



INSPIRE compliant international standards for the SeaDataNet marine metadata

Enrico Boldrini¹, Dick Schaap², Stefano Nativi¹

ESSI-Lab of CNR-IIA¹, Maris²

23rd September 2013

IMDIS 2013

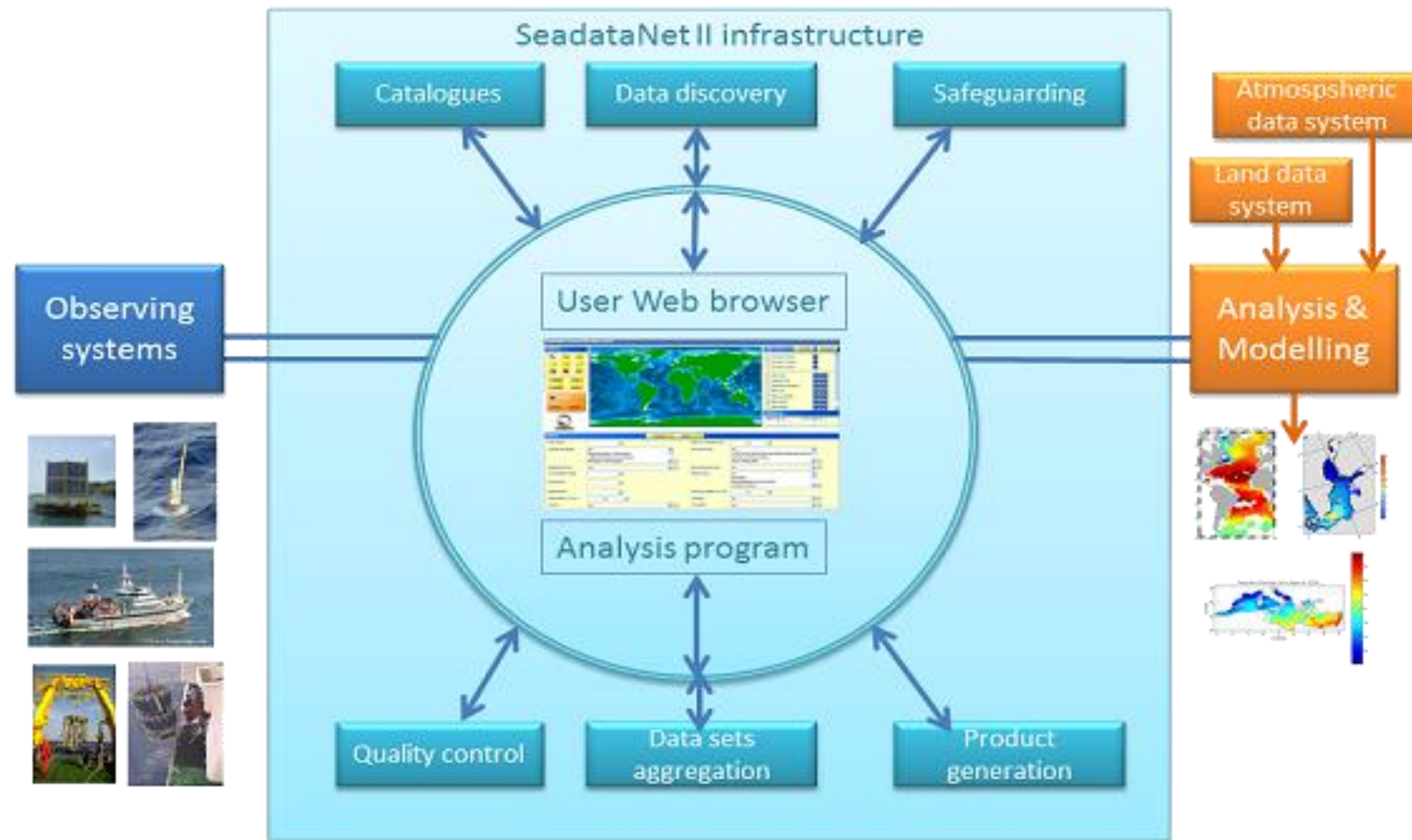
Lucca, Italy

enrico.boldrini@cnr.it

EU SeaDataNet project



- Marine data management infrastructure:
 - A Pan-European network for marine data
 - 29 countries, 88 national data centers

