

# 3 Combining RDF and XML Schemas to Enhance Metadata Interoperability

Lecture Slides  
Steffen Staab

With grateful acknowledgements to  
Jane Hunter, University of Queensland, Australia

## Why we need both

- The Problem
- One Approach
- A Simple Example
- RDF <-> XML Schema Interface mechanisms
- Mapping using XSLT
- Metanet - Ontology of metadata terms
- Linking MetaNet to XSLT to perform mappings

Slide 2

## Metadata Standards

- Bibliographic - MARC, IFLA, DublinCore
- Archival – ISAD
- Museum – CIDOC CRM
- Educational - IMS, GEM, EDNA
- Government - GILS, AGLS
- Geospatial – FGDC
- News – NewsML
- Image - VRA Core
- Audio - ID3
- Audiovisual - MPEG-7, SMPTE, EBU, FIAF
- Copyright - <indecs>, XrML, ODRL

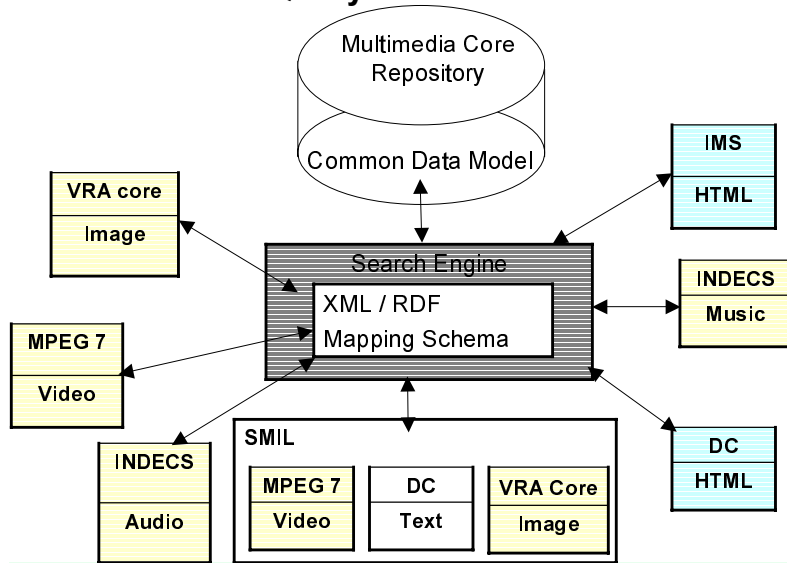
Slide 3

## New Requirements

- MPEG-21, TV-Anytime, OAI, ISLE
- Extend existing domain-specific metadata models/schemas
- Refine existing metadata models/schemas
- Mix elements from different metadata domains
- Define local application-specific encodings of existing metadata standards
- Maintain Interoperability

