

CEOP Standardization Metadata Development and Applications

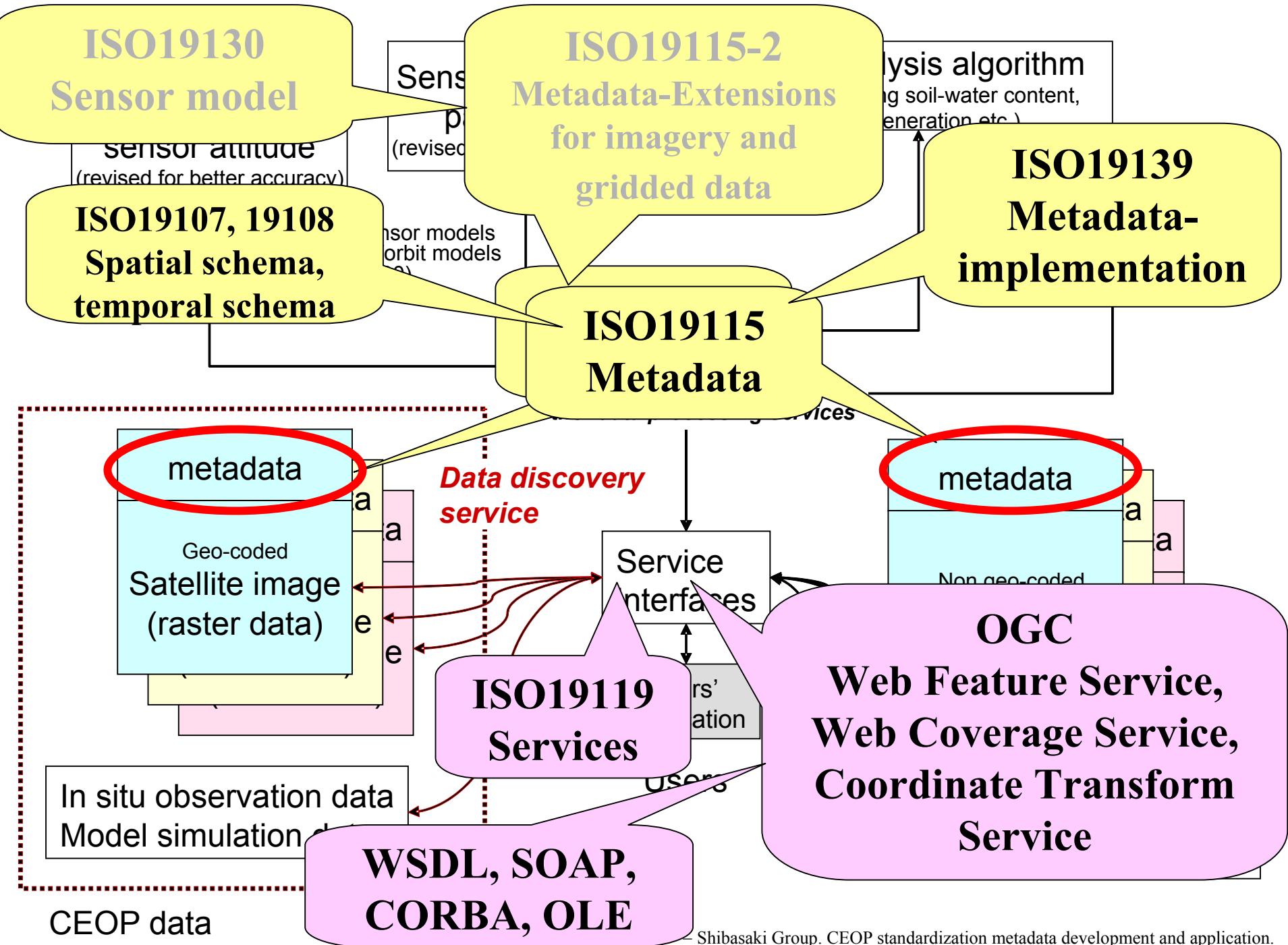
Prof. Ryosuke Shibasaki, Rong Xie
{shiba, xierong}@iis.u-tokyo.ac.jp

Outline of the presentation

1. Background
2. A view of our contributions to CEOP in the phase I
3. Advanced satellite image metadata development in the phase II
4. Conclusions

Outline of the section

1. Background

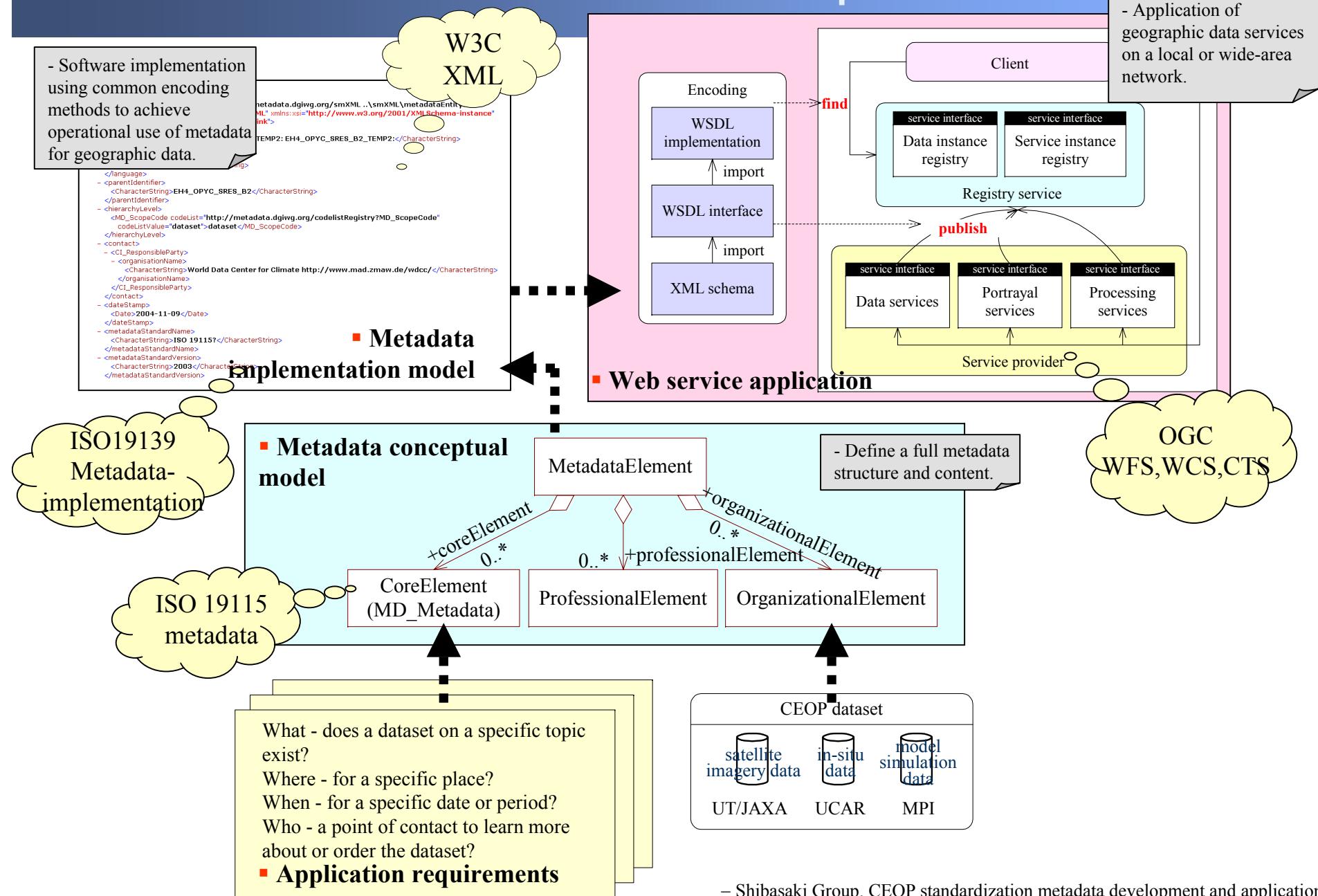


Outline of the section

2. A view of our contributions to CEOP in the phase I

- 2-1 Metadata standard development framework
- 2-2 Metadata design for integrating CEOP satellite imagery, reference site data and simulation result data
- 2-3 Metadata application architecture

2-1 Metadata standard development framework



2-2 Metadata design for integrating CEOP satellite imagery, reference site data and simulation result data

Header Information of Each data file (Ver.02)			
	Revision date of this table(Ver.02):	Aug. 11, 2003	
	First version of this table(Ver.01):	Nov. 18, 2002	
No.	Description	Type	Size
1	Filename	Character	64
2	Sensor	Character	64
3	Product	Character	64
4	Observation Date and Time	Character	64
5	Image Size	Character	64 データサイズ
6	Data Type	Character	64
7	Data Unit	Character	64
8	Scale Factor	Character	64 ピクセル値
9	Observation Channel	Character	256
10	Reference Site	Character	64
11	LAT./LON. in center of Lower Left Pixel	Character	64
12	Grid Size	Character	64 地上でのグリッド間隔 (緯度・経度単位)
13	Missing Value	Character	
14	Observation area ratio (Observed Pixel/All Pixel)	Character	
15	Subset software version	Character	
16	Processing Date	Character	
17	Processing Center	Character	
18	Input Original Filename	Character	
19	Original File Processing Center	Character	
20	HDF library version	Character	
21	Blank	Character	576
Total Size:		2048 Byte	

for “Satellite imagery data”

CEOP Reference Site Data Set Metadata Procedures - Microsoft Internet Explorer

CEOP Reference Site Data Set Metadata Procedures

The documentation (i.e. the "Readme" file) that accompanies each CEOP Reference Site Data Set is as important as the data itself. This information permits collaborators and other analysts to become aware of the data and to understand any limitations or special characteristics of data that may impact its use elsewhere. The data set documentation should accompany all data set submissions and contain the information listed in the outline below. The following outline (and content) should be adhered to as closely as possible. Documentation file submission must accompany e

TITLE: This should match the data set name

CONTACT(S):

Name(s) of Reference Site Contact(s)
Complete mailing address, telephone/fax/fax
E-mail address and WWW address (if applicable)
Similar contact information for data questions (if different than above)

1.0 DATA SET OVERVIEW:

Introduction or abstract
Time period covered by the data
Physical location (including lat/lon/elev) of the measurement or platform, landscape, and soil characteristics
Data source if applicable (e.g. for operational data include agency)
Any World Wide Web address references (i.e. additional documentation such as Project WWW site)

2.0 INSTRUMENTATION DESCRIPTION:

for “Reference site data”

WMO Core Metadata Standard - Microsoft Internet Explorer

for “GRIB data”

This standard provides a general definition for directory searches and exchange that should be applicable to a wide variety of WMO datasets. It does not specify how these metadata should be archived or presented to users. It also does not specify any particular implementation and could be implemented as a database, a flat file, or any other suitable mechanism. However, XML is recommended as the standard for exchange. The comprehensive and technical details required for implementation of this standard in XML are provided in the WMO Core Metadata XML Schema, which is given in the annex to paragraph 4.2 below. In the XML Schema, Class names (such as MD_Metadata) are reserved for data types. These data type names do not appear in the XML instance. Instead the classes are referenced by the name of an object (e.g. metadata) which is an instance of the class (MD_Metadata).

