Methodology Supporting Metadata for Cultural Heritage Collections
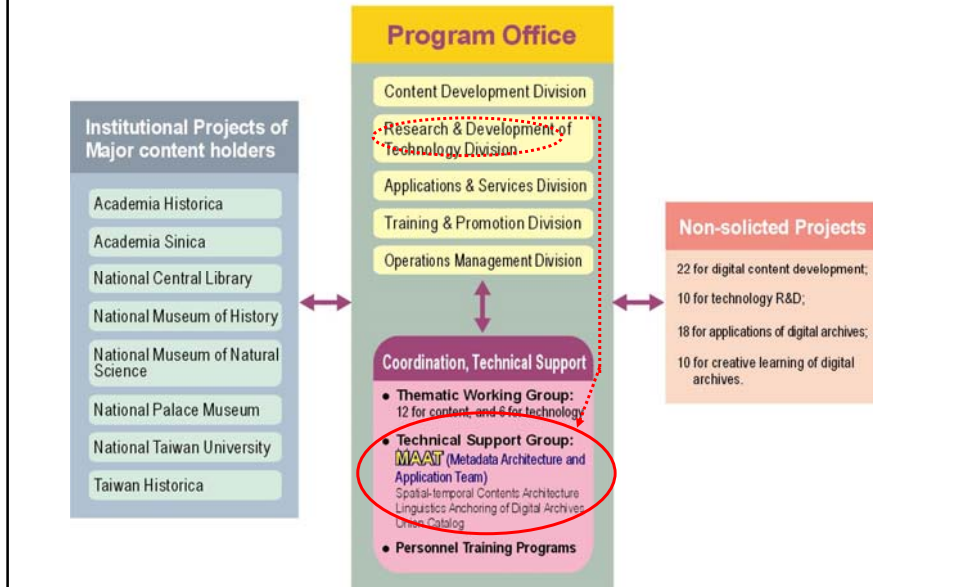## A Taiwan Perspective

Shu-Jiun (Sophy) Chen
MAAT, National Digital Archives Program, Taiwan

MCN 2005, Boston

Outline

I. Overview of the NDAP and MAAT

II. The Metadata Challenges

III. The Metadata Strategies

IV. The Methodologies

V. The Outcomes

VI. Conclusion Remarks

MAAT and the NDAP：Organization Chart

Program Office

- Content Development Division
- Research & Development of Technology Division
- Applications & Services Division
- Training & Promotion Division
- Operations Management Division

Institutional Projects of Major content holders

- Academia Historica
- Academia Sinica
- National Central Library
- National Museum of History
- National Museum of Natural Science
- National Palace Museum
- National Taiwan University
- Taiwan Historica

Coordination, Technical Support

- **Thematic Working Group:**
  12 for content, and 6 for technology
- **Technical Support Group:**
  MAAT (Metadata Architecture and Application Team)
  Spatial-temporal Contents Architecture
  Linguistics Anchoring of Digital Archives
  Union Catalog
- **Personnel Training Programs**

Non-solicted Projects

22 for digital content development;

10 for technology R&D;

18 for applications of digital archives;

10 for creative learning of digital archives.

MAAT Matadata Architecture and Application Team

Metadata Architecture and Application Team

News | Best Practice | Introduction | Projects | Getting Start | Management Systems | Documents | Events | Resource Sharing |

Projects	Attributes			
	Data Types	Subjects	Communities	Metadata Standards
The Digital Archives of Rubbings and Archaic Texts	Rubbings	History	Museum	CDWA
Digital archives for the Grand Secretariat Archives	Archives	Politics/History	Archives	EAD
Digital Archives of Chinese Archaeological Data	Artifacts	Archaeology	Museum	CDWA, MIDAS
Digital Archive for Rare Books of Fu Ssu-Nien Library	Rare Books	History	Archives	TEI
Digital Library Project for Official Economic Archives and Diplomatic Archives	Archives	Diplomacy	Archives	EAD
Zoological Research of Taiwan: Fish and Mollusks	Specimen	Herbarium	Herbarium	SPECIES 2000 Standard Data
Digital Library of Taiwan Herbarium	Specimen	Herbarium	Herbarium	HISPID
Linguistics Anchoring and Language Archive of Digital Archives	Corpus	Language	N/A	OLAC
GIS Archive for Modern Chinese Historic Atlas and Aerial Photos	Map, Photography	GIS	N/A	FGDC
Taiwan Aborigine: Ping-pu	Video, Costume	Race/Aborigine	Museum	DC
Digital Archives for Ethnological Specimen, Photo and Archives	Report	Race/Aborigine	Museum	CDWA

