

Current Progress of SMART

Sinica Metadata Architecture and Research Task

Ya-ning Chen

Shu-jiun Chen

Computing Centre

Academia Sinica

6 September 2000

Outline

- **Metadata Workflow**
 - **Interview**
 - **Metadata Analysis Model**
 - **Common Metadata Structure**
 - **Criteria for selecting an appropriate metadata format to mapping**
 - **XML DTD for own and interchange with others**
- **Case Study**
- **Current Progress**

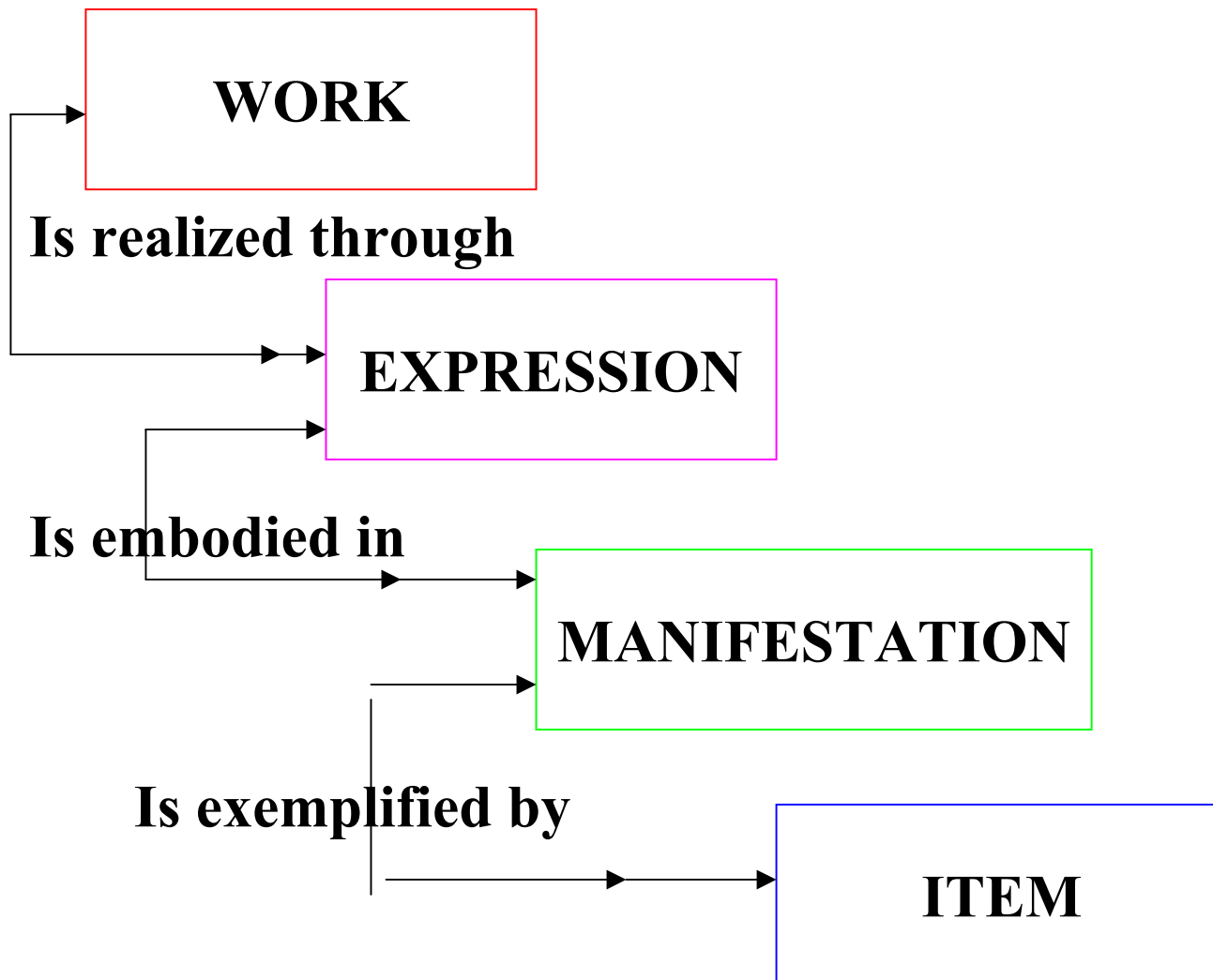
Metadata Workflow - 1

- **Acquisition of user needs and data requirements.**
- **Metadata Analysis**
 - **Structure and relationship – horizontal and vertical**
 - **Analysis and level of elements**
- **A common metadata structure**

Metadata Analysis Model

- **IFLA FRBR model**
- **INDECS model**
- **SMART model abridged from IFLA model**

