

# Z39.50 and other standards and protocols

---

Sally H. McCallum  
Library of Congress  
smcc@loc.gov

# What I plan to cover

---

- Retrieval protocols – Z39.50, OAI, OpenURL
- Metadata formats – MARC, DC, MODS, XML
- Emerging digital standards – METS

# Retrieval standards

---

- Z39.50 (classic)
  - ZING
- OpenURL
- OAI

# Z39.50

---

- International Standard
  - ISO 23950 – Information Retrieval (Z39.50): Application Service Definition and Protocol
  - ANSI/NISO Z39.50
    - Available from NISO free on the web
    - [www.niso.org](http://www.niso.org)

# Z39.50 characteristics

---

- Supports information retrieval and associated activities
- Supports use of local interface for all external searching; local displays for responses
- Client/server model
- Software and hardware independent searching and record transfer
- Communicates peer-computer to peer-computer
- Protocol defines each computer's tasks
- Modular – simple but rich

# Z39.50

---

- What it is not in its classic form
  - Common command language
  - User interface
  - Database specification
  - Network
- Z39.50 is important component in building networks

# Z39.50 search

---

- Typical (simple) communication
  - User enters search in local syntax, indicates target
  - Z39.50 Client parses query into components and sends query to Z39.50 Server at target
  - Z39.50 Server converts query into search logic of server site database
  - Results obtained from server database
  - Z39.50 Server packages results as requested by user and sends them to Z39.50 Client site
  - Results received by Z39.50 Client and screened in local result set displays for user

# Z39.50 searches

---

- Search term
- Search attributes
  - Use – subject, author, name, title, ...
  - Relation – less than, equal, ...
  - Position – 1<sup>st</sup> in field, anywhere, ...
  - Structure – word, phrase, date, string, ...
  - Truncation – end, beginning, ...
  - Completeness – total field, part of field, total subfield
  - Boolean – and, or, and\_not
  - Proximity – near, within XX characters



# Z39.50 results

---

- What comes back?
  - Full bibliographic records
  - Brief bibliographic records
  - Circulation information
  - Holdings information
  - Combinations

# Z39.50 results

---

- Record syntaxes
  - Generic
    - SUTRS (unstructured)
    - GRS-1 (structured)
  - Content specific
    - MARC formats, e.g., MARC 21 (specific MARC specified with search)
    - OPAC

# Z39.50 services

---

- Initialize – Search – Response – Present
  - Basic, most used and implemented
- Scan – Explain – Access control – Resource control
- Save result sets – Sort – Export
- Order an item – Update a database

# Z39.50 implementation

---

- Some implementation choices
  - Fully integrated client embedded in local search system – most integrated library systems
  - Integrated client with new user interface
  - Stand-alone client
  - Record incorporation capability
  - Web gateway
    - Web interface (http) to Z gateway (Z39.50) to Z39.50 target

# Z39.50 uses

---

- Common uses
  - Reference and ILL searching of other system(s) – single or multiple (virtual union catalog)
  - Copy cataloging – using the record incorporation capability
  - New front end to existing system
  - Resource files and other local file access
  - Bibliographies

# Z39.50 maintenance

---

- Maintenance
  - ZIG (Z3950 Implementers Group)
  - Library of Congress maintenance agency [www.loc.gov/z3950/agency/](http://www.loc.gov/z3950/agency/)
  - Listserv - [www-zig@w3.org](mailto:www-zig@w3.org)

# Z39.50 software

---

- [www.loc.gov/z3950/agency/resources/software.html](http://www.loc.gov/z3950/agency/resources/software.html)
- Free
  - YAZ Toolkit
  - VB Zoom
  - ICONE
  - Many more
- Commercial
  - Bookwhere
  - Endnote
  - Procite
  - Most integrated library systems

# Z39.50 sites

---

- Who has it?
  - Most larger libraries in the US – over 500 sites
  - Many European libraries
  - Strong in Australia, South Africa



# Z39.50 issues

---

- Highly flexible
- Different indexing in systems
- Different formats for records
  
