


# Metadata in the Framework of Geospatial Information Interoperability

Stefano Nativi

Unidata/UCAR  
Boulder, January 2002



# Metadata

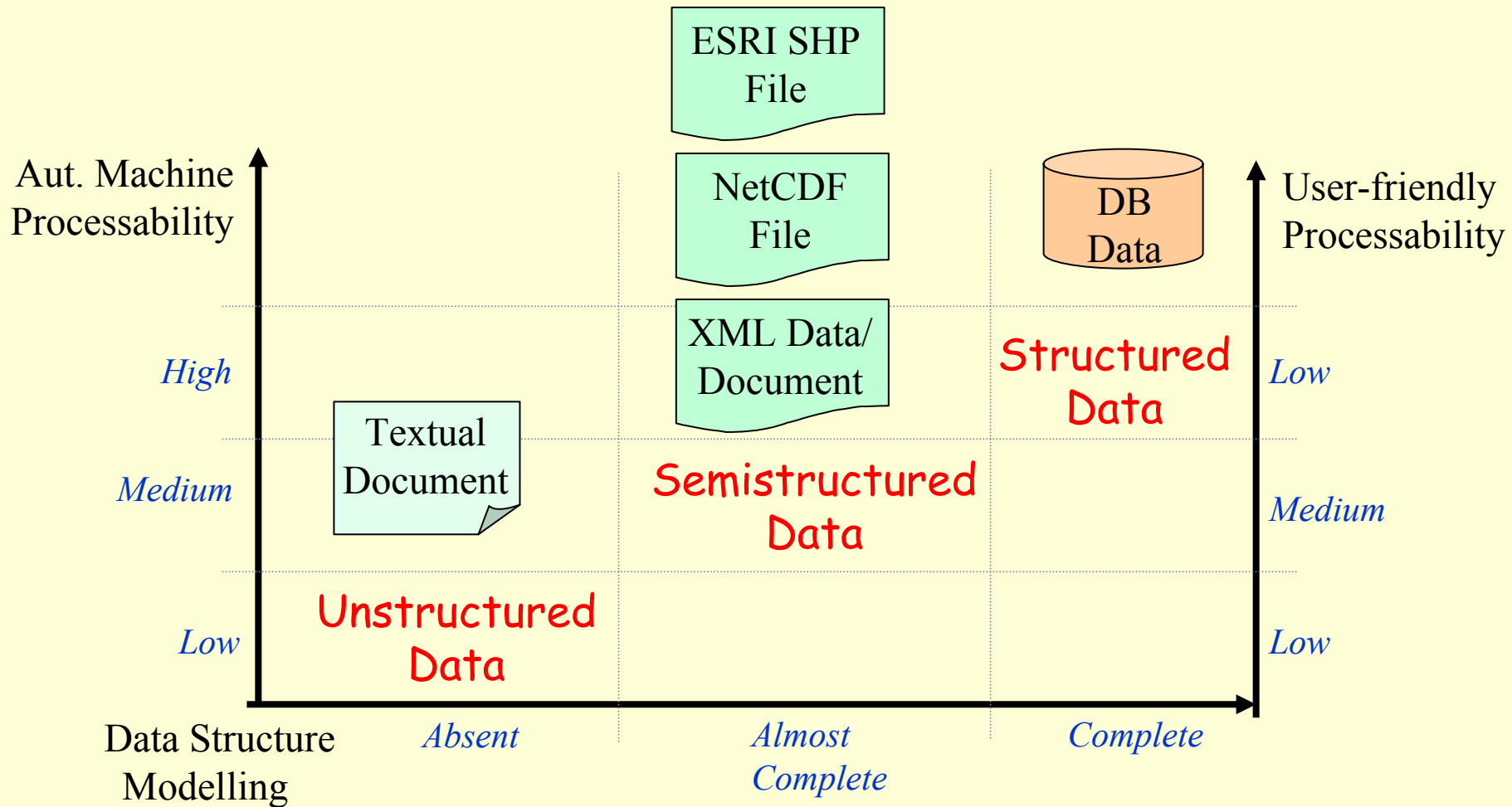
- Metadata is data about characteristics of data  
(General)
  - Metadata is data documenting other data, and is not contained in the data itself (DL, Data Mining)
  - Data describing other data or system elements to support their documentation, reusability and interoperation (Systems Interoperability)
- 



# Metadata Purpose

- Documentation → User-friendly Processability
    - User Browsing and Queries
    - Information Retrieval
    - Information Discovery
  
  - Interoperation & Reusability → Automatically Machine Processability
    - Heterogeneous Data Integration
    - Heterogeneous Semantic Integration
- 

# Data Structure Modelling



■ Metadata supplemented  $\Rightarrow$  Processability level

# Metadata Supplementation

Metadata Purpose	Examples of <u>Metadata Model</u>
Documentation	DC-ED, ADL, .... ISO 19115, OpenGIS Topic 11, FGDC specs., CEN/ENV 12657, .... DC, ...
Interoperation & Documentation	IEEE/LTSC/LOM, ARIADNE, IMS, TEI, ... ISO TC 211 models, OpenGIS models, ESML model, ...
n	DAML+OIL, MOF, RDF, DWC, IEEE/BIDM, ...

▽ Metadata Model Heterogeneity

# Metadata Model Heterogeneity

## Information Community needs

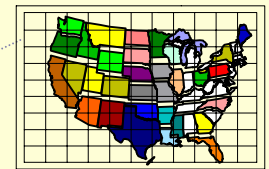
### Data facet

- Concepts (jargon)
- Semantic Abstractions (data granularity)

