

迎著晨曦

談

XML 暨其延伸 與我們圖資領域的關係

飛資得資訊有限公司

溫達茂

中華民國九十三年初夏於台北

XML-related Issues Covered:

- ◆ MARC 的另一思維 -- Metadata
- ◆ MARC 提昇 -- Web-based MARC
- ◆ Applications for Other Technical Services
 - OAI-PMH
 - FRBR
 - RSS

Who are We?

- ◆ Pure Librarian
- ◆ Metadata Specialist
- ◆ Hybrid Librarian
- ◆ Knowledge Facilitator

WWW對圖書館的衝擊

- ◆ 傳統圖書館 -- Hybrid Library -- Digital Library
- ◆ Access Catalog -- Holding Catalog
- ◆ Person-to-Person Reference Service -- Online Services (online help, tutorial, Electronic Cooperative Reference Service -- Question Point)

虛擬資源之特色

- 以多種形式出現
- 提供全文式的資訊
- 具有超文件或超媒體的能力，可以連結相關電子文件或其他電子資源
- 互動式多媒體形式
- 電腦與網路通訊、資訊內容業結合將促進多媒體業的發展
- 資訊以多層次方式提供
- 電子資源採用各種電子文件編碼標準，如TEI、SGML、HTML、XML
- 電子資源有些在本地電腦可直接取用，有些是透過遠端取用
- 電子資源的組織與資訊系統的檢索功能密切關連

http://public.ptl.edu.tw/publish/suyan/45/text_05.htm

Digital Library 館藏建置之來源

- ◆ 電子化資源
- ◆ 數位化資源
- ◆ 網路資源

Digital Library 館藏建置之來源 – 電子化資源

◆ 電子化資源: 電子化出版品

- 類型
 - ◆ 電子期刊
 - ◆ 電子圖書
 - ◆ CD-ROM 資料庫
 - ◆ 全文資料庫

Digital Library 館藏建置之來源 -- 數位化資源

◆ 數位化資源: 將傳統館藏加以數位化之後之資源

- Surrogate(載體) -- 書目資料之紀錄 (亦即metadata之描述)
- 內容之數位化
 - ◆ 文字部份:
 - 影像檔: Tiff, Gif, Pdf
 - 文字檔: txt檔, 註記語言(mark-up language)檔
 - ◆ 圖表部份
- 資料庫核心之選擇: 檢索引擎? 資料庫管理系統?

Digital Library 館藏建置之來源

-- 網路資源

◆ 靜態網路資源：

- Html
- Pdf

◆ 動態網路資源

- Java & JavaScript
- CGI
- PHP
- ASP

虛擬資源編目的難點

◆ 虛擬資源主要著錄根據難以確定

- 虛擬資源的各項目是變化的，任何一項變動都會打破書目記錄與虛擬資源的關係
- 虛擬資源極易複製，同一物件可能被保存在不同的網址上

◆ 虛擬資源的格式多樣

- 格式不同、文獻類型不同，因此所需之軟體、硬體都將有其限制，需註明否則將影響利用
- 多種格式的文件在目錄中將如何處理？

◆ 對於虛擬資源的檢索方式描述

- 需有特定的通訊地址(主機的IP位置)說明資源的位置
- 需提供通訊功能，便於直接登錄檢索有關資源

MARC對於虛擬資源描述所做的修訂

- ◆ 1970年代，編目規則加入「電腦檔」之資料類型
- ◆ 1980年代，由於唯讀光碟及個人電腦的成長，MARC加入了相關的欄位以處理這些媒體的屬性及硬體需求
- ◆ 1990年代，由於線上電子資源的增加，透過OPAC連結線上資源有立即的需求，因此MARC新增856欄位

陳昭珍教授

適用於虛擬資源編目的MARC屬性

- ◆ Leader/06: Type of Record
 - m (Computer File -- computer software (including programs, games, fonts), numeric data, computer-oriented multimedia, online systems or services)
- ◆ 006: Additional Material Characteristic
 - 00: Form of material(m--Computer File)
 - 09 - Type of computer file (008/26)

適用於虛擬資源編目的MARC屬性

- ◆ 008: Fixed Length Data Element
 - 26 (type of computer file)
 - ◆ a (Numeric Data)
 - ◆ b (Computer Program)
 - ◆ e (Bibliographic Data)
 - ◆ j (Online System or Service)

適用於虛擬資源編目的MARC屬性

- ◆ 256: COMPUTER FILE CHARACTERISTICS
 - \$a Computer file characteristics

Example :

- 256 \$a Computer data (2 files : 876,000, 775,000 records).
- 256 \$a Computer programs (2 files : 4300, 1250 bytes).
- 256 \$a Data (1 file : 350 records).

適用於虛擬資源編目的MARC屬性

◆ 516: TYPE OF COMPUTER FILE OR DATA NOTE

- \$a Type of computer file or data note

Example :

- 516 \$a Computer programs
- 516 \$a Numeric (Summary statistics)
- 516 \$a Numeric (Spatial data: Point)
- 516 \$a Text (Law reports and digests)

適用於虛擬資源編目的MARC屬性

◆ 538: SYSTEM DETAILS NOTE

Technical information about an item, such as software programming language, computer requirements, peripheral requirements

- \$a System details note

Example :

- 538 \$a Data in extended ASCII character set.
- 538 \$a Written in FORTRAN H with 1.5K source program statements.
- 538 \$a System requirements: IBM 360 and 370; 9K bytes of internal memory; OS SVS and OSMVS.
- 538 \$a Disk characteristics: Disk is single sided, double density, soft sectored.
- 538 \$a VHS
- 538 \$a Compact disc

適用於虛擬資源編目的MARC屬性

◆ 856: ELECTRONIC LOCATION AND ACCESS

- ◆ \$k Password
- ◆ \$l User Id
- ◆ \$u Uniform Resource Locator (NR)

Example :

856 \$u <http://www.flysheet.com.tw/>\$kquest\$lguest

適用於虛擬資源編目的MARC屬性 Holdings Data

- ◆ 5xx & 84x: Note Fields
- ◆ 852 & 856: Location & Access Field
- ◆ 853-855: Caption & Pattern Field
- ◆ 863-865: Enumeration & Chronology Field
- ◆ 866-868: Textual Holding Statement Field
- ◆ 876-878: Item Information Field