IFLA FRBR Model

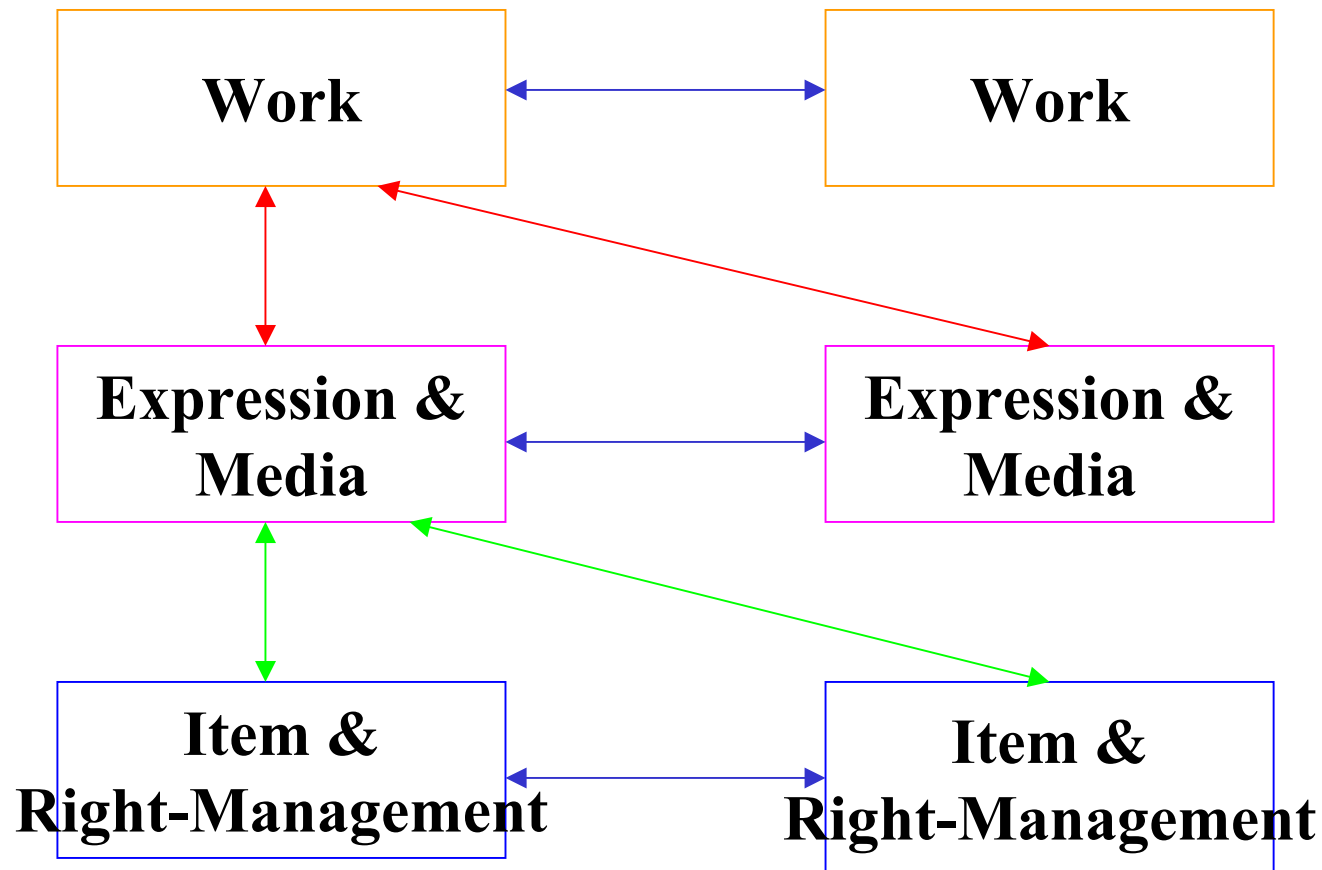


IFLA Case Study

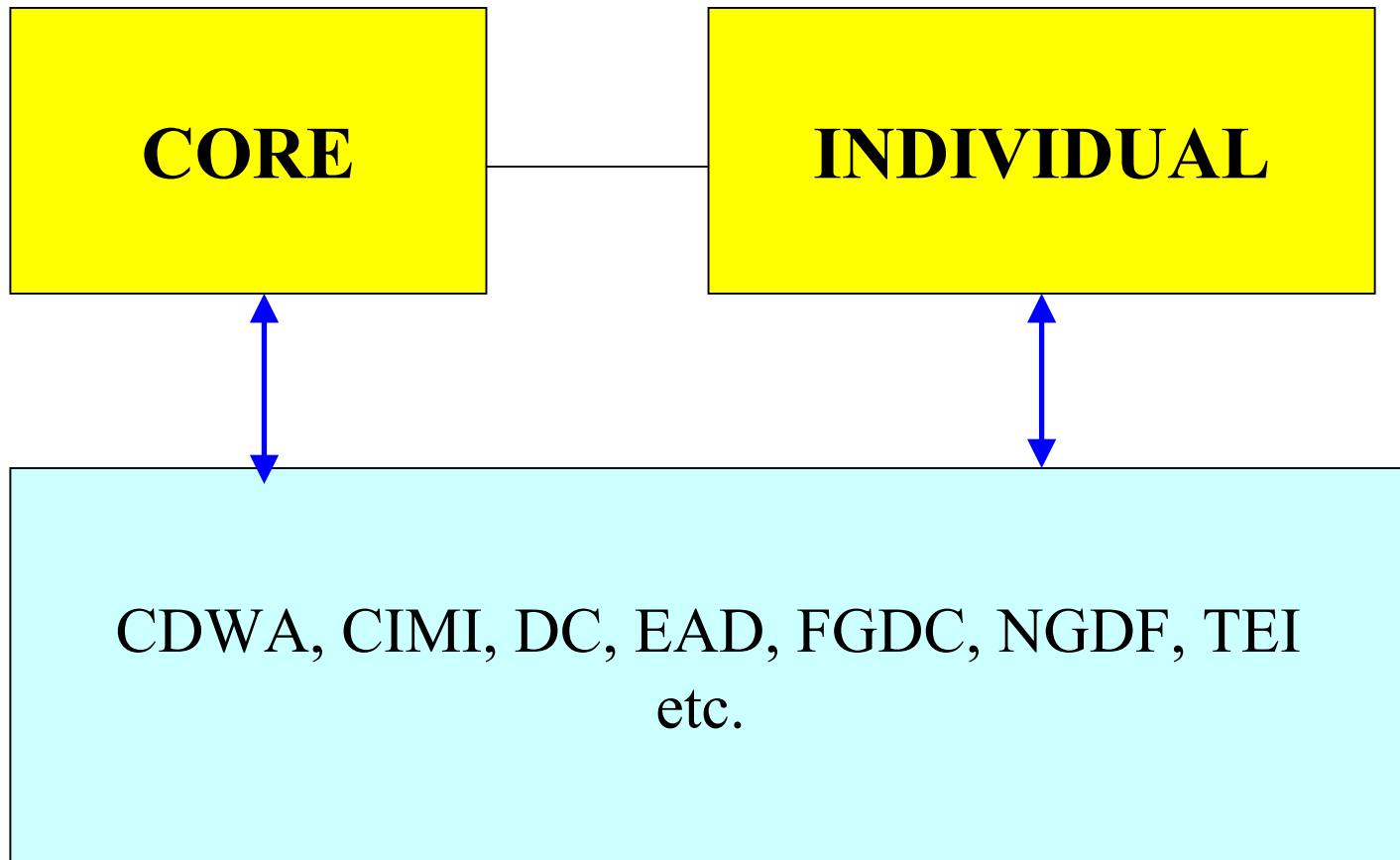
- **INDECS Model**
- **An experiment is taken by DC/DOI Communities.**
- **Application - Copyright management, and e-commerce.**

SMART Model

abridged from IFLA FRBR



A Common Metadata Structure



Metadata Workflow-2

- **Metadata Mapping**
 - **Mapping table among project's elements, standard metadata elements, examples, and functions.**
 - **Criteria for selecting an appropriate metadata format to mapping**
 - **XM DTD table - both for internal use and interchange**
 - **A document is composed of Standards documents application, comparison and suggestion**