- Bath profile
  - Identifies two set of searches (attributes) - for *basic* and for *more precise* library searches

# ZING

---

- Z39.50 International Next Generation
- End-user orientation
- Simpler and more limited in functionality than Z39.50 classic - Bath profile oriented
- Uses web technology – XML, http
- Builds from Z39.50, thus enables gateways between ZING and Z39.50

# ZING functionality

---

- Search, retrieval, and explain are key functions
- XML for protocol messages
- XML record syntax
  - MARCXML for MARC 21
- Query can be end user (common query language) or in an XML encoding
- Keeps Z39.50 search attributes, flattened

# ZING components

- Search/Retrieve web service (SRW)
  - Information retrieval protocol
- Search/Retrieve URI service (SRU)
  - Search embedded in a URL
  - Example:
    - `http://www.kb.nl/cgi-zoek/srw.pl?query=(dc.titleWord=%22gone%22and dc.titleWord=%22wind%22)&sort=d,,,dc_record,date&maxRec=1`

# OpenURL

---

- Redirect facility
  - Click on citation
  - URL containing metadata from the citation is sent to a resolver service
  - Resolver sends back service options menu for user, examples of services:
    - Full text of article from location XX
    - Alternative delivery services
    - Abstract of article
    - Reviews

# OpenURL

---

- Status: Draft standard in process of finalization
- Will go out as “Draft Standard for Trial Use” soon
  - [www.niso.org](http://www.niso.org)
- Example
  - `http://LinkFinderPlus.library.edu?genre=article&issn=12345678&volume=99&issue=1&date=20020101&spage+27&atitle=What Is An OpenUDL&title=Harry's White Papers`

# OAI

---

- Open Archive Initiative (OAI)
- Protocol for *harvesting* metadata
  - Simple harvest criteria
    - date ranges
    - database names
    - single record by ID
  - Specify the record format
- Not for searching

# OAI

---

## ■ Metadata formats

- OAI is XML-based; records must be in XML
- Version 2.0 specifies
  - MARC21's MARCXML for MARC records
  - DCMI's DC XML schema for Dublin Core records

## ■ Example

- <http://www.perseus.tufts.edu/cgi-b:in/pdataprov?verb=ListRecords&from=2002-05-01T14:15Z&until=2002-05-01T14:20:00Z&metadataPrefix=marcxml>



# Protocol relationships

---

- Z39.50 and ZING (SRW/SRU)
  - Support information retrieval across different systems - metadata and/or resources
- OpenURL
  - Supports retrieval of known items - resource itself, locations for the item, reviews of the item, etc.
- OAI
  - Supports harvesting metadata for a union database

# Protocol relationships

---

- For general search by name, topic and keyword
  - use Z39.50 or ZING (SRU/SRW)
- For citations in hand
  - use OpenURL for pathways to item and related information
- For union catalog-type projects
  - use OAI to harvest catalog records into one file, then use Z39.50 or SRU/SRW to search the union file

# Retrieval standards summary

---

- Established and basic: Z39.50 (classic)
  - well developed
  - widely implemented
  - many tools available
  - crosses dissimilar systems
  - supports reference and cataloging
- Experimental
  - ZING (SRU/SRW), OAI, OpenURL

# Questions??

---

# Metadata formats

---

- Format components
- Format structures
- MARC
  - MARC 21
  - Unimarc
- CCF (UNESCO CDS/ISIS)
- Dublin Core
- EAD
- MARC in XML

# Format components

---

- Content - the data
  - for example, MARC 21 cataloging data
    - 100 1#\$a**Stevanovic, Vesna**
- Markup - data tagging
  - for example, MARC21 tags and subfield codes
    - **100 1#\$a**Stevanovic, Vesna
- Structure - overall record framework
  - for example MARC leader, directory, indicators, variable fields, etc.