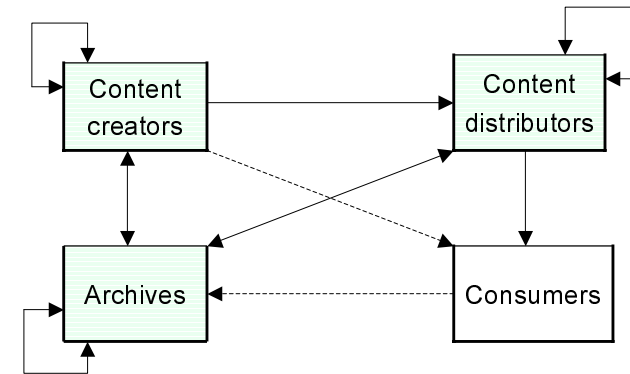
Slide 4

## Problem 1 – Query Mediation



Slide 5

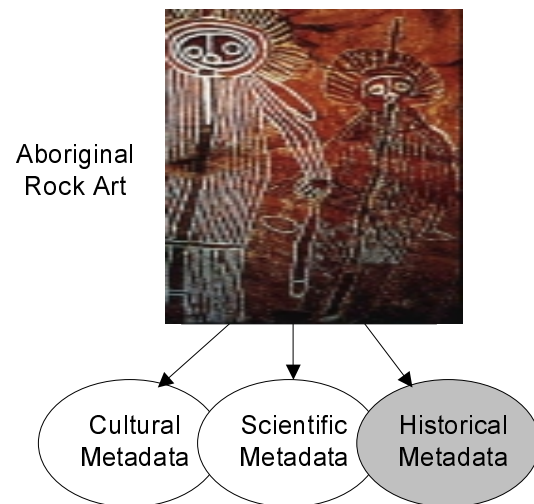
## Problem 2 – Harmonization



MPEG-21, TV-Anytime

Slide 6

## Problem 3 – Transportation



Slide 7

## Current Situation

- DC – RDF Schema, Draft XML Schema
- MPEG-7, MPEG-21 – XML Schema
- TV-Anytime – XML Schema
- OAI – XML Schema
- Prism Project – RDF Schema
- SCHEMAS Project – RDF Schema
- CIDOC CRM – RDF Schema
- Others – don't know? But need to define their own application-specific encodings of existing metadata standards

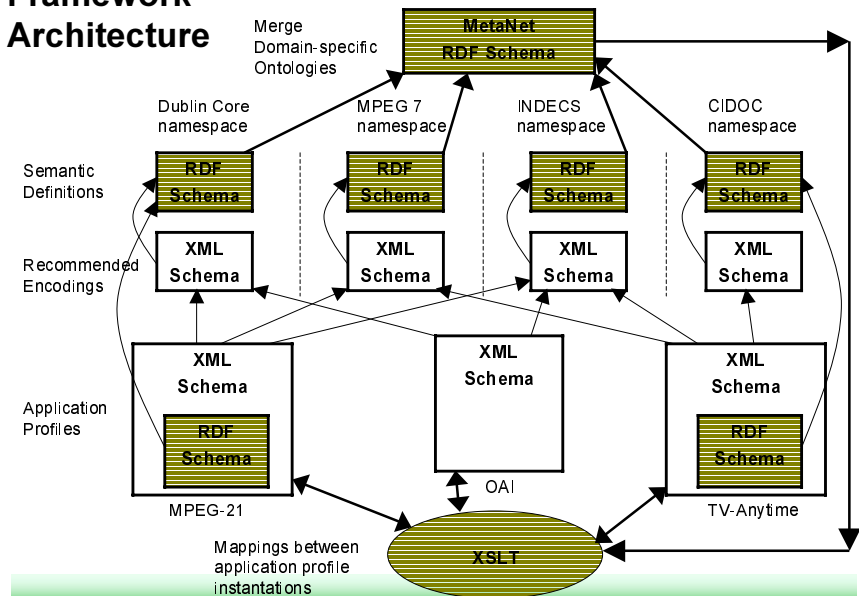
Slide 8

## Desired

- Metadata domains register authorized namespaces containing both an XML and an RDF Schema
- Separation of semantics from local structural, datotyping and cardinality encoding constraints
- Local “application profiles” in XML Schema
  - Mix and match complementary metadata elements from different domain-specific namespaces
  - Apply local application-specific constraints

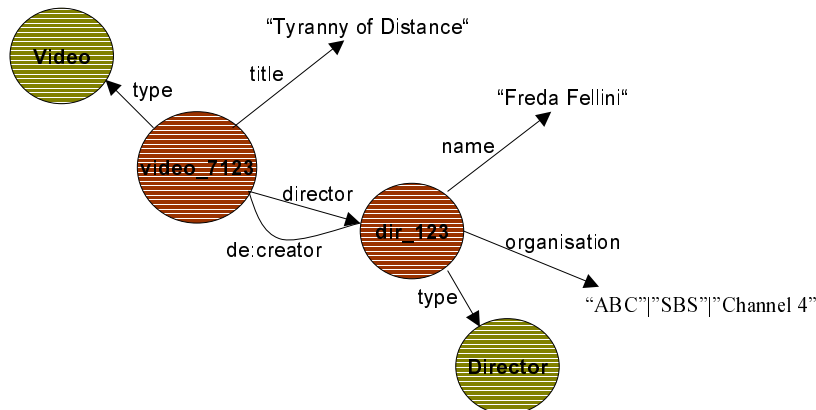
Slide 9

## Framework Architecture



Slide 10

## Simple Video Example



Slide 11

```
<?xml version='1.0'?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1"/>

<rdfs:Class rdf:ID="Video" >
  <rdfs:label>Video</rdfs:label>
  <rdfs:comment>The class of videos</rdfs:comment>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
</rdfs:Class>

<rdf:Property rdf:ID="title" >
  <rdfs:label>Title</rdfs:label>
  <rdfs:comment>The name given to the resource</rdfs:comment>
  <rdfs:domain rdf:resource="#Video"/>
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal"/>
</rdf:Property>
```