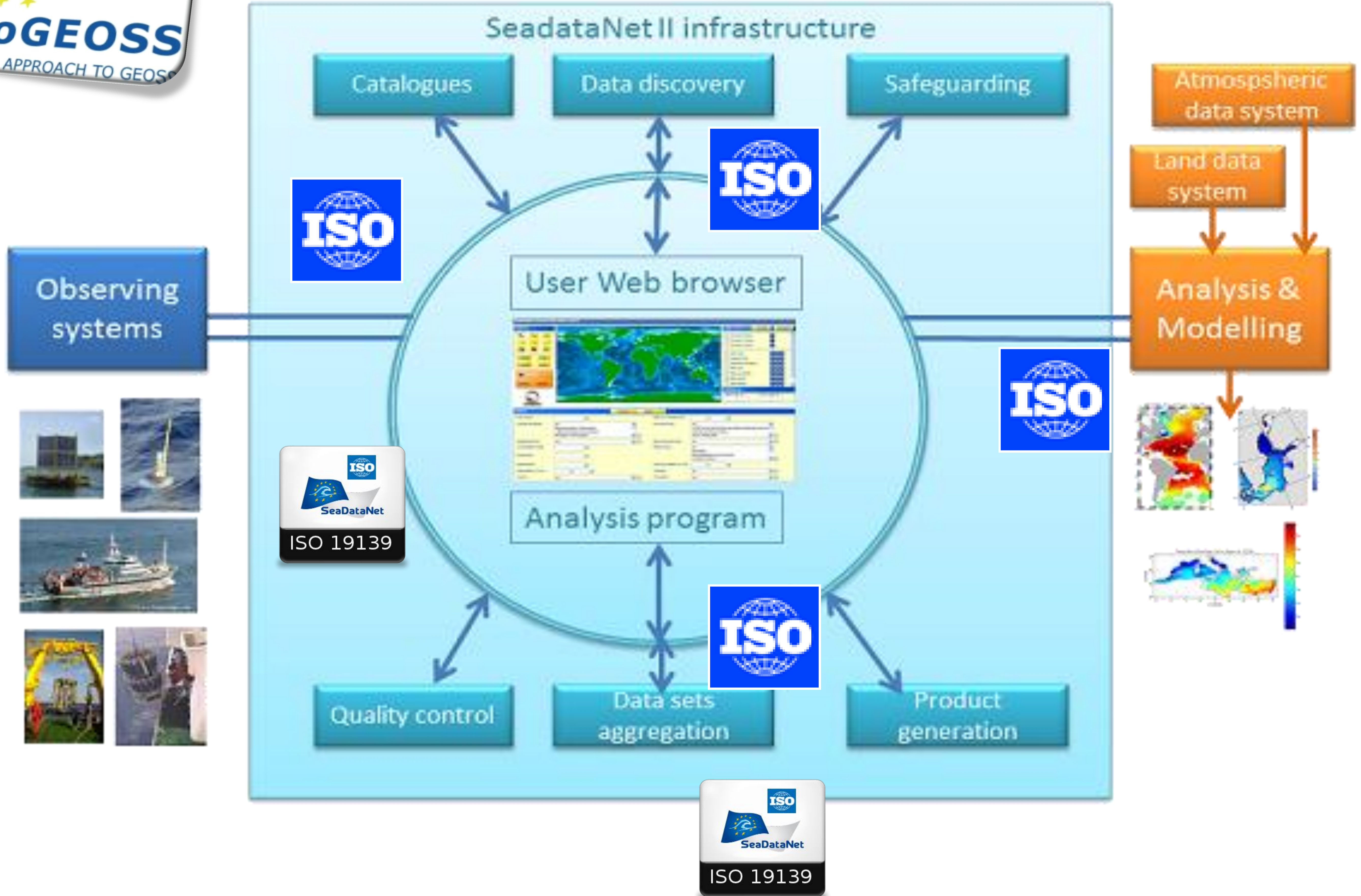


- Different tools:
- Meta data editor
 - Central portal (discovery & access)
 - Catalogues
 - Visualization
 - ...

EU FP7 SeaDataNet (2nd phase)



- Infrastructure upgrade
 - Setting, adoption, promotion of international standards (e.g. ISO compliant metadata)
 - Outreach to international communities
 - Technical and semantic interoperability with existing (and even future) tools

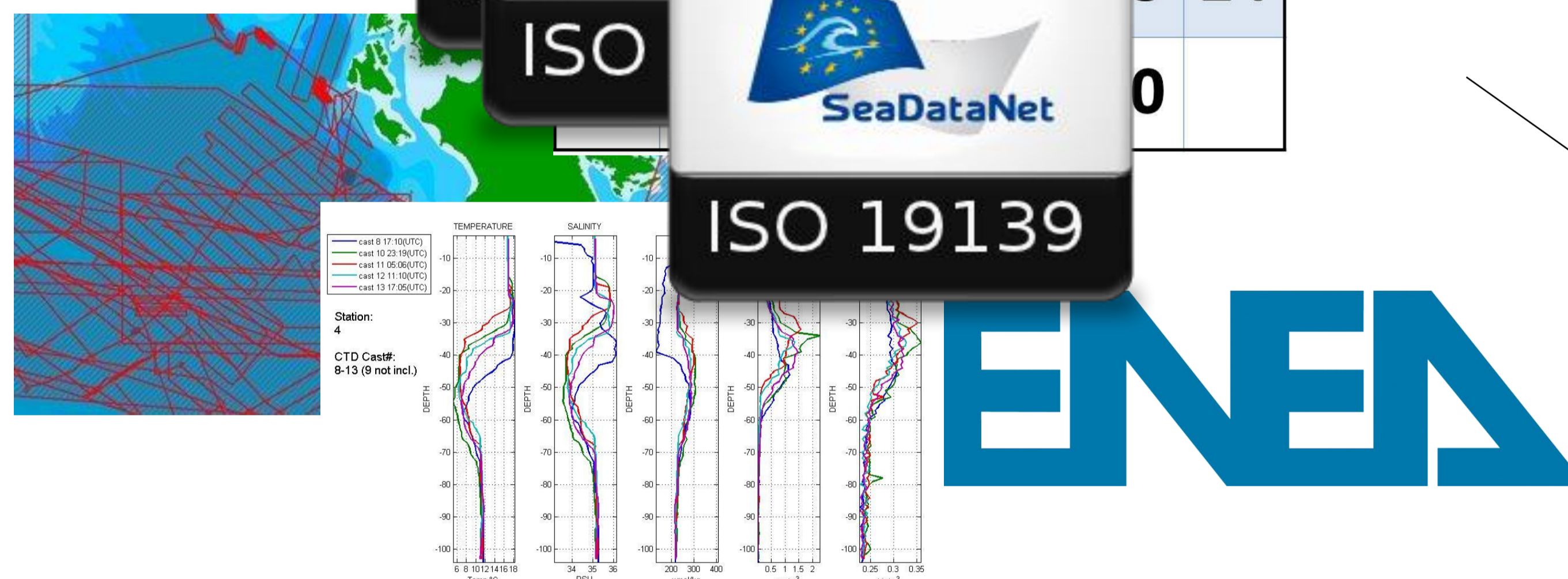
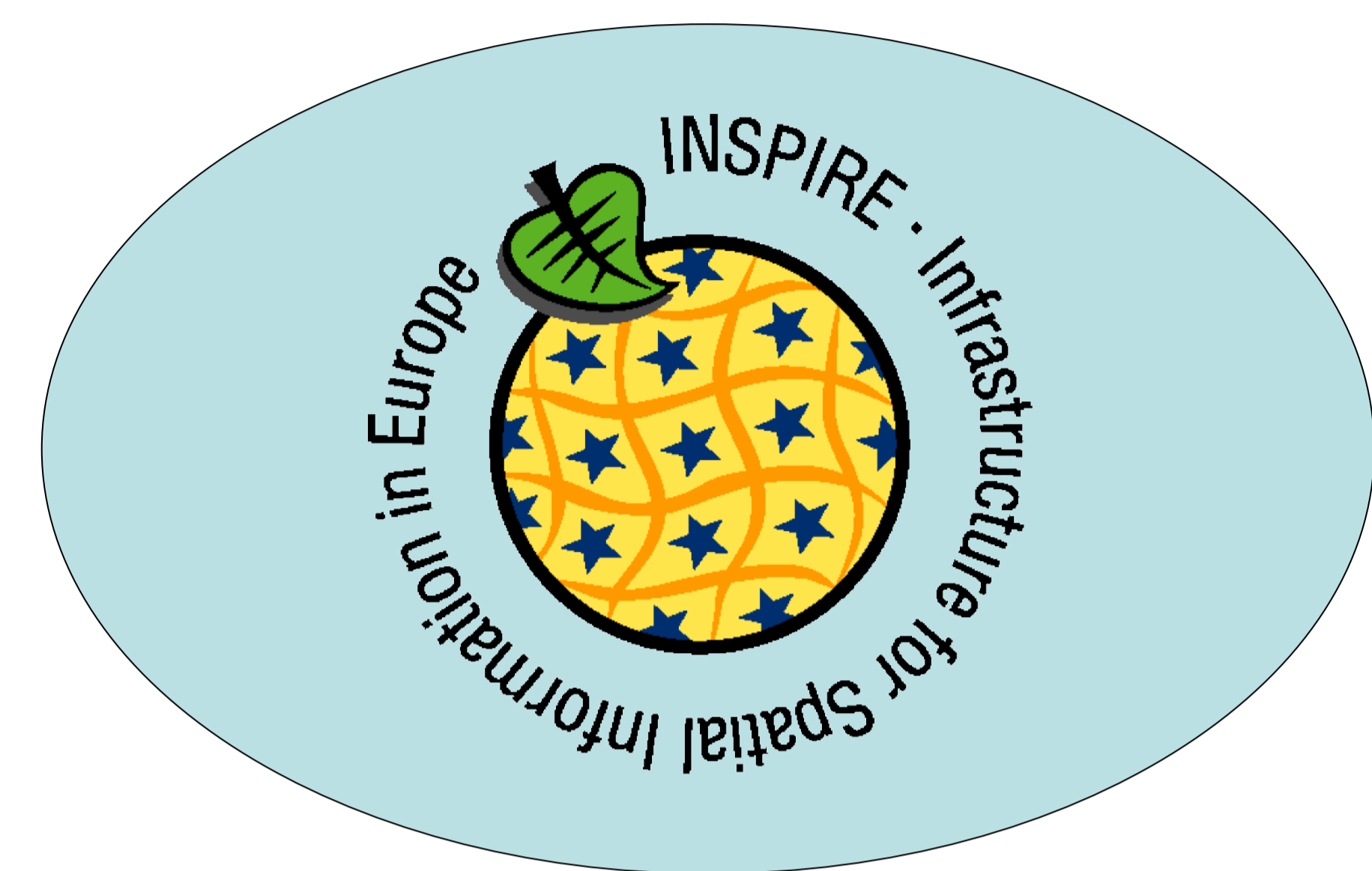