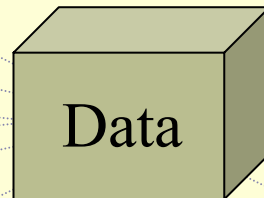
Digital Library Community



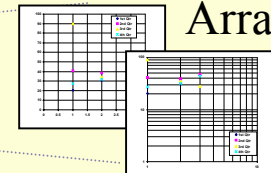
Set of  
Gridded features



NetCDF File



Counts  
Array



Meteo  
Community

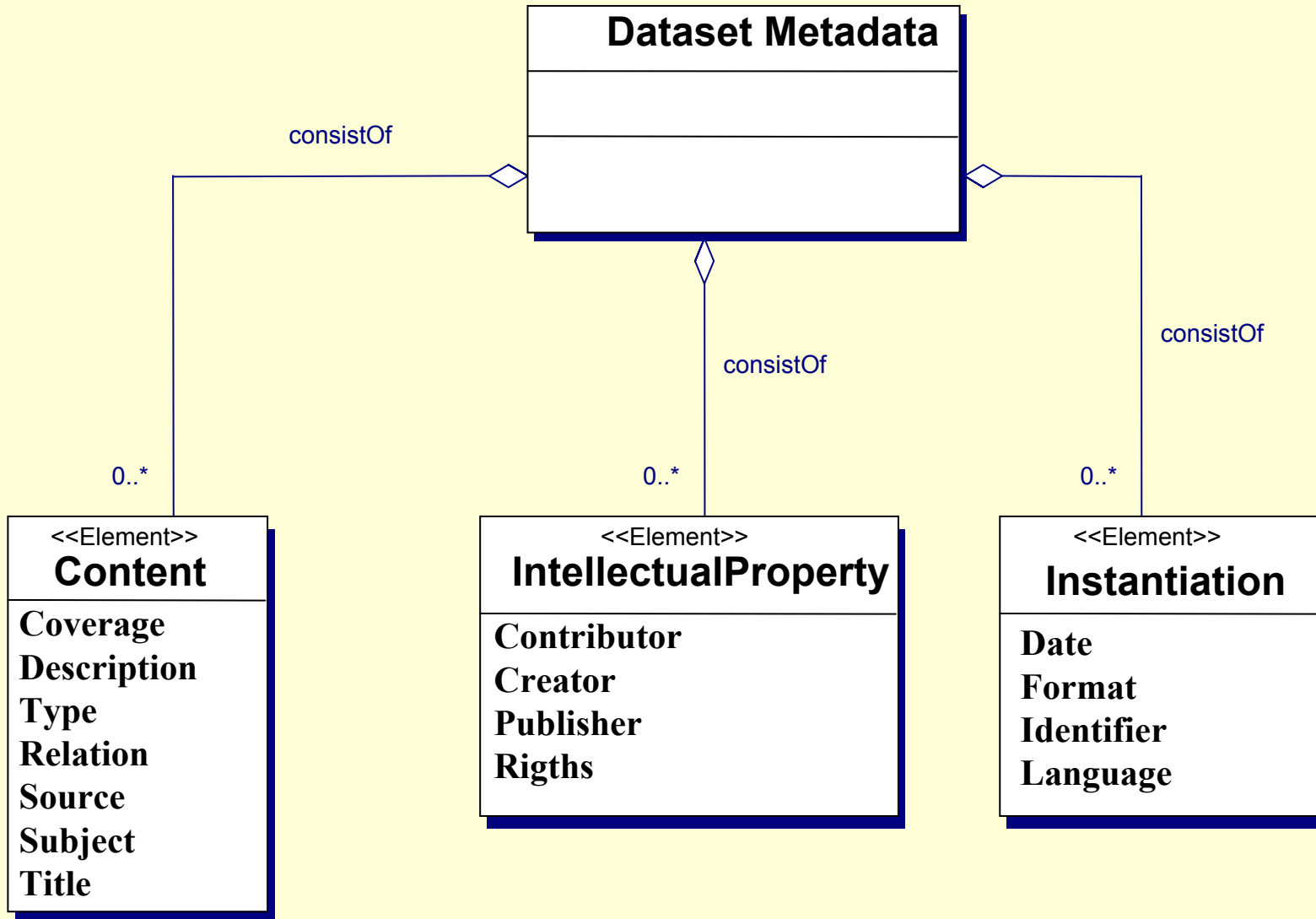


Geography Community

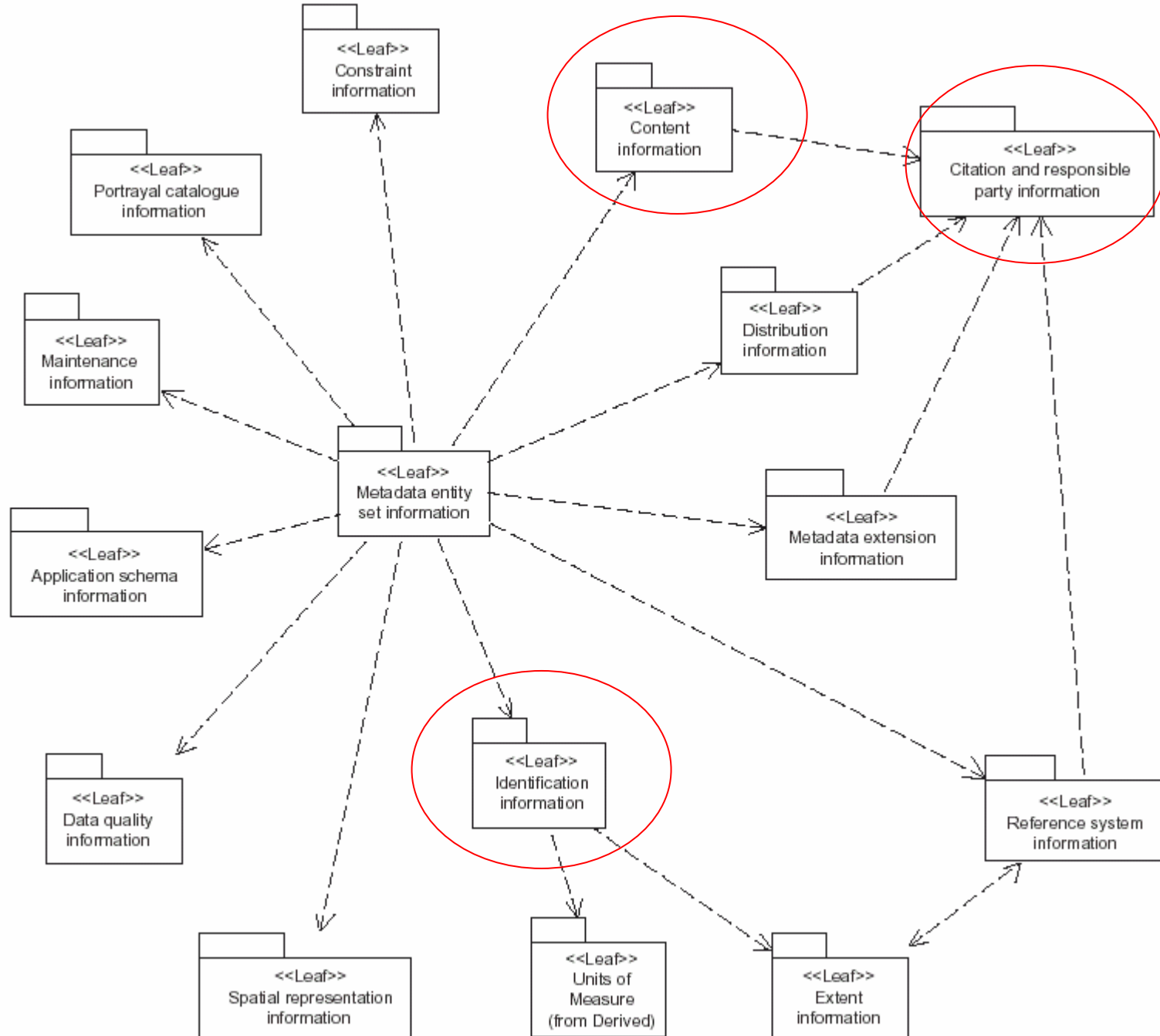
Map Documents



# Dublin Core Model



# ISO 19115 Model



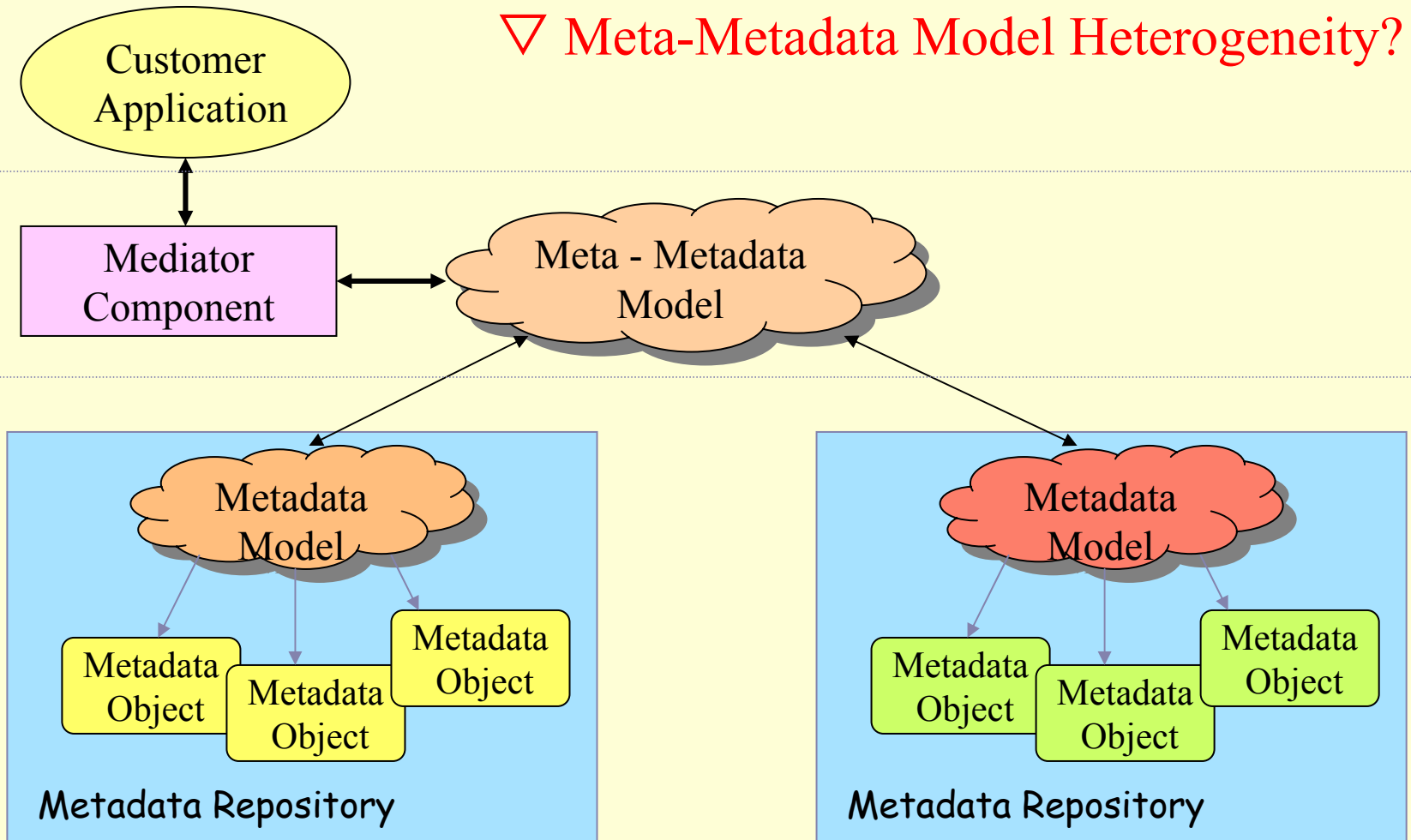
# Metadata Model Heterogeneity

■ IEEE BIDM

■ ISO/IEC 11179

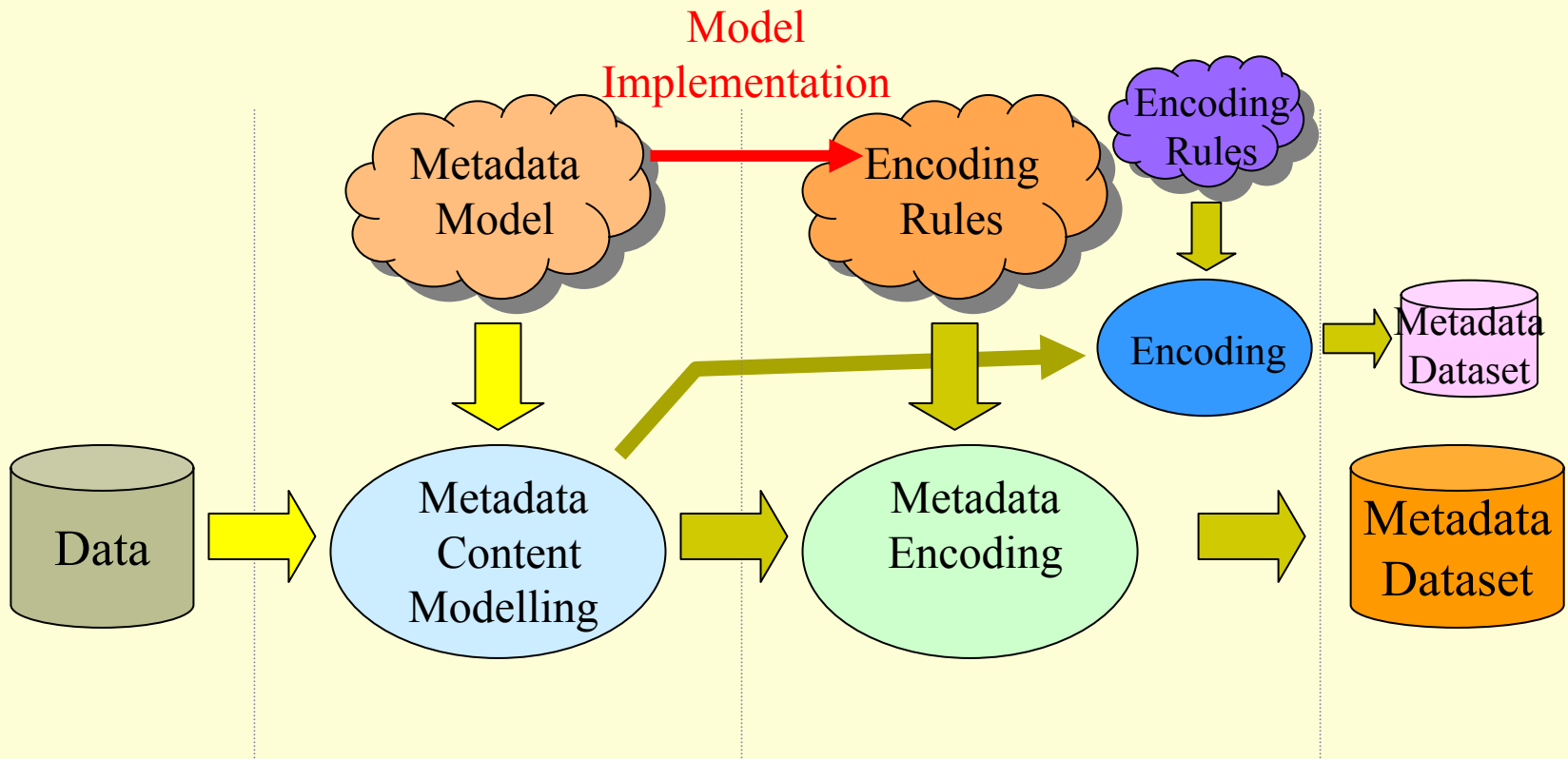
■ DESIRE/RDF/XML

▽ Meta-Metadata Model Heterogeneity?



# Metadata Supplementation: Encoding

- To distinguish two steps:
  - Metadata Content Modelling
  - Metadata Encoding
    - system-independent data structures suitable for transport or storage.



# Examples of Metadata Model Encoding

## ▽ Metadata Encoding Heterogeneity

Metadata Model	Examples of Encoding
Dublin Core	RDF/XML; HTML 4.0, DCSV (Dublin Core Structured Value)/XML, MS Office file header encoding, ISO 19118/XML, GML 3.0, ...
ISO 19115	ISO 19118 rules/XML, DB dump encoding, GML 3.0, Java object serialisation,....
OpenGIS models	GML 1.0, GML 2.0, GML 3.0, legacy DDLs, CORBA IDL, ....
ISO TC211	ISO 19118/XML
Unidata VDC	VDC/XML encoding
ESML model	ESML

# Interoperability Process

- For IT domain, *Interoperability* has different meanings
  - The process to make “*integratable*” systems to be compatible

- "Compatibility" – The quality of systems to cooperate and/or to coexist without interfering with one another.