MARC的先天與後天

- ◆ MARC不同字符有其特定涵義，不依據相同格式標準很難分析出所代表的具體內容
- ◆ 10萬位元之限制
- ◆ 網路環境中的物件本體, 物件本體層次, 及物件與物件之關連性無法於856段中精準的加以呈現
- ◆ 856段需維護以確保資源的準確性
- ◆ 對一般使用者來說，MARC過於複雜，不易理解
- ◆ 著錄項目繁多，導致著錄工作速度慢、效率低
- ◆ 大部分應用侷限在書目資訊之描述
- ◆ MARC著錄結果用在HTML語言加以標示，過程複雜
- ◆ 結構複雜管理不便(以修改記錄為例)
- ◆ 三段式結構造成檢索上的不利
- ◆ 只有符合MARC規範的OPAC、Z39.50，才能取得MARC數據，無法透過搜尋引擎提供MARC給其他平台用戶

Alternatives to the Issues:

- ◆ *MARC vs. Metadata (Semantic -- 語意層面)*

- ◆ *Web-based MARC (Structure -- 檔案架構層面)*

DC與MARC21間的對應關係

- ◆ 一對一關係
- ◆ 一對多關係
- ◆ 多對一關係
- ◆ 無對應關係

一對一關係

DC	MARC21
Publisher	260\$b
Publisher place	260\$a

一對多關係

DC	MARC21
Subject	6XX段，如： 650普通論文主題 651地理名稱主題 600個人名稱主題
Language	546語種附註 041語種代碼 008的第35-37字元位元

多對一關係

DC	MARC21
< Relation > Is Format Of Has Format Of	776(其他載體形態款目)
Contributor CorporateName	710(附加款目、團體名稱)\$a \$b

無對應關係

- ◆DC中有些限定詞找不到適當的對應關係
- ◆如要強制對應，用附註字段來表示，如 Date、Modified對應到500段。

DC到MARC21轉換問題

- ◆DC的值未遵守MARC21的著錄規則
- ◆DC元素在轉換過程中會產生無法對應之處
- ◆非控制主題到控制主題的轉換需人工干預
- ◆DC元素轉入MARC21的非檢索欄位後，易失去具有檢索意義的資訊
- ◆DC的子元素轉換成MARC之後，無從判別特定資訊，如修改日期

MARC21到DC轉換問題

- ◆MARC21著錄中的標點符號轉到DC後會造成資料顯示的問題
- ◆MARC21中有說明作用的子欄位被忽略
- ◆MARC21與DC有些對應缺乏專指性
- ◆MARC21欄位的值在轉換中大量流失

MetaData格式的多樣化

- ◆由於DC與MARC21本身結構與句法定義的不同，Metadata間的轉換即使可透過自動轉換程序進行，仍無法避免資料的損失，
- ◆要達到理想的資料轉換，需要大量的人工干預。

Metadata格式需整合

- ◆資源類型繁多、內容龐雜、結構複雜、分佈廣泛
- ◆任何一種Metadata都無法對所有類型資進行描述
- ◆必須先解決Metadata之間的對照和轉換的問題

RDF/XML對Metadata多樣式的影響

- ◆RDF / XML結構清楚、定義嚴謹，可對任何Metadata進行描述
- ◆XML可透過各種資料著錄格式自動設置標記，比MARC21的著錄方式更為簡單
- ◆XML將內容的描述與顯示分開，而且可任意擴展
- ◆RDF / XML將不同Metadata整合在一起，可解決MARC與DC的互換問題

MARC vs. Dublin Core

MARC vs. XML

◆ semantic & structure perspectives:

- 差異性的比較應該不在於element vs. tag; qualifier vs. subfield 之比較.

◆ Syntax perspectives:

- Dublin Core: W3C XML
- MARC: ISO/2709.

Metadata 應用層面 -- 資源整合

◆ 單一平台: 內部資源的連結

◆ 跨平台: 外部資源整合

- Source 本身先metadata化 (標準化), 同樣的外在資源(target)也採用相同的metadata來作為彼此溝通依據:
 - ◆ Z39.50 (metadata: Bib-1 Attribute Set)
 - ◆ OAI (metadata: Dublin Core)
 - ◆ OpenURL (metadata: self-defined)

Alternatives to the Issues:

- ◆ MARC vs. Metadata (Semantic -- 語意層面)

- ◆ *Web-based MARC (Structure -- 架構層面)*

HTML Display -- MARC Display -- ISO-2709

- ◆ [A Bibliographical Record's HTML Display](#)

- ◆ [MARC Display](#)

- ◆ [ISO-2709](#)

檔案格式所造成的困擾

◆ This is a test.txt

◆ This is a test.rtf

◆ This is a test.doc

MARC 檔案格式所造成的困擾

◆ Test.txt -- test.rtf -- test.doc

◆ Iso-2709

◆ WWW (XML)

What is ISO-2709

◆ Plain Text_1

◆ Plain Text_2

◆ Where

◆ What & How?

WWW is -----

◆ Independent of

- Operating system
- Database Structure
- File format

WWW markup language

◆ Content:

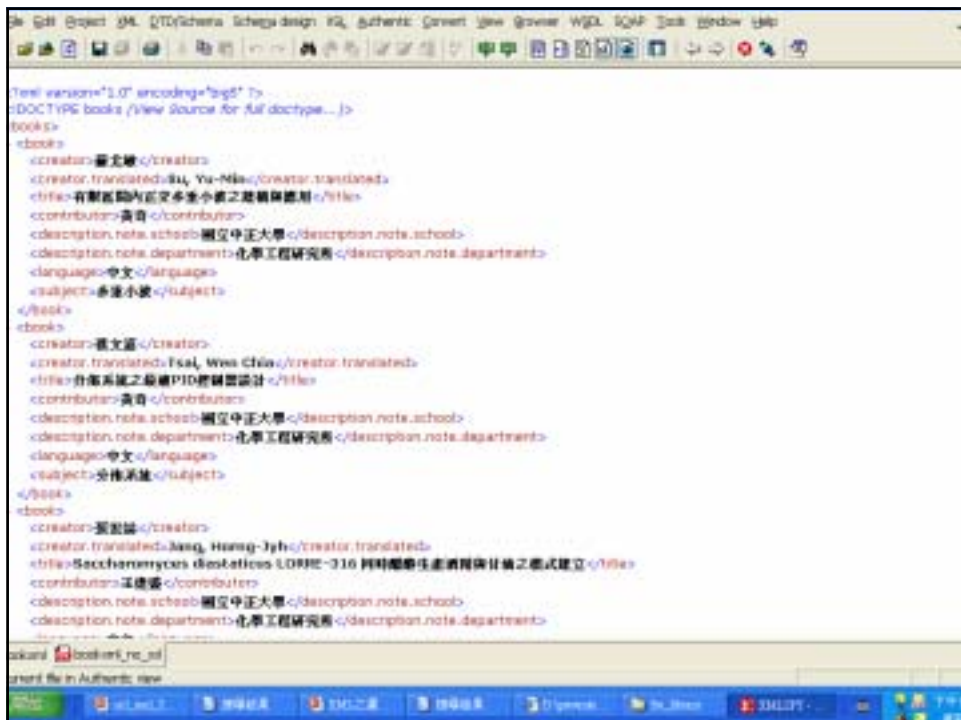
- 從網頁(Web Page)到網路資源(Web Resources)
- 從Presentation (Appearance) 到Representation
- 從網頁內容的單純呈現(Display)到網頁內容的語義暨主題分析之後的表達暨呈現

◆ Markup Language:

- HTML:
 - ◆ (A Single, Predefined markup language)
 - ◆ A language for specifying the layout of web pages.
 - ◆ Markup specifies how content appears
 - Level of headings
 - Emphasize (with bold or italic) portions of the content
- XML
- SGML

What is XML

- ◆ XML, a formal recommendation from the World Wide Web Consortium ([W3C](http://www.w3.org))
- ◆ A MetaLanguage -- A Language for Describing Other Language -- which lets you design your own customized markup languages for limitless different types of documents
- ◆ XML (Extensible Markup Language) is a flexible way to create common [information](#) formats and share both the format and the [data](#) on the World Wide Web
- ◆ XML is "extensible" because, unlike HTML, the markup symbols are unlimited and self-defining
- ◆ XML describes the content in terms of what data is being described logically & semantically



```
Test: version="1.0" encoding="big5" ?>
<DOCTYPE books [view source for fulldoctype... ]>
books
<books>
  <creator>蕭文雄</creator>
  <creator.translated>Hsi, Yu-Miao</creator.translated>
  <title>有關系統內嵌多邊小波之結構與應用</title>
  <contributor>黃奇</contributor>
  <description.note.school>國立中正大學</description.note.school>
  <description.note.department>化學工程研究所</description.note.department>
  <language>中文</language>
  <subject>多邊小波</subject>
</book>
  <creator>蕭文雄</creator>
  <creator.translated>Fsal, Wen-Chia</creator.translated>
  <title>分散系統之最優PID控制與設計</title>
  <contributor>黃奇</contributor>
  <description.note.school>國立中正大學</description.note.school>
  <description.note.department>化學工程研究所</description.note.department>
  <language>中文</language>
  <subject>分散系統</subject>
</book>
  <creator>蕭文雄</creator>
  <creator.translated>Jiang, Hong-Jyh</creator.translated>
  <title>Saccharomyces cerevisiae LDPRE-316 同時態產生酒精與甘糖之模式建立</title>
  <contributor>江建豪</contributor>
  <description.note.school>國立中正大學</description.note.school>
  <description.note.department>化學工程研究所</description.note.department>
</books>
```

◆ 屬性的定義 (Semantic)

- MetaData

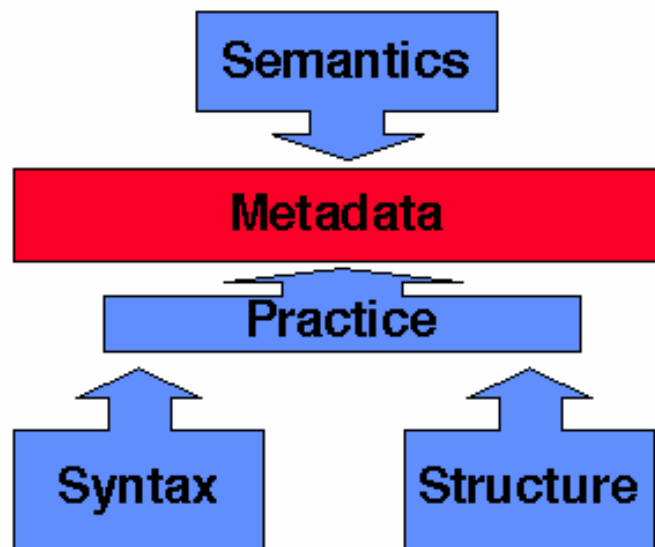
◆ 屬性的架構 (Structure)

- DTD; XML/Schema

◆ 著錄的儲存 (Storage, Syntax)

