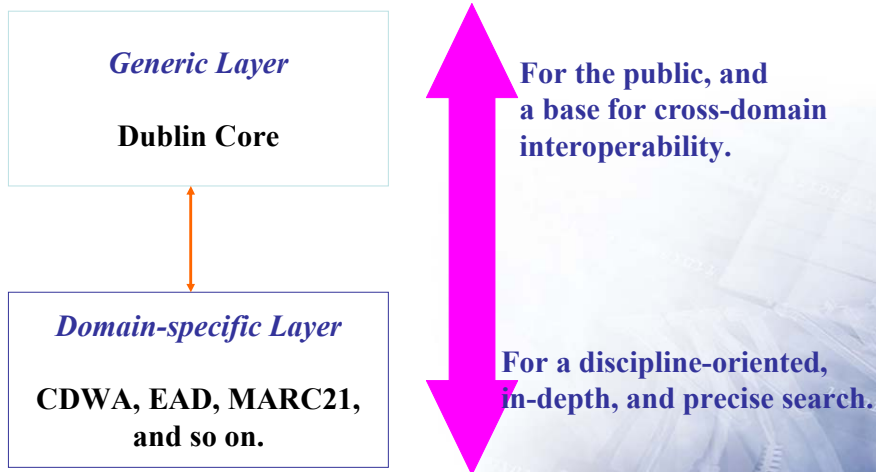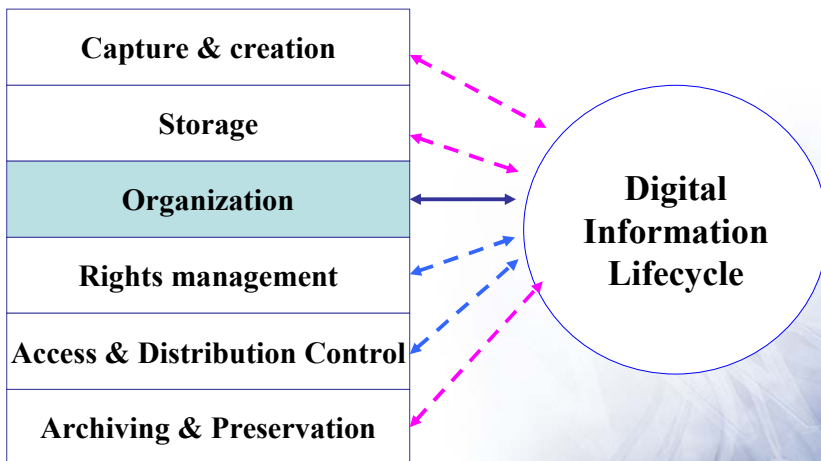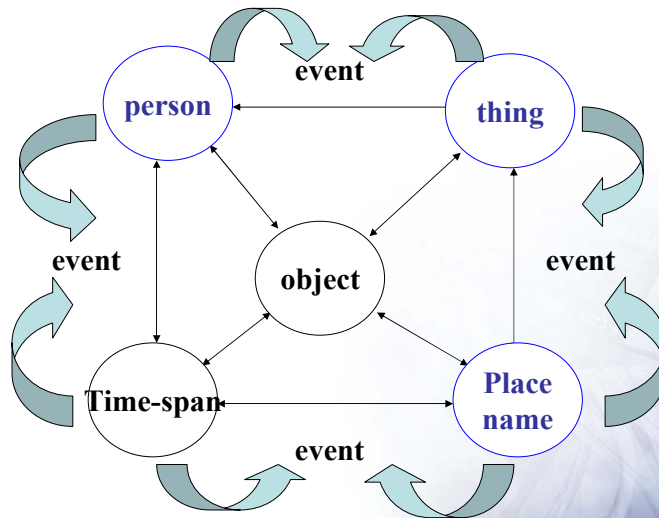| *Community Layer*	Library, Archives, Museum, Herbarium
Data Type Layer	Rubbing, Rare Book, Photo, Audio, Film Specimen, Archive, Field Report, etc.
Discipline Layer	Art, Archaeology, History, Ethnology Linguistics, Biodiversity, Genomics, etc.
Function Layer	Resource Discovery, Description, Exhibition, Management, E-Commerce

Generic Layer

Dublin Core

Domain-specific Layer

CDWA, EAD, MARC21, and so on.

For the public, and a base for cross-domain interoperability.

For a discipline-oriented, in-depth, and precise search.

Capture & creation

Storage

Organization

Rights management

Access & Distribution Control

Archiving & Preservation

Digital Information Lifecycle

研究發現

- 多元化－資料、社群、學科與功能
  - 資料
  - 社群－標本館
  - 學科－生物多樣性
  - 功能－電子商務
- 整合化－多種標準的整合應用
- 應用領域－電子政府

結語與建議

- 應用領域(domain) － eLearning, & E-Commerce
- 類型(attribute) －時間與事件
- 功能(function) － 長久保存(long-term preservation and archiving)
- 產生方式(creation) －從人工轉換至半自動或全自動
- 描述深度(level) －Collection －Item －Full-text
- 語意(semantics) －描述規則與索引典
- 資訊交換與共享(interchange)－XML Schema
- 多種標準的整合應用

縮寫字對照表[1]

- AAT: Art & Architecture Thesaurus
- CDWA: Categories for the Description of Works of Art
- CIMI: Computer Interchange of Museum Information
- CLD: Collection Level Description
- CSDGM: Content Standards for Digital Geospatial Metadata
- CURL: Consortium of University Research Libraries
- DCMI: Dublin Core Metadata Initiative
- DISA: Data Interchange Standards Association
- EAD: Encoded Archival Description
- ECML: Electronic Commerce Modeling Language
- EELS: Engineering Electronic Library

縮寫字對照表[2]

- EEVL: Edinburgh Engineering Virtual Library
- FGDC: Federal Geographic Data Committee FIAF: International Federation of Film Archives
- FRBR: Functional Requirements for Bibliographic Records
- GCS: Gazetteer Content Standard
- GILS: Government Information Locator Service
- HISPID: Herbarium Information Standards and Protocols
- IAFA: Internet Anonymous FTP Archives
- IAML: International Association of Music Libraries
- IANA: Internet Assigned Number Authority
- ICA: International Council of Archives
- ICANN: Internet Corporation for Assigned Names and Numbers
- ICOM: International Council of Museums

縮寫字對照表[3]

- ICPSR: Inter-university Consortium for Political and Social Research
- IFLA: International Federation of Library Associations
- IMC: Internet Mail Consortium
- ISOC: Internet Society
- LDAP: Lightweight Directory Access Protocol
- LDIF: LDAP Data Interchange Format
- LOM: Learning Object Metadata
- MARC: Machine-Readable Catalogue
- MDRs: Metadata Registries
- OAI-PMH: Open Archives Initiative-Protocol for Metadata Harvesting
- OAIS: Open Archival Information System
- OASIS: Organization for the Advancement of Structured Information Standards

縮寫字對照表[4]

- OLAC: Open Language Archives Community
- RDF: Resource Description Framework
- SGML: Standard Generalized Markup Language
- SOIF: Summary Object Interchange Format
- TEI: Text Encoding Initiative
- UFBD: USMARC for Bibliographic Data
- URC: Uniform Resource Characteristics
- VRA: Visual Resources Association
- XrML: Extensible Rights Markup Language
- XML: eXtensible Markup Language

報告完畢，敬請賜教與指正！