Of the core elements listed, those in **bold** are required, with all others being optional.

If must be remembered that this list defines a minimum set of information to describe data for WMO exchange and is not exhaustive. To fully meet the requirements of WMO Programmes for metadata, application of far more comprehensive standards would be required. The development of these comprehensive standards should be pursued by the individual programmes.

Generic Name	ISO Field/Class Name and Reference Lines	Definition
Metadata ID	MD_Metadata (1) fieldIdentifier (2)	Unique identifier for this metadata item
Metadata language	language (3)	Language of this metadata item
Metadata char. set	characterSet (4)	Character set of this metadata item (Default of ISO 10646-1)
Metadata contact	contact (5) CI_ResponsibilityParty	Party responsible for this metadata item
Metadata date	dateStamp (9)	Date that this metadata item was created
Metadata name	metadataStandardName (10)	Name of the metadata standard (including profile name) used
Metadata Version	metadataStandardVersion (11)	Version (profile) of the metadata standard used
Data Reference System	referenceSystemInfo (13)	Description of the data temporal and spatial reference system
	MD_ReferenceSystem (14)	Information about the reference systems used (temporal, coordinate and geographic)

for “MOLTS data”

Data Set Creation

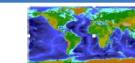
Dataset_Creator: NASA GSFC Global Modeling and Assimilation Office
Dataset_Title: GEOS3 CEOP EOP-1
Dataset_Release_Date: 2003
Dataset_Release_Place: Greenbelt, MD
Dataset_Publisher: NASA/GSFC
Online_Resource: http://oma.gsfc.nasa.gov/sci_research/CEOP/README1st.php

Temporal Coverage

Start Date: 2001-07-01
Stop Date: 2004-09-30

Spatial Coverage

Southernmost Latitude: -90.0
Westernmost Longitude: -180.0
Northernmost Latitude: 90.0
Easternmost Longitude: 180.0



Location Keywords

GLOBAL

Data Resolution

Latitude Resolution: 1 degree
Longitude Resolution: 1 degree

Parameters

HYDROSPHERE > SNOWICE > SNOW DEPTH 
CRYOSPHERE > SEA ICE > SNOW DEPTH 

ISO Topic Category

CLIMATE LOGIC/METEOROLOGY/ATMOSPHERE
BIOTA
OCEANS
INLAND WATERS
GEOSCIENTIFIC INFORMATION
ELEVATION

Platform

GCM = General Circulation Model
MODELS

Project

CEOP > Coordinated Enhanced Observing Period
GEWEX > Global Energy and Water Cycle Experiment

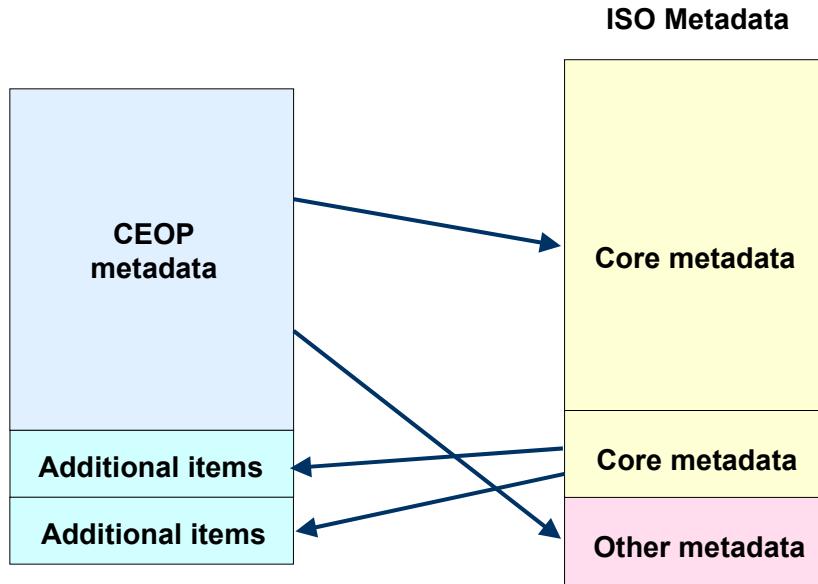
Entry ID

CEOP.DAO

Related URL

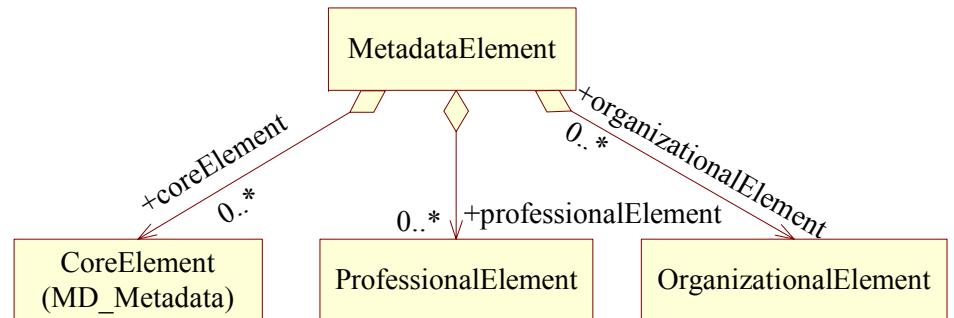
Content Type: RELATED DATA SET DESCRIPTION
URL: http://teman.refer.nasa.gov/sci_research/EOPlIndex.php

Metadata Conceptual Model



(a) Metadata content and structure

- Coordination with the global metadata directories through the core metadata of **ISO 19115**.
- Additions for the specific professional needs.
- Additions for the specific need of different research communities.

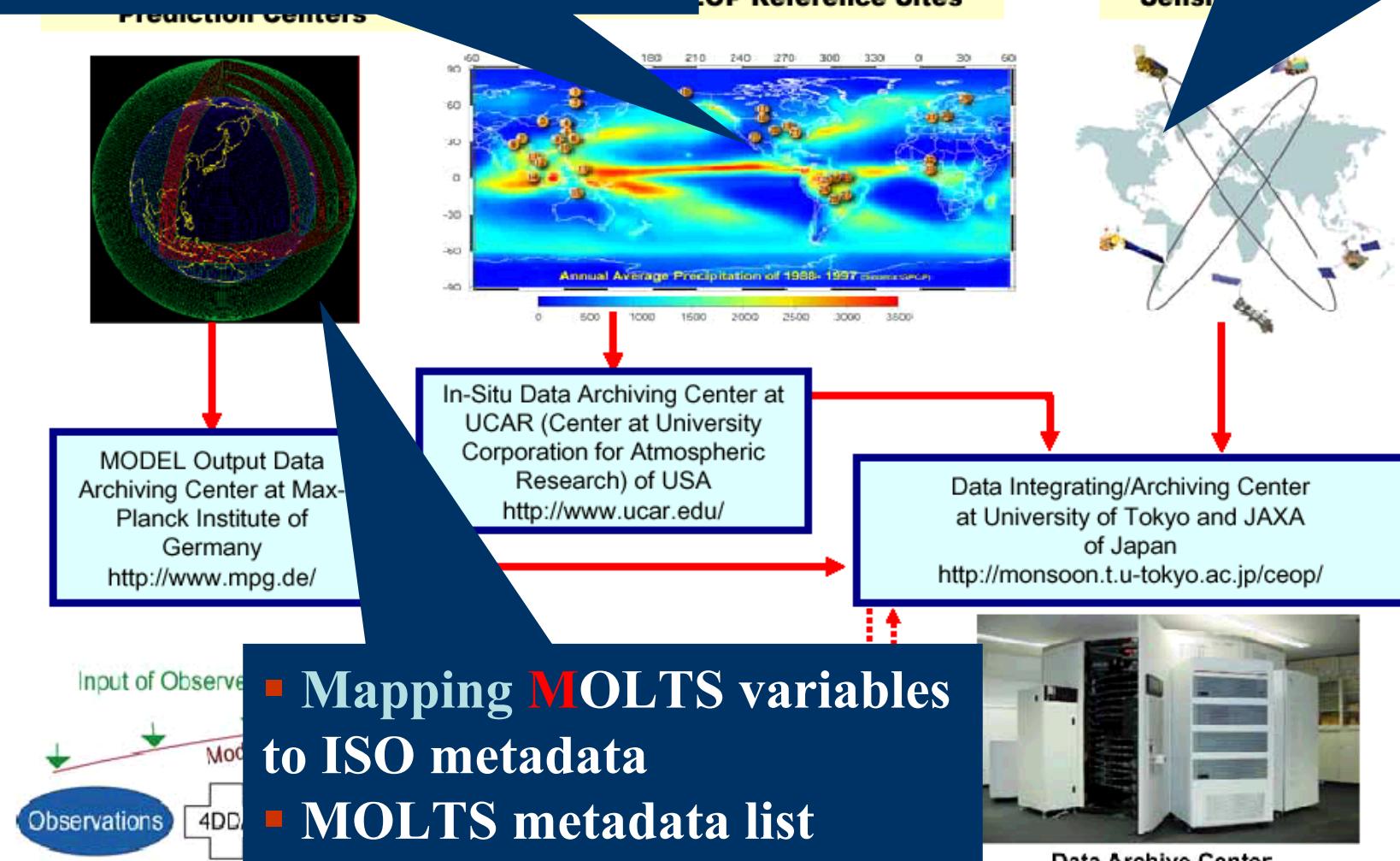


(b) CEOP/ISO metadata standard harmonization

Concerns on the standardization

- Mapping CEOP reference site metadata and ISO metadata
- Reference site metadata list