Slide 12

```
<rdfs:Property rdf:ID="director" >
  <rdfs:label>Director</rdfs:label>
  <rdfs:comment>A person responsible for directing a video</rdfs:comment>
  <rdfs:subPropertyOf
    rdf:resource="http://purl.org/dc/elements/1.1/dcmes.rdf#creator"/>
  <rdfs:domain rdf:resource="#Video"/>
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-
    schema#Director"/>
</rdfs:Property>

<rdfs:Class rdf:ID="Director"/>
<rdf:Property rdf:ID="organisation" >
  <rdfs:label>Organisation</rdfs:label>
  <rdfs:comment>"The director's affiliation."/>
  <rdfs:domain rdf:resource="#Director"/>
  <rdfs:range rdf:resource="#OrgName"/>
</rdf:Property>
<rdfs:Class rdf:ID="OrgName"/>
<OrgName rdf:ID="ABC"/>
<OrgName rdf:ID="Channel 4"/>
<OrgName rdf:ID="SBS"/>
</rdf:RDF>
```

Slide 13

## RDF Schema

- RDF Data model
- Richer semantic descriptions
  - Class/property relationships
  - Type hierarchies – subClasses & subProperties
  - Human readable descriptions (label,comment)
- Very limited structural, cardinality, datatyping constraints at the encoding level

Slide 14

```
<schema xmlns="http://www.w3.org/1999/XMLSchema"
  targetNamespace="http://www.dstc.edu.au/"
  xmlns:dstc="http://www.dstc.edu.au/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

  <import namespace=" http://purl.org/dc/elements/1.1/">
  <element name="Video" >
    <annotation>
      <documentation>The class of videos</documentation>
    </annotation>
    <sequence>
      <element name="title" type="string" minOccurs="1" maxOccurs="1"/>
      <element name="director" type="director" maxOccurs="4"/>
    </sequence>
    <attribute name="id" type="uriReference"/>
  </element>
```

Slide 15

```
<complexType name="director">
  <extension base="dc:creator" >
    <element name="organisation" type="OrgNames"/>
  </extension>
</complexType>
<simpleType name="OrgNames" >
  <restriction base="string" >
    <enumeration value="ABC"/>
    <enumeration value="Channel 4"/>
    <enumeration value="SBS"/>
  </restriction>
</simpleType> </schema>
```

Slide 16

## XML Schema

- Rich support for structural, cardinality and datatype encoding constraints
- Poor support for semantics  
<annotation> <documentation>

Slide 17

## Link from application-specific XML Schemas to a single RDF Schema

Semantic Definitions in RDF Schema

DCMES.rdf - RDF Schemata for Dublin Core

```
<rdf:RDF>
  <rdf:Property ID="title">
    <rdf:label="Title"/>
    <rdf:comment="The name given the resource"/>
  </rdf:Property>
  <rdf:Property ID="creator">
    <rdf:label="Creator"/>
    <rdf:comment="The entity primarily responsible for making the content of the resource"/>
  </rdf:Property>
  .....
</rdf:RDF>
```

Application-specific Encodings using XML Schemata

```
<schema...>
  <simpleType name="title">
    <restriction base="string"/>
  </simpleType>
  <simpleType name="creator">
    <restriction base="string">
      <enumeration value="Brickley"/>
      <enumeration value="Swick"/>
      <enumeration value="Weibel"/>
    </restriction>
  </simpleType>
  .....
</schema>
```

Schema1.xsd

```
<schema...>
  <simpleType name="title">
    <restriction base="anyType"/>
  </simpleType>
  <simpleType name="creator">
    <restriction base="anyType"/>
  </simpleType>
  .....
  <element name="title" type="title"
    maxOccurs="2"/>
  .....
</schema>
```