# Record components

---

- Record content controlled by cataloging rules or other specifications
- Data tagging is format-specific
- Data structures can change without change to tagging or content rules

# Record structures

---

- ISO 2709
  - Used for all MARC formats
  - Compact
  - Enables efficient random access
  - Many systems built around 2709 record “look”
  - Many tools handle 2709 information
  - Most libraries use 2709



# Record structures

---

- Document Type Definition (DTD)
  - SGML and XML “formats”
  - Special DTD language
  - Well developed but complex
  - Some inflexibility
- Schema
  - XML “formats”
  - Uses XML to define formats
  - Simpler than DTDs

# MARC formats - briefly

---

## ■ Characteristics

- Layout: Introduction - Table of contents - Data
  - Parts called: Leader - Directory - Fields
- variable length records
- variable length fields
- cataloging rules “independent” (ISBD “bias”)
- coded data for specialized retrieval

# MARC record (machine view)

00684cam 2200193 a 4500

001000900000008004100009020001500050040001800065

050002100083082001500104100003700119245010600156

260006000262300003700322504003400359651004600393

655003100439700002000470

□00265359□000301r19991830enka b 000 1 eng □

□a019283567X □ □aDLC□cDLC-00□aPR5317□b.H4 1999-00□a823/.7□221□

1 □aScott, Walter,□cSir,□d1771-1832. □

14□aThe heart of Midlothian /□cSir Walter Scott ; edited with an introduction and notes by Claire Lamont. □ □aOxford [Oxfordshire] :□bOxford University Press,□c1999.□

□axxviii, 583 p. :□bill. ;□c20 cm.□ □aBibliography: p. [xxiv]-xxvi.□ 0□aScotland□xHistory□y18th century□vFiction.□

7□aHistorical fiction.□2gsafd□1 □aLamont, Claire.□□

# MARC

## ■ Sample cataloger display:

Leader 00684cam 2200193 a 4500  
008 00265359□000301r19991830enka b 000 1 eng □  
020 ##□a019283567X □  
050 00\$aPR5317\$b.H4 1999□  
082 00\$a823/.7□221□  
100 1#□aScott, Walter,□cSir,□d1771-1832. □  
245 14□aThe heart of Midlothian /□cSir Walter Scott ; edited  
with an introduction and notes by Claire Lamont. □  
260 ##□aOxford [Oxfordshire] :□bOxford University  
Press,□c1999.□  
300 ##□axxviii, 583 p. :□bill. ;□c20 cm.□  
504 ##□aBibliography: p. [xxiv]-xxvi.□  
651 #0□aScotland□xHistory□y18th century□vFiction.□  
655 #7□aHistorical fiction.□2gsafd□  
700 1#□aLamont, Claire.□□

# MARC

## ■ Sample user display:

Author: Scott, Walter, Sir, 1771-1832.

Title: The heart of Midlothian / Sir Walter Scott ; edited with an introduction and notes by Claire Lamont.

Publisher: Oxford [Oxfordshire] : Oxford University Press, 1999.

Extent: xxviii, 583 p. : ill. ; 20 cm.

Note: Bibliography: p. [xxiv]-xxvi.

Language: English

Subject: Scotland--History--18th century--Fiction.

Genre: Historical fiction.

Name: Lamont, Claire.

ISBN: 019283567X

Call number: PR5317\$b.H4 1999

# MARC potential

---

- Local online catalog
- Supports development of of union catalogs
- Exchange of bibliographic data
- Supports shared record building
- Promote consistency across systems

# MARC 21

---

- MARC 21
  - Content - supports any cataloging rules
  - Tagging - MARC 21 specific tagging
  - Structure - ISO 2709 (classic)
    - also XML version
- Users
  - USA, Canada, Latin America, etc.
  - UK, Spain, Sweden, Finland, Norway, Poland, Switzerland, Russian State Library, etc.
  - Australia, New Zealand, Hong Kong, China, Philippines, Thailand, Vietnam, etc.