News | Best Practice | Introduction | Projects | Getting Start | Management Systems | Documents | Events | Resource Sharing |

Project	Type	Discipline	Community	Metadata
Digital Archives for Ethnological Specimen, Photo and Archives	Report	Race/Aborigine	Museum	CDWA
Knowledge Base of Taiwan's Earthquake	Report, Photo	Earthquake	N/A	DC
Digital Archives Project of the Office of Governor-General in Taiwan	Archives	Politics/History	Archives	EAD
Digital Archives Project of Chinese Antiquities at the National Palace Museum	Artifacts	Arts	Museum	CDWA
Digital Archives Project of Chinese Painting and Calligraphy at the National Palace Museum	Painting and Calligraphy	Arts	Museum	CDWA
Digital Library Project of National Museum of History	Artifacts/Painting	Arts	Museum	CDWA
Digital Video Library Project	Video Tape	Multimedia	Museum	IFLA FRBR /ECHO
Taiwan Momery: Digital Photo Museum	Photography	Multimedia	Museum	DC/CDWA
National Digital Archives Program (2002-2006): Academia Historica (Taiwan.	Archives/Photo	Politics/History	Archives	EAD

Layers of Knowledge Management for Metadata in the NDAP

Data Type Layer	Rubbings, Files, Rare Books, Full Text, Specimens, Corpus, Maps, Photos, Painting and Calligraphy, Audio-Visual, Score, Dress, Reports, Investigation, etc.
Discipline Layer	Arts History, Politics, Humanities, Archeology, Ethnology, Diplomacy, Zoology, Botany, Genomics, Linguistics, Geology, Geography, etc.
Functionality Layer	Resource Discovery, Description, Exhibition, Preservation, Rights Management for e-Commerce, e-Learning, etc.
Community Layer	Museum, Archives, Library, Herbarium, etc.

The Metadata Challenges

- The NDAP faces issues of
  - A lot of collection projects
  - A diversity on communities, disciplines, data types, and functions simultaneously
  - The limitation of metadata practitioners
- Can a systematic approach for various projects to achieve:
  - Cost-effectiveness ?
  - Quality assurance ?
  - Consistency ?
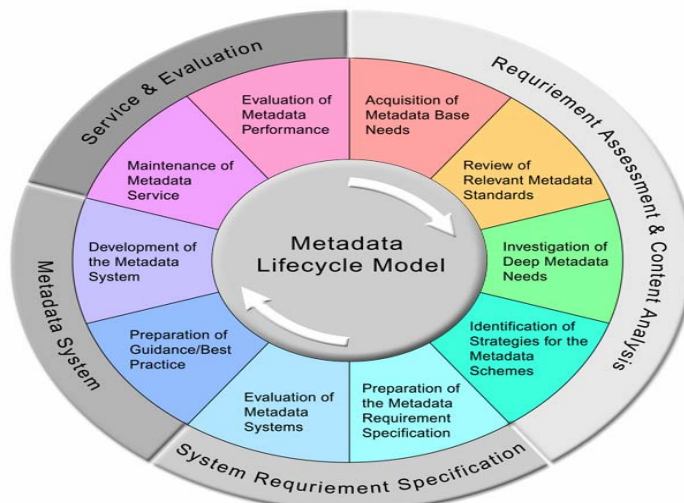 - Interoperability ?

The Metadata Challenges
from view point of the collection projects

- How to getting started?
- No idea of metadata standards
  - A-state-of-the-art metadata development
    - Museum Community
    - Archives Community
    - Biodiversity Community…etc.
- How to decide the metadata standards?
  - A universal standard?
  - Adopt one set or multiple sets of standards?

- To what extent of standards can help?
  - description, retrieval, exhibition, preservation…
- No idea of metadata system development
  - Content experts vs. system developers
- Domain specific vs. generic Metadata
  - For domain experts' research
  - Union catalogue For publics
  - Cross domains
- ……