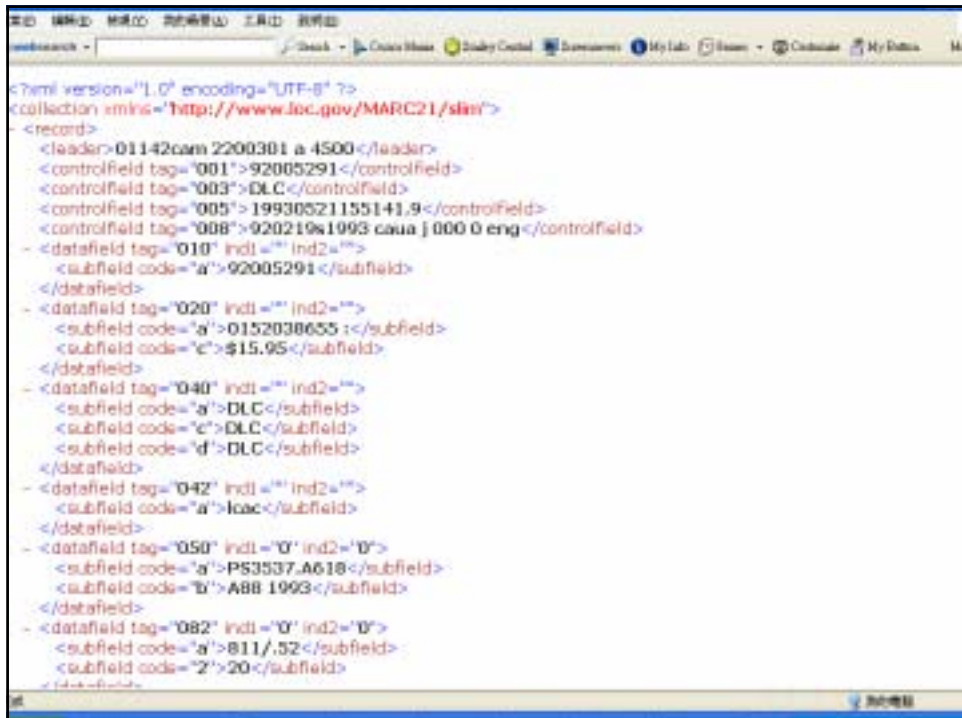
- XML

Overall



XML機制之流程

- DTD -- [XML](#) -- [XSL](#) -- [OUTPUT](#)
- [XML](#) -- [Database](#)



```
<?xml version="1.0" encoding="UTF-8" ?>
<collection xmlns="http://www.loc.gov/MARC21/slim">
  <record>
    <leader>01142cam 2200301 a 4500</leader>
    <controlfield tag="001">92005291</controlfield>
    <controlfield tag="003">DLC</controlfield>
    <controlfield tag="005">19930521155141.9</controlfield>
    <controlfield tag="008">920219s1993 caua j 000 0 eng</controlfield>
    <datafield tag="010" ind1="1" ind2="1">
      <subfield code="a">92005291</subfield>
    </datafield>
    <datafield tag="020" ind1="1" ind2="1">
      <subfield code="a">0152030655</subfield>
      <subfield code="c">$15.95</subfield>
    </datafield>
    <datafield tag="040" ind1="1" ind2="1">
      <subfield code="a">DLC</subfield>
      <subfield code="c">DLC</subfield>
      <subfield code="d">DLC</subfield>
    </datafield>
    <datafield tag="042" ind1="1" ind2="1">
      <subfield code="a">lcac</subfield>
    </datafield>
    <datafield tag="050" ind1="0" ind2="0">
      <subfield code="a">P53537.A618</subfield>
      <subfield code="b">A88 1993</subfield>
    </datafield>
    <datafield tag="082" ind1="0" ind2="0">
      <subfield code="a">811/.52</subfield>
      <subfield code="2">20</subfield>
    </datafield>
  </record>
</collection>
```


ISO-2709 -- MARC -- XML

◆ XML

- [OAI_MARC](#) -- [OAI_MARC - DTD](#) -- [xml schema](#)
- [MARCXML](#) -- [MARCXML - DTD](#) -- [xml schema](#)
- [MODS](#) -- [MODS Schema](#)
 - ◆ [What is MODS](#)

為何要從ISO 2709跨越到XML

上游 -- 圖書館 -- 下游

- 轉成ISO 2709比轉成XML麻煩
- ISO 2709的資料無法直接呈現在網頁上與data的再利用
- 書目資料(或metadata)無法直接與標誌過的全文資料結合
- 機讀格式較繁瑣，有些欄位未發揮功能
- 適合固定、已出版的資料
- 由館員描述資料

XML Schema (DTD)?

◆ What:

- A Schema describes what one or more XML documents can look like, and defines:
 - ◆ The elements the document contains, and the order in which they appear
 - ◆ The element content, and element attributes if any

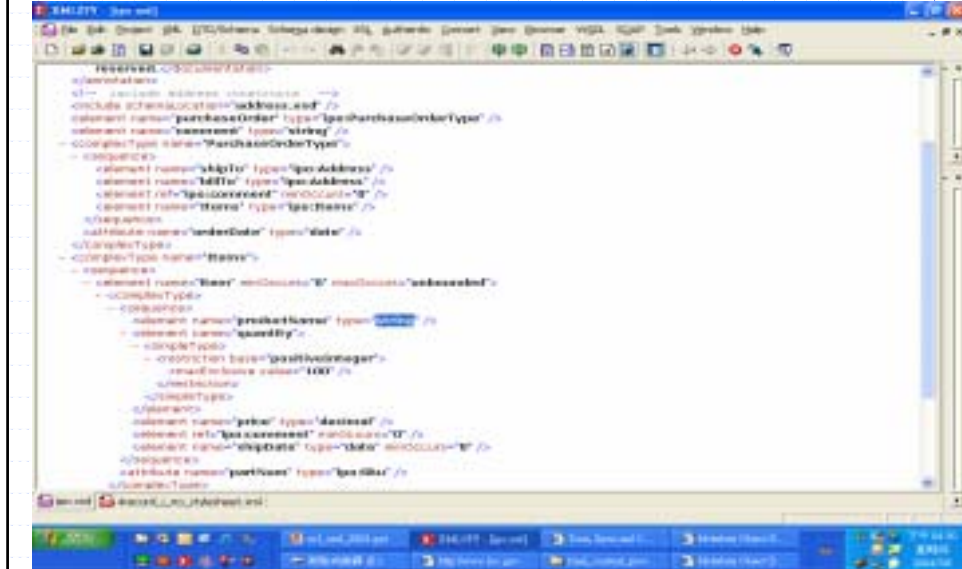
◆ Why:

- The purpose of a schema is **to allow machine validation of document structure**. Instead of using the syntax of XML 1.0 DTD declarations, schema definitions use XML element syntax. A correct XML schema definition is, therefore, a well-formed XML document.