- "Integratability" – The quality of possibly incompatible systems to adapt them or their exchanged data to cooperate

[David Flater]

- In such process Metadata plays a key role

# Geo-Information Systems Interoperability

## ■ Three knowledge to share

### ● Source data schema

### **Data Model Interoperability**

- Metamodel (e.g. OO metamodel, etc.)

### Schema Interoperability

- Data schema (e.g. data dictionary, class diagrams, etc.)

- Data semantics (e.g. thesauri, ontology, etc.) **Semantic Interoperab.**

### ● Dataset encoding schema

### **Encoding Interoperability**

- Encoding rules (e.g. ISO 19118, OGC GML 2.0, DTDs, XSDs)

- Encoding syntax (e.g. XML syntax, HTML syntax, ...)

### ● Message transfer schema

### **Protocol Interoperability**

- Communication protocol (e.g. SOAP, HTTP, FTP, SMTP, etc.)

- Data interchange service interface (e.g. File transfer service interface, Transactional service interface)

High

Low

Semantics

# Data Model Interoperability

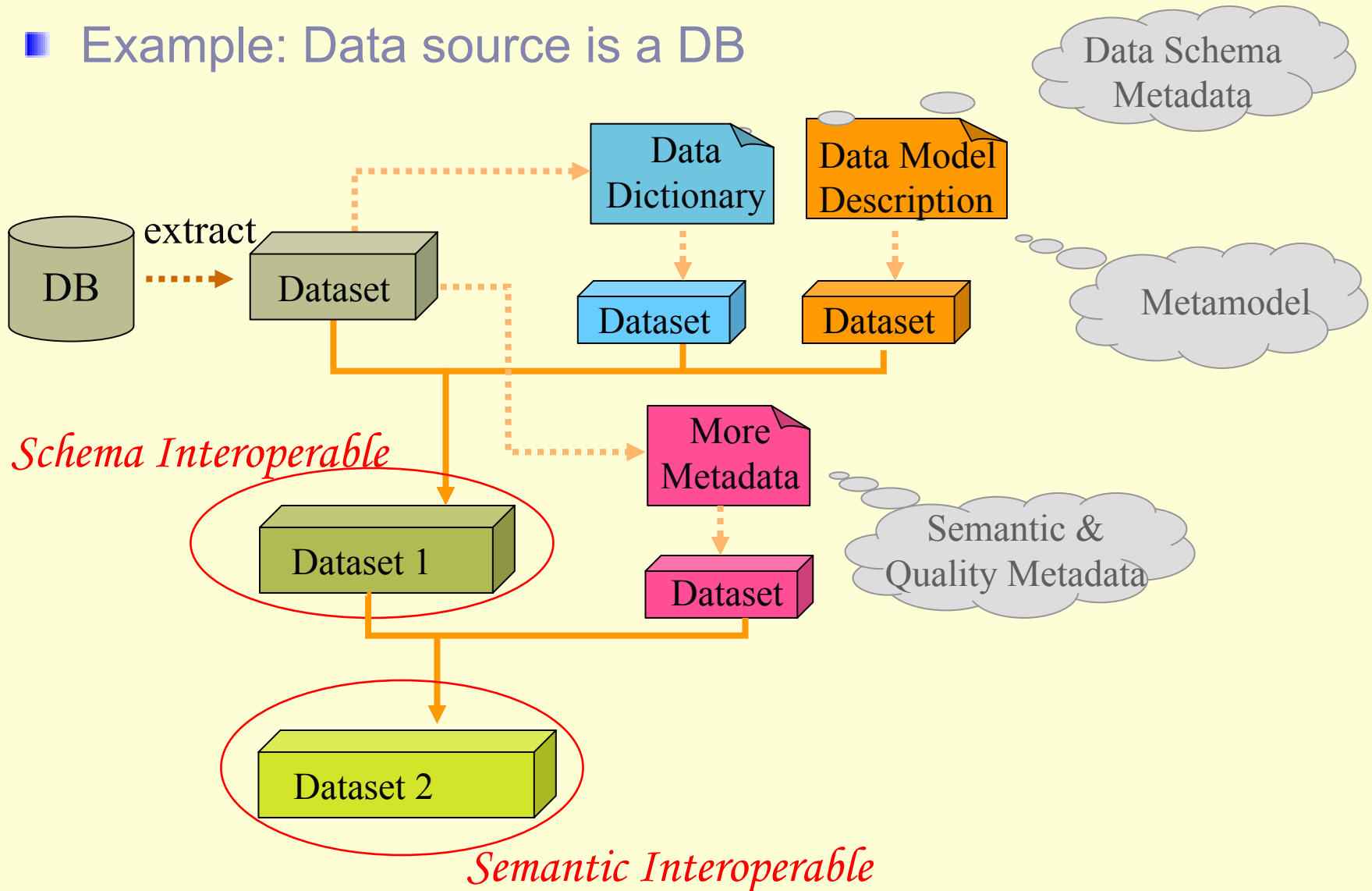
## ■ Three main solutions

- The Universal Schema Solution (USS)
  - A Universal Data Schema
- The Metadata Model Solution (MMS)
  - Metadata to describe sources Data Model heterogeneities
    - Schema
    - Semantics
- The Combo Solution (CS)
  - A common and extensible Reference Data Model
  - A metadata model to describe source data schema and data semantics, on the top of the Common Reference Model