# MARC 21 maintenance

---

- Maintenance process
  - Internationally open change process and participation - Web postings, listserv
  - Expert input
  - Library of Congress commitment to support staff and documentation



# MARC 21 documentation

---

- MARC 21 Full
  - Detailed explanations, examples, usage guidelines
  - Currently available in print and CDROM
- MARC 21 Concise
  - Brief explanations, many examples
  - Available on Web and in print
- MARC 21 LITE
  - Subset of MARC 21 Full
  - Core data elements covering all forms of material
  - Available on Web only

# MARC 21 web site

■ [www.loc.gov/marc](http://www.loc.gov/marc)

■ Documentation

- Concise and Lite format versions
- All code lists: language, country, relators, etc.
- Character sets specifications
- *Understanding MARC Bibliographic*

■ Change proposals

■ MARC in XML

■ Lists of MARC record services, systems, tools

# Unimarc

---

- Same structure as MARC 21 (ISO 2709)
- Different tagging, subfield codes
  - MARC 21 - subfielding primarily for access
  - Unimarc - subfielding for punctuation
  - Different parsing of data
- Cobiss format is similar to Unimarc

# CCF

- Common Communications Format
  - Content - supports any cataloging rules
  - Tagging - CCF tags
  - Structure - ISO 2709
- CDS/ISIS native format is derived from CCF
  - Simple format, but library-oriented
- Used in many developing countries

# Dublin Core

---

- Characteristics
  - Content: no rules
  - Tagging: DC tags
  - Structure: XML
- Targeted for use in electronic document headers
  - author input?
- Enables cross domain searching (museums, archives, libraries, other cultural institutions)
  - Different library data formats have more detail - can reduce to DC
  - Other domains have different detail and formats - can reduce to DC

# DC Elements

---

- 15 Data Elements

- Title
- Creator
- Contributor
- Publisher
- Subject
- Description
- Date
- Type
- Format
- Identifier
- Source
- Language
- Relation
- Coverage
- Rights

# Dublin Core

---

- What is “right” with DC:
  - The element set is small and basic
  - The set is core enough to be understood by different communities and by untrained staff
  - There are tools available for reducing sets and manipulating data
  - There is material that may not be described in more detail

# Dublin Core

---

- Problems with DC:
  - So basic users tend to make unique extensions
  - Lack of content rules means data is heterogeneous
  - More detail is needed within a specific user community



# EAD

---

- Encoded Archival Description - EAD
  - Content: archival finding aids
  - Tagging: EAD tags
  - Structure: SGML
- Finding aids - documents that describe an archival collection, usually largely composed of lists of "box" contents, file names, etc.
- SGML DTD for making searchable electronic versions of finding aids
- Each EAD finding aid document has a MARC record.

# EAD

---

- Increasing international use for archival control
- Library of Congress maintains the official standard web site:
  - [www.loc.gov/ead](http://www.loc.gov/ead)

# Format relationships

---

- MARC
  - Supports library interoperability
- Dublin Core
  - Supports cross domain access
- EAD
  - Supports access to contents of archival collections