The Strategies

- User-centered approaches
- Metadata Methodologies
  - The Metadata Lifecycle Model
- Generic & Domain-specific Parallel Strategy
- Metadata Application profiles
- Domain-specific Common Core
- Translation of Metadata Standards
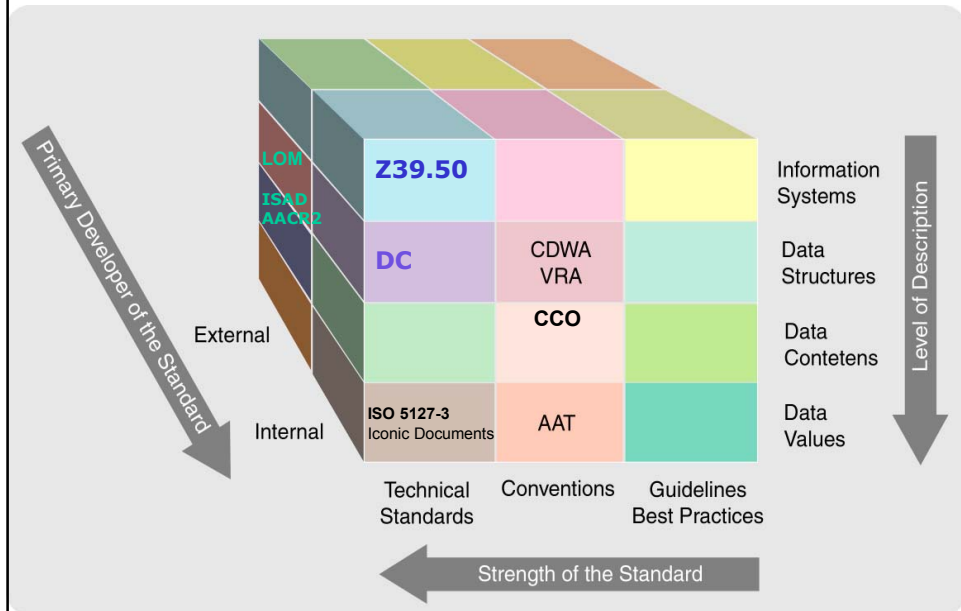
The Methodology
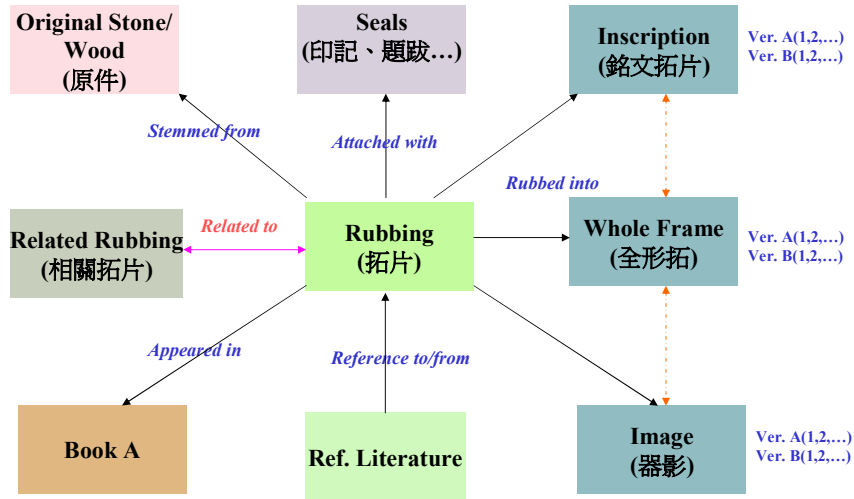# The Metadata Lifecycle Model

Chen, Y.-N., Chen, S.-J., & S.C. Lin. (2003). A metadata lifecycle model for digital libraries: Methodology and application. Paper presented at *The World Library and Information Congress: 69th IFLA General Conference and Council*, 1-9 Aug. 2003

The Concept of Metadata Lifecycle Model (MLM)

- A continuous event served as a lifecycle to trigger responsive activities for various requirements when new needs come up.
- A systematic workflow to conduct content analysis and system analysis.
- A service for content experts and system designers.

A close look at the Metadata Lifecycle Model

- **Step 1**
  Acquisition of Metadata Base Needs
  - Acquire preliminary information and contacts
    - Contact window, scope, legacy system…
  - Establish a better understanding among the team work
    - Goal, schedule…
- **Step 2**
  Review of Relevant Metadata Standards and Projects
  - Identification of potential metadata standards
    - The four-layer for metadata selection
  - Current metadata trends and issues
    - (Example) Biodiversity - GBIF, TDWG, BioCASE, Species 2000
  - Examination of use cases and their experiences

Types of Standards

Types of Knowledge Organization systems

- **Term Lists**
  - **Authority files**
  - Glossaries
  - Dictionaries
  - Gazetteers
- **Classification and Categories**
  - Subject headings
  - Classification Schemes, Taxonomies, and Categorization Schemes
- **Relationship Lists**
  - **Thesauri**
  - Semantic Networks
  - Ontologies