The same approaches are valid for the  
Encoding and Protocol Interoperability

# MMS: Data Model Interoperability

- Example: Data source is a DB



# MMS: Metadata encoding for Interoperability

Interoperability Type	Examples of Metadata Encoding
Data Model Interoperability	XMI, RDF/XML, DAML+OIL/XML, ....
Encoding Interoperability	ISO 19118/XML, DOM/XML, ...
Protocol Interoperability	CORBA IIOP+IDL, SOAP+WSDL, Java RMI+Java Interface spec,....

- The OMG Model Driven Architecture (MDA) uses such approach to address SW components reusability, deployment and interoperability

▽ Is MMS ready to face Geo-inf. data model interoperability?

# CS: Geo-Inf. Systems Interoperability

## ■ Data Model Interoperability

- A Common and extensible Reference Data Model
- A common metadata model to describe heterogeneous data schemas on the top of the Common Reference model
- Source data models are mapped on the top of the Reference Model and Metadata are provided for interoperability

## ■ Encoding Interoperability

- DOM/XML
- Encoding rules are fixed ▼



## ■ Protocol Interoperability

- XML over HTTP
- Data exchange service interface is fixed ▼



# Computational Solutions for Geo-Inf. Systems Compatibility

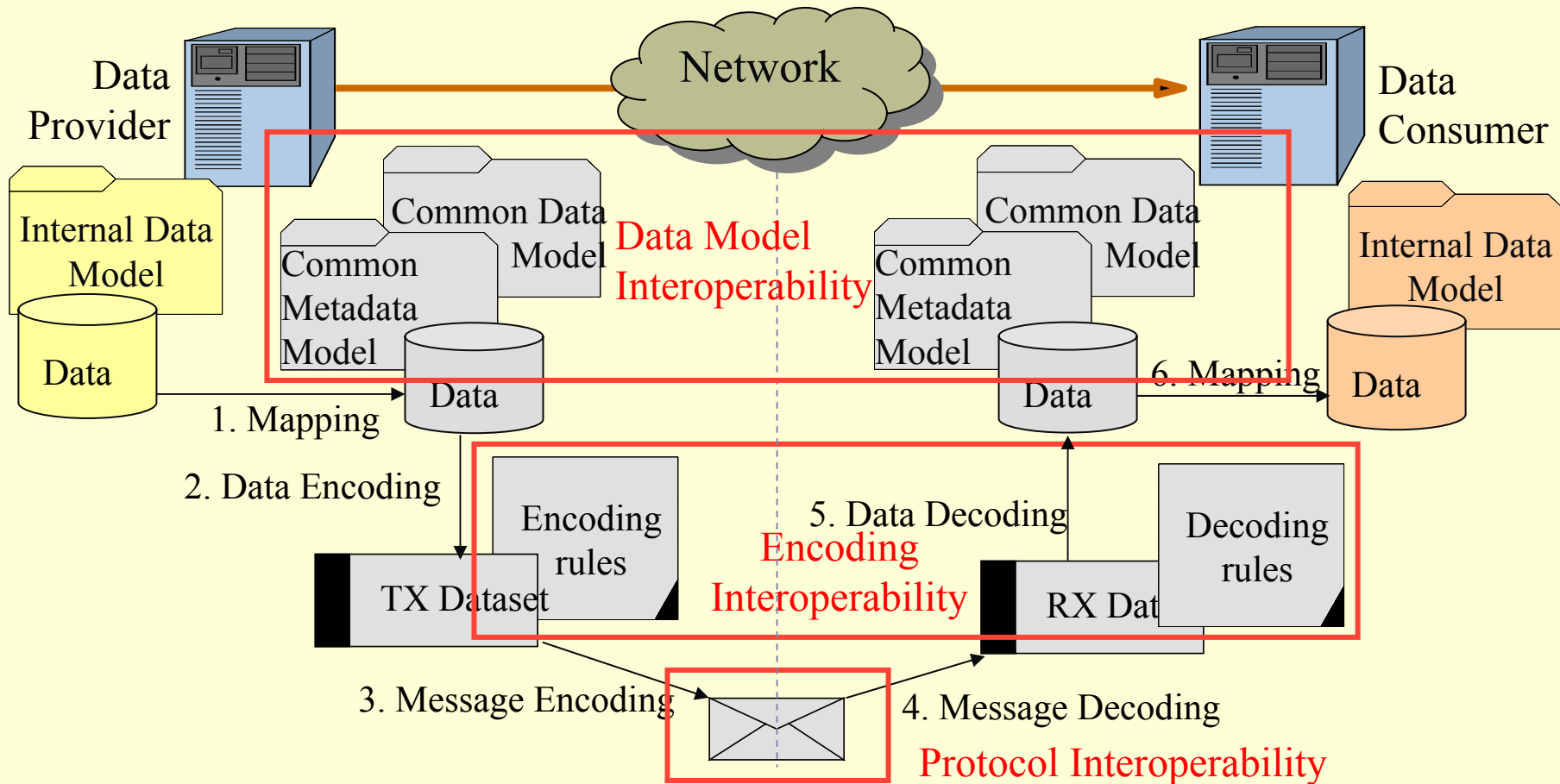
- Data Model Interoperability solution

Compatibility Solution	Data Model Interoperability	Extensibility/ Evolution	Semantic Integration
Cataloguing	Introduces a set of resource documentation	High	Low 
Gateways	Metadata Heterogeneity not supported (USS)	Low 	High
Data Warehousing	Introduces a common data schema (CS)	Medium	High
Mediating	Introduces a common query result schema (CS)	High	High

# DW: Geo-Inf. Systems Interoperability

## ■ Knowledge Sharing

- Common Metadata Model over a Common Reference Model
- Common Encoding Model
- Common Transfer Model