Metadata Criteria - 1

- **Application environment**
 - **Characteristics of user group audience, subject domain**
 - **material genre**
 - **accordance and compatibility of existing international standards : focus on Internet, GUI(WWW), and data format**
 - **Multilingual capability**
 - **User-friendly focus on system developer, manager, and end-user.**

Metadata Criteria - 2

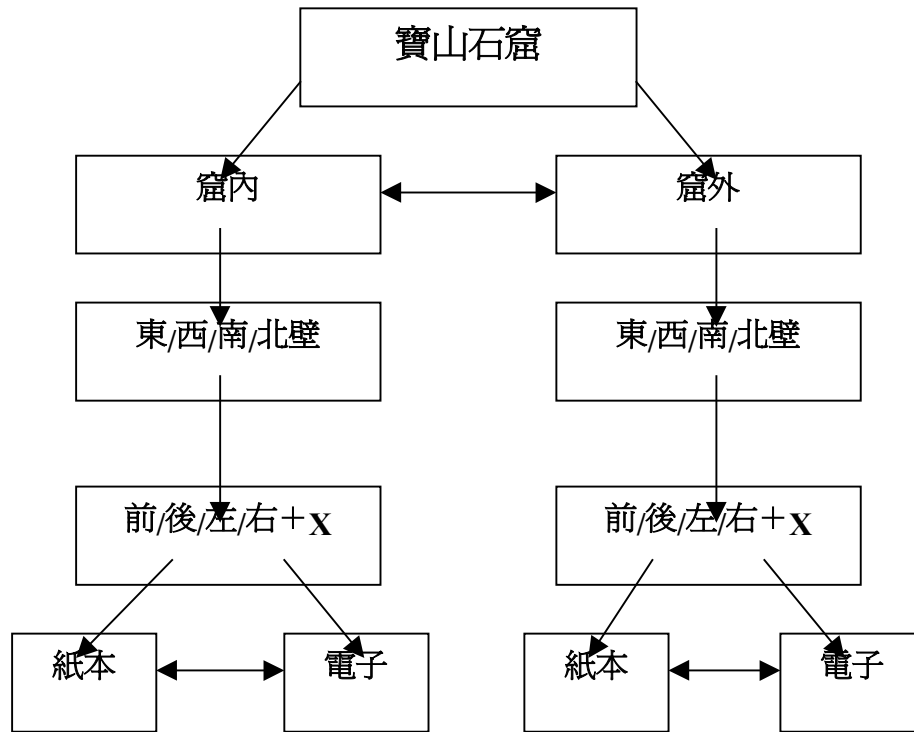
- **Format**
 - data structure
 - interchange and exchange
 - establishment and management of interactive linkage relationship among distributed documents and resources on network space
 - content of data elements