DTD vs. Schema

- | | |
|---------------------|-----------------------|
| ◆ 採用不同於XML的語言 | ◆ 採用和XML相同的語言 |
| ◆ 一個XML文檔對應一個DTD | ◆ 一個XML文檔可對應多個SCHEMA. |
| ◆ 對子類和繼承的支持有限(層級描述) | ◆ 對子類和繼承支持度高(層級描述) |
| ◆ 沒有資料類型的概念 | ◆ 有資料類型的概念 |

DTD vs. XML Schema



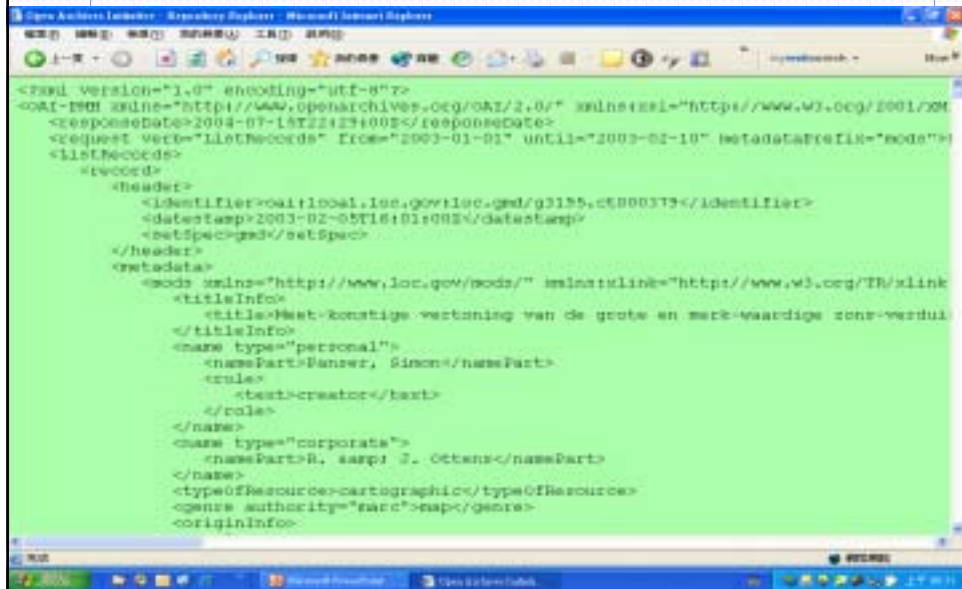
```
<?xml version="1.0" encoding="UTF-8" ?>
<store xmlns="http://schemas.xml.org/store" ?>
  <include href="http://schemas.xml.org/store/online" ?>
    <element name="purchaseOrder" type="xsd:string" />
    <element name="customer" type="xsd:string" />
    <complexType base="PurchaseOrderType" />
  </include>
  <element name="shipTo" type="xsd:string" />
  <element name="billTo" type="xsd:string" />
  <element ref="purchaseOrder" minOccurs="0" />
  <element name="Items" type="xsd:string" />
  <complexType base="OrderData" type="data" />
</store>
<complexType name="Items">
  <sequence>
    <element name="Item" minOccurs="0" maxOccurs="unbounded">
      <complexType>
        <sequence>
          <element name="productId" type="xsd:string" />
          <element name="quantity">
            <complexType>
              <restriction base="xsd:float">
                <minInclusive value="1.00" />
              </restriction>
            </complexType>
          </element>
          <element name="price" type="xsd:string" />
          <element ref="purchaseOrder" minOccurs="0" />
          <element name="shipData" type="data" minOccurs="0" />
        </sequence>
      </complexType>
    </element>
  </sequence>
</complexType>
</store>
```

Schema -- XML Document

or

XML Document -- Schema

So what! And Then?



```
<?xml version="1.0" encoding="utf-8"?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:dc="http://purl.org/dc/terms/">
  <responseDate>2004-07-15T22:23:00Z</responseDate>
  <request verb="ListRecords" from="2003-01-01" until="2003-02-10" metadataPrefix="mods">
    <listRecords>
      <record>
        <header>
          <identifier>oai:loc.gov:loc.gov:gd/g3135.ct000373</identifier>
          <datestamp>2003-02-05T16:01:08Z</datestamp>
          <setSpec>gd</setSpec>
        </header>
        <metadata>
          <mods xmlns="http://www.loc.gov/mods/" xmlns:xlink="http://www.w3.org/TR/xlink" >
            <titleInfo>
              <title>Niet-konstige vertoning van de grote en merkwaardige zons-verdui</title>
            </titleInfo>
            <name type="personal">
              <namePart>Panzer, Simon</namePart>
              <role>
                <text>creator</text>
              </role>
            </name>
            <name type="corporate">
              <namePart>R. sasp; J. Ottens</namePart>
            </name>
            <typeOfResource>cartographic</typeOfResource>
            <genre authority="sarc">map</genre>
            <originInfo>

```

Application of XML for Technical Service Librarian

◆ Cataloging:

- Bibliographical Information Resources for both Physical and Virtual Collections
 - ◆ Z3950 and XML
 - ◆ Z3950 and OpenURL
- Controlled Vocabularies (Classification, Thesaurus, and Subject Headings)

◆ Selection & Acquisition:

- Publisher -- Aggregator -- Dealer -- Trader -- Library
- Acquisition Module -- Cataloging Module

Application of XML

◆ OAI:

- Reader Services:
- Technical Services:

◆ RSS:

- Reader Services:
- Technical Services

◆ FRBR:

Application of XML: OAI-PMH

◆ What is OAI-PMH?

◆ How is OAI-PMH working?

◆ Why do we need to know about OAI-PMH?

Definition

◆ Open

- Not “free”, “unlimited”
- Facilitate the availability of content from a variety of providers

◆ Archive

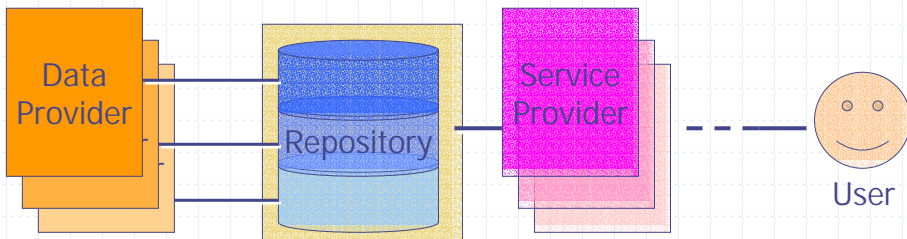
- A repository of scholarly papers
- A repository for stored information

Mission

- ◆ To facilitate the efficient dissemination of content
- ◆ To enhance access to e-print archives
- ◆ Opening up access to a range of digital materials