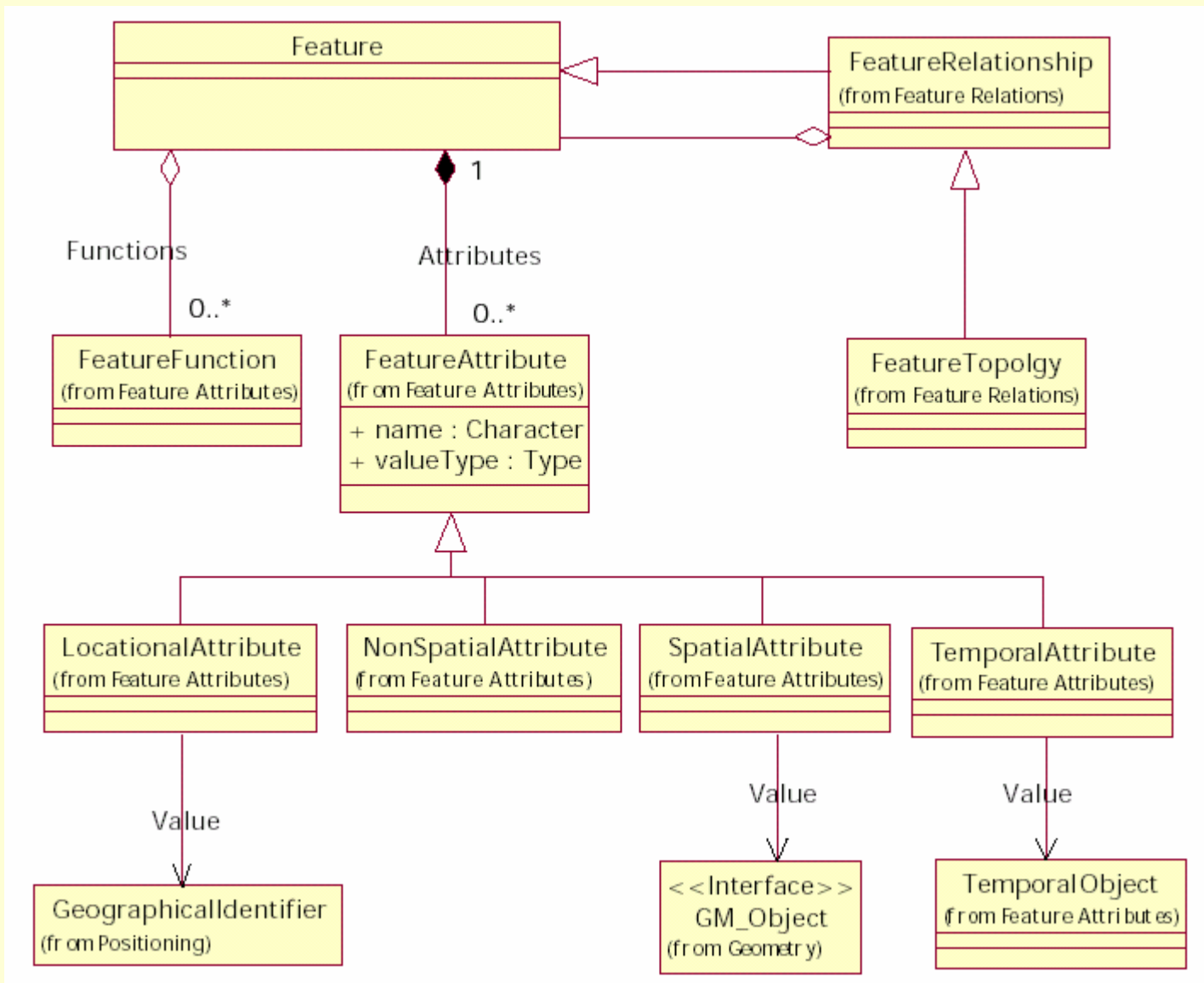




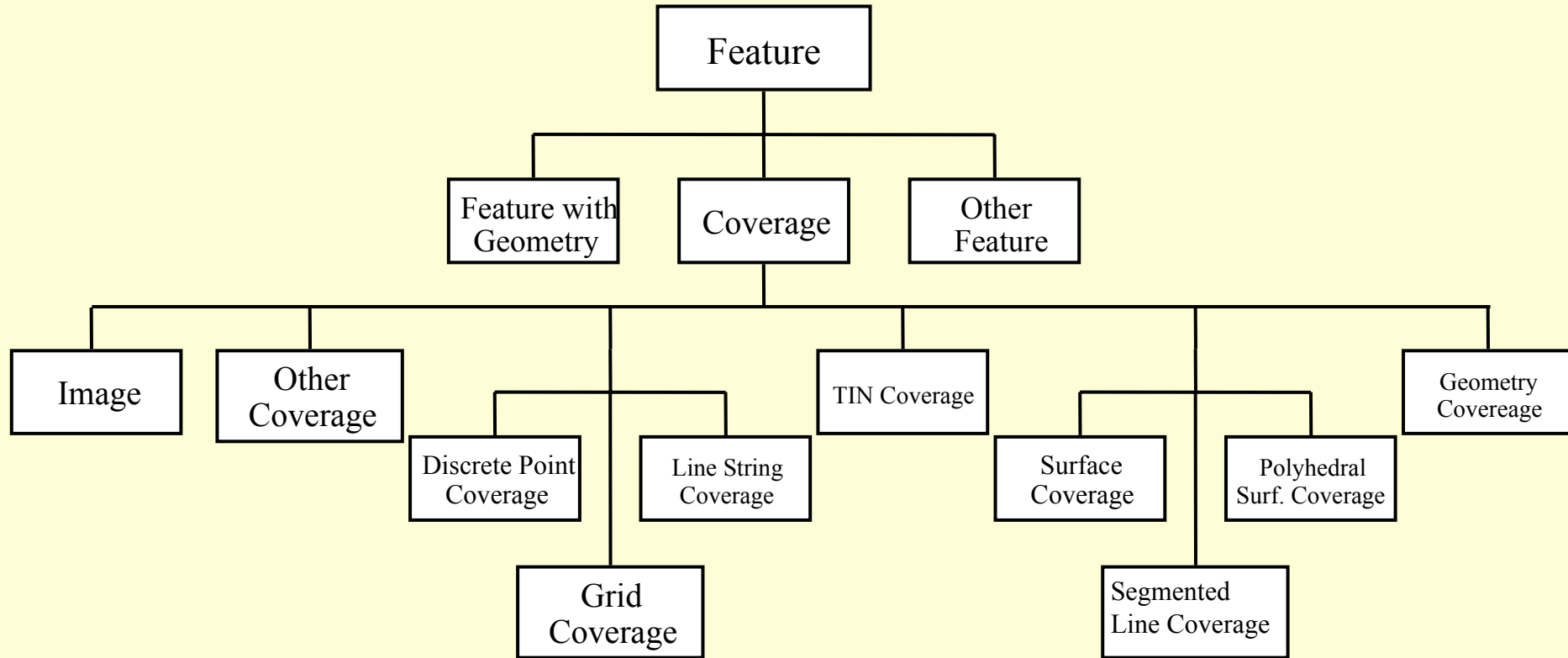
# Shared Main Concepts

- Information Community
    - Geospatial Information Community
  - Feature
    - Abstraction of real world phenomena
  - Coverage (Feature with no shape)
    - Abstraction of
      - Thematic Layer, Image, Multi-dimensional gridded dataset, ...
- 