The methodologies
selection of metadata standards
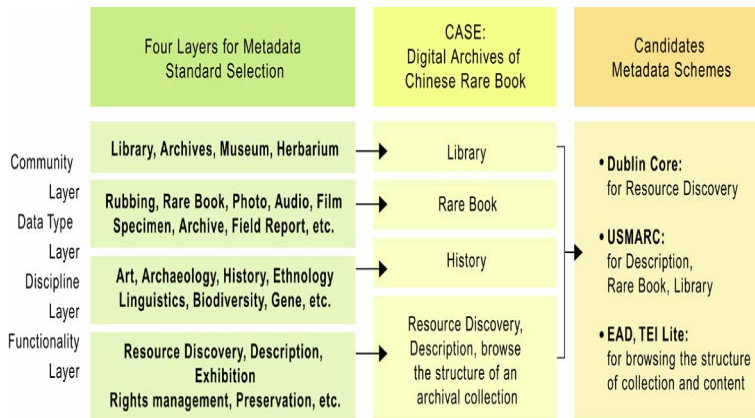
Chen, Y.-N., Chen, S.-J., & S.C. Lin. (2003). A metadata lifecycle model for digital libraries: Methodology and application. Paper presented at *The World Library and Information Congress: 69th IFLA General Conference and Council*, 1-9 Aug. 2003.

A close look at the Metadata Lifecycle Model

- **Step 3**
  Investigation of Deep Metadata Needs
  - Cover data structures, data contents, data values and system requirements
  - A concept of content analysis is deployed.
  - Granularity identification
  - Relationship identification
    - Internal
    - External

- **Step 4**
  Identification of Strategies for the Metadata Schemes and Interoperability Mechanism
  - One size can not fit for all.
  - The concept of metadata "application profile"
    - Adopting one or several existing metadata standards
    - (Example) Rare book : MARC21＋EAD＋TEI Lite

Association Diagram
An example of Chinese Rubbings

Original Stone/Wood (原件)

Seals (印記、題跋…)

Inscription (銘文拓片)
Ver. A(1,2,…)
Ver. B(1,2,…)

Stemmed from

Attached with

Rubbed into

Related Rubbing (相關拓片)

Related to

Rubbing (拓片)

Whole Frame (全形拓)
Ver. A(1,2,…)
Ver. B(1,2,…)

Appeared in

Reference to/from

Book A

Ref. Literature

Image (器影)
Ver. A(1,2,…)
Ver. B(1,2,…)

Model of the levels of arrangement of a fonds
A Case of the NDAP

Administrative fonds

Fonds

Sub-founds

Series

Series

Series

Sub-series

Sub-series

Sub-sub-series

Sub-sub-series

12 archival collection projects
4~8 levels

2 projects
1 project
2 projects
2 projects
1 project
2 projects
2 projects

Files

Files

Files

Files

Items

Items
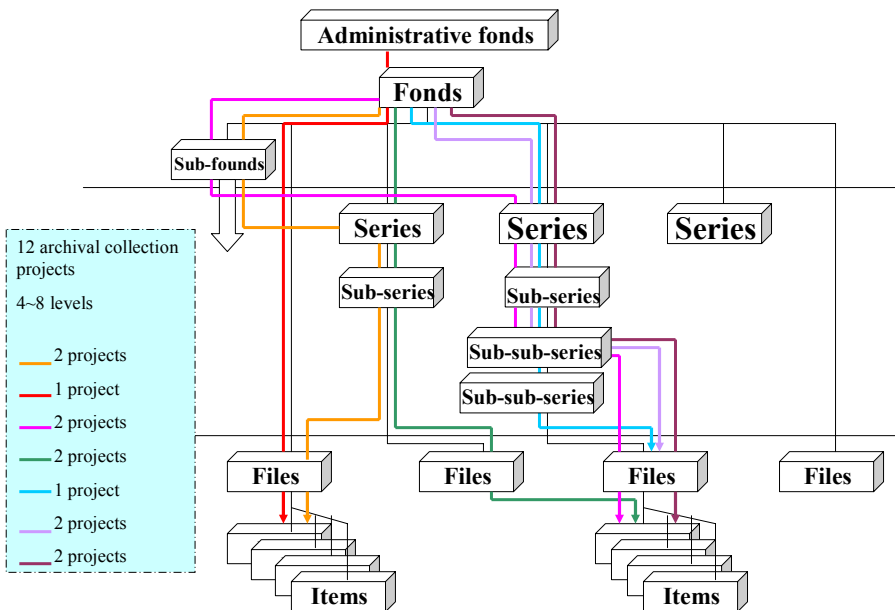
A FRBR-based Metadata format (CDWA) for the case of NPM
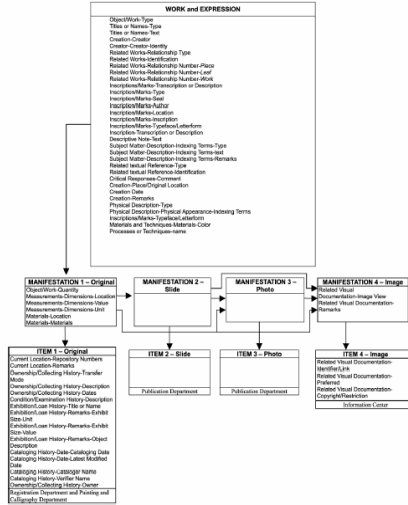
Figure 1.
A FRBR-based metadata
format (CDWA) for the
NPM