Schema2.xsd

Slide 18

## Annotated XML Schema

```
<schema xmlns= http://www.w3.org/1999/XMLSchema"
  targetNamespace= http://purl.org/dc/elements/1.1/"
  xmlns:dc= http://purl.org/dc/elements/1.1/"
  xmlns:rdf= http://www.w3.org/1999/02/22-rdf-syntax-ns# >
  xmlns:rdfs= http://www.w3.org/2000/01/rdf-schema# >
<annotation>
  <documentation>
    Draft XML Schema for the Dublin Core Element Set, V 1.1
  </documentation>
</annotation>
<simpleType name= title >
  <annotation>
    <appinfo>
      <rdf:Property ID= title >
        <rdf:label= Title />
        <rdf:comment= The name given to the resource. />
      </rdf:Property>
    </appinfo>
  </annotation>
  <restriction base= string"/>
</simpleType>
.....
</schema>
```

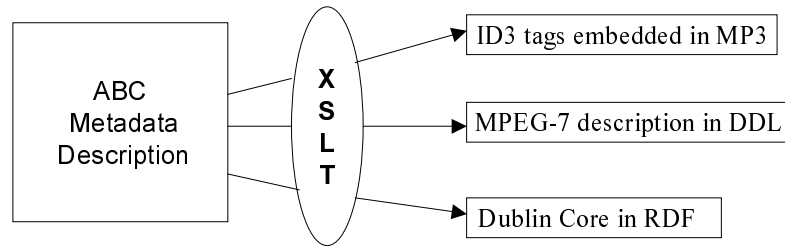
Slide 19

## Linked Schemas

```
<schema xmlns= http://www.w3.org/1999/XMLSchema"
  targetNamespace= http://purl.org/dc/elements/1.1/"
  xmlns:dc= http://purl.org/dc/elements/1.1/"
  xmlns:xx="http://www.example.org/XMLRDFSchemaBridge">
<annotation>
  <documentation>
    Draft XML Schema for the Dublin Core Element Set, V 1.1
  </documentation>
</annotation>
<simpleType name= title xx:semantics="http://purl.org/dc/elements/1.1/dcmes.rdf#title">
  <restriction base= string"/>
</simpleType>
<simpleType name= creator xx:semantics=
  "http://purl.org/dc/elements/1.1/dcmes.rdf#creator">
  <restriction base= string"/>
</simpleType>
.....
</schema>
```