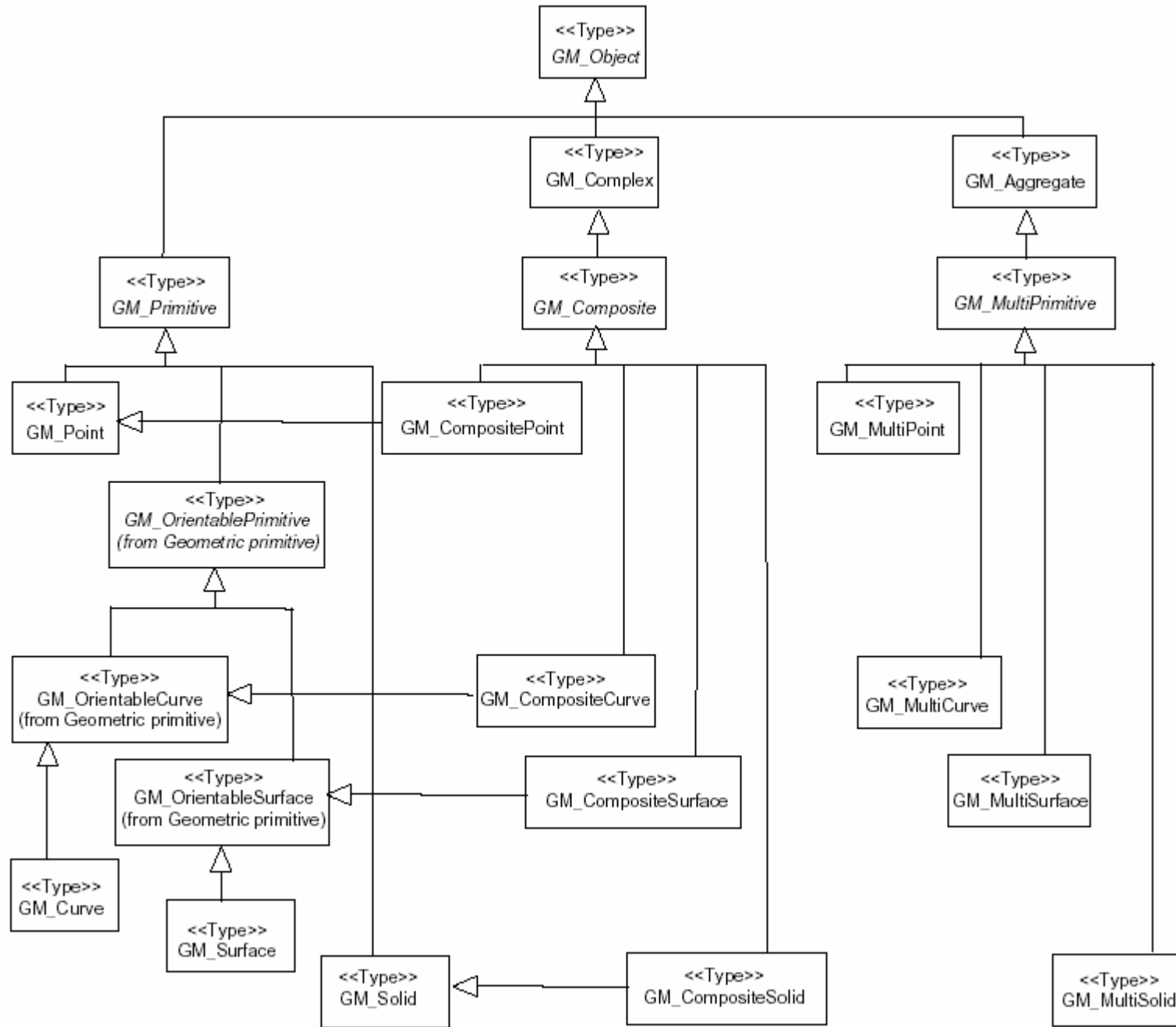
# Example: OpenGIS Feature Model



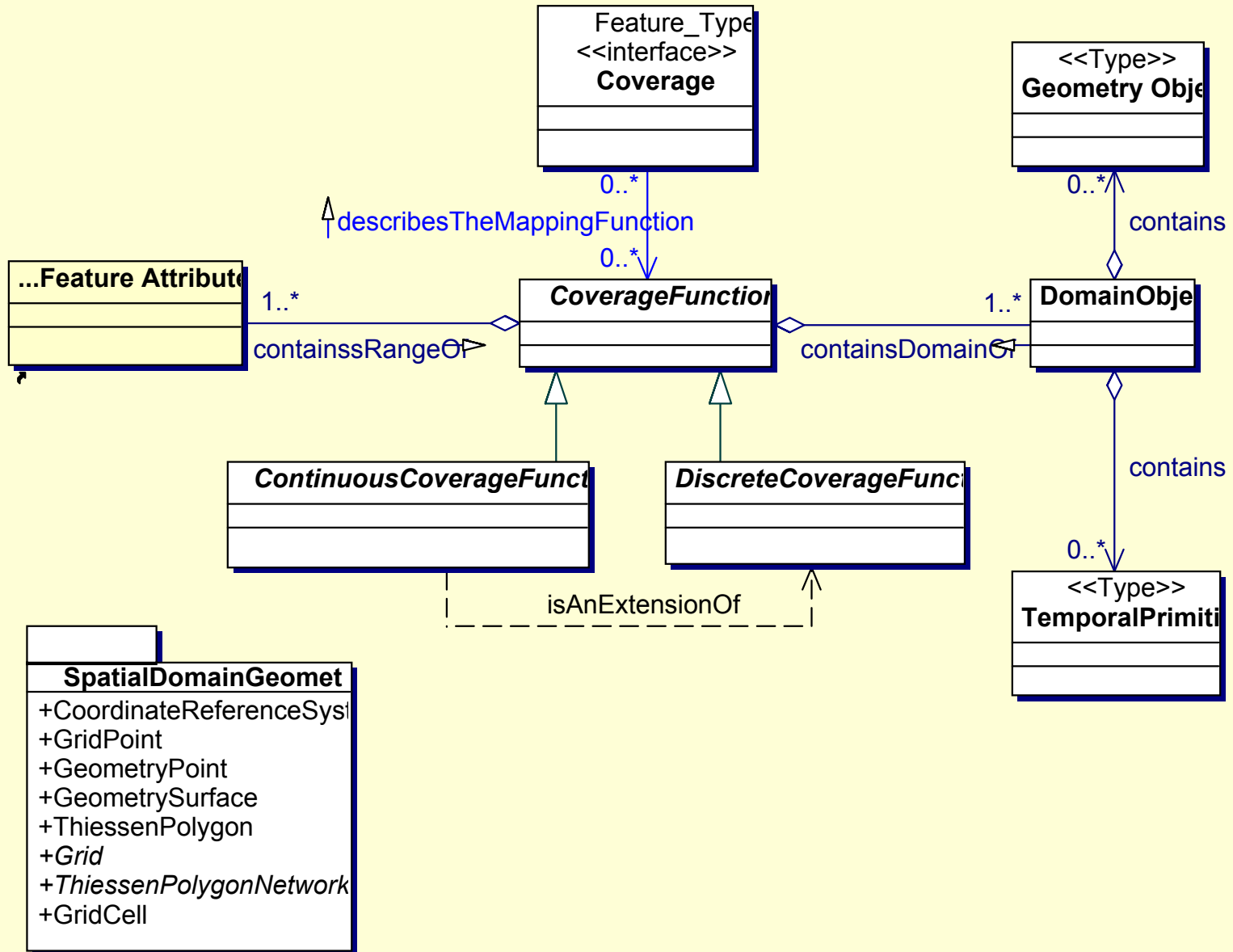
# Example: OpenGIS Coverage types



# ISO 19107: Geometry basic types

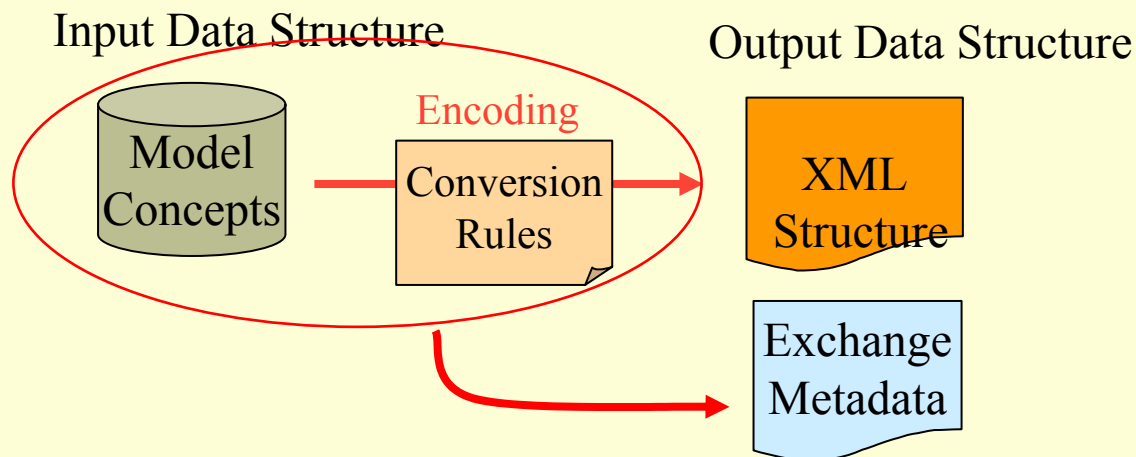


# ISO 19123: Coverage Geometry



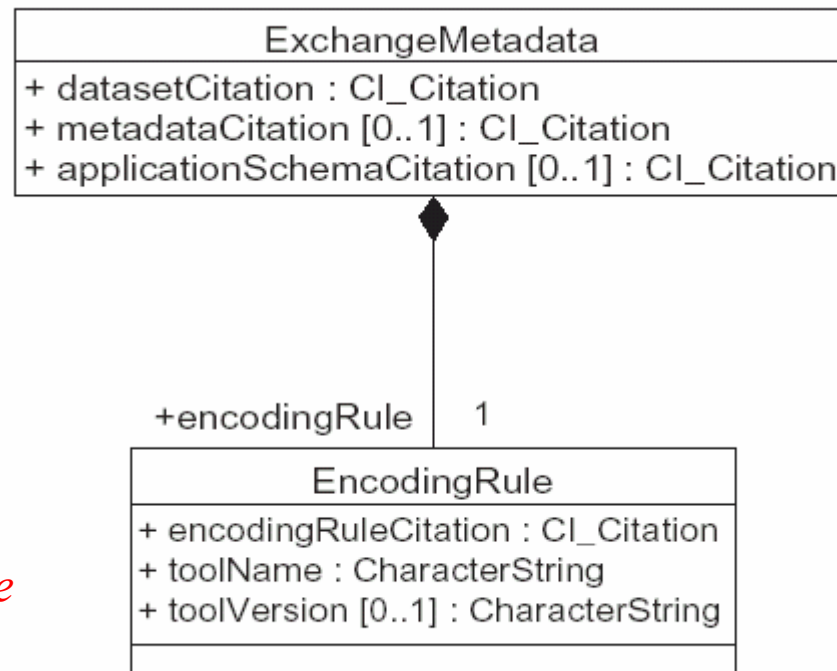
# Encoding Model

- OpenGIS: GML  $\Rightarrow$  an XML-dialect
- ISO 19118: Encoding Rules
  - Encoding Rules (Rules for converting data into a series of codes)
    - Input data structure
    - Output data structure
    - Conversion rules
    - Exchange Metadata