# Questions?

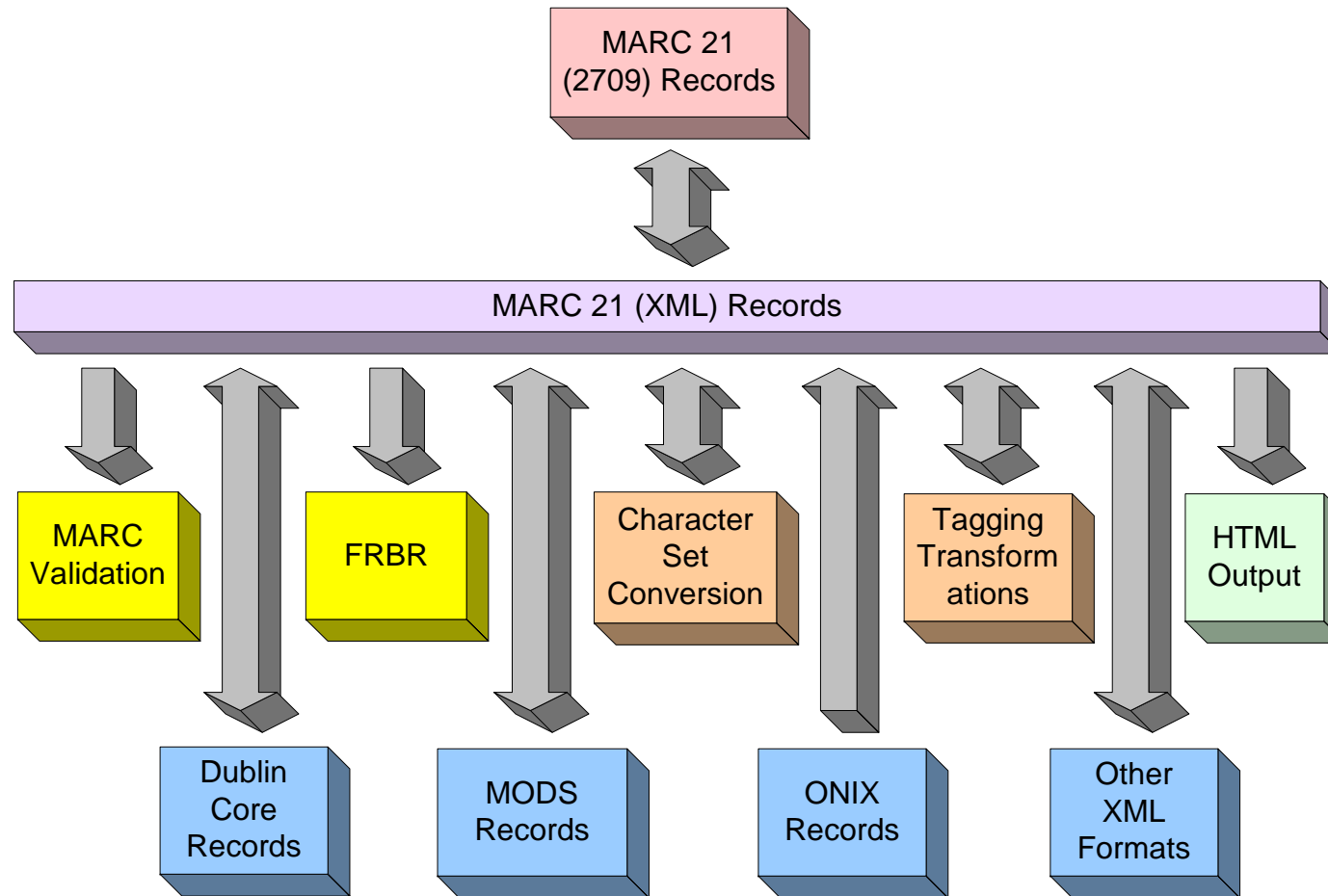
---

# MARC format **future**s

---

- MARC 21 in an XML structure
  - Content: Same as MARC 21
  - Tagging: MARC 21
  - Structure: XML
- Coordinated set of tools
- Flexible transition options
- Continued commitment to *standards* and change by *evolution*

# MARC 21 Tool Kit



# MARC 21 in XML

---

- MARCXML record
  - XML exact equivalent of MARC 21 (2709) record – base for further transformations
  - Lossless/roundtrip conversion to/from MARC 21 record
  - Simple flexible schema, no need to change when MARC 21 changes
  - Presentations using XML stylesheets
  - Converters provided by LC, open source
  - Use with Z39.50, ZING, OAI, METS

```
<?xml version="1.0" encoding="UTF-8"?>
  <collection xmlns="http://www.loc.gov/MARC21/slim">
    <record>
      <leader>00000cam 2200000 4500</leader>
      <controlfield tag="001">69017649</controlfield>
      <controlfield tag="008">700409r19681948nyu          001 1 eng
        </controlfield>
      <datafield tag="040" ind1=" " ind2=" ">
        <subfield code="a">DLC</subfield>
        <subfield code="c">DLC</subfield>
      </datafield>
      <datafield tag="050" ind1="0" ind2="0">
        <subfield code="a">PZ3.S43</subfield>
        <subfield code="b">H43</subfield>
        <subfield code="a">PR5317</subfield>
      </datafield>
      <datafield tag="082" ind1="0" ind2="0">
        <subfield code="a">823/.7</subfield>
      </datafield>
      <datafield tag="100" ind1="1" ind2=" ">
        <subfield code="a">Scott, Walter,</subfield>
        <subfield code="c">Sir,</subfield>
        <subfield code="d">1771-1832.</subfield>
      </datafield>
      <datafield tag="245" ind1="1" ind2="4">
        <subfield code="a">The heart of Midlothian.</subfield>
        <subfield code="c">Introd. by David Daiches.</subfield>
      </datafield>
```