Metadata Criteria - 3

- **System functional cells**
 - **input**
 - **indexing and retrieval**
 - **representation**
 - **storage**
 - **interchange and exchange**
 - **others**
- **Implementations**

Metadata Workflow-4

- **A manual for Metadata Guidelines**
 - **end-user**
 - **system developer**
 - **related standards**
- **Tutorial**



計畫元素名稱	物品主體	物品展現形式	物品管理	說明
品名	◎			傅斯年圖書館之題名
別稱	◎			歷代不同稱呼名稱
年代	◎			1.中曆：含朝代、年號 2.西曆
出土地	◎			出土地
位置	◎			在石窟的所在地點方位
釋文	◎	◎		1.全文，共有三種形式： 全文(text)、照片、影像 2.檔名(有影像與 ASCII 兩種) 3.DPI 4.URL
撰者	◎			撰文者
書者	◎			書法書寫者
書體	◎			
收藏印	◎			標示用
圖像	◎			標示用，圖像合一
著錄	◎			參考資料
件數	◎			
高廣	◎			
收藏地點	◎			現典藏地點
收藏編號			◎	傅斯年圖書館收藏號
收藏筒號			◎	傅斯年圖書館收藏位置
電腦檔案			◎	1.檔案名稱 2.DPI 3.URL
附註	◎			
語文	◎			

計畫元素名稱	Metadata Elements (Dublin Core)	實 例	功能需求 (D/I/L/A/S/H)
品名	Title	大住聖窟大集經月藏 分人摩訶摩耶經中	D`I`L`S
別稱	Title.Alternative		D`I`L
年代	Date Coverage.Temporal	隋開皇九年 589	D`I`L`S
出土地	Coverage.Spatial	河南安陽寶山	D`I`L`S`H
位置	Coverage.Spatial	刻於窟內南壁西側北 圖 9.61	
釋文	Description		D`I`L`S`H
撰者	Contributor		D`I`L`S
書者	Contributor		D`I`L`S
書體	Description	楷書	D`I`L`S
收藏印			D`I`L`S`H
圖像	Description		D`I`L`S`H`A
著錄	Source & Relation		D`H
件數	Description	1	
高廣	Description		D
收藏地點	Rights-Management	傅斯年圖書館	D`I`L`S
收藏編號	Identifier	11017 之 1	D`I
收藏筒號	Identifier	1108	D
電腦檔案	Identifier	300 DPI	D`H
備註	Description		

Findings-1

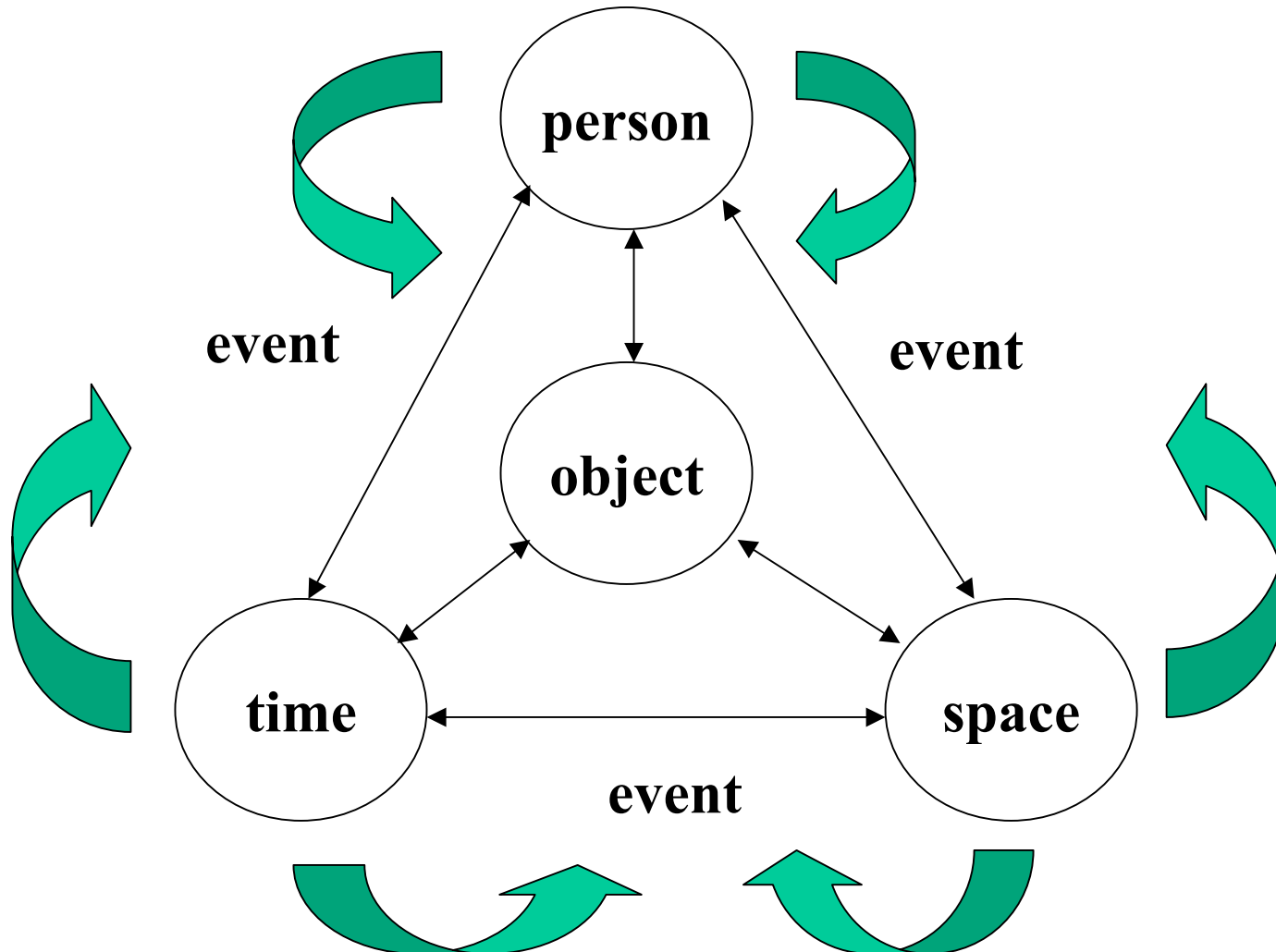
- **Collaboration Model**
 - **Content and research**
 - **Management**
 - **Metadata**
 - **System**
- **Why is not one of specified metadata standards**
 - **Approaches – content base, context focus, function approach, and user needs orientation**