# Metadata Exchange

- Description of
  - Input data structure
  - Encoding process



*Citation ⇔ Reference*



# Conclusions

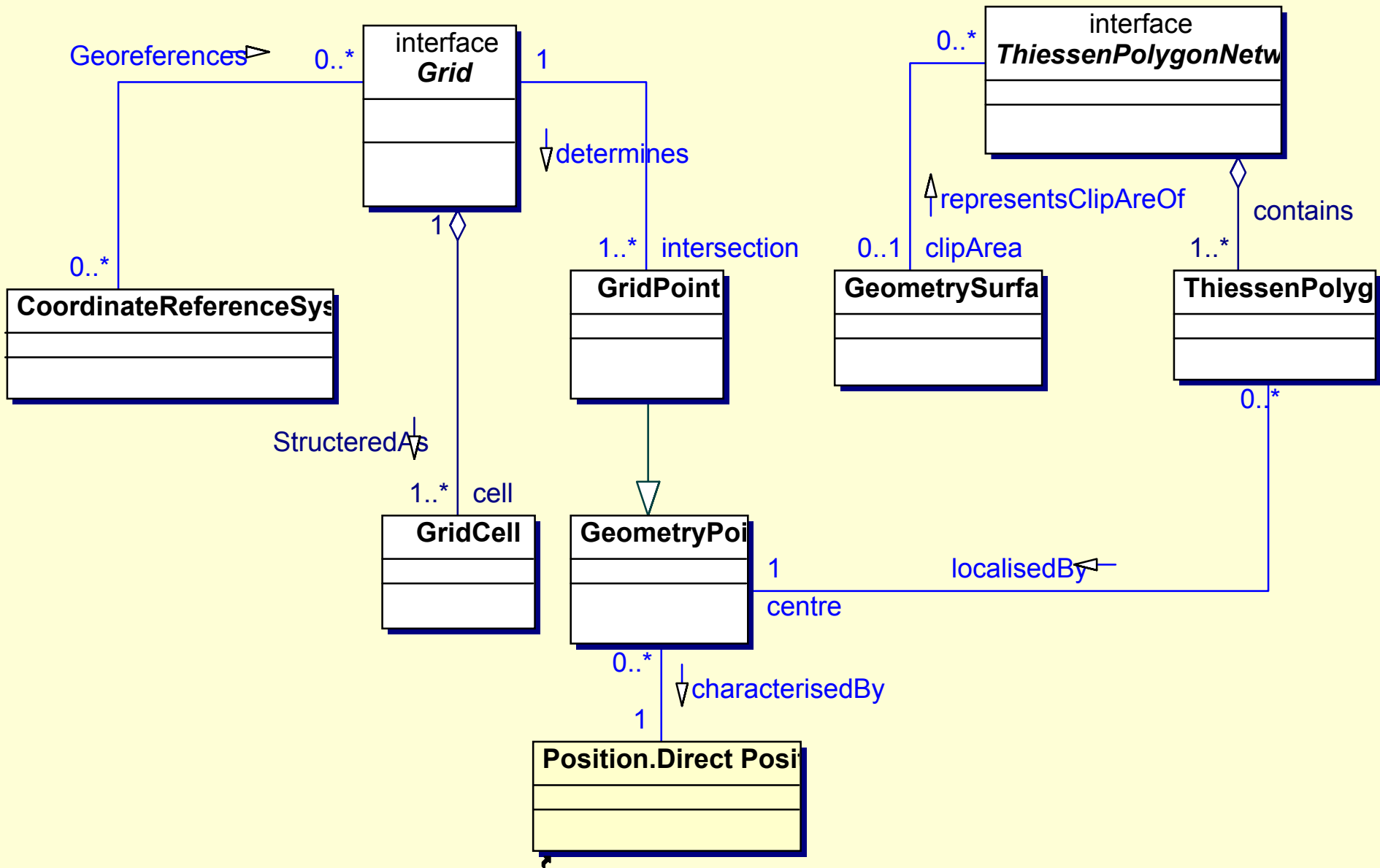
- Metadata is important to
    - Document data
    - Address heterogeneity
      - Data
      - SW components
    - Address distribution
      - Data
      - SW Components
  - It helps to develop
    - Flexible solutions
    - Evolvable solutions
- 



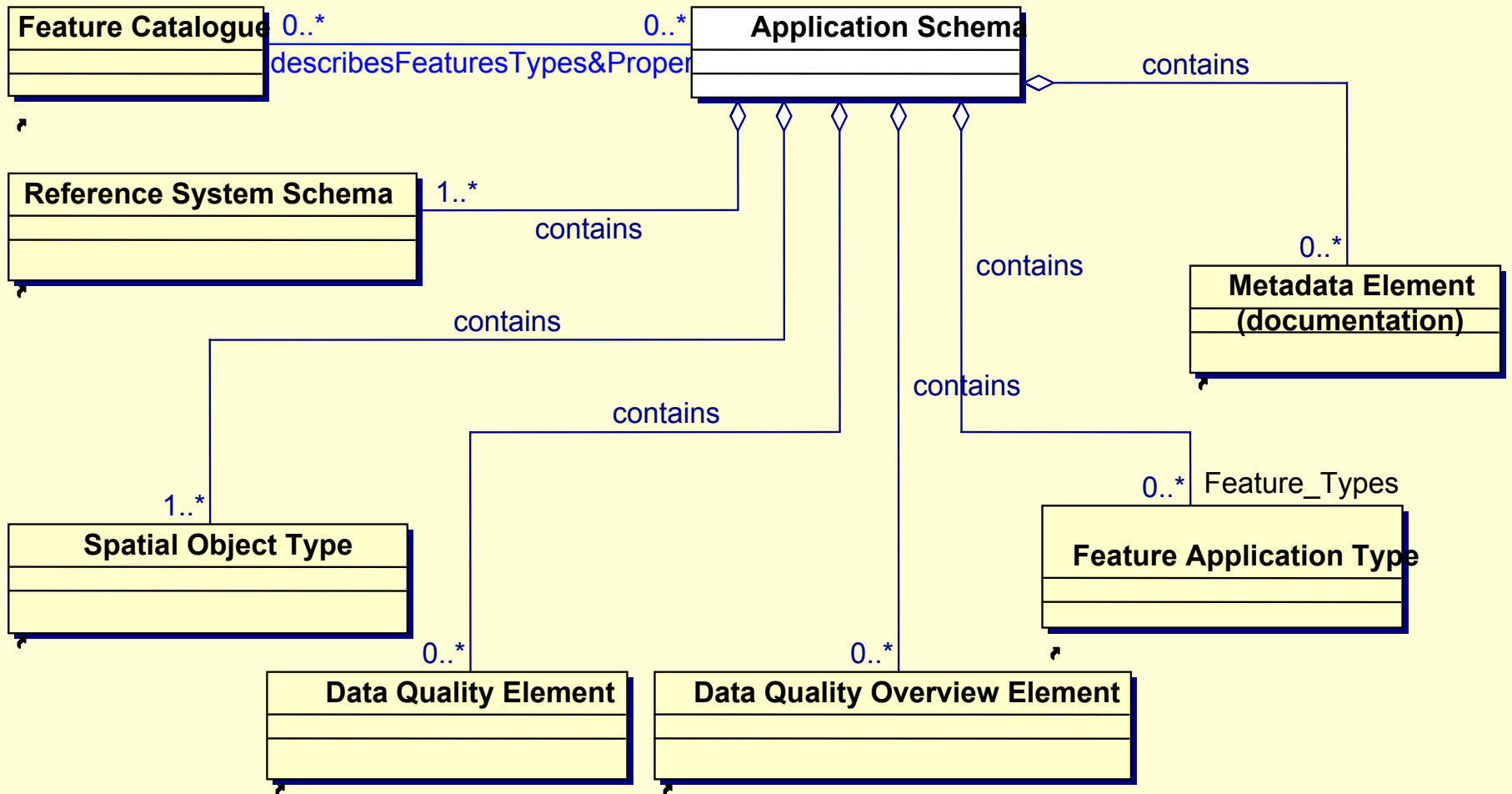
# Interoperability Metadata Types

Metadata Type	Heterogeneity addressed	Components Flexibility/Evolv.
Technical Metadata	System Syntax	Middleware
Logical Metadata	Data Schema	<u>comp.s</u> <u>Data Sources</u>
Metamodels	Data Model	Middleware comp.s
Semantic Metadata	Data Concept Semantics	<u>Data Sources</u> <del>Data Sources</del> Middleware
<b><i>User Support Metadata</i></b>		
Quality-related Metadata	Data Reliability	<u>Data Sources</u> Presentation
Infrastructure Metadata	System structure	Middleware comp.s
User-related Metadata	Users	Presentation comp.s

# ISO 19123: Coverage Geometry



# Source Data Model: Application Schema



# THREDDS Possible Solution

## ■ Data Model Interoperability

- OpenGIS/ISO model for GIS-like data;
- THREDDS Extension for complex gridded-based data; such extension is a specification of the OpenGIS/ISO Coverage model for gridded data.
- OpenGIS/ISO model for Spatial and Temporal Reference Systems, including also coordinate transformations

## ■ Encoding interoperability

- To support a generic encoding standard (XML-based), such as:
  - GML 2.0/3.0
  - ISO 19118
  - ESML

## ■ Protocol Interoperability

- SOAP + WSDL
- (to be ready to leverage UDDI standard for service discovery). UDDI is based on WSDL.