ISO Geographic metadata

Interoperability with International communities:

- Semantic int. (data model definition)
- Technology int. (software, tools, encoding)



...

Discovery



What?
Title: CTD casts

When? Time extent: 11/12

NOVEMBER 2012						
SUN	MON	TUE	WED	THU	FRI	SAT
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

ISO Geographical metadata



Evaluation

Resolution: 50 metres

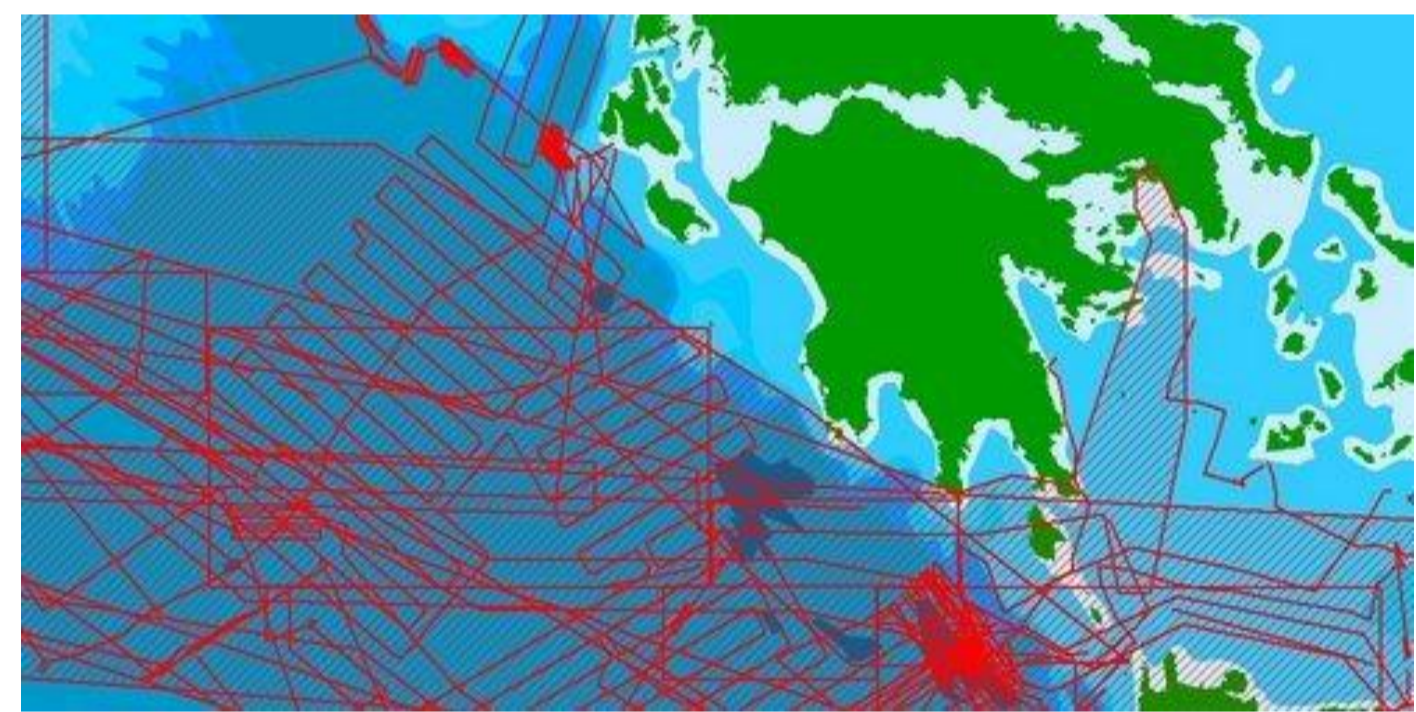
Legal constraints: none

Lineage

Access & use

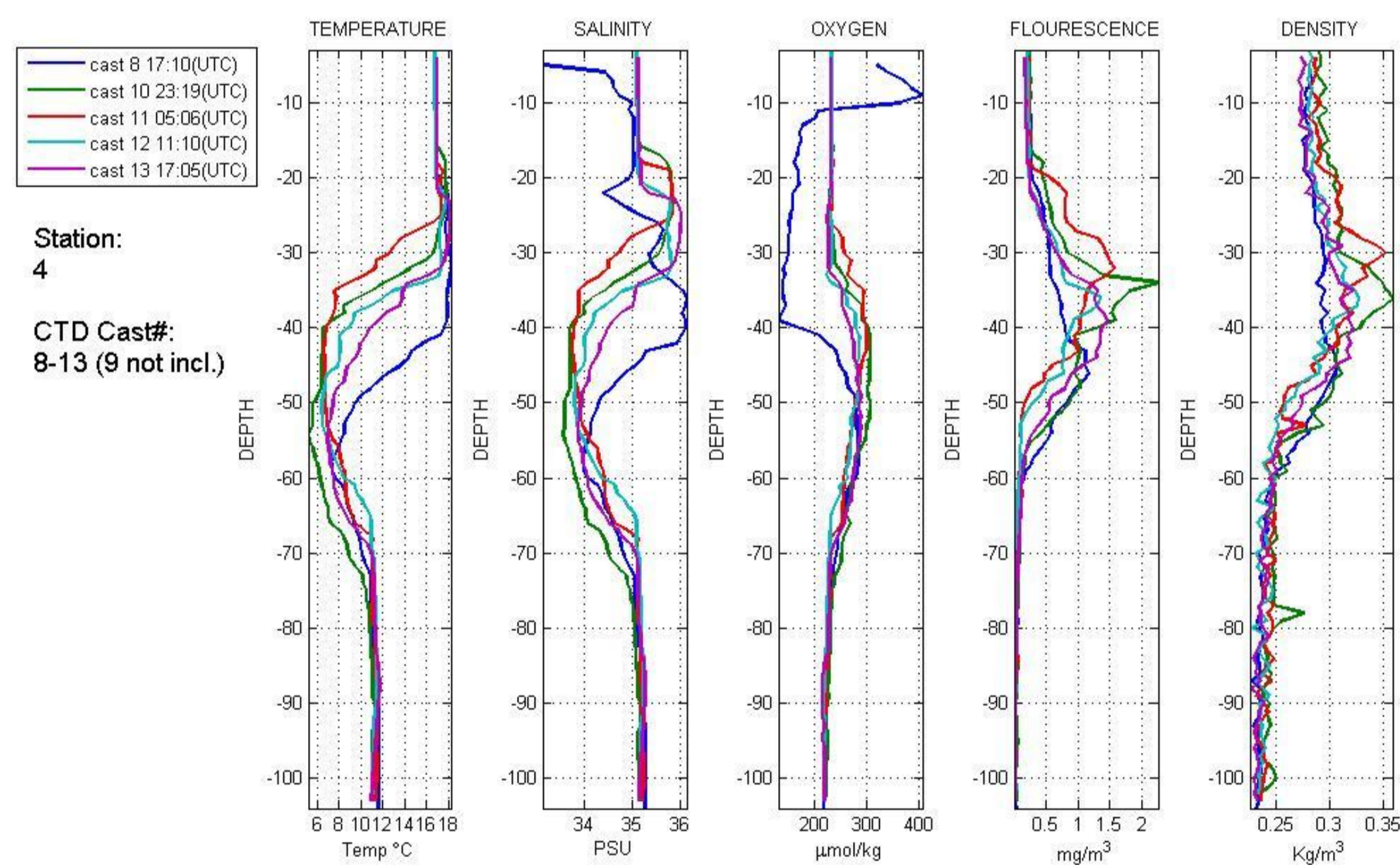
Online distributor: endpoint

Online resource protocol: FTP



Where?

Geo area: Ionian sea



Dataset (e.g. NetCDF)

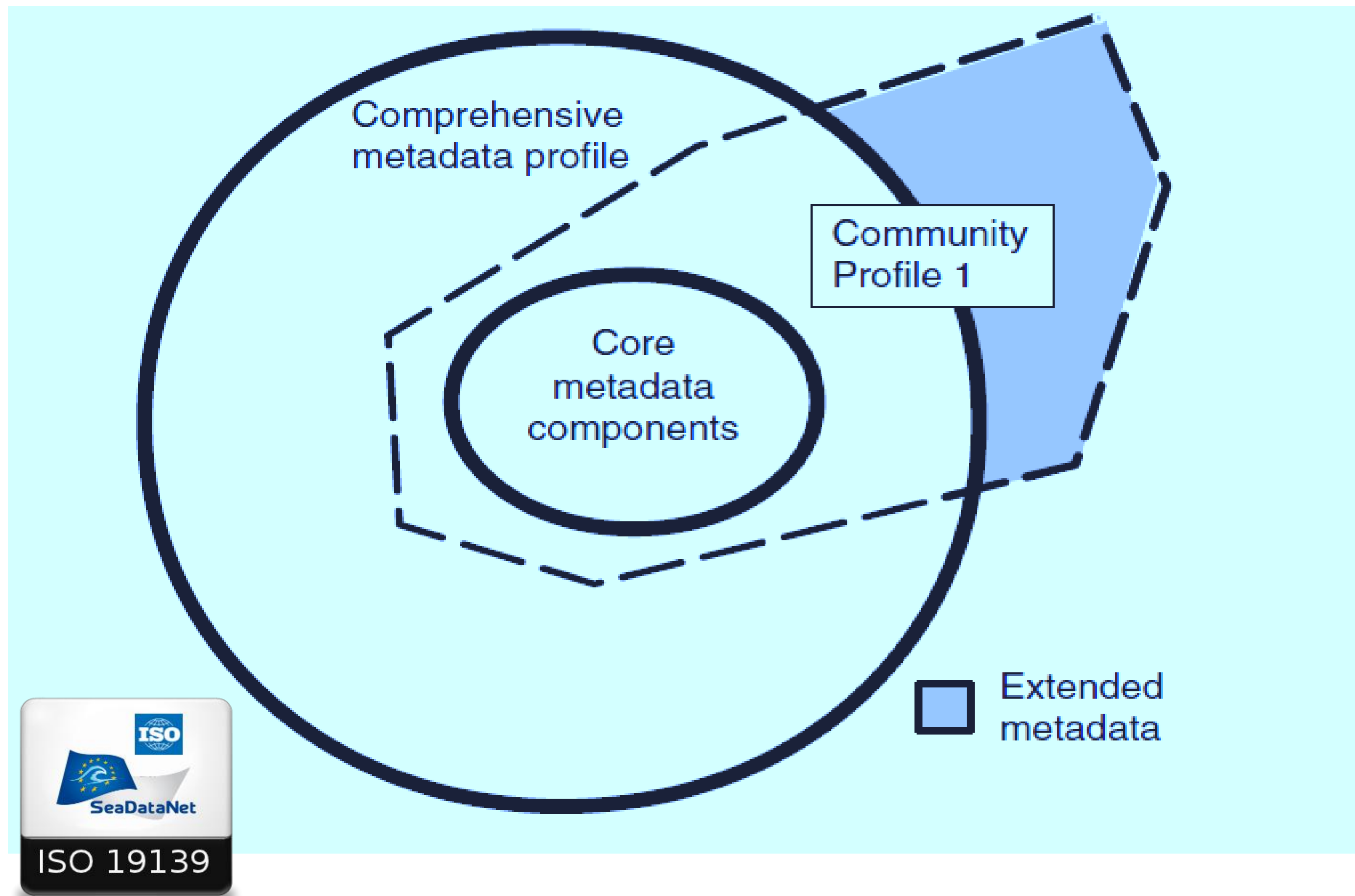


Who?