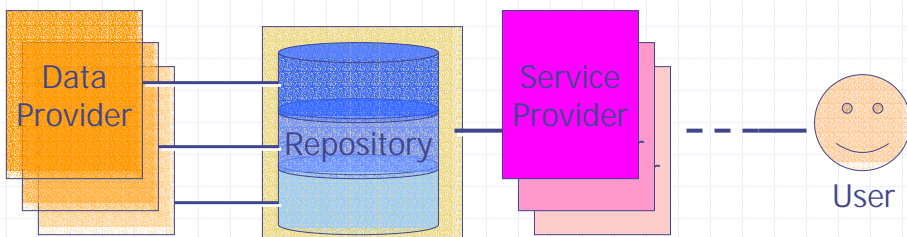
Data Provider

- ◆ Maintain one or more repositories (Web server) that support the OAI protocol as a means of exposing metadata about their content



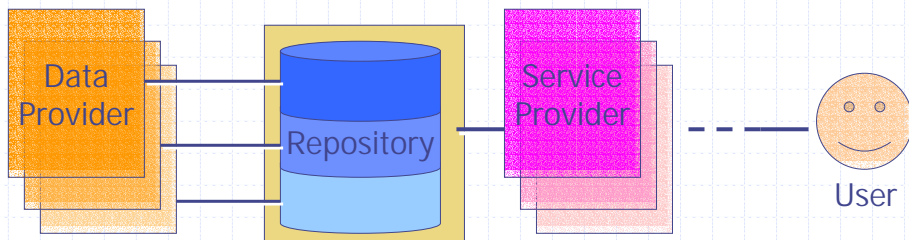
Service provider

- ◆ Issue OAI protocol requests to data providers and use the metadata as a basis for building value-added services



Repository

- ◆ A Network accessible to which OAI protocol requests, embedded in HTTP



OAI Metadata Harvesting Protocol

◆ Verbs

Identify: return administrative information about a repository

ListMetadataFormats: return list of metadata formats supported by repository, or for specific record in repository

ListSets: return a list of sets supported by the repository

GetRecord: return 1 record given an identifier & format desired

ListIdentifiers: return a list of record identifiers, optionally filtered by date or set

ListRecords: return a list of records in a given metadata format, optionally filtered by date or set

- ◆ Can't Retrieve/Filter by Subject or Keyword

聯合目錄工作流程

OAI-- Metadata Harvesting

The screenshot displays the Open Archives Initiative (OAI) website interface. At the top, the text "Open Archives Initiative" is prominently displayed, followed by the Chinese translation "我會得數位典藏計畫專家網站". Below this, there is a blue box with the text "seamless information flow". A search bar is present with the placeholder text "請選擇您查詢單位" and a dropdown menu showing "Check Latest Items".

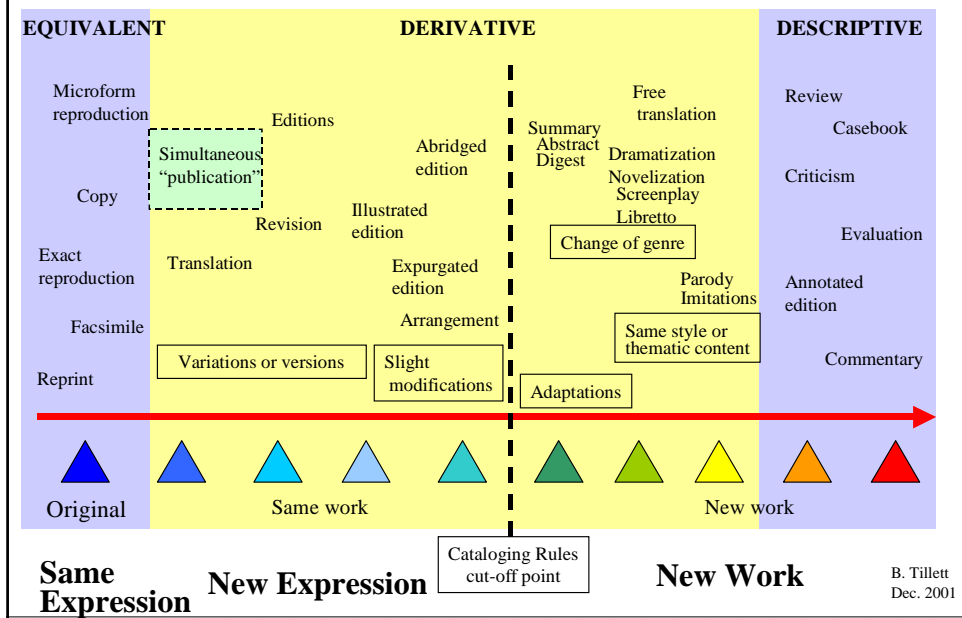
The main content area is divided into two sections by arrows pointing downwards. The left section is titled "[Verb]" and lists several options: "RAW XML", "PARSED", "Metadata", "List Sets", "List Members", "List Records", and "Get Records". The right section is titled "[Parameters]" and contains several input fields: "from (qy... YYYY-MM-DD):", "until (qy... YYYY-MM-DD):", "metadataPrefix:", "identifier:", "set:", and "metadataType:". The website is viewed in a Microsoft Internet Explorer browser window.

Why do I, as a Technical Services Librarian, need to care about OAI?

Application of XML FRBR

- ◆ What is FRBR
- ◆ How it works!
- ◆ Why?