- Mapping header information of satellite data and ISO metadata
- Satellite imagery metadata list



Mapping header information of satellite data and ISO metadata

Index	CEOP Header Data Item Description	ISO_Metadata package	ISO_Metadata Class	ISO_Metadata attribute	Data type	Explanations
1	Filename	Identification Information	MD_Identification	citation	CI_Citation	Filename can be title (an attribute of CI_Citation class). See 19115_No.360
2	Sensor		MD_Sensor			If MD_Sensor class suggested in 19115-2 is not applied, this will be handled as "title".
3	Product		MD_DataIdentification	topicCategory	Class	High-level geographic data thematic classification to assist in the grouping and search of available geographic data sets. Can be used to group keywords as well. Listed examples are not exhaustive. "imageryBaseMapsEarthCover" seems to be the most appropriated among the examples given. But, "CEOP_Satellite_Imagery" is suggested as a better term.
			or MD_Keywords	keywords	CharacterString	Keywords can be used.
4	Observation Date and Time	Extent information	EX_TemporalExtent	Class	TM_Primitive	Date and time for the content of the dataset
5	Image Size	Spatial representation information	MD_GridSpatialRepresentation			Spatio grid dimension
			numberOfDimensions	Integer		
			axisDimensionProperties	Sequence<MD_Dimension>		Row, column, vertical etc. see B.5.14 MD_DimensionNameTypeCode <<CodeList>> Resolution of Grid is also represented here
			cellGeometry	class (MD_CellGeometryCode)		Point or area
		Content information	MD_RangeDimension	sequenceIdentifier	MemberName	Number of spectral/frequency bands
6	Data Type	Content information	MD_Band	cellValueType		e.g. 2 byte Integer Proposed in 19115-2: mandatory (if applicable) bit representation of data value in raster cell
7	Data Unit	Content information	MD_CoverageDescription	attributeDescription	record type (see. 19103)	Unit of physical measurement Def. Record type (see. 19103): "A Record is used as an implementation representation for features, by keeping a list of (name, value) pairs in a dictionary. This represents a generic storage structure for features."

..... (omit)

CEOP Satellite imagery metadata list

+metadataEntitySetInformation

MD_Metadata:

fileIdentifier: (unique identifier for this metadata file)

language: (language used for documenting metadata)

characterSet: (full name of the character coding standard used for the metadata set)

contact:

CI_ResponsibleParty:

individualName: (name of the responsible person- surname, given name, title separated by a delimiter)

organisationName: (name of the responsible organization)

positionName: (role or position of the responsible person)

contactInfo:

CI_Contact:

phone:

CI_Telephone:

voice: (telephone number by which individuals can speak to the responsible organization or individual)

facsimile: (telephone number of a facsimile machine for the responsible organization or individual)

address:

CI_Address:

deliveryPoint: (address line for the location)

city: (city of the location)

administrativeArea: (state, province of the location)

postalCode: (ZIP or other postal code)

country: (country of the physical address)

electronicMailAddress: (address of the electronic mailbox of the responsible organization or individual)

onlineResource: (on-line information that can be used to contact the individual or organization)

role: (function performed by the responsible party)

dateStamp: (date that the metadata was created)

metadataStandardName: (name of the metadata standard (including profile name) used)

metadataStandardVersion: (version (profile) of the metadata standard used)

+identificationInformation

MD_Identification:

abstract: (brief narrative summary of the content of the resource(s))

pointOfContact:

CI_ResponsibleParty:

individualName: (name of the responsible person- surname, given name, title separated by a delimiter)

organisationName: (name of the responsible organization)

positionName: (role or position of the responsible person)

contactInfo:

CI_Contact:

phone:

CI_Telephone:

voice: (telephone number by which individuals can speak to the responsible organization or individual) data development and application.

Satellite imagery metadata list

..... (omit)

Mapping CEOP reference site metadata and ISO metadata

CEOP content	Item	ISO package	Class	Attribute	Description	
Contact(s)	Name(s) of reference site Contact(s)	Citation and responsible party information	CI_ResponsibleParty	organisationName	Name of the responsible organization	
	Complete mailing address, telephone/facsimile Nos.	Citation and responsible party information		address	Physical and email address at which the organization or individual may be contacted.	
	E-mail address and WWW address (if applicable)	Citation and responsible party information	CI_Contact	phone	Telephone numbers at which the organization or individual may be contacted.	
	Similar contact information for data questions (if different than above)	Citation and responsible party information	CI_Contact	onlineResource	On-line information that can be used to contact the individual or organization.	
Data set overview	Introduction or abstract	Identification Information	MD_Identification	abstract	Brief narrative summary of the content of the resource(s)	
	Time period covered by the data	Identification information	MD_Usage	useageDateTime	Date and time of the first use or range of uses of the resource and/or resource series	
	Physical location (including lat/lon/elev) of the measurement or platform, landscape, and soil characteristics	Identification	MD_DataIdentification	geographicBox	Minimum bounding rectangle within which data is available	
	geographicDescription			Description of the geographic area within which data is available		
	environmentDescription			Description of the dataset in the producer's processing environment, including items such as the software, the computer operating system, file name, and the dataset size		
	Data source if applicable (e.g. for operational data include agency)	LI_Lineage	source	source	Identification information	
			LI_Source	description	Information about the source data used in creating the data specified by the scope	
		Lineage information		sourceCitation	Recommended reference to be used for the source data	
				sourceExtent	Information about the spatial, vertical and temporal extent of the source data	

..... (omit)

– Shibasaki Group. CEOP standardization metadata development and application.

Additional items for CEOP metadata (**Reference site data**)

ISO package	Class	Attribute	Description
Metadata entity set information	MD_Metadata	fileIdentifier	Unique identifier for this metadata file
	MD_Metadata	language	Language used for documenting metadata
	MD_Metadata	characterSet	Full name of the character coding standard used for the metadata set
	MD_Metadata	contact	Party responsible for the metadata information
	MD_Metadata	dateStamp	Date that the metadata was created
	MD_Metadata	metadataStandardName	Name of the metadata standard (including profile name) used
	MD_Metadata	metadataStandardVersion	Version (profile) of the metadata standard used
	MD_Metadata	referenceSystemInfo	Description of the spatial and temporal reference systems used in the dataset
	MD_Metadata	dataQualityInfo	Provides overall assessment of quality of a resource(s)
Identification information Identification information	MD_Identification	citation	Basic information required to uniquely identify a resource or resources
	MD_Identification	pointOfContact	Identification of, and means of communication with, person(s) and organizations(s) associated with the resource(s)
	MD_Identification	resourceFormat	Provides a description of the format of the resource(s)
	MD_DataIdentification	spatialRepresentationType	Method used to spatially represent geographic information
	MD_DataIdentification	spatialResolution	Factor which provides a general understanding of the density of spatial data in the dataset
	MD_DataIdentification	language	Language(s) used within the dataset
	MD_DataIdentification	topicCategory	Main theme(s) of the dataset
	MD_DataIdentification	extent	Additional extent information including the bounding polygon, vertical, and temporal extent of the dataset
	MD_ServiceIdentification	typeProperties	Attributes that can be queried for this service type. Standardized in a service type registry
Data quality information	DQ_DataQuality	lineage	Non-quantitative quality information about the lineage of the data specified by the scope
Distribution information	MD_Distribution	distributionFormat	Provides a description of the format of the data to be distributed
	MD_Format	name	Name of the data transfer format(s)
	MD_Format	version	Version of the format (date, number, etc.)
Citation and responsible party information	CI_Citation	title	Name by which the cited resource is known
	CI_Citation	date	Reference date for the cited resource
	CI_Citation	citedResponsibleParty	Name and position information for an individual or organization that is responsible for the resource
	CI_ResponsibleParty	contactInfo	Address of the responsible party
	CI_OnLineResource	linkage	Location (address) for on-line access using a Uniform Resource Locator address or similar addressing scheme such as http://www.statkart.no/isotc211

..... (omit)

Reference site metadata list

+metadataEntitySetInformation
MD_Metadata:
 fileIdentifier: (unique identifier for this metadata file)
 language: (language used for documenting metadata)
 characterSet: (full name of the character coding standard used for the metadata set)
 contact:
 CI_ResponsibleParty:
 individualName: (name of the responsible person- surname, given name, title separated by a delimiter)
 organisationName: (name of the responsible organization)
 positionName: (role or position of the responsible person)
 contactInfo:
 CI_Contact:
 phone:
 CI_Telephone:
 voice: (telephone number by which individuals can speak to the responsible organization or individual)
 facsimile: (telephone number of a facsimile machine for the responsible organization or individual)
 address:
 CI_Address:
 deliveryPoint: (address line for the location)
 city: (city of the location)
 administrativeArea: (state, province of the location)
 postalCode: (ZIP or other postal code)
 country: (country of the physical address)
 electronicMailAddress: (address of the electronic mailbox of the responsible organization or individual)
 onlineResource: (on-line information that can be used to contact the individual or organization)
 role: (function performed by the responsible party)
 dateStamp: (date that the metadata was created)
 metadataStandardName: (name of the metadata standard (including profile name) used)
 metadataStandardVersion: (version (profile) of the metadata standard used)

+identificationInformation
MD_Identification:
 abstract: (brief narrative summary of the content of the resource(s))
 pointOfContact:
 CI_ResponsibleParty:
 individualName: (name of the responsible person- surname, given name, title separated by a delimiter)
 organisationName: (name of the responsible organization)
 positionName: (role or position of the responsible person)
 contactInfo:
 CI_Contact:
 (omit)
 application.