# THREDDS Possible Solution

## ■ Data Model Interoperability

- Top-down approach
  - To start from User needs and generate a User Needs Common Model
  - To conceive the User Needs Common Model at the conceptual level (NOT at the structural level)
  - To conceive the User Needs Common Model according to a hierarchical model (it is the way people conceptualise complex reality, mentally)
- To conceive mediating services to get query results from data sources (e.g. DODS) and to map them according to the User Needs Common Model. Three main steps:
  1. Wrap data source to encode their data model into a semi-structured format
  2. To query the data semi-structured model
  3. To Implement simple mediating rules to map query results onto the User Needs Common Model (i.e. to solve structural and semantic mismatching)
- To implement Application-level mediating services which semantically integrate the distributed and already mapped and present as a unique result.

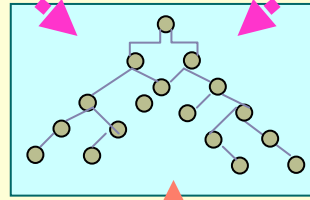
Digital Library  
Community



Meteo  
Community



Encoded Dataset

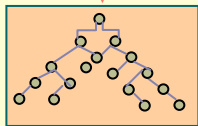


Hyerarchical conceptual model  
of the User seeds  
(User Needs Common Model)

Application  
Mediating Service

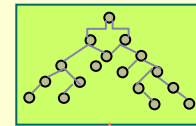
Source  
Mediating Service

Source  
Mediating Service



Semistructured  
view of  
data schema

Encoded Dataset



Semistructured  
view of data schema

Data  
Source

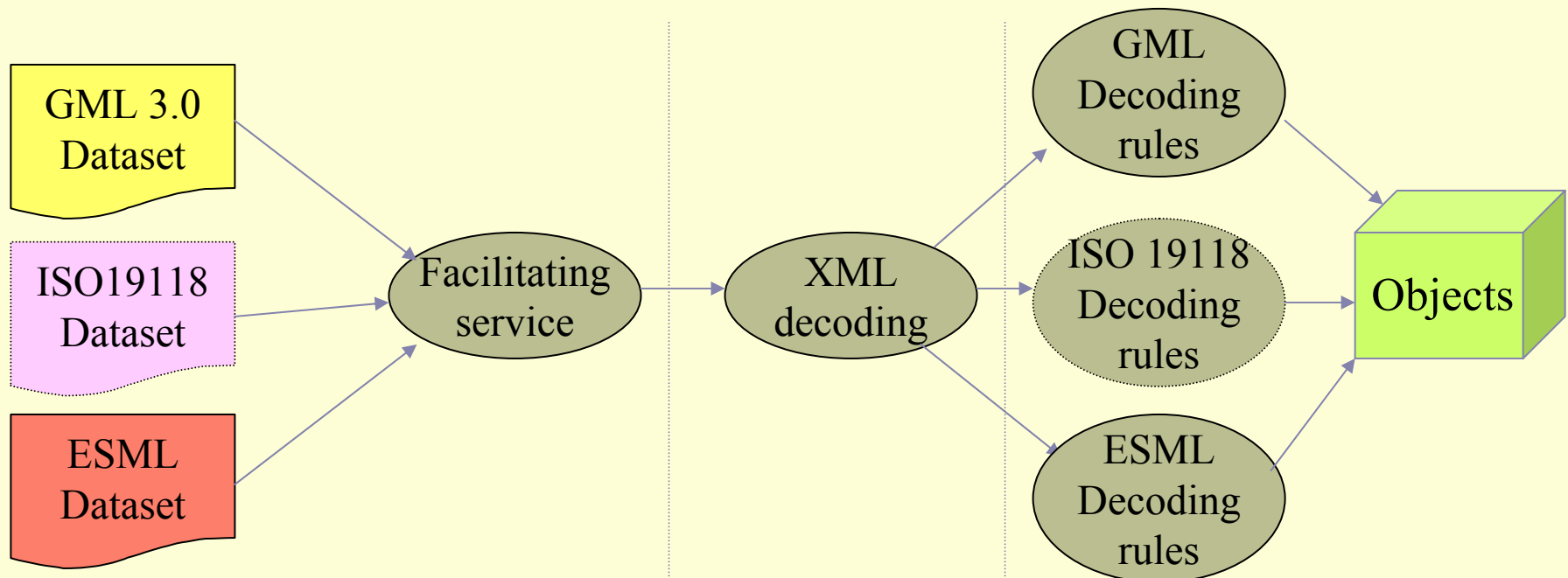
Data  
Source

# THREDDS Possible Solution

## ■ Encoding Interoperability

### ● Bottom-up approach

- Starting from a set of finite existing encoding formats, to conceive a way to cope with their heterogeneity. Three main steps:
  1. To implement a general service for XML syntax decoding
  2. To implement customised services for encoding rules decoding
  3. To implement a facilitating service which is able to recognise the encoding language and activate the appropriate decoding process

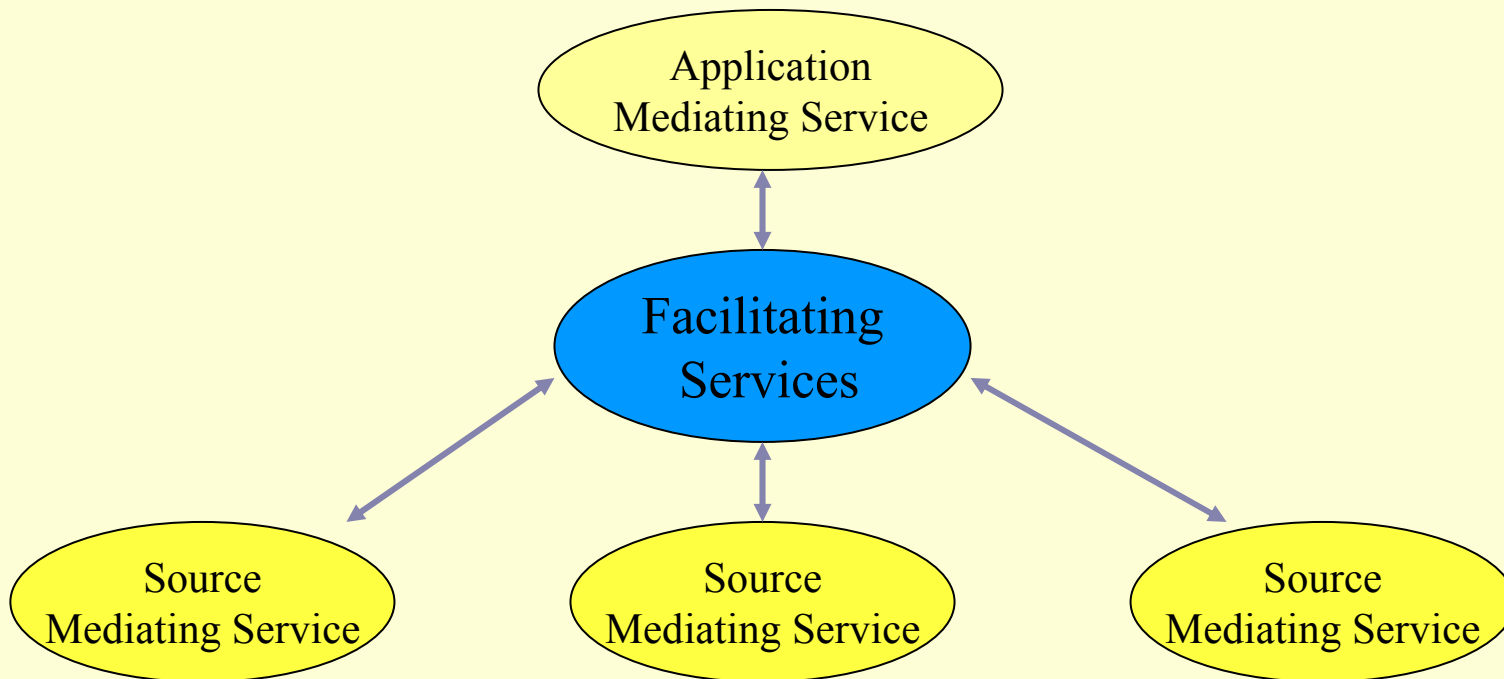



# THREDDS Possible Solution

## ■ Protocol Interoperability

### ● Facilitating approach

- In order to make mediator components work together, resolving routing, interface heterogeneities, and security issues, a facilitator component is introduced
- The technology used is Web-enabled tech. (SOAP/XML/HTTP)



- 
- In particular SINOTS use the same solution;
  - To use SINOTS architecture, theoretically, it should be necessary only:
    - To change the models
      - User Needs Common model
      - Sources semistructured view models
    - To synchronise the mediators
      - Application mediator
      - Sources Mediators
  - Such solution is extensible
    - Change/introduction of data sources
    - Change/introduction of dataset encoding rules
    - Change/introduction of User needs
    - Introduction of new Information Communities
- 