Organisation: ENEA

400+ metadata elements ...
 ISO 19115 -> Data model
 ISO 19139 -> XML encoding

Community metadata profile



- A subset of the ISO metadata elements (mandatory elements must be included)
- Extended metadata elements (e.g. community codelists)
- Additional constraints amongst elements (e.g. for INSPIRE)

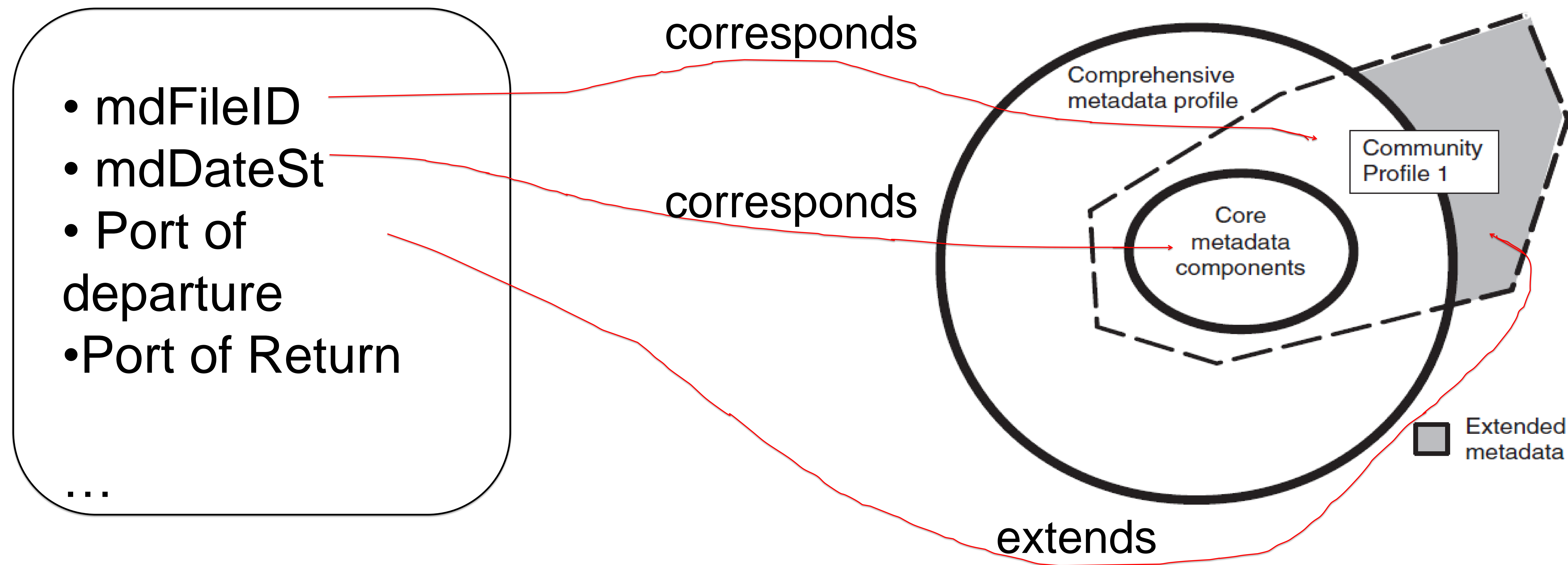
Reference : ISO 19115 Annex C (Metadata extensions and profiles)

23rd September 2013, IMDIS 2013

enrico.boldrini@cnr.it

Seadatanet CSR Metadata profile of ISO 19115

All original CSR elements must fit in the ISO 19115 metadata model!



• CSR Metadata Model

□ ISO 19115 Geographic metadata model

SeaDataNet ISO metadata profiles



Common Data Index

✓ v.10 Stable, in use

23rd September 2013, IMDIS 2013



Cruise Summary Report

✓ v. 3 Available for comments

enrico.boldrini@cnr.it

Common Data Index (CDI) METADATA

1 <?xml version="1.0" encoding="UTF-8"?>

2

3 Mikado 3.2 SDN V2 Manual / New CDI

4 Manual Automatic Options Tools ?

5

6 Where to find the data Cruise/Station Documentation Quality Others

7 Identification Where When What How Who

8

9 Geog seadatanet.maris2.nl/v_cdi_v2/result.asp

10 SEADATANET COMMON DATA INDEX (CDI) V2

11 Tools ?

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

Horizontal

Vertical

Depth

Minimum

Maximum

Water

Unit

SeaDataNet

Add to basket

Summary

#	Data set name	Country
<input type="checkbox"/>	77AR2010_00135_H10	Sweden
<input type="checkbox"/>	77ZZ2012_00012_H09	Sweden

Details

WHAT?

Data set name 77AR2010_00135_H10

Discipline Physical oceanography

Category Acoustics
Other physical oceanographic measurements
Water column temperature and salinity

Variables measured Density of the water column
Salinity of the water column
Sound velocity and travel time in the water column
Temperature of the water column

Abstract CTD measurements from the seas surrounding Sweden

Data format Ocean Data View ASCII input Version 0.4

Data set creation date 20130918

WHERE?

Map

Latitude 1 57.62979889

Longitude 1 8.80679989

Datum World Geodetic System 84

Measuring area type point

Water depth (m) 01

23rd September 2013, IMDIS 20

-sampling of cod and/or haddock for CERAS

ADDITIONAL INFORMATION	
Parameters measured	
Instruments used	
Linkage / Report / Station list	
PRINCIPAL INVESTIGATORS	
A	Asgeir AGLÉN - Institute of Marine Research
B	Oeivind OESTENSEN - Institute of Marine Research
C	Merete FONN - Institute of Marine Research
D	Jarle NORDEIDE - Bodø University College
E	Eva TORGILSTVEIT - National Institute of Nutrition and Seafood Research (NIFES)
F	T. JEFFS - Centre for Environment, Fisheries and Aquaculture Science, Lowestoft Laboratory

Summary of measurements and samples taken

Biology & Fisheries

PI Number	Type	Unit	Type of measurement	Description	Reference date
A 23	B14	stations	Pelagic fish	Pelagic trawl stations for species identification and biological sampling.	
A 142	B19	stations	Demersal fish	Bottom trawl stations for species identification and biological sampling.	
A 3500	B28	stations	Acoustic reflection on marine organisms	Continuous acoustic recordings by echo sounder (Simrad ER60) and echo integration (LSSS) of pelagic and demersal fish species.	
C 202	B90	samples	Other biological / fishery measurements	Gonad sampling of cod.	
D 183	B90	samples	Other biological / fishery measurements	Genetic sampling of cod (muscle).	