Example of reference site metadata – taking BALTEX Lindenberg reference site as an example

```
+metadataEntitySetInformation
MD_Metadata:
    fileIdentifier: BALTEX_LINDENBERG_20021001_20030930
    language: en
    characterSet: utf8
    contact:
        CI_ResponsibleParty:
            individualName: Dr. Frank Beyrich
            organisationName: Meteorologisches Observatorium Lindenberg
            positionName: Deutscher Wetterdienst (DWD)
            contactInfo:
                CI_Contact:
                    phone:
                        CI_Telephone:
                            voice: +49 33677 60228
                            facsimile: +49 33677 60280
                    address:
                        CI_Address:
                            deliveryPoint: Am Observatorium 12
                            city: OT Lindenberg
                            administrativeArea: Tauche
                            postalCode: D - 15848
                            country: Germany
                            electronicMailAddress: frank.beyrich@dwd.de
                            onlineResource: http://www.dwd.de/en/FundE/Observator/MOL
                    role: CEOP Reference Site Manager
    dateStamp: 2004-11-12
    metadataStandardName: CEOP reference site data set metadata
    metadataStandardVersion: 1.0
```

..... (omit)

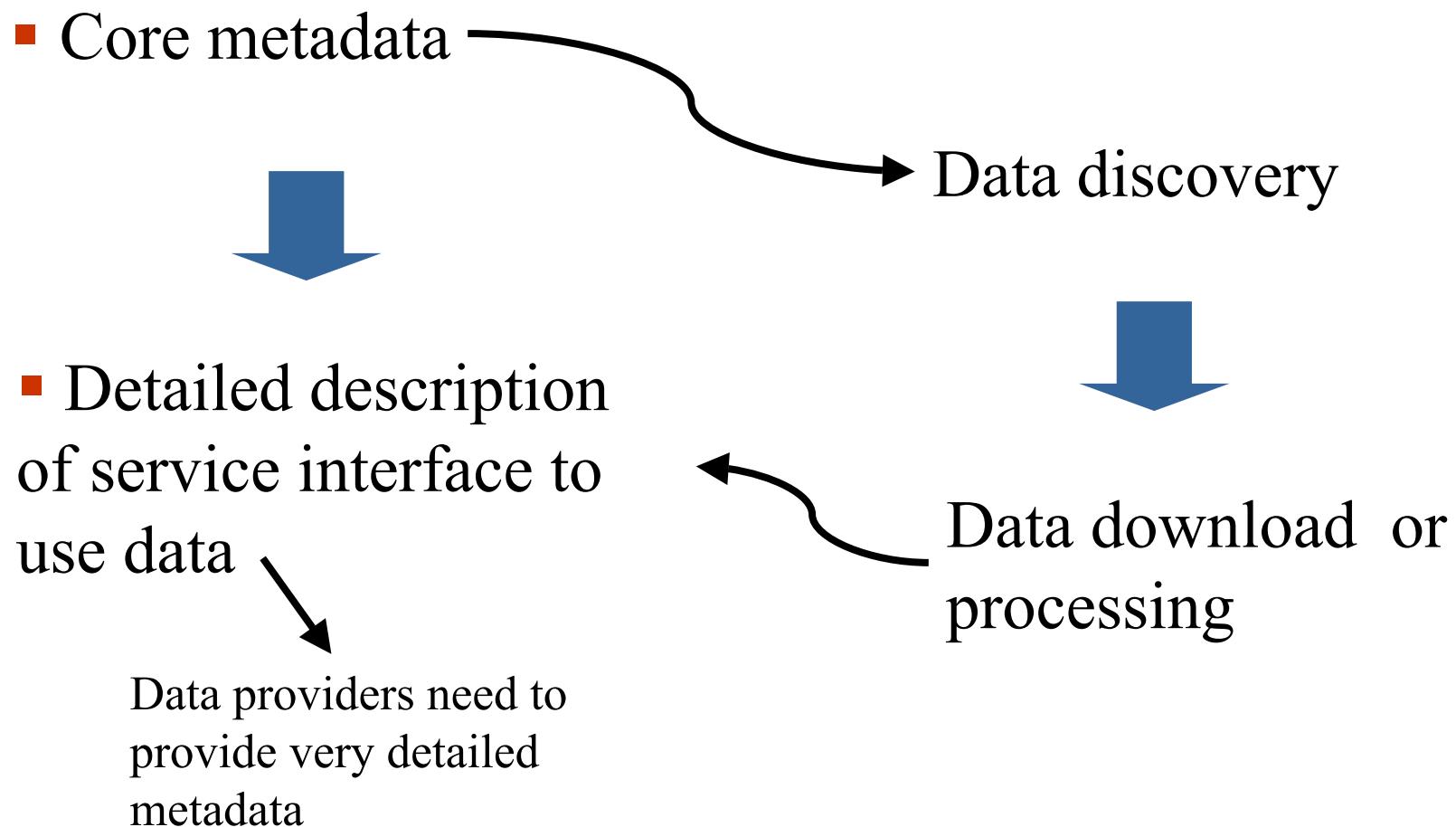
MOLTS variables to ISO mapping

GMAO record	Record content/example	ISO Class	Attribute	Description
Data set citation	Dataset Creator	MD_Identification	citation	Citation data for the resource(s)
	Dataset Title	MD_Identification	abstract	brief narrative summary of the content of the resource(s)
	Dataset Release Date			
	Dataset Release Place			
	Dataset Publisher	MD_Identification	pointOfContact	identification of, and means of communication with, person(s) and organization(s) associated with the resource(s)
	Online Resource	MD_Identification	resourceFormat	provides a description of the format of the resource(s)
Temporal coverage	Start Date Stop Date	EX_TemporalExtent	extent	Date and time for the content of the dataset
Geographic coverage	Southernmost Latitude	EX_GeographicBoundingBox	southBoundLatitude	Southern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north)
	Westernmost Longitude	EX_GeographicBoundingBox	westBoundLongitude	Western-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)
	Northernmost Latitude	EX_GeographicBoundingBox	northBoundLatitude	Northern-most coordinate of the limit of the dataset extent expressed in latitude in decimal degrees (positive north)
	Easternmost Longitude	EX_GeographicBoundingBox	eastBoundLongitude	Eastern-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)
Location keywords	GLOBAL	EX_GeographicDescription	geographicIdentifier	Identifier used to represent a geographic area
Data resolution	Latitude Resolution Longitude Resolution	MD_DataIdentification	spatialResolution	Factor which provides a general understanding of the density of spatial data in the dataset
Parameters	HYDROSPHERE > SNOW/ICE > SNOW DEPTH CRYOSPHERE > SEA ICE > SNOW DEPTH	MD_DataIdentification	topicCategory	Main theme(s) of the dataset
ISO topic category	CLIMATOLOGY/METEOR OLOGY/ATMOSPHERE BIOTA OCEANS INLAND WATERS GEOSCIENTIFIC INFORMATION ELEVATION	MD_TopicCategoryCode <<Enumeration>>	Climatology Atmosphere, inlandWaters, oceans etc.	High-level geographic data thematic classification to assist in the grouping and search of available geographic data sets. Can be used to group keywords as well. Listed examples are not exhaustive.
Platform	GCM > General Circulation Model MODELS	MD_CoverageDescription	attributeDescription	Description of the attribute described by the measurement value

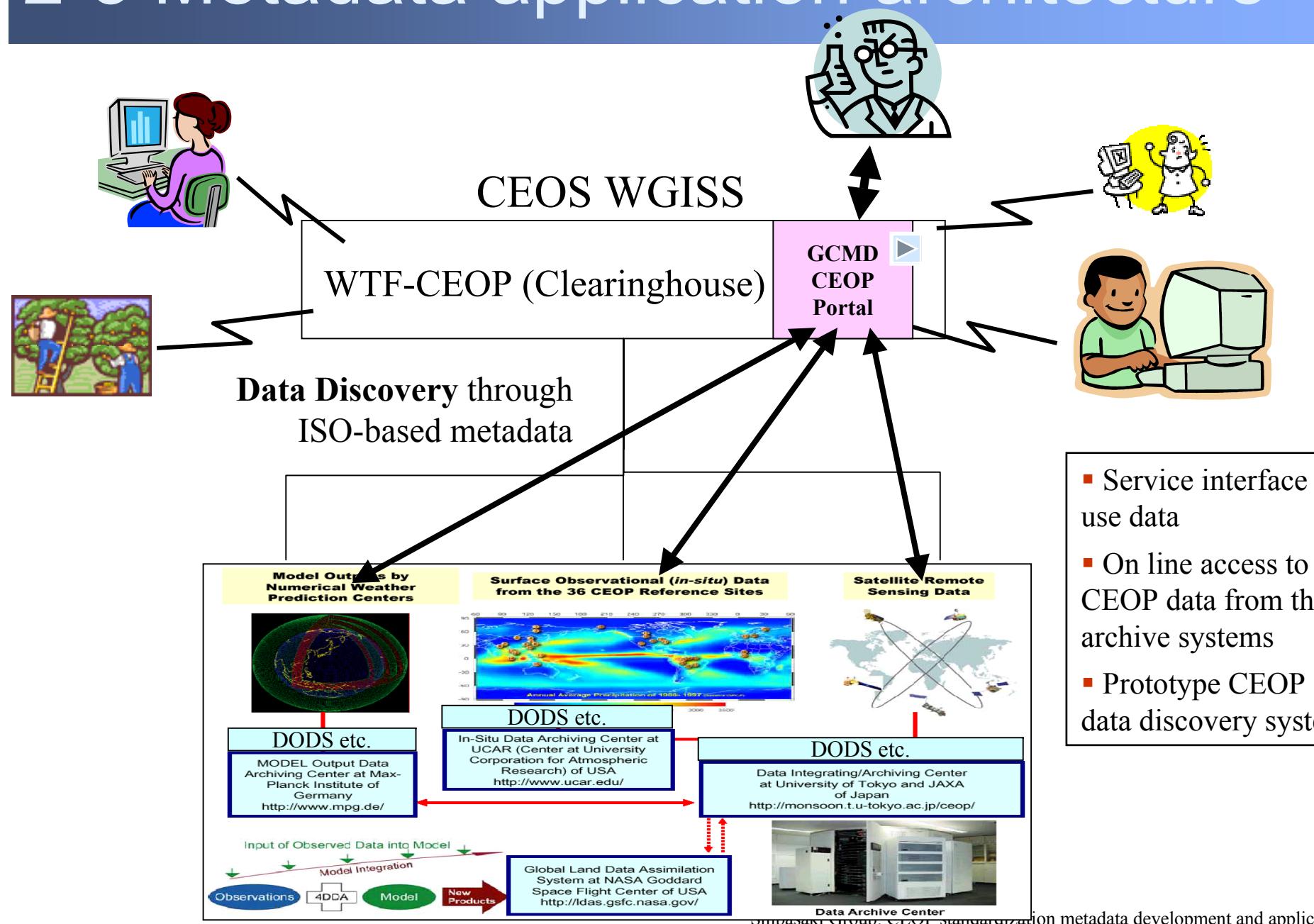
..... (omit)