IFLA FRBR model

135

A re-examination of workflow for the digital library
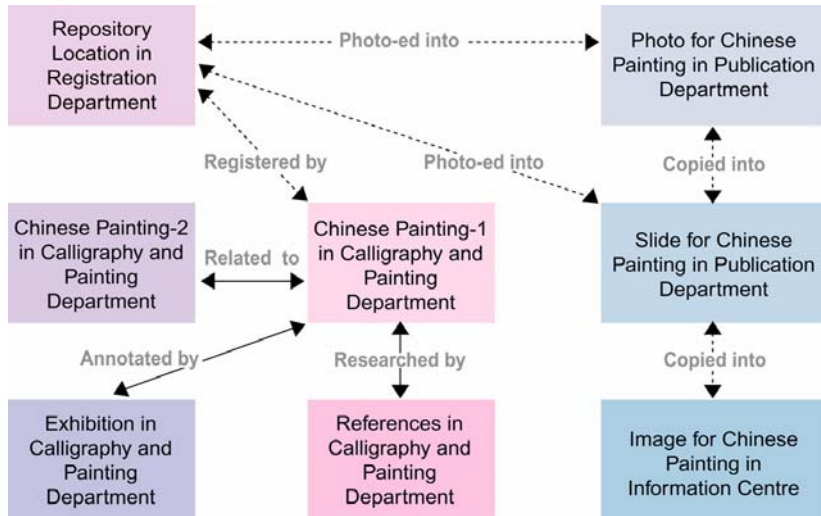
Granularity Issues
Object/Work: COMPONENTS

Source: National Palace Museum
Round miniature curio-box with bamboo-strip veneer and twined Indian-lotus décor
清 竹絲纏枝番蓮多寶格圓盒

A close look at the Metadata Lifecycle Model

- **Step 5**
  Preparation of the Metadata Requirement Specification
  - A common agreement / bridge among collection project participants, metadata specialists, and system designers.

  - The content includes:
    - Background information of the project
    - Objectives and scope of the metadata system
    - Statement of adopted metadata standards
    - Metadata elements and structure
    - Attributes of metadata elements
    - Input template and related standards crosswalking

- Functional requirements
- XML DTD

- **Step 6**
  Evaluation of Metadata Systems
  - EVALUATION
    - Assessment of potential metadata systems
    - As a reference for
      - Revision of metadata requirements
      - Outsourcing
      - System design, or
      - Collaboration of system development

The methodologies
functional requirements

- Documentation
- Access
- Display
- Administration

The English title of the element		Basic search	Advanced search	Brief Display	Display order	Detail Display	Pay to access
Element	Sub-element						
Object Type		Y	Y	Y		Y	
Accession Number			Y	Y	Y	Y	
Main Object		Y	Y	Y	Y	Y	
Other Name		Y		Y		Y	
Creator		Y	Y			Y	
Condition			Y			Y	
Grade						Y	
Quantity						Y	
Storage Department						Y	
Cataloging Date							

A close look at the Metadata Lifecycle Model

- **Step 7**
  Preparation of Best Practice Guidance
  - To ensure the quality control of the metadata records in the collection project
  - The content includes:
    - Element name
    - Definition
    - Description rules
    - System suggestion
    - Examples
    - Mapping to relevant standards