Family of Works

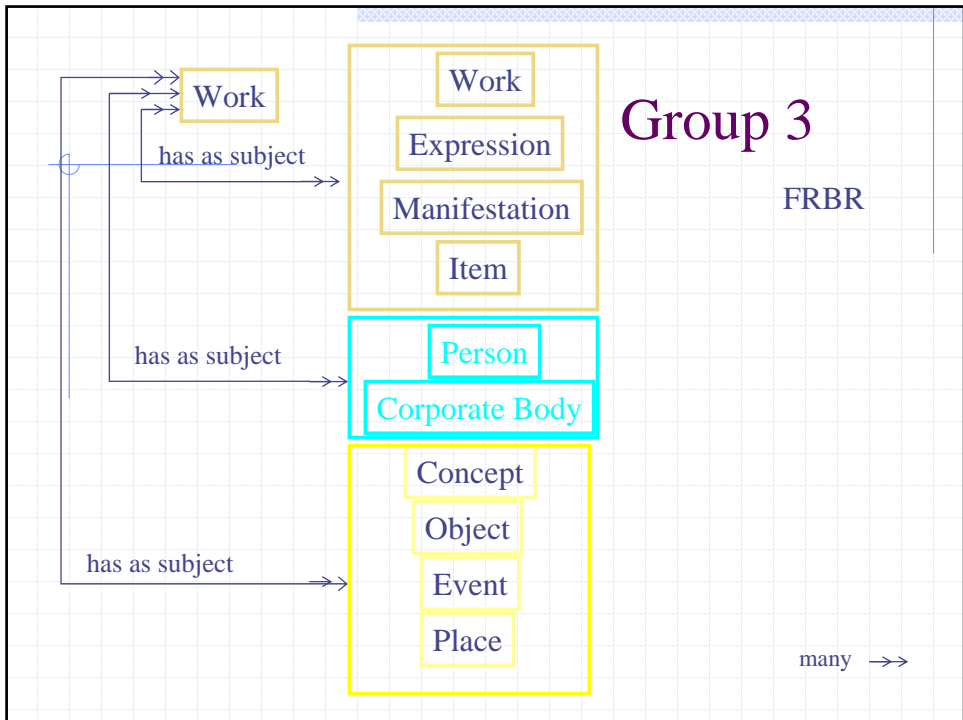
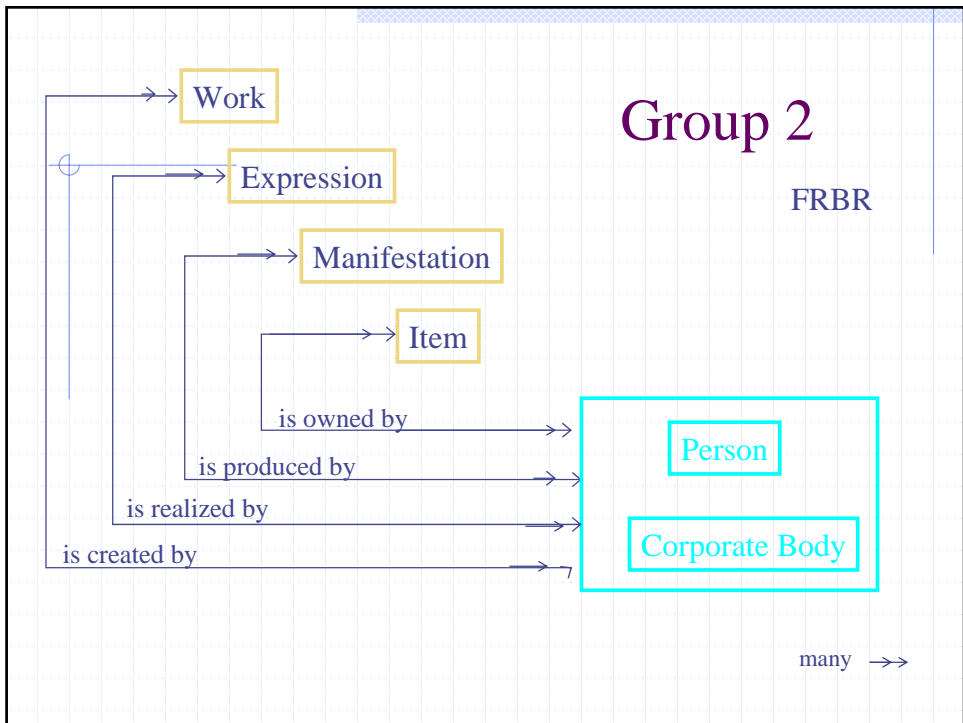


FRBR Group 1 Content Relationships

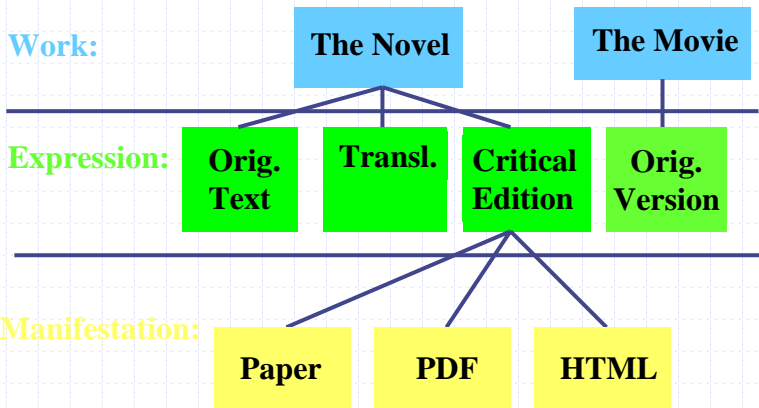
- ◆ Equivalent
- ◆ Derivative
- ◆ Descriptive

Work to work relationships are inherited by hierarchically related

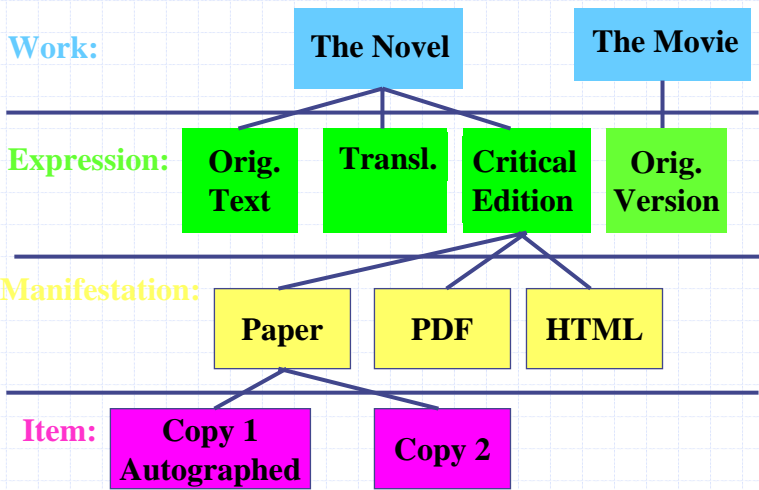
- Expressions
- Manifestations
- Items



FRBR Entity Levels



FRBR Entity Levels



FRBR Entity Levels

Family of works

Work:

The Novel

The Movie

Expression:

Orig.
Text

Transl.