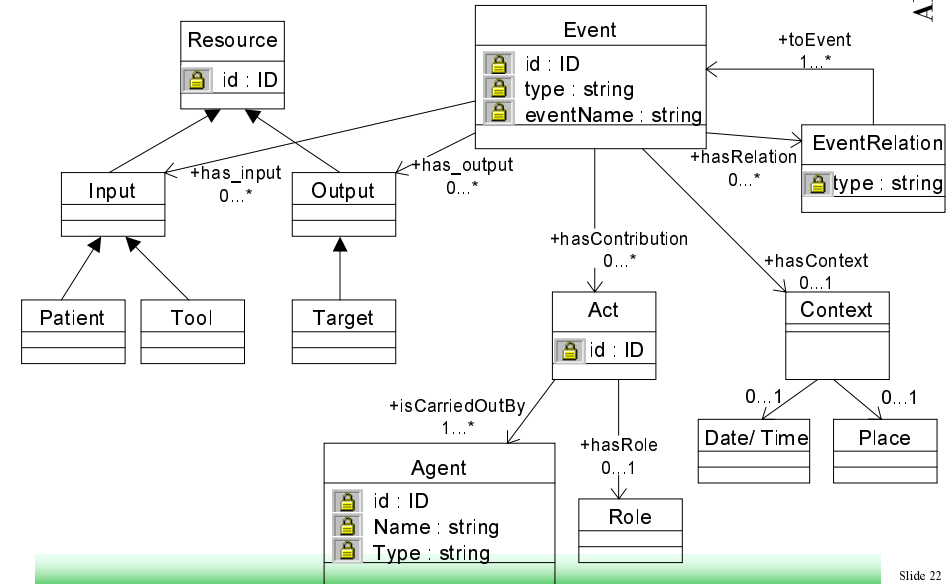
Slide 20

## Metadata Mappings



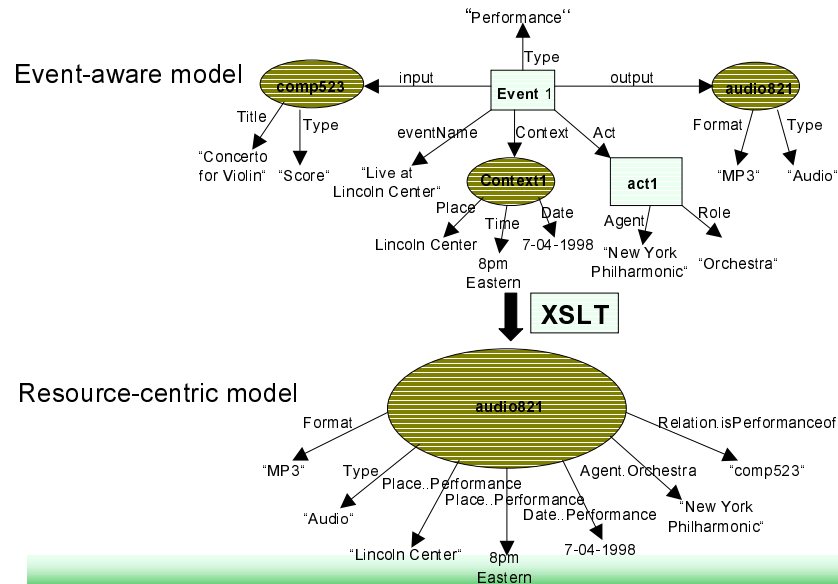
Slide 21

## Harmony – ABC Model



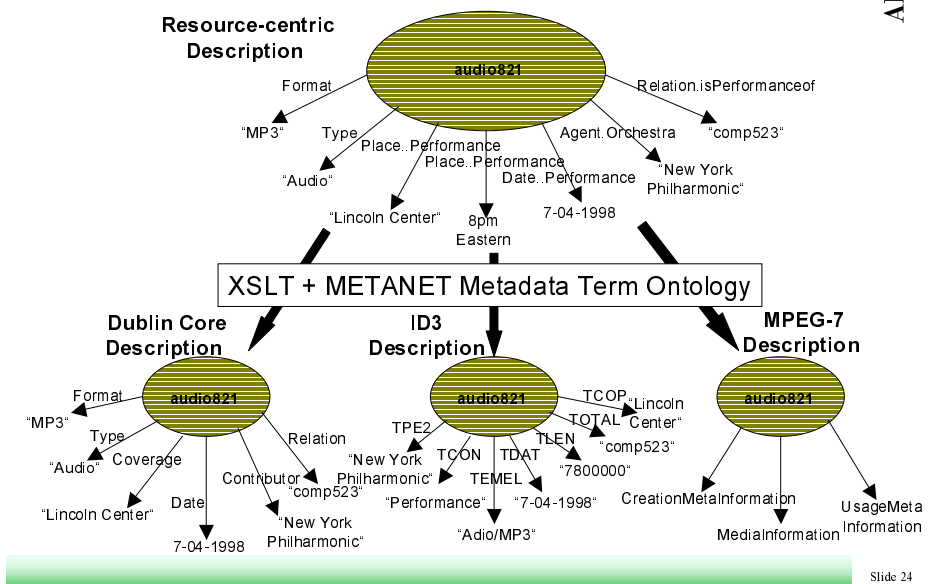
Slide 22

## Step 1 – Structural Mapping



Slide 23

## Step 2 – Semantic Mapping



Slide 24

## Evaluation of XSLT

- Extensible Style Language Transformation
- Works well for structural and syntactic mapping between metadata descriptions
- Semantic mappings need to be hardcoded
- Unsuitable for loosely constrained or variable input
- Flexible semantic mappings require additional knowledge:
  - Mapping Matrix or
  - Metadata Term Ontology - MetaNet