– Shibasaki Group. CEOP standardization metadata development and application.

2-3 Metadata application architecture



2-3 Metadata application architecture





Find Data Sets by Topic:

- [About Portals](#)
- [Portal Listings](#)
- [CEOP Website](#)



[Agriculture](#)

soils , ...



[Land Surface](#)

soils , topography ...

Data Set Text Search



[Map/Date Search](#)

[Search tips](#)



[Atmosphere](#)

atmospheric temperature ,
atmospheric winds ...



[Oceans](#)

ocean temperature , ocean
winds ...



[Biosphere](#)

vegetation , ...



[Spectral/Engineering](#)

microwave , ...



[Cryosphere](#)

snow/ice , sea ice ...



[Data Centers](#) - [Locations](#) -

[Instruments](#) -

[Platforms/Sources](#) - [Projects](#)



[Hydrosphere](#)

snow/ice , ...



[Search the entire GCMD database](#)



NASA Website Privacy and Security Statement, Disclaimer and Accessibility Certification
Responsible NASA Official: [Lola Olsen](#) · Editor: [Gene Major](#)
Contact [GCMD User Support](#) for assistance

Internet

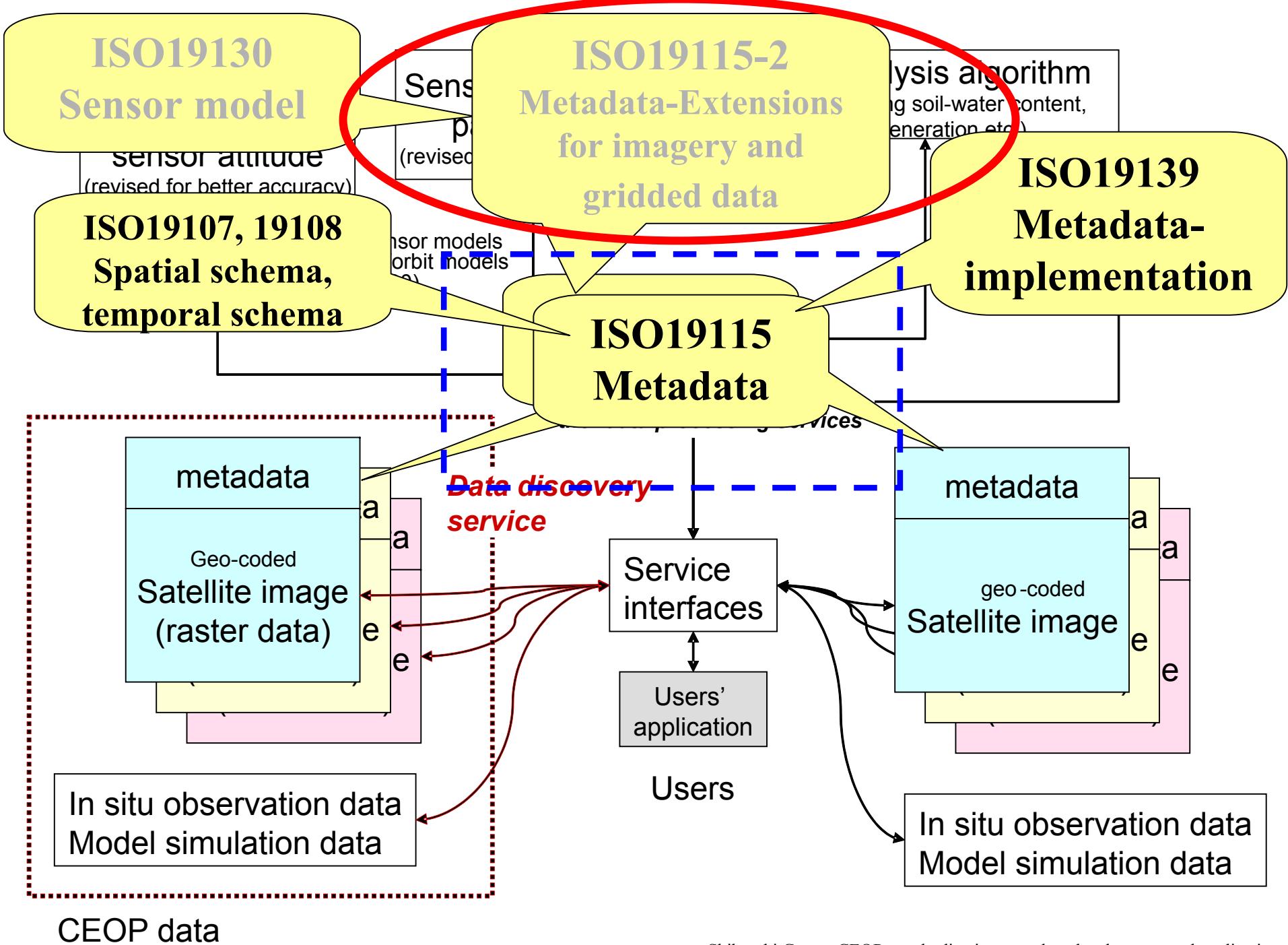
Outline of the section

3. Advanced satellite image metadata development in the phase II

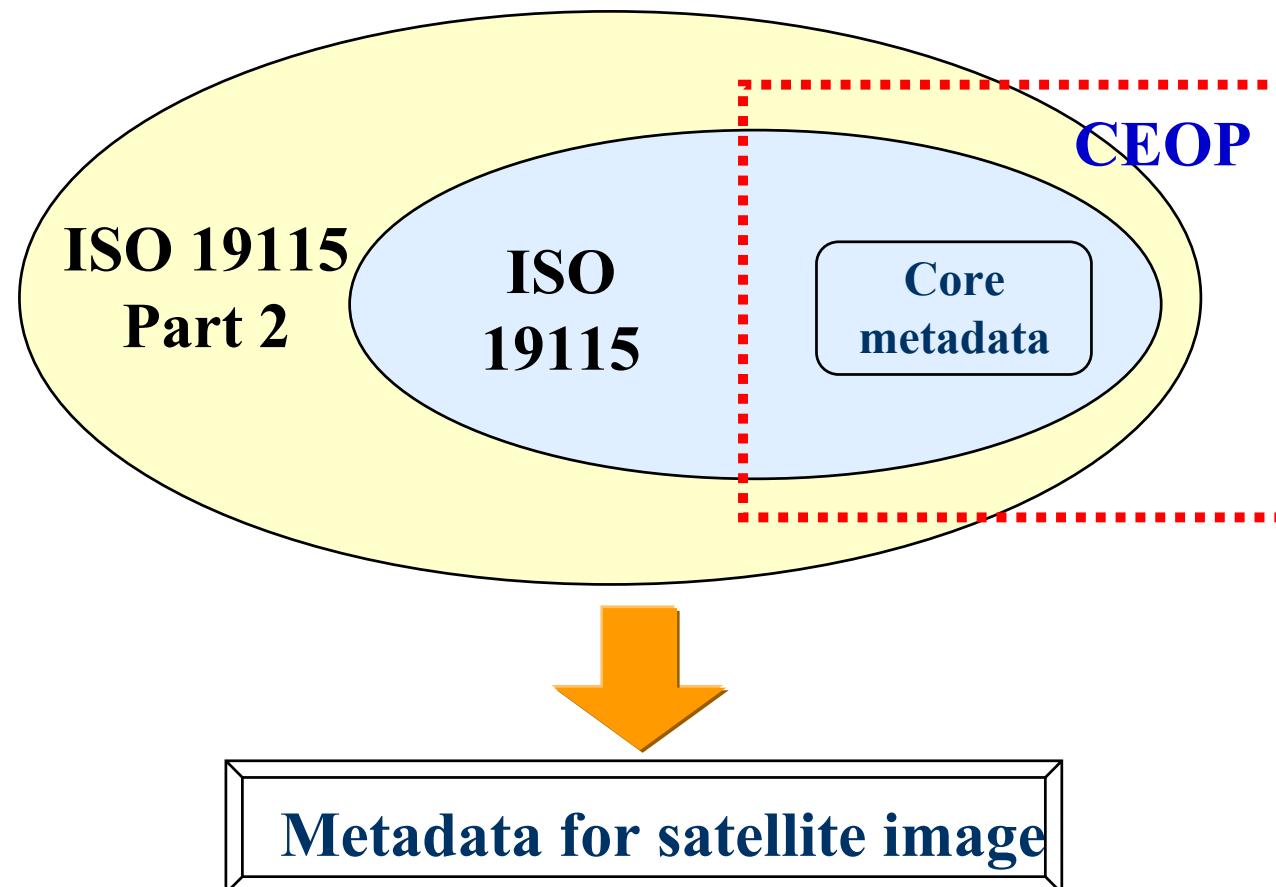
3-1 Advanced satellite image metadata development