# MARC 21 to DC in XML

---

- Dublin Core transformation
  - DC application target for cross domain applications, metadata in document headers
  - Transformation software maintained by LC, open source
  - LC already maintains DC $\leftrightarrow$ MARC 21 mapping

```
<?xml version="1.0"?>
<rdf:Description
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#">
  <dc:title>The heart of Midlothian /</dc:title>
  <dc:creator>Scott, Walter, Sir, 1771-1832.</dc:creator>
  <dc:creator>Lamont, Claire.</dc:creator>
  <dc:type>text</dc:type>
  <dc:type>Historical fiction.</dc:type>
  <dc:publisher>Oxford [Oxfordshire] : Oxford University
Press,</dc:publisher>
  <dc:date>1999.</dc:date>
  <dc:language>eng</dc:language>
  <dc:description>Bibliography: p. [xxiv]-
xxvi.</dc:description>
</rdf:Description>
```

# MARC 21 in XML companion

---

- MODS - Metadata Object Description Schema – a MARC 21 companion
  - Simpler element set than full MARC, but MARC semantics
  - Richer element set than DC
  - More compatible with MARC than others
  - “Friendly” schema and tagging, no coded values
  - Special accommodation of electronic resources

# MODS

---

- Potential uses of MODS
  - Describing electronic resources
  - Technician input
  - Incorporation with XML resources
  - Use with ZING, OAI, METS

```
<?xml version="1.0"?>
<mods xmlns="http://www.loc.gov/mods/"
  <titleInfo>
    <nonSort>The </nonSort>
    <title>heart of Midlothian </title>
  </titleInfo>
  <name type="personal">
    <namePart>Scott, Walter, Sir,</namePart>
    <namePart type="date">1771-1832.</namePart>
    <role>creator</role>
  </name>
  <name type="personal">
    <namePart>Lamont, Claire.</namePart>
  </name>
  <typeOfResource>text</typeOfResource>
  <genre authority="marc">bibliography</genre>
  <genre authority="gsafd">Historical fiction.</genre>
  <publicationInfo>
    <placeCode authority="marc">enk</placeCode>
    <place>Oxford [Oxfordshire] ;</place>
    <publisher>Oxford University Press,</publisher>
    <dateIssued>1999.</dateIssued>
    <dateIssued encoding="marc">1999</dateIssued>
    <issuance>monographic</issuance>
  </publicationInfo>
  <language authority="iso639-2b">eng</language>
```

# MARC 21 in XML - other tools

---

- Transformations
  - Unicode standard for XML: UTF-8
  - ONIX format
  - Various presentations - HTML outputs
  - tagging language translations
- Tools
  - Validation
  - FRBR experimentation tool

# More information

---

- [www.loc.gov/marcxml](http://www.loc.gov/marcxml)
- [www.loc.gov/mods](http://www.loc.gov/mods)
- [www.loc.gov/marc](http://www.loc.gov/marc)

# Digital standards

---

- Digital resources “era”
  - More types of metadata needed
  - Bibliographic data ***and*** resources in electronic form
  - Internet and web technologies = immediate access and communications + user interfaces and linking