Slide 25

## Metadata Mapping Matrix

ABC	DC	ID3	MPEG-7
Resource/Title	Title	TIT2	CreationMetaInformation/Creation/Title/TitleText (@TitleType="original")
Event/Act/Agent	Creator Publisher Contributor	TPE1 TPUB IPLS	CreationMetaInformation/Creation/Creator UsageMetaInformation/Publication/Publisher CreationMetaInformation/Creation/Creator
Resource/Subject	Subject	TIT1	CreationMetaInformation/Creation/Classification/PackagedType
Resource/Description	Description	TIT3	
Event/Context/Date	Date.Creation Date.Publication Date.Recording	TRDA	CreationMetaInformation/Creation/CreationDate UsageMetaInformation/Publication/ PublicationDate
Event/Context/Place	Coverage.Place		
Resource/Type	Type	TCO	CreationMetaInformation/Classification/Genre
Resource/Format	Format Format.length Format.size	TFLT TLEN TSIZ	MediaInformation/MediaProfile/MediaFormat/ FileFormat
Resource/Identifier	Identifier	UFID	MediaInformation/MediaIdentification/Identifier
Event/Input	Source	TOAL	

Slide 26

## MetaNet

- Metadata term Ontology
- Expressed in RDF Schema
- Defines equivalence and hierarchical (NT and BT) relationships between metadata terms from different domains
- Generate by merging domain-specific RDF Schemas

Slide 27

## MetaNet Browse Interface

Results of Search for metadata term: <i>resource</i>			
Core Term	Synonyms/Equivalent Terms	Hyponyms/Narrower Terms	Hypo-hyponyms/Narrowest Terms:
resource	object, entity, thing, product, manifestation, creation, expression, work, item, stuff, artefact	physical_object, man-made_object, iconographic_object, program, programme, series, episode, clip, album, soundtrack, track, audioclip, videoclip, document, artifact, film, lecture, novel, play, poem, essay, biography, symphony, concerto, sonata, map, drawing, painting, etching, lithograph, photograph, advertisement, script	-

Slide 28

## MetaNet Search Interface

Results of Search for metadata term: *author*

Core Term

agent

Synonyms/ Equivalent Terms:

actor, contributor, player, doer, worker, performer

Hyponyms/ Narrower Terms:

*author*, writer, wordsmith

Hypo-hyponyms/ Narrowest Terms:

novelist, playwright, dramatist, essayist, poet,  
scriptwriter, copywriter, journalist, columnist

Slide 29

```
<?xml version="1.0"?>
<rdf:RDF xml:lang="en"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

  <rdfs:Class rdf:ID="Agent" >
    <rdfs:comment xml:lang="en" >The resources which contribute to or act in an event.
    Typically
    agents are people, groups of people, organisations or instruments.</rdfs:comment>
    <rdfs:label xml:lang="en" >Actor</rdfs:label>
    <rdfs:label xml:lang="en" >Contributor</rdfs:label>
    <rdfs:label xml:lang="en" >Player</rdfs:label>
    <rdfs:label xml:lang="en" >Doer</rdfs:label>
    <rdfs:label xml:lang="en" >Worker</rdfs:label>
    <rdfs:label xml:lang="en" >Performer</rdfs:label>
    <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
  </rdfs:Class>
```

Slide 30

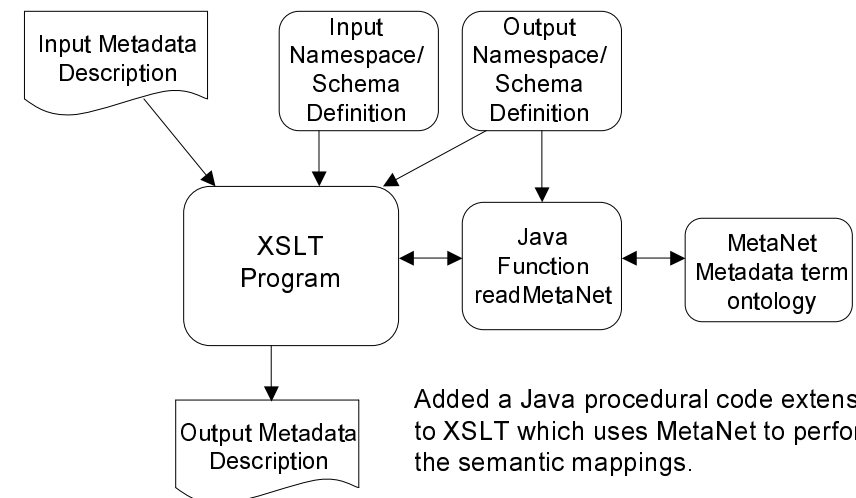
```
<rdfs:Class rdf:ID="Author" >
  <rdfs:label xml:lang="en" >Writer</rdfs:label>
  <rdfs:label xml:lang="en" >Wordsmith</rdfs:label>
  <rdfs:subClassOf rdf:resource="#Agent"/>
</rdfs:Class>

<rdfs:Class rdf:ID="Journalist" >
  <rdfs:label xml:lang="en" >Columnist</rdfs:label>
  <rdfs:label xml:lang="en" >Reporter</rdfs:label>
  <rdfs:subClassOf rdf:resource="#Author"/>
</rdfs:Class>

</rdf:RDF>
```