- **Step 8**
  Development of the Metadata System
  - Approaches include:
    - Metadata Requirement Specification (MRS)
      - Develop metadata systems according to the final version of MRS that are agreed between collection project team and metadata team
      - The case is usually happened in the situation that the collection project well understands/expresses their own requirements
    - Prototyping
      - Develop metadata systems according to both the draft of MRS and the direct interactive communication between collection project team and system development team
      - This case is usually happened in the situation that the collection project only have rough/unclear ideas about their own requirements

A close look at the Metadata Lifecycle Model

- **Step 9**
  Maintenance of Metadata Service
  – Guarantee the quality assurance of metadata

  - The service model

  ←——→ **Direct relationship**

  ←·—·→ **Indirect relationship**

  End User

  Metadata Team

  System Designer ←·—·→ **Content Provider**
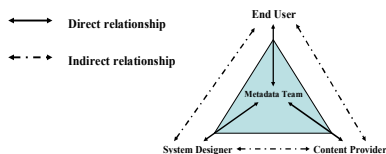
- **Step 10**
  Evaluation of Metadata Performance
  – ASSESSMENT
    - The whole metadata process and performance.
    - The content includes to evaluate:
      – Metadata record quality
      – Effectiveness of adopting a metadata scheme for retrieval
      – Use of metadata creation tools within the collection project
      – How well each stage of the metadata lifecycle model has been performed

---

The Outcomes
Metadata Applications of Museum-wide Community

- Data Structures
  – CDWA(16)，VRA(1)，IFRB-FRBR(1)，CIDOC(1)，MARC(1)
- CDWA Application Profile
  – 22~ 210 elements
  – Application of Main Categories：96% (26:27)
  – Application of Sub-categories：66%
  – Fully Adoption：6 elements

- Data Contents
  – Local rules
  – CCO
- Data Value
  – 2~58 elements use knowledge organization systems (controlled vocabularies)
  – Types of the knowledge organization systems
    - Lists of controlled vocabularies
    - Authority Control (Personal name, place name, time-span, subject, archaeological period

The Diagram of Chinese Antiquities Database & Applied CDWA Categories

『Reference Database』mapping to
【Related textual References】

『Inscription Database』mapping to
【Inscription/Marks】

Time Span Authority File

『Visual Documentation Database』
mapping to：
【Related Visual Documentations】

Place Name Authority File

『Artifacts Database』

[Accession Number]

『Exhibition Database』mapping to：【Exhibition/Loan History】

Personal Name Authority File

『Motif Database』mapping to：
【Subject Matter】

『』NPM ；【】CDWA; proceeding；*in progress*

Metadata Mappings

Metadata Mappings

Metadata Mappings

The Union Catalog of NDAP

Dublin Core-base
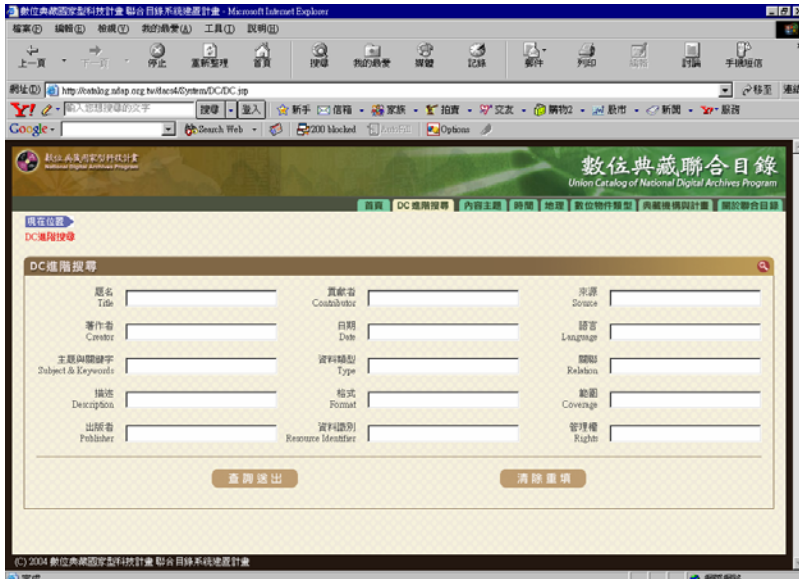
CDWA Lite
An Empirical Study