3-2 Introduction to ISO 19115 part 2

3-3 Satellite image metadata models



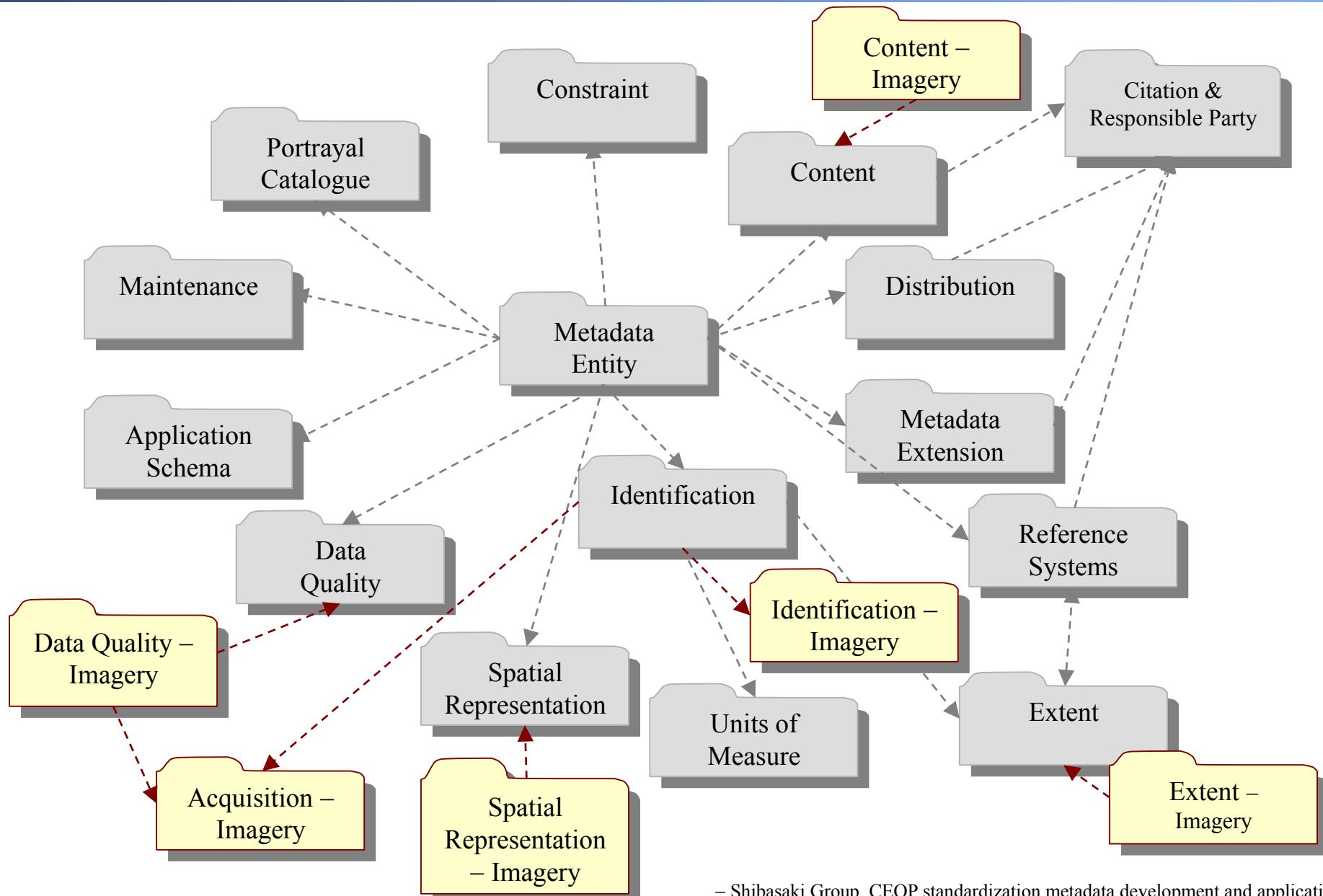
3-1 Satellite image metadata development



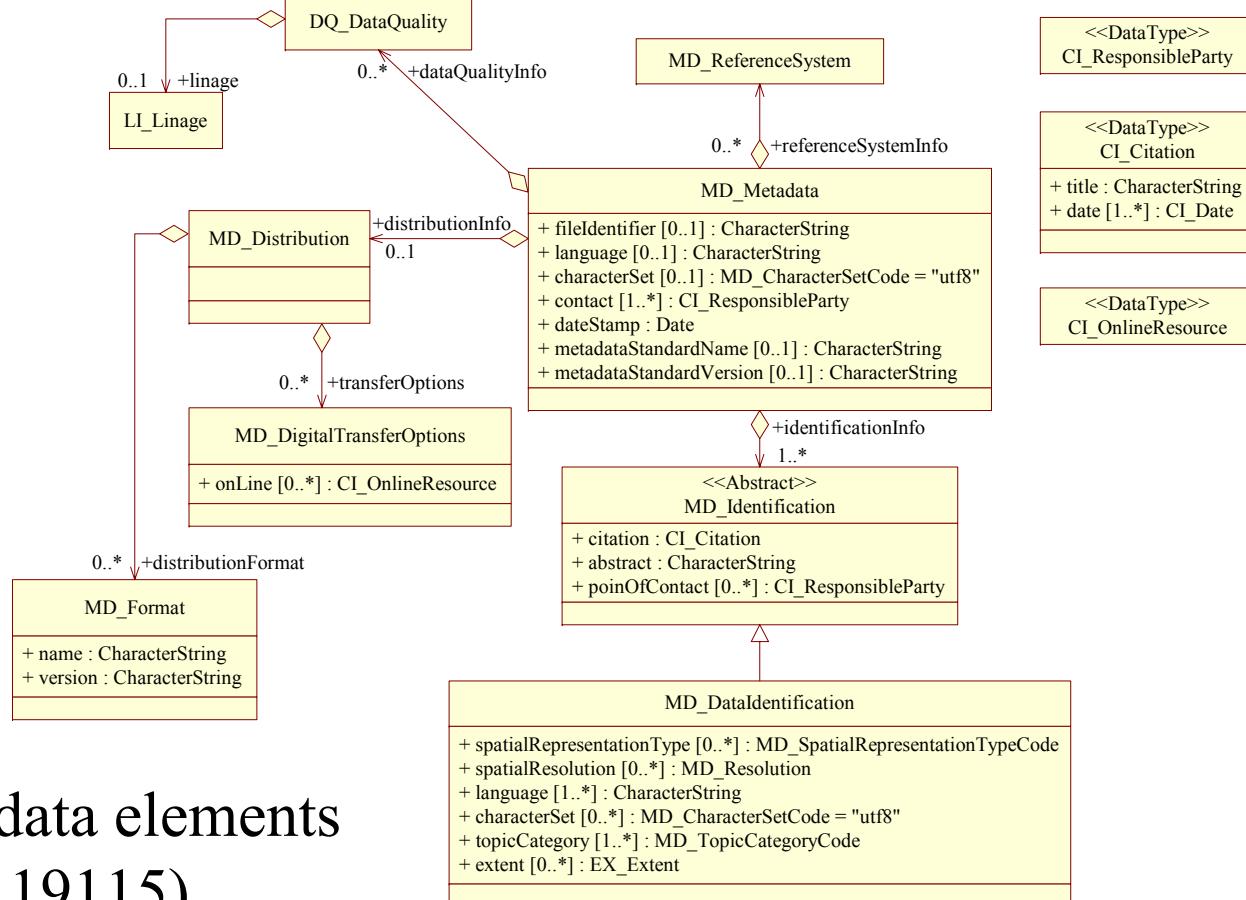
3-2 Introduction to ISO 19115 part 2

- **Title:** Geographic Information - Metadata - Part2: Extensions for imagery and gridded data
- **Scope:** This International Standard extends ISO 19115:2003 Geographic Information – Metadata by defining the schema and additional metadata required for imagery and gridded data