Contamination

PI Number	Type	Unit	Type of measurement	Description	Reference date
E 223	P13	samples	Contaminants in organisms	Sampling of fish for analysis of contamination.	
F 223	P13	samples	Contaminants in organisms	Sampling of fish for analysis of contamination.	

Physical Oceanography

PI Number	Type	Unit	Type of measurement	Description	Reference date
-----------	------	------	---------------------	-------------	----------------

OBJECTIVES

Description	Description
23 rd Sep	Annual combined acoustic and bottom trawl survey in the Barents Sea -map the distribution and estimate acoustic and bottom trawl abunda



XML
Det
GEN
Platf
Cruis
Cruis
Port
Port
Chief
Resp
LOC
Gene
Mars
Boun
Spec
Link
PRO
Proje
Coor
OBJ
Desc
ADD
Para

CDI & CSR profile products

- Documentation:

 - 🔗 **ISO 19115 metadata profile documentation**

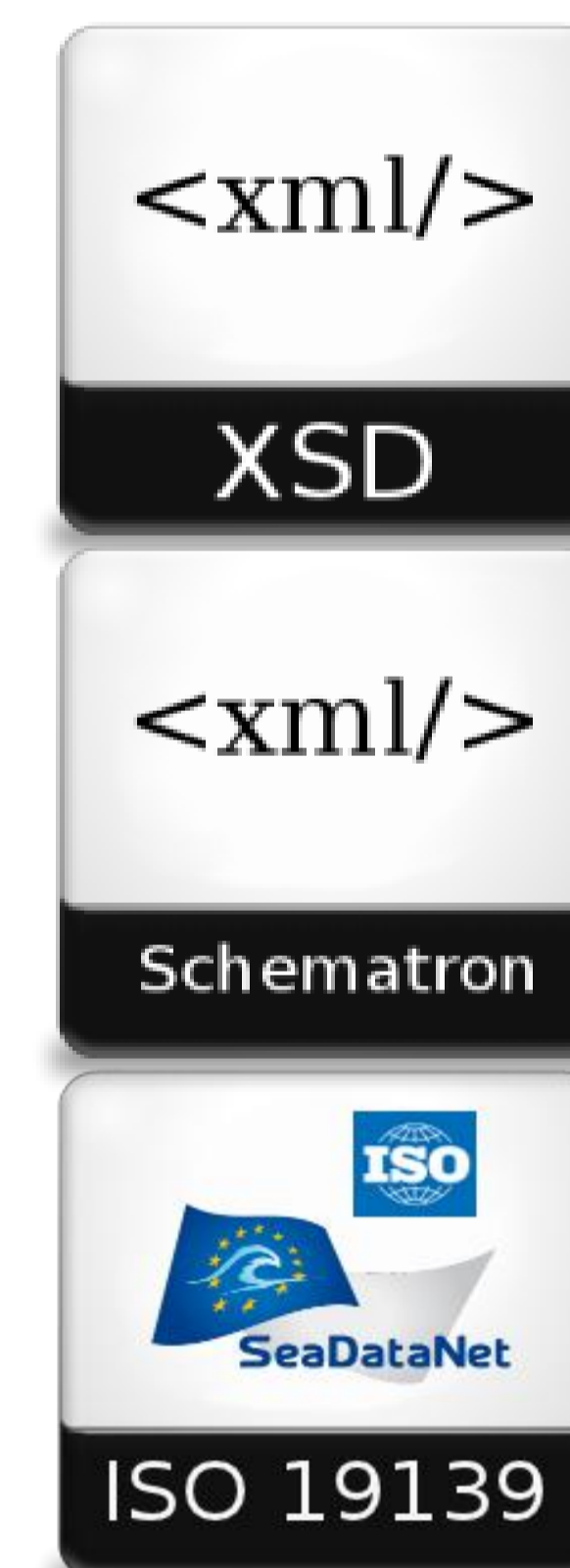
 - 🔗 **XML implementation documentation**

- XML encoding implementation:

 - ✓ **XML Schema definition**

 - ✓ **Schematron rules**

 - ✓ **Sample XML metadata**



 - ✓ **ISO 19139 compliant**



www.seadatanet.org/Standards-Software/Metadata-formats

- Tools update:
 - Metadata editors (e.g. MIKADO 3.2)
 - Discovery services (e.g. CSW ISO)
 - Central portal, CSR Inventory portal



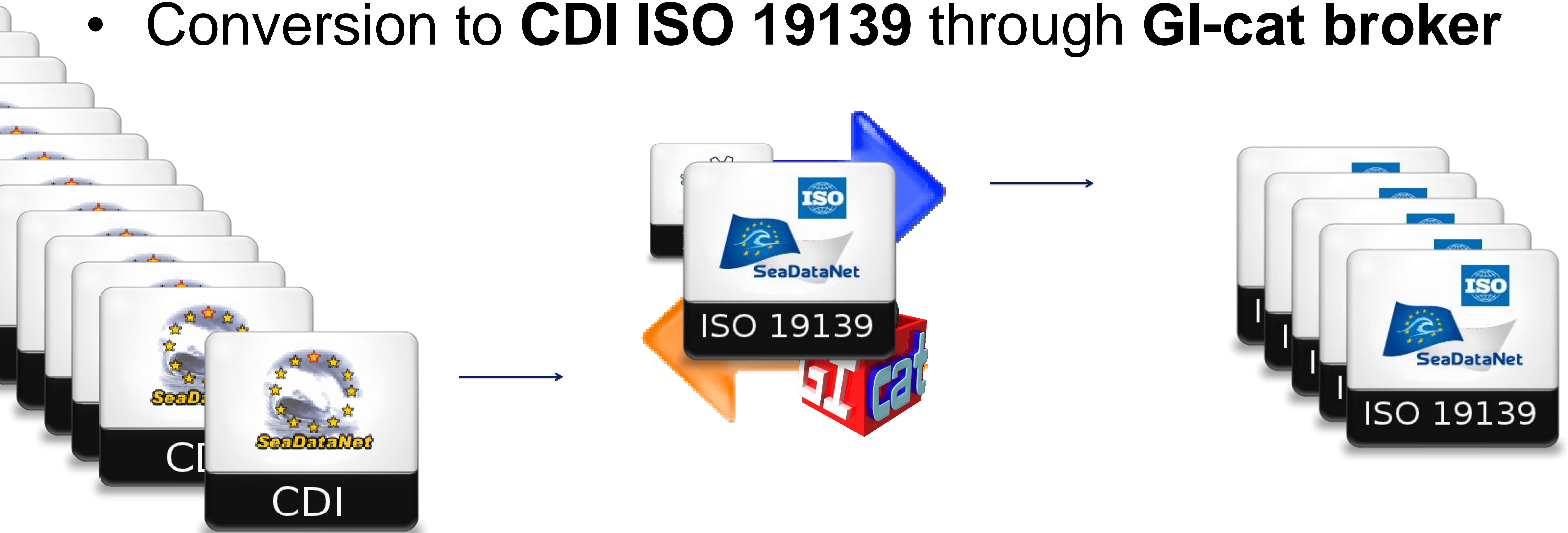
SeaDataNet Cruise Summary Report Inventory (CSR)