CDWA Lite
An Empirical Study

	CDWA Lite	NPM	Dublin Core	NDAP-Union
1.	Object/Work Type Wrapper	「Media」 Cloisonné, Glass	Type	✓
2.	Title Wrapper	「Object Title」 Cloisonné box with lotus-spray decoration, Ching-t'ai reign (1450-1456), Ming dynasty	Title	✓
3.	Display Creator	「Creator」	Creator	✓
4.	Indexing Creator Wrapper	「Creator」	Creator	✓
5.	Display Measurements	「Dimension and Weight」 6.3cm (height of body), 12.4cm (diameter of mouth), 634.6gram (weight)	Format	✓
6.	Indexing Measurements Wrapper	「Dimension and Weight」 6.3cm (height of body), 12.4cm (diameter of mouth), 634.6gram (weight)	Format	✓
7.	Display Materials/Techniques	「Formal Description」 Ware(胎質): copperware Glaze color(釉色): multi-color glaze (多彩釉)-Cloisonné (technique)	Description	
8.	Indexing Materials/technique Wrapper	「Formal Description」 Ware(胎質): copperware Glaze color(釉色): multi-color glaze (多彩釉)-Cloisonné (technique)	Description	
9.	Display State/Edition Wrapper		Description	
10.	Style Wrapper	「Kiln System」	Type/Description	
11.	Culture Wrapper	「Archaeological Period」	Type/Description	
12.	Display Creation Date	「Time」 Begin: 1450AD (Ching-t'ai Period) End: 1456AD	Date	✓
13.	Indexing Dates Wrapper	「Time」 Begin: 1450AD (Ching-t'ai Period) End: 1456AD	Date	✓

	CDWA Lite	NPM	Dublin Core	NDAP-Union
14.	Location/Repository Wrapper	「Identify Number」 「Storage Department」 「Storage Place」 ID：中趈000714N000000000 National Palace Museum Antiquities Department of the National Palace Museum	ID	✓
15.	Subject Indexing Wrapper	「Decoration」 Plant-Flower-Passionflower (Indian-lotus)	Subject	✓
16.	Class	「Function」 container	Type	✓
17.	Description/Descriptive Note	「Introduction」 Carved in low relief, this copper body box and its lid were cast in the shape of a lotus blossom. ～～Each petal of lotus blossom inside the box and the lid is decorated with one multi-colored passionflower (Indian-lotus). ～～This object was exhibited in the United States in the tour exhibition "Possessing the Past: Treasures from the National Palace Museum, Taipei" in 1996.	Description	✓
18.	Inscriptions Wrapper	「Inscription」 incision in a horizontal row, from right to left—semi-cursive script, carved with knife, a reign title, Chinese character, standard script	Description	
19.	Related Work Wrapper	「Catalog Level」	Relation	
20.	Rights for Work	「Copyright」 National Palace Museum	Right	✓
21.	Record Wrapper	「Cataloging History」	Relation	✓
22.	Resource Wrapper	「Image」	Relation	✓

Interoperability Strategy[1]

Generic

Dublin Core
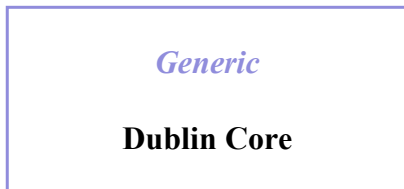
For the public, serve as a base for cross-domain interoperability

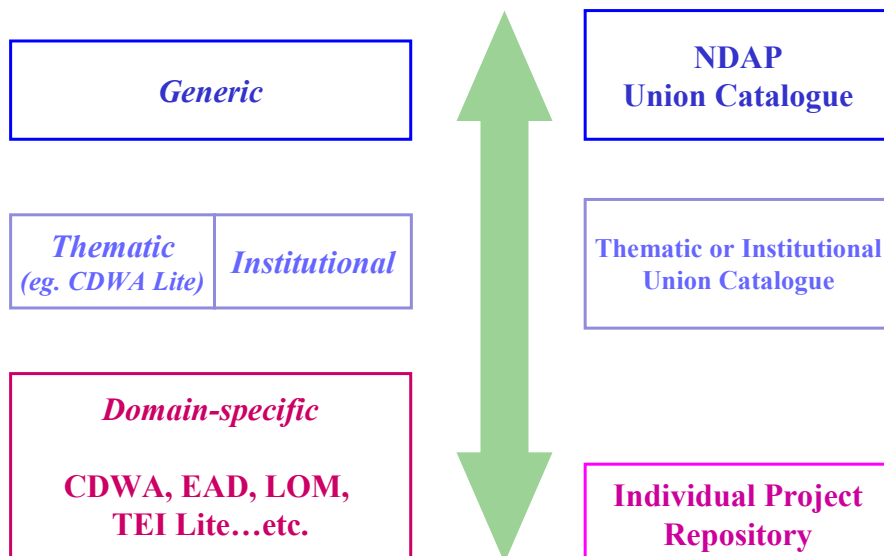*Domain-specific*

CDWA, EAD, LOM, TEI Lite…etc.