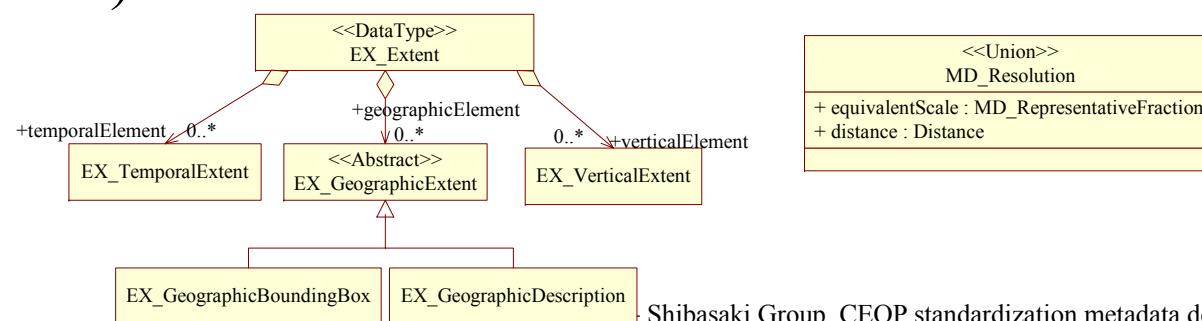
Metadata packages and its extension



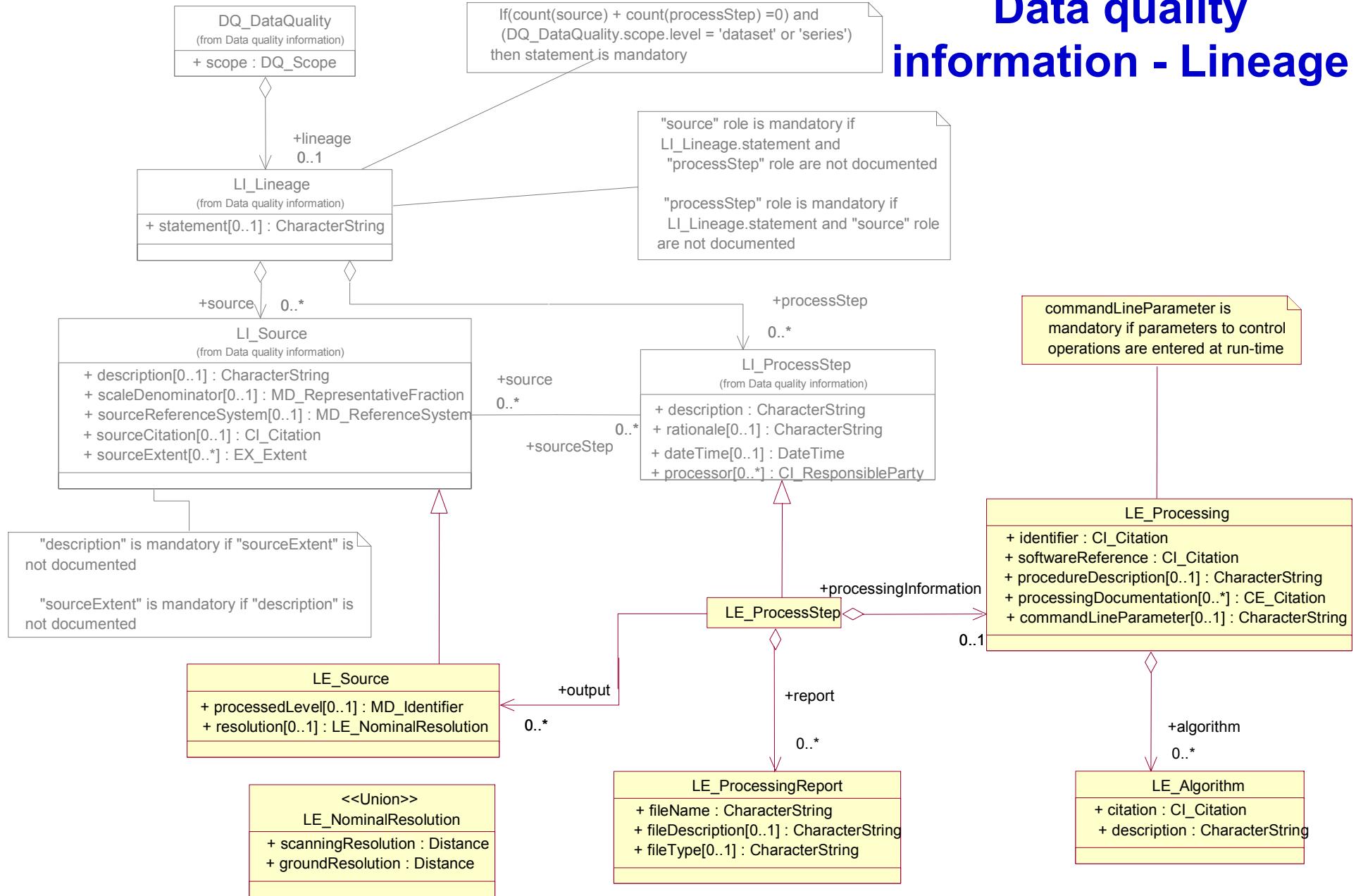
3-3 Satellite image metadata models

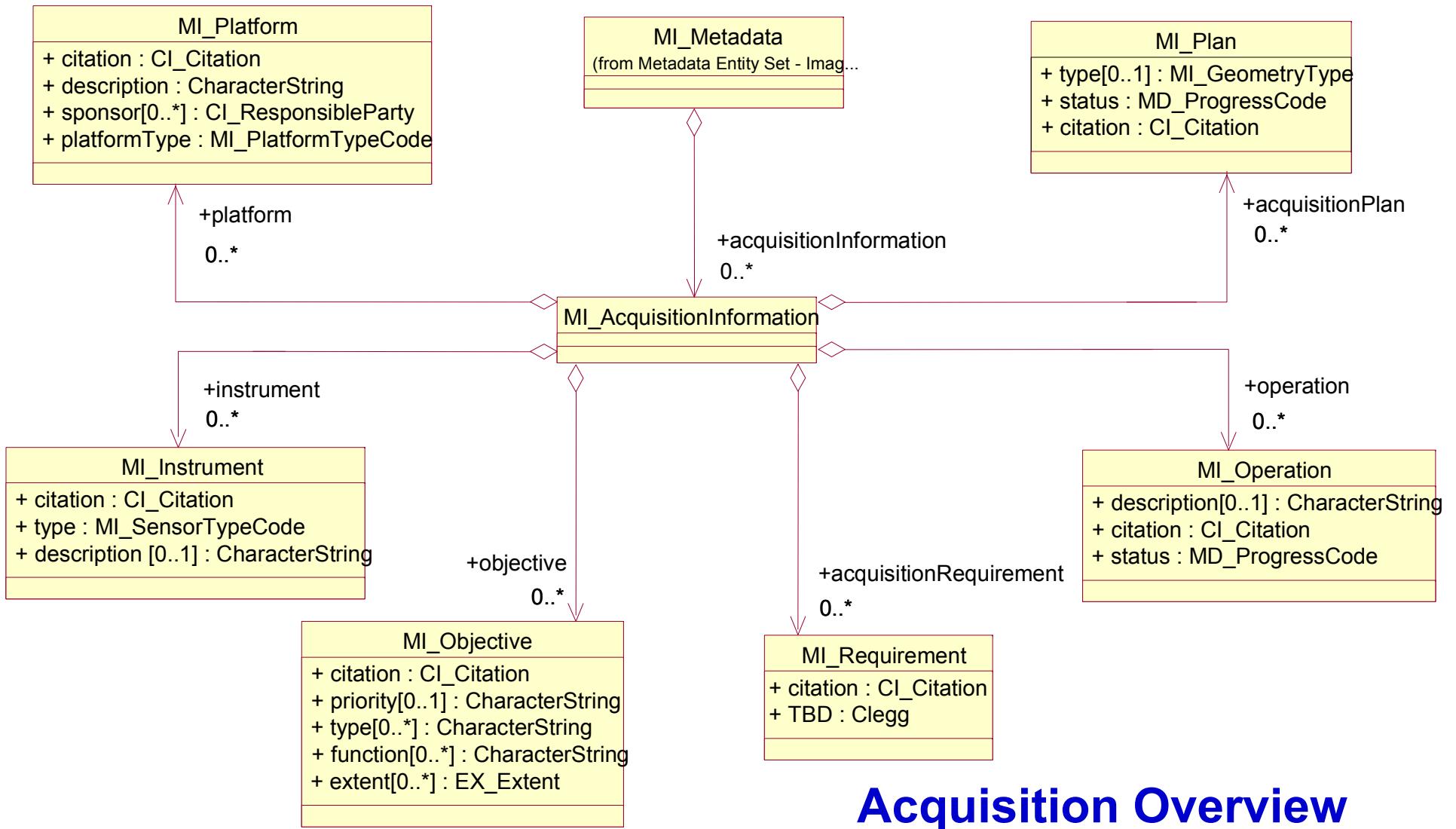


Core metadata elements
(from ISO 19115)