Slide 31

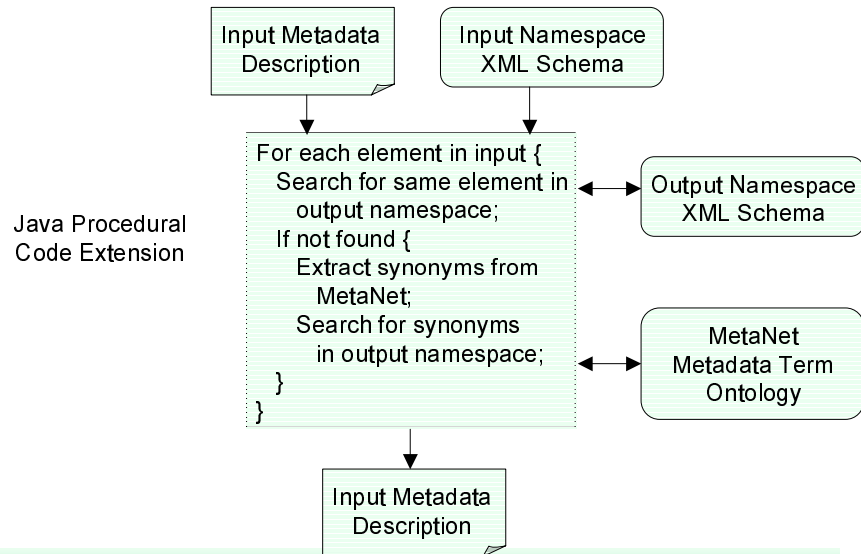
## Linking MetaNet to XSLT



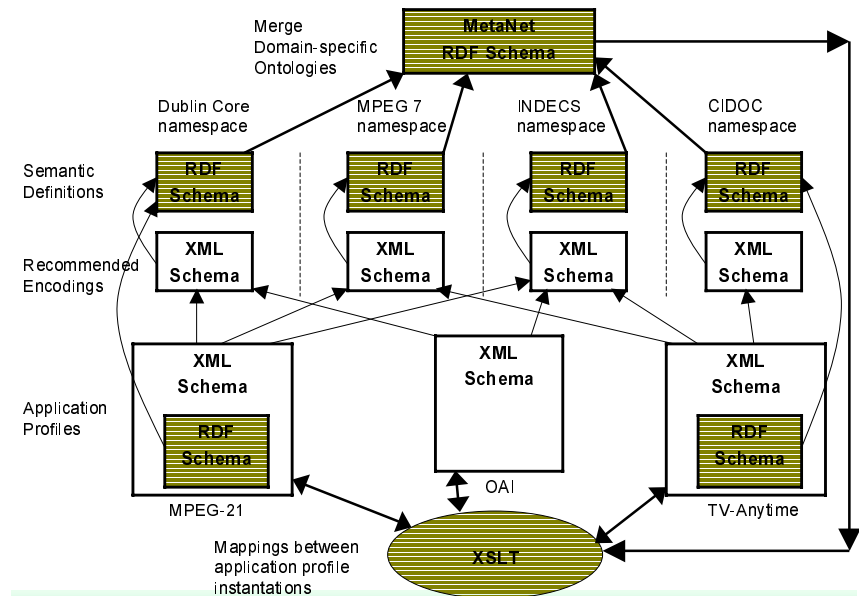
Slide 32



## Semantic Mapping Using MetaNet



Slide 33



Slide 34

## Future Work

- Ontology harmonization – CIDOC CRM
- Schematron and RDF Schema
- Express MetaNet in DAML/OIL
- More schema languages coming out – RELAX, TREX – how do these fit in

Slide 35