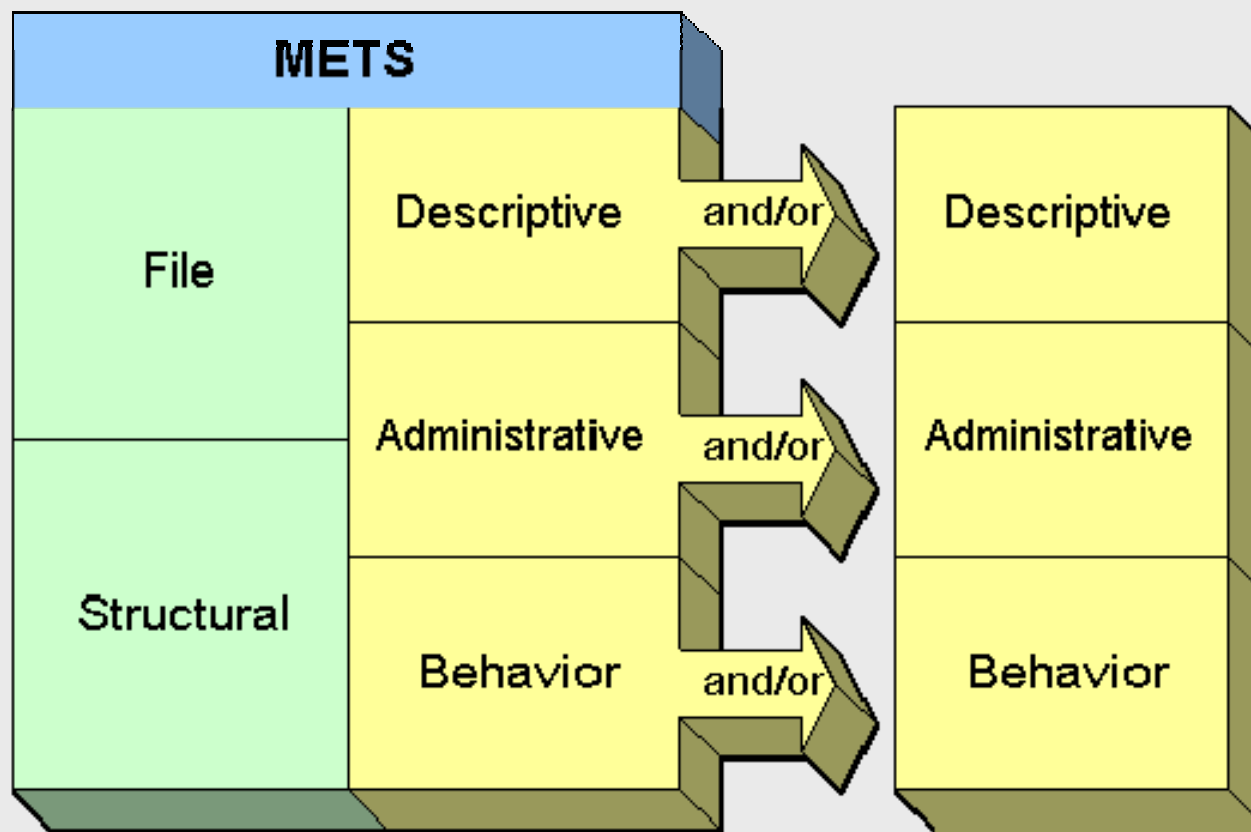


# One direction

---

- METS - Metadata Encoding and Transmission Standard
  - XML “package”
    - Metadata associated with electronic resource
    - Structure of electronic resource
    - Name and location of files for electronic resource
- [www.loc.gov/mets](http://www.loc.gov/mets)

# METS Schema



# METS

---

- Descriptive Metadata
  - MARCXML for MARC 21
  - DC
  - MODS
  - Prefer XML but can use other structures
  - Internal to METS or pointer

# METS

- Administrative metadata
  - Technical
    - Still image - NISO schema
    - Text - draft
    - Moving image - draft
    - Audio - draft
  - Provenance and rights
  - Derivatives
  - Source and transformations

# Summing up

---

- I hope I have some information on some key standards
  - Where they are
  - What they are for
  - What is implemented
  - What is new
  - What is evolving