Search	
Free Search	Cruise date from 1 Jan 1975 to 1 Jan 2014
Ship	Select the Ship:
Select country: Algeria Australia Belgium	
Ocean/Sea Area	
Adriatic Sea (Mediterranean Sea), IHB28Bg Aegean Sea (The Archipelago) (Mediterranean Sea), IHB28Bh Alboran Sea (Mediterranean Sea), IHB28Ab	
Discipline	Select data type(s):
Select data group: All Physical Oceanography Chemical Oceanography	
Responsible Laboratory	Select the Institute:
Select country: Albania Algeria Australia	



Legacy datasets update

- About 1.3 million legacy **CDI** entries
- About 30000 legacy **CSR** entries
- Conversion to **CDI ISO 19139** through **GI-cat broker**



Success story: GEO DAB integration

The screenshot displays the EuroGEOSS Discovery and Access Broker web application. The browser window shows the URL www.eurogeoss-broker.eu and several open tabs. The application header features the GEOSS and EuroGEOSS logos, along with the title "GEOSS DISCOVERY AND ACCESS BROKER".

The interface is divided into several panels:

- Sources selection:** A list of data sources including One Geology, South Africa Environmental Observator, eHabitat WCS, WIS GISC DWD, SeaDataNet (checked), PANGAEA, Web Accessible Folder, Webservice-Energy Catalogue, INPE, NASA GLobal Change Master Directory, GEOSS Clearinghouse, EuroGEOSS, Knossos OAI-PMH, and IODE.
- Map:** A satellite map of the Mediterranean region with a red rectangular selection box over the sea. Below the map, it shows "Search results: 102 - Elapsed time: 3 seconds".
- Search results table:**

Source	Description
MARINE)	Chemical oceanography from UNIVERSITE DE LA MEDITERRANEE (U2) / CENTRE D'OCEANOLOGIE DE MARSEILLE
	Terrestrial from UNIVERSITE DE ROUEN / LABORATOIRE DE GEOLOGIE
	Environment from UNIVERSITE DE ROUEN / LABORATOIRE DE GEOLOGIE
- Query constraints selection:** A panel for refining search results.
 - Keyword:** seadatanet
 - Location:** A text input field with a placeholder "Enter a location name (case is ignored), e.g.: europe,italy,rome,etc...".
 - Selected area:** A numeric display showing a bounding box with values: -7.926, 52.862, 28.285, and 31.065.
 - Overlaps/Contains/Disjoints:** Radio buttons for selecting the relationship type.
 - Time:** Input fields for "From:" and "To:".

Thank you for your attention!

`enrico.boldrini@cnr.it`

Legacy CDI encoding (ISO19115 DTD based)

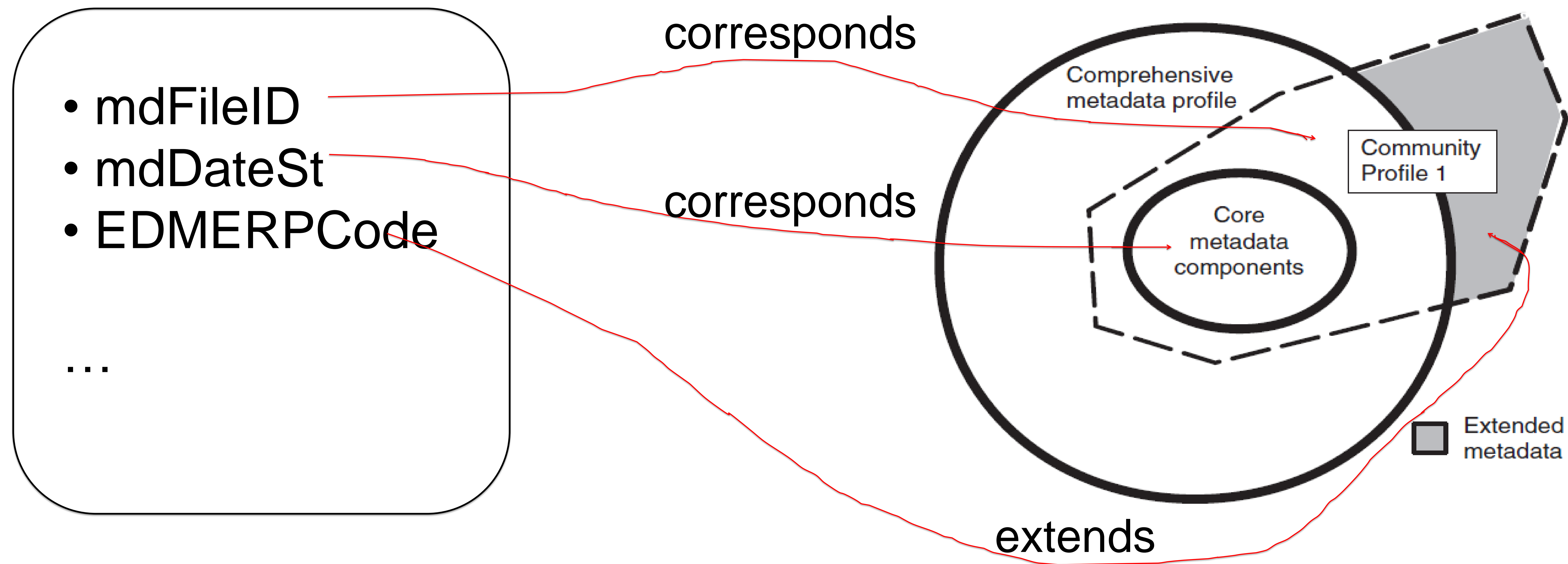
```
<?xml version="1.0" encoding="UTF-8"?>
<Metadata>
  <mdFileID>SDN:CDI:LOCAL:MARIS-TEST</mdFileID>
  <mdLang>
    <languageCode value="en" />
  </mdLang>
  <mdChar>
    <CharSetCd value="utf8" />
  </mdChar>
  <mdHrLv>
    <ScopeCd value="dataset" />
  </mdHrLv>
  <mdHrLvName SDNIdent="SDN:L231:3:CDI" >Common Data
Index record</mdHrLvName>
  <mdContact>
    <rpOrgName SDNIdent="SDN:EDM0::1" >University of
Birmingham, Department of Geological Sciences</rpOrgName>
    <rpCntInfo>
      <cntPhone>
        <voiceNum></voiceNum>
        <faxNum></faxNum>
      </cntPhone>
      <cntAddress>
        <delPoint>School of Earth Sciences Edgbaston
</delPoint>
        <city>Birmingham</city>
      </cntAddress>
    </rpCntInfo>
  </mdContact>
</Metadata>
```