Findings-2

- **Principles – content centric, standards interoperability, system implementation and integration**
- **Benefits – data attributes, research needs, system development, beyond the limitations of current metadata standards**

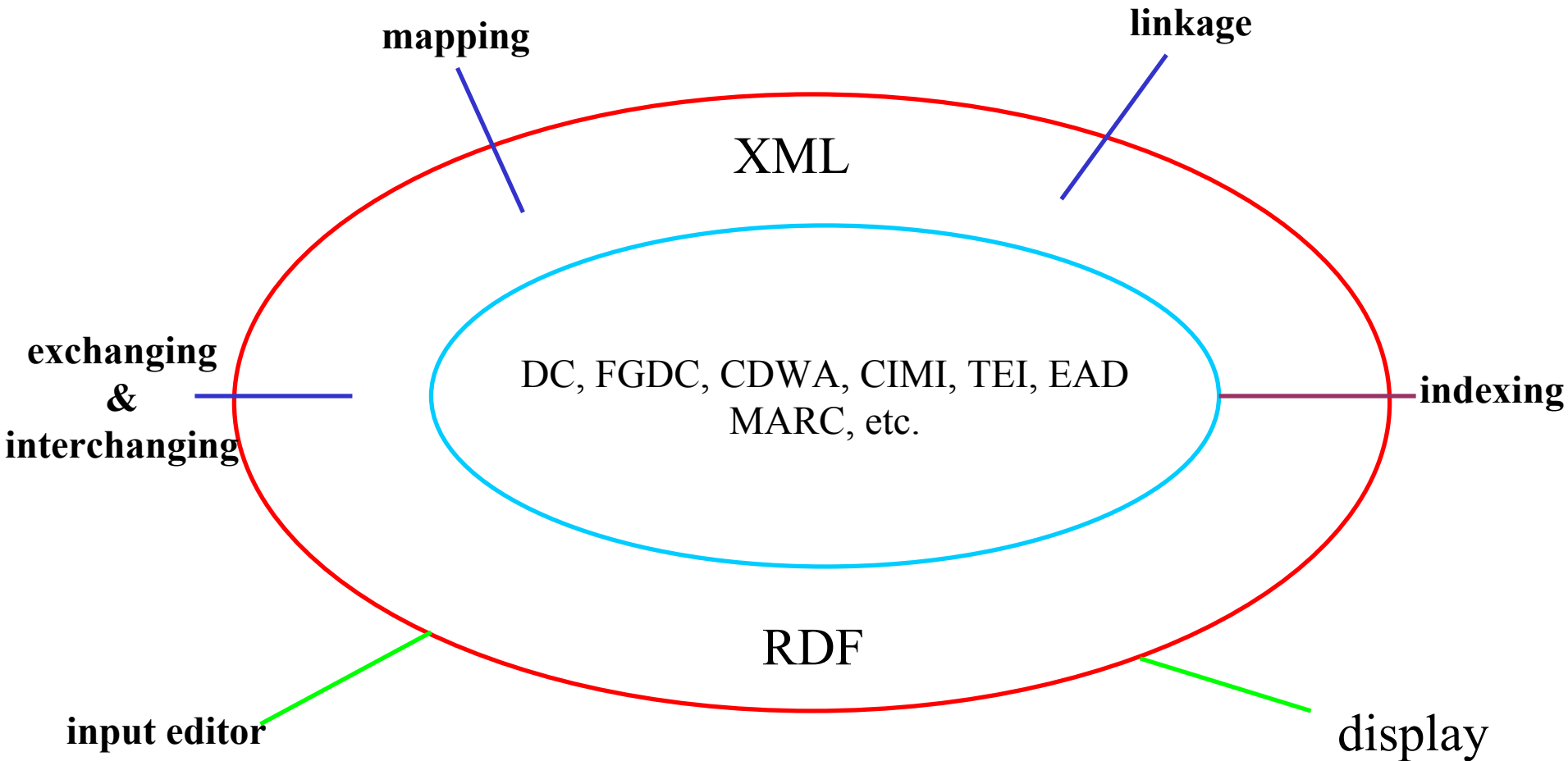
Findings-3

Metadata Relationship



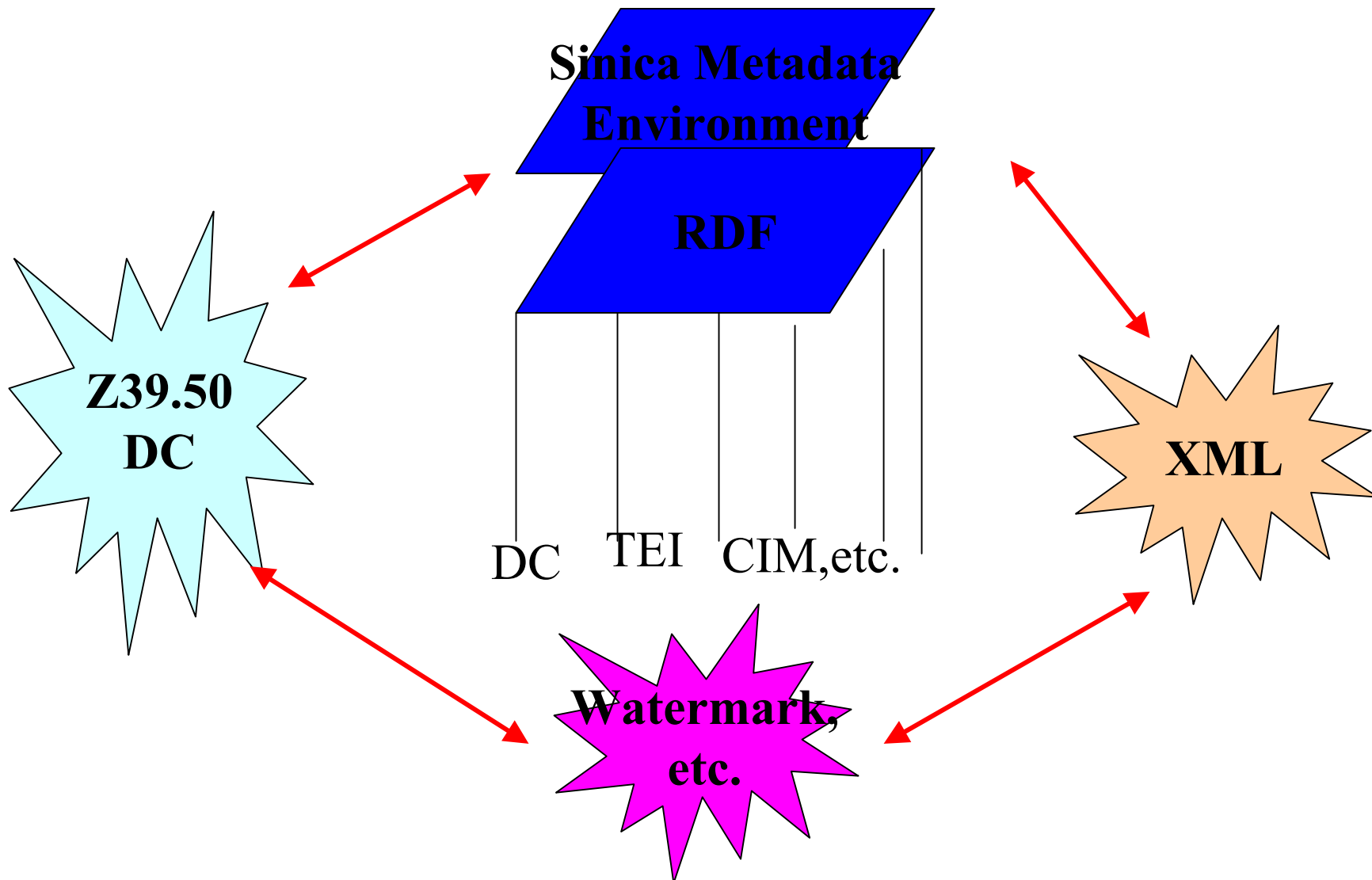
Findings-4