Critical
Edition

Orig.
Version

Manifestation:

Paper

PDF

HTML

FRBR---以Sister Carrie為例

Sister Carrie 基本書目資料

1	Sister Carrie / Theodore Dreiser (1871-1945). Johnson Reprint Corp, 1969
2	Sister Carrie / Theodore Dreiser (1871-1945). Doubleday, 1997
3	Sister Carrie / Theodore Dreiser (1871-1945). University of Pennsylvania Press, 1981
4	Sister Carrie / Sheldon Norman Grebstein Everett. Edwards, 1970

The screenshot displays a bibliographic record for 'Sister Carrie' with the following details:

- Author:** Dreiser, Theodore, 1871
- Works:** Sister Carrie
 - Form:** text - English (expression)
 - Edition:** 1st New York Public Library collector's ed.
 - Title: Sister Carrie
 - Statement of responsibility: Theodore Dreiser.
 - Imprint: Doubleday, 1997 (manifestation)
 - Physical Description: 280 p. : ill. ; 20 cm
 - ISBN: 0383487205 (alk. paper)
 - Edition:** The Pennsylvania ed.
 - Title: Sister Carrie
 - Statement of responsibility: Theodore Dreiser.
 - Imprint: University of Pennsylvania Press, 1981 (manifestation)
 - Physical Description: 24, 679 p. : ill. ; 24 cm.
 - ISBN: 08122778-8
 - ISBN: 0812211103 (pbk.)
 - Form:** cartographic - English (expression)
 - Edition:** An abridged ed. by Theodore Dreiser and Arthur Henry. With a new introd. by Jack Selman.
 - Title: Sister Carrie
 - Imprint: Johnson Reprint Corp, 1969
 - Physical Description: 2, 357 p. ; 21 cm.
- Author:** Grebstein, Sheldon Norman
- Works:** Sister Carrie (Theodore Dreiser)
 - Form:** text - English
 - Edition:**
 - Title: Sister Carrie (Theodore Dreiser)
 - Imprint: Everett/Edwards, [1970] 165
 - Physical Description: p. 1 cassette, 2 1/2 x 4 in.

Sister Carrie之MARC

FRBR Alogrithm

1.Title and Author: Work:

- author (100段、110段、111段)
- title(240段、243段、245段)

2.Title only:

- when no 100 or 110 or 111
- Match display 240、243、245

3.Expression level : match

- Leader/06 (type of record)
- 008/35-37 (language)

4.Manifestation level : 排序依據 008/07-10 (publication date)

Where is XML in FRBR?

◆ Iso-2709 -- XML ([slimfrbr.xml](#))

◆ [Clean.xml](#)

◆ [Match.xml](#)

◆ [Result.html](#) [Result.xml](#)

<http://tw.news.yahoo.com/rss/>



<http://www.kclibrary.org/rss/>



http://lxml1.ust.hk/na/na_display.pl

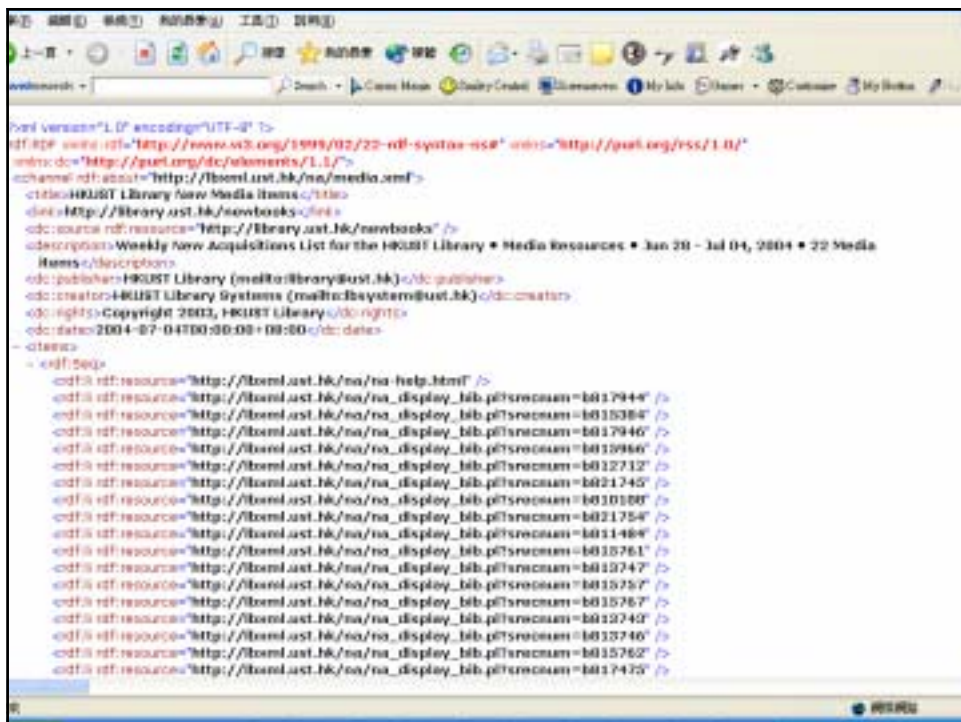


RSS(Really Simple Syndication)

全名：Really Simple Syndication

是一種用來分發和匯集網頁內容(例如新聞標題)的XML格式。





Latest Research from OCLC

◆ ERROL

- [OAI Registry](#)
- [Extensible Repository Resource Locators \(ERROLs\) for OAI Identifiers](#)

◆ "*info*" URIs

- Naming
- Addressing

◆ [LC name authority file](#)

◆ [FAST](#)

Finally! But, What do we Have? Why do we Have These? How do we Use These?

◆ XML is not HTML (Present vs. Represent)

- We make a HTML File, but We Generate (systematically) an XML file

◆ 3s for XML (Semantic, Structure, Syntax)

- Semantic-- Attributes for the Objects
 - ◆ Semantically (MODS) & Structured (MARCXML) Perspectives
- Structure-- ComplexType – Element – Attribute – Type – Value (Defined from DTD or XML Schema)
- Syntax-- Reference to each other for Elements (XMLNameSpace)

◆ XML is Displayable but not Readable -- Stylesheet

◆ XML & Database are Highly Related to Each Other

No more! That's It.

Q & A