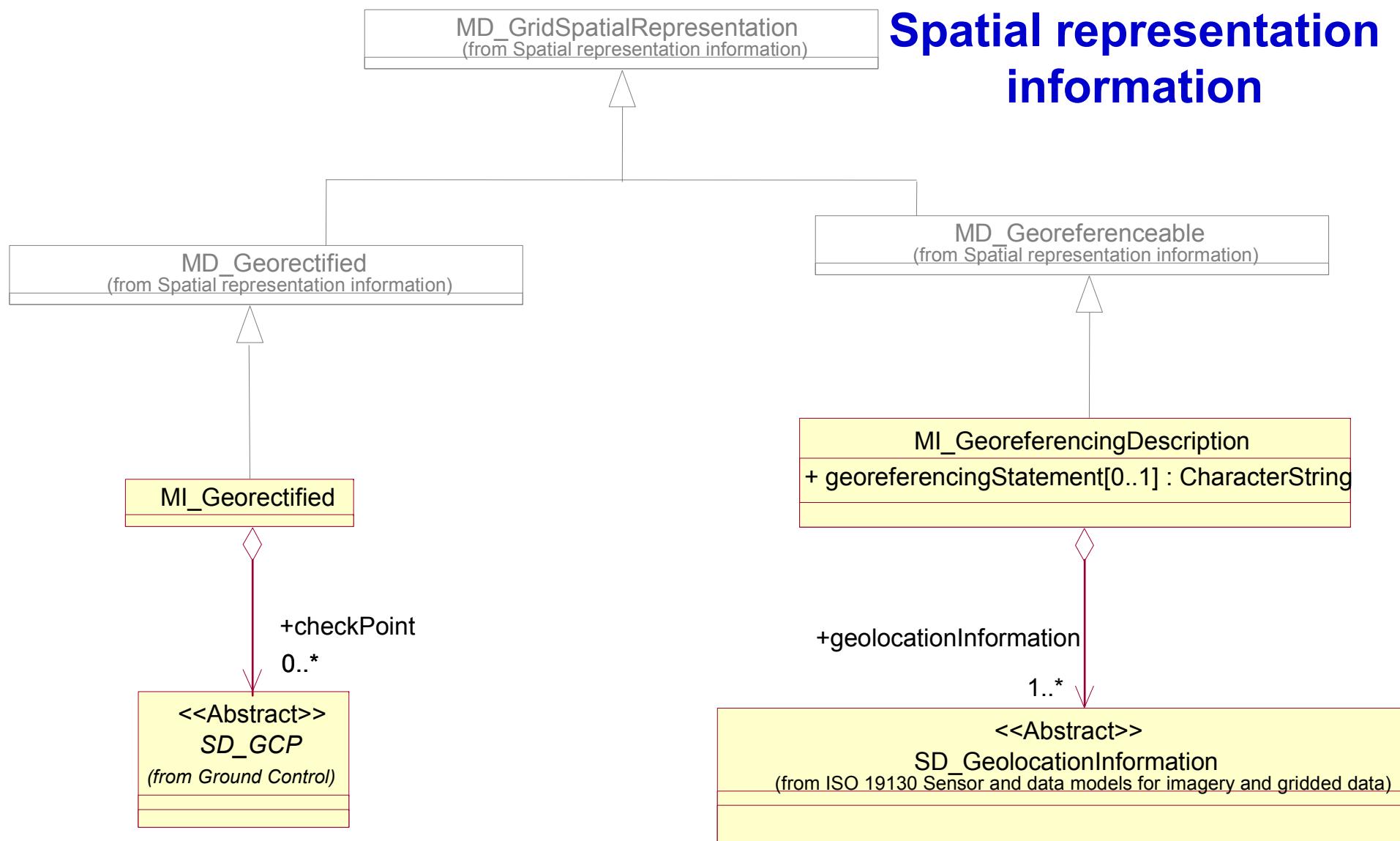
Data quality information - Lineage

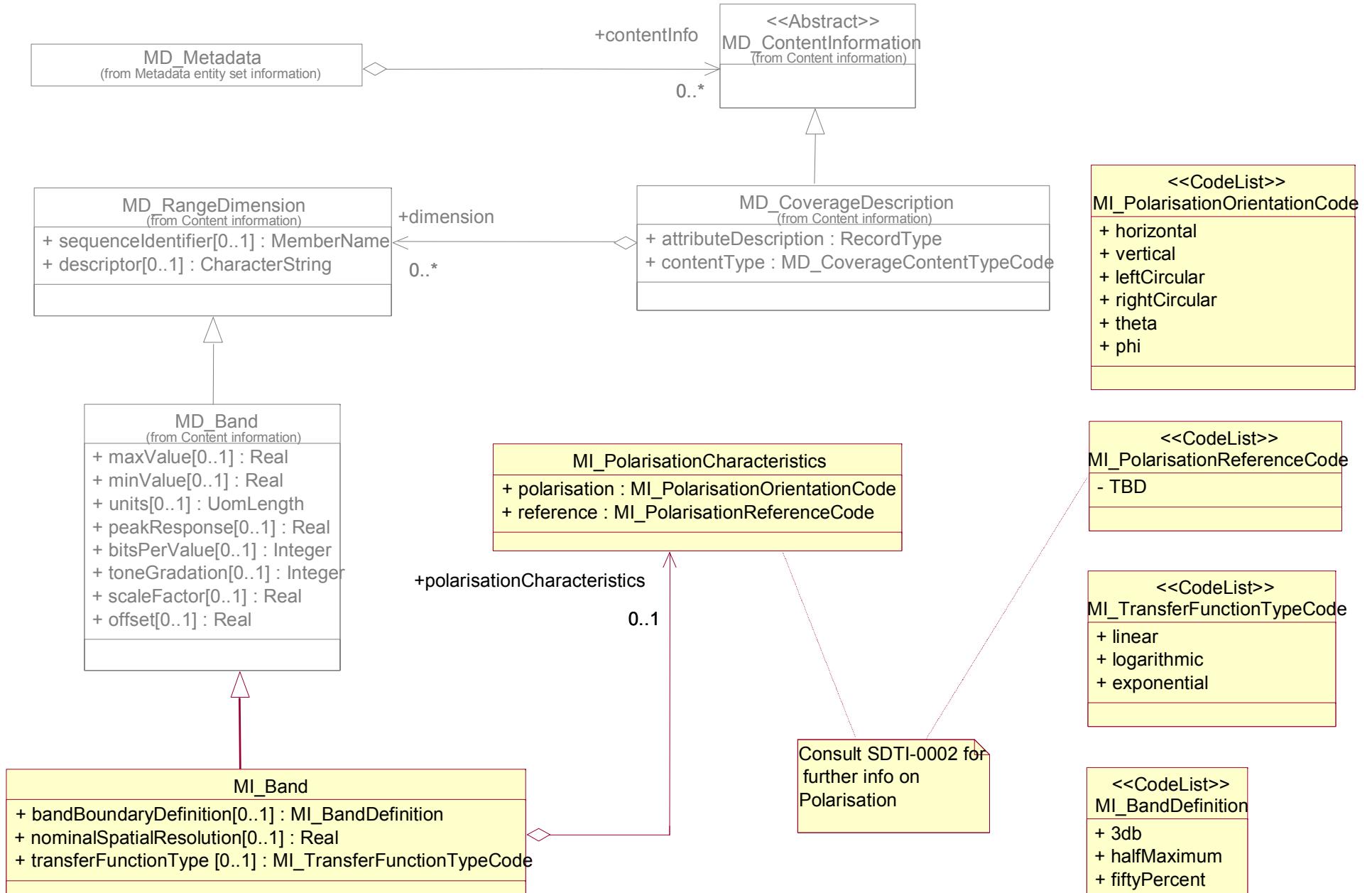




Acquisition Overview

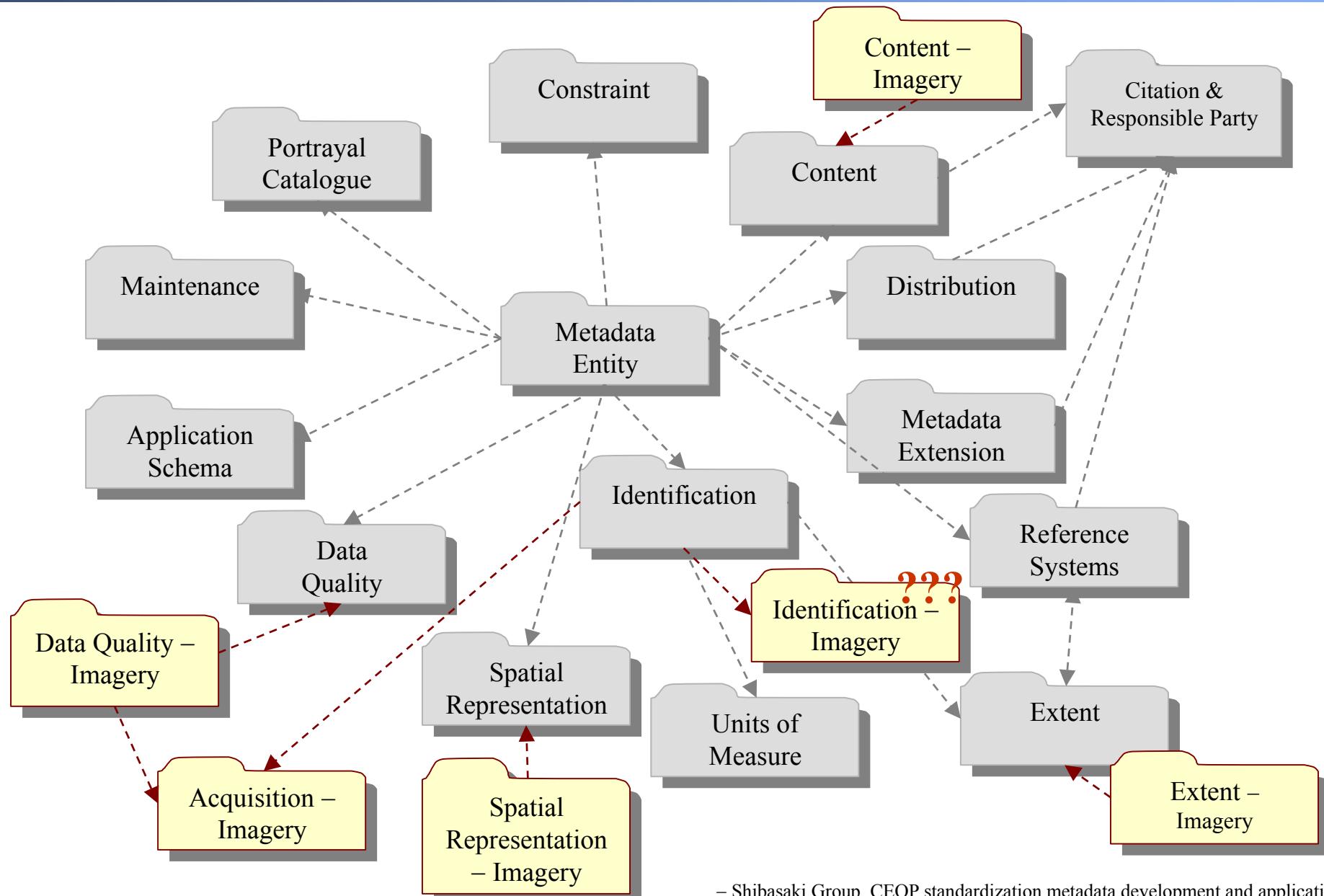
Spatial representation information





Content information

Metadata packages and its extension



Outline of the section

4. Conclusions

4 Conclusions

- In the first phase, we propose a **framework** for CEOP metadata development and application.
- We develop **satellite imagery** metadata list, **reference site** metadata list, **MOLTS** metadata list based on ISO 19115.
- In the second phase, we are revising CEOP satellite image metadata and image processing service based on **new ISO 19115-2** (extensions for **imagery/gridded** data).
- The proposed metadata can provide a basis for the discovery and the integration of various CEOP data, like satellite imagery data, reference site data and model simulation results (**GRIB, MOLTS**).

The end of the presentation

*Thank you very much
for your attention !*