Metadata Framework



Findings-5

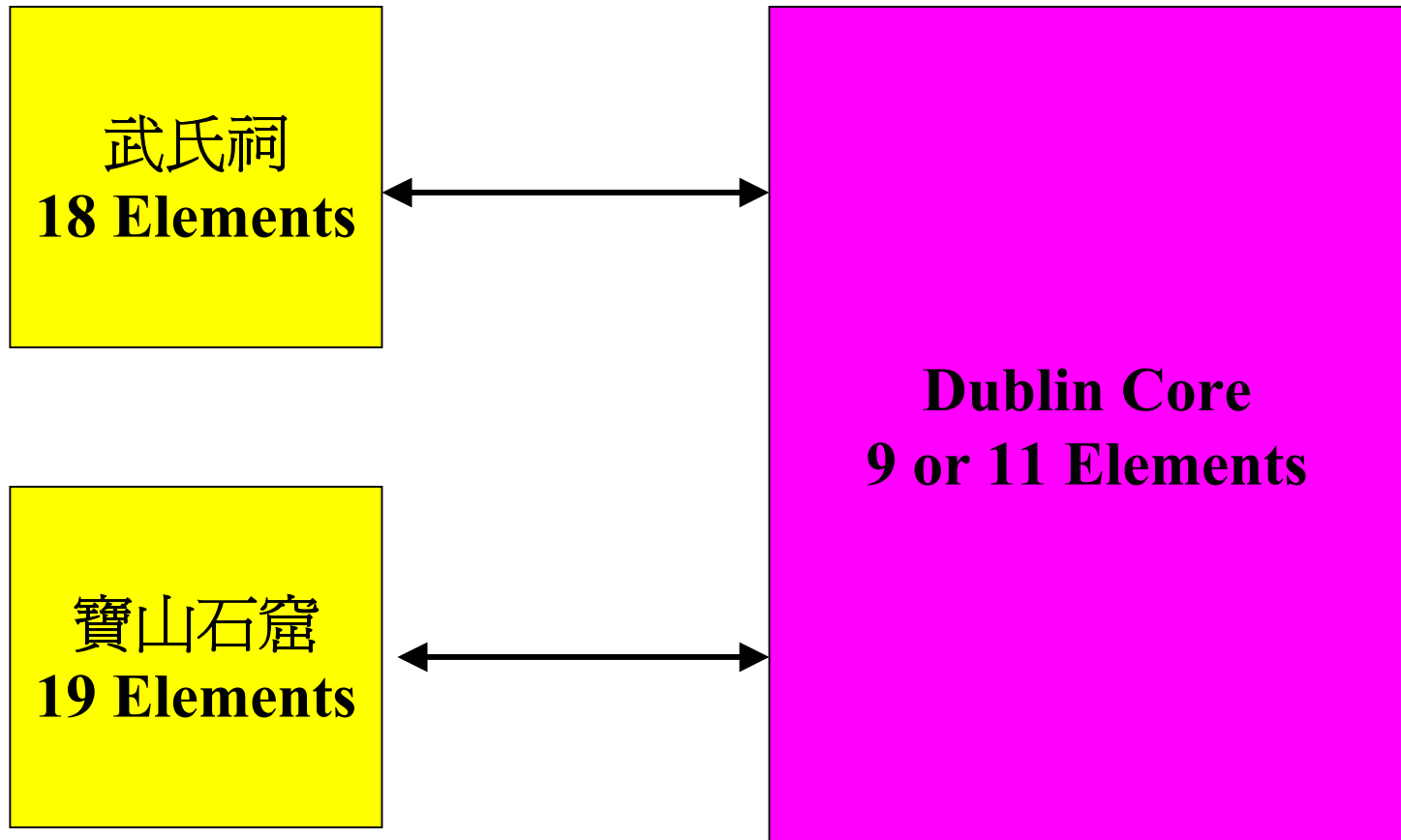
Metadata Environment



Conclusion and Suggestion-1

- **Metadata partnership – one total solution service from digitization to system and security management**
- **Metadata project management**
- **A question about Dublin Core – purpose and position of metadata**

Conclusion and Suggestion-2



計畫元素名稱	物品主體	物品展現形式	物品管理	說明
品名	◎			傅斯年圖書館之題名
別稱	◎			歷代不同稱呼名稱
年代	◎			1.中曆：含朝代、年號 2.西曆
出土地	◎			出土地
位置	◎			在石窟的所在地點方位
釋文	◎	◎		1.全文，共有三種形式： 全文(text)、照片、影像 2.檔名(有影像與 ASCII 兩種) 3.DPI 4.URL
撰者	◎			撰文者
書者	◎			書法書寫者
書體	◎			
收藏印	◎			標示用
圖像	◎			標示用，圖像合一
著錄	◎			參考資料
件數	◎			
高廣	◎			
收藏地點	◎			現典藏地點
收藏編號			◎	傅斯年圖書館收藏號
收藏筒號			◎	傅斯年圖書館收藏位置
電腦檔案			◎	1.檔案名稱 2.DPI 3.URL
附註	◎			
語文	◎			