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

Chen, Y.-N., Chen, S.-J., Sum, H.-C., & S.C. Lin. (2003). Functional requirements of metadata system: From user needs perspective. Paper presented at *2003 Dublin Core: Supporting Communities of Discourse and Practice: Metadata Research and Application*, 28 Sep. – 2 Oct. 2003.
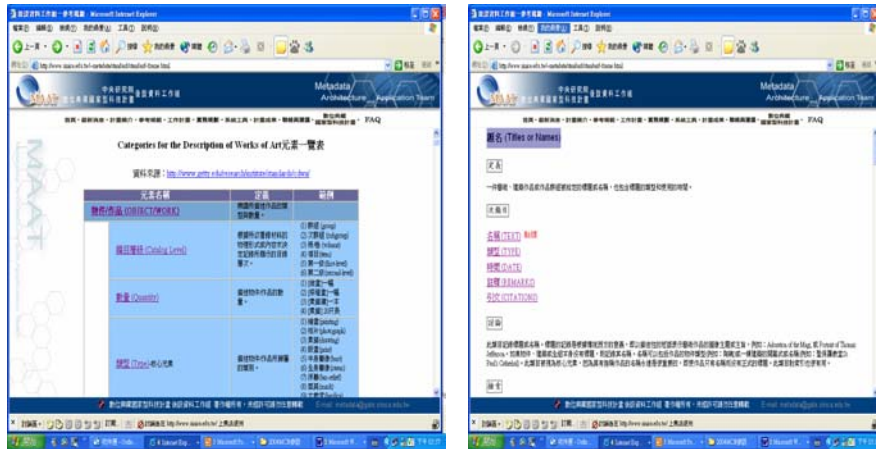
Interoperability Strategy[2]

Thematic Common Core	Institutional Common Core

For the specific community, serve as a base for cross-project interoperability

Domain-specific

CDWA, EAD, LOM, TEI Lite…etc.

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

Interoperability Strategy[3]

Generic

NDAP Union Catalogue

Thematic (eg. CDWA Lite)	Institutional

Thematic or Institutional Union Catalogue

Domain-specific

CDWA, EAD, LOM, TEI Lite…etc.

Individual Project Repository

Translation of Metadata Standards

Ongoing and Future Work

- Localization of Metadata Standards
  - CCO in Chinese
  - The Common Core vs. CDWA Lite
- Knowledge Organization System
  - Ontology (Conceptual Model)
    - CIDOC CRM (Conceptual Reference Model)
    - IFLA FRBR (Functional Requirement of Bibliographical Record)
- Preservation Metadata
  - OAIS Reference Model

References

- Chen, Y.-N. , Chen, S.-J.(2005) . Metadata Lifecycle Model and Metadata Interoperability. 5th International Conference on Conception of Library and Information Science, June 4-8, 2005, Glasgow, UK.
- Chen, S.J. & Cheng, C.J. (2005). The metadata development and application in the museum community. Paper presented at The Metadata Research and Development for Digital Libraries Conference, 20 Nov. 2004, Taipei.
- Chen, S.J., Chen, Y.N., Cheng, C.J. and Chiu,Y.C. (2004). Metadata for the National Digital Archives Program (NDAP) of Taiwan. Presented at the Getty Center, 15 Nov., 2004, Los Angeles, CA, US.
- Chen, Y.N.& Chen, S.J. (2004). A metadata practice of the IFLA FRBR model: A case study for the National Palace Museum. Journal of Documentation, 60(2), 128-143.
- Chen, Y.-N., Chen, S.-J., & S.C. Lin. (2003). A metadata lifecycle model for digital libraries: Methodology and application. Paper presented at The World Library and Information Congress: 69th IFLA General Conference and Council, 1-9 Aug. 2003.
- Chen, Y.-N., Chen, S.-J., Sum, H.-C., & S.C. Lin. (2003). Functional requirements of metadata system: From user needs perspective. Paper presented at 2003 Dublin Core: Supporting Communities of Discourse and Practice: Metadata Research and Application, 28 Sep. – 2 Oct. 2003.

Thank You

sophy@sinica.edu.tw

Metadata Architecture and Application Team
NATIONAL DIGITAL ARCHIVES PROGRAM
Taiwan