計畫元素名稱	物品主體	物品展現形式	物品管理	說明
品名	◎			
榜題	◎			1.方向或層級 2.內容
題跋	◎			1.方向 2.內容
年代	◎			
出土地	◎			
原石存佚	◎			1.標示用 2.流傳記錄
收藏印	◎			1.印文字內容 2.顏色 3.收藏編號
著錄	◎			與出版品的關係
高廣	◎			
件數	◎			
備註一		◎		資料內在外觀的比較註
備註二		◎	◎	1.收藏方式 2.拓製方式
目錄號		◎?		
收藏編號			◎	
收藏簡號			◎	包含蒐錄的收藏編號
電腦照片檔號		◎	◎	1.檔名 2.DPI 3.URL
收藏地點			◎	
語文	◎			

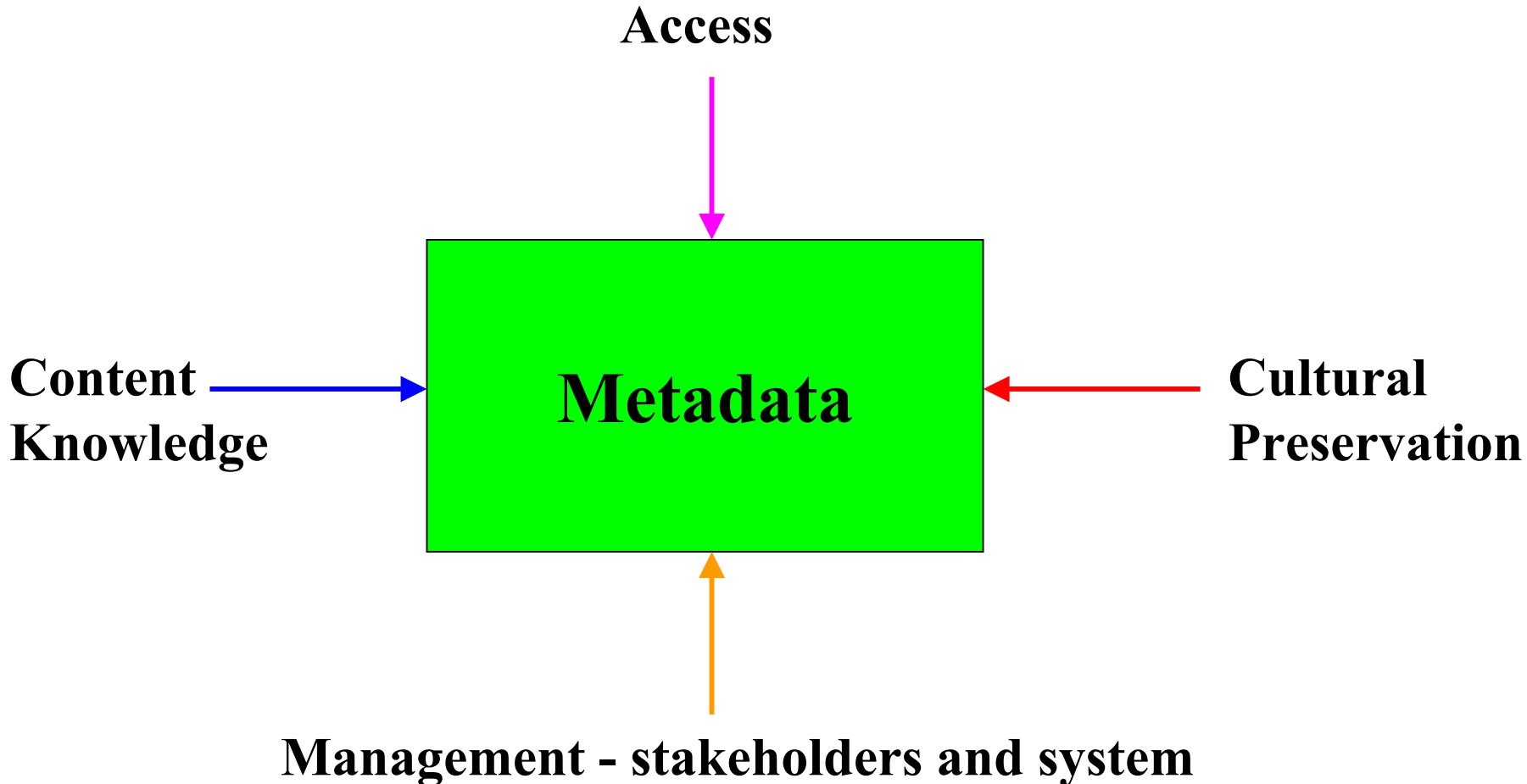
Conclusion and Suggestion-3

- **Metadata as DB Shema Approach**
 - **Standardization**
 - **DB's kernel is forced to change while standard revision occurs.**
 - **Metadata centric, not content centric**
 - **Overhead is high.**
 - **Precision is low.**

Conclusion and Suggestion-4

- **Metadata as Exchange and OO DB Approach**
 - **Content centric approach.**
 - **Overhead is low.**
 - **Standard can be implemented fully and connected smoothly.**
 - **Flexibility is high.**
 - **Precision is high.**

A right metadata approach



Current Status

- **Chinese rubbings - phase I**
- **Formosan plain aborigines - phase II**
- **Chinese Rare Books - TEI from Nov. 2000**
- **Metadata Clearinghouse - Sept. 2000**
- **Workshop - Oct. 2000**
- **Others**