CDI encoding (ISO19139 based)

```
<?xml version="1.0" encoding="UTF-8"?>
<gmd:MD_Metadata
xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <gmd:fileIdentifier>
    <gco:CharacterString>SDN:CDI:LOCAL:MARIS-
TEST</gco:CharacterString>
  </gmd:fileIdentifier>
  <gmd:language>
    <gmd:LanguageCode
codeList="../schema/resources/Codelist/ML_gmxCodeLists.xml
#LanguageCode"
codeListValue="eng">English</gmd:LanguageCode>
  </gmd:language>
  <gmd:characterSet>
    <gmd:MD_CharacterSetCode
codeList="../schema/resources/Codelist/gmxCodeLists.xml#MD
_CharacterSetCode" codeListValue="utf8"
codeSpace="ISOTC211/19115">utf8</gmd:MD_CharacterSetCode>
  </gmd:characterSet>
  <gmd:hierarchyLevel>
    ...
  </gmd:hierarchyLevel>
</gmd:MD_Metadata>
```

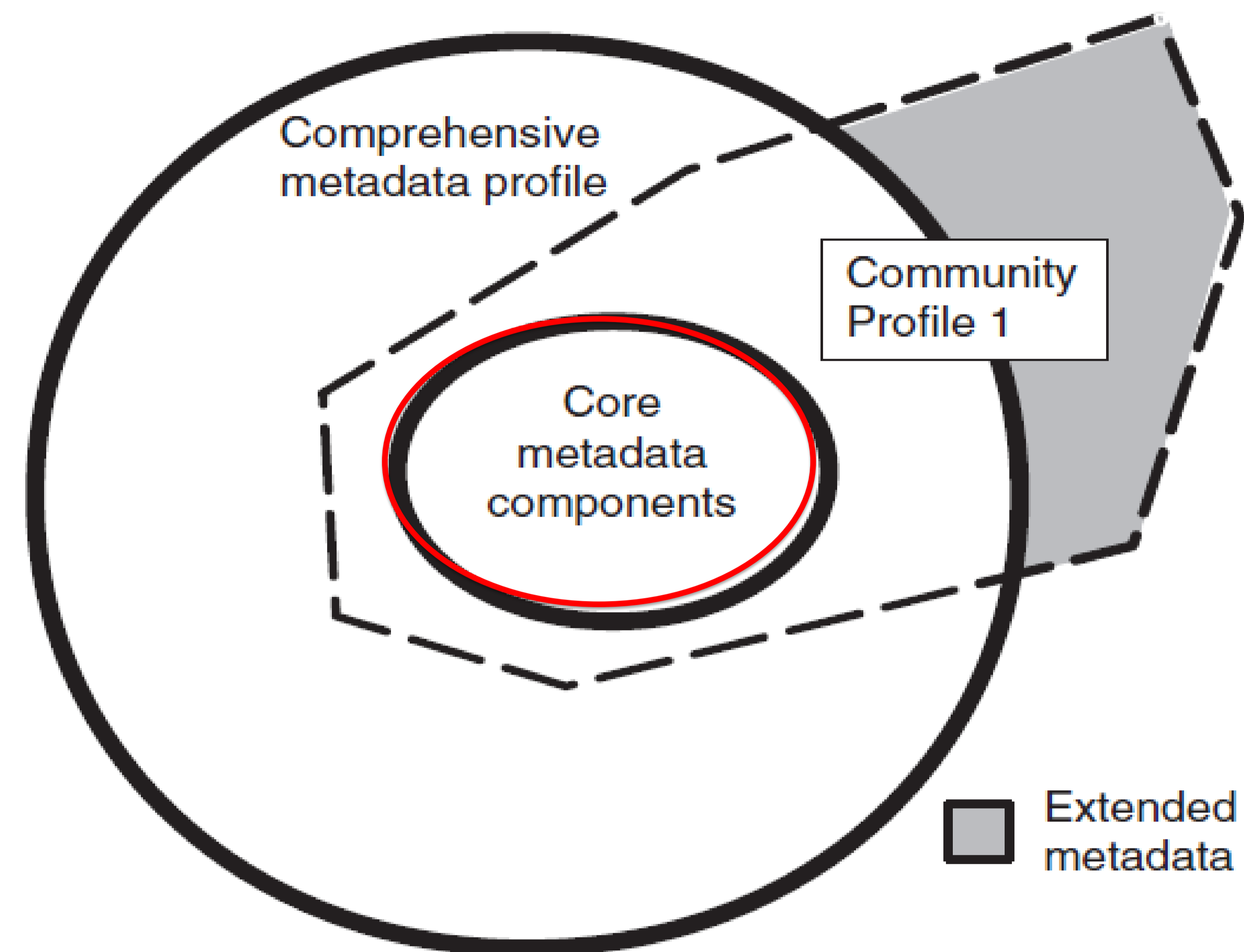

CDI Metadata profile of ISO 19115

All the CDI elements must fit in the ISO 19115 metadata model!



- CDI Metadata Model
- ISO 19115 Geographic metadata model

CDI Metadata profile of ISO 19115



- **Mandatory (M) requirements** in ISO 19115 **shall remain M** in the profile
- **Options (O)** in ISO 19115 may remain O or may be changed to:
 - mandatory (M)
 - conditional (C)
 - out of scope
 - prohibited

ISO 19115 Annex C (Metadata extensions and profiles) is the reference for the allowed extensions and methodology.

CDI profile products

- Abstract metadata **model** specification (7.0):

- ✓ SeaDataNet metadata profile of **ISO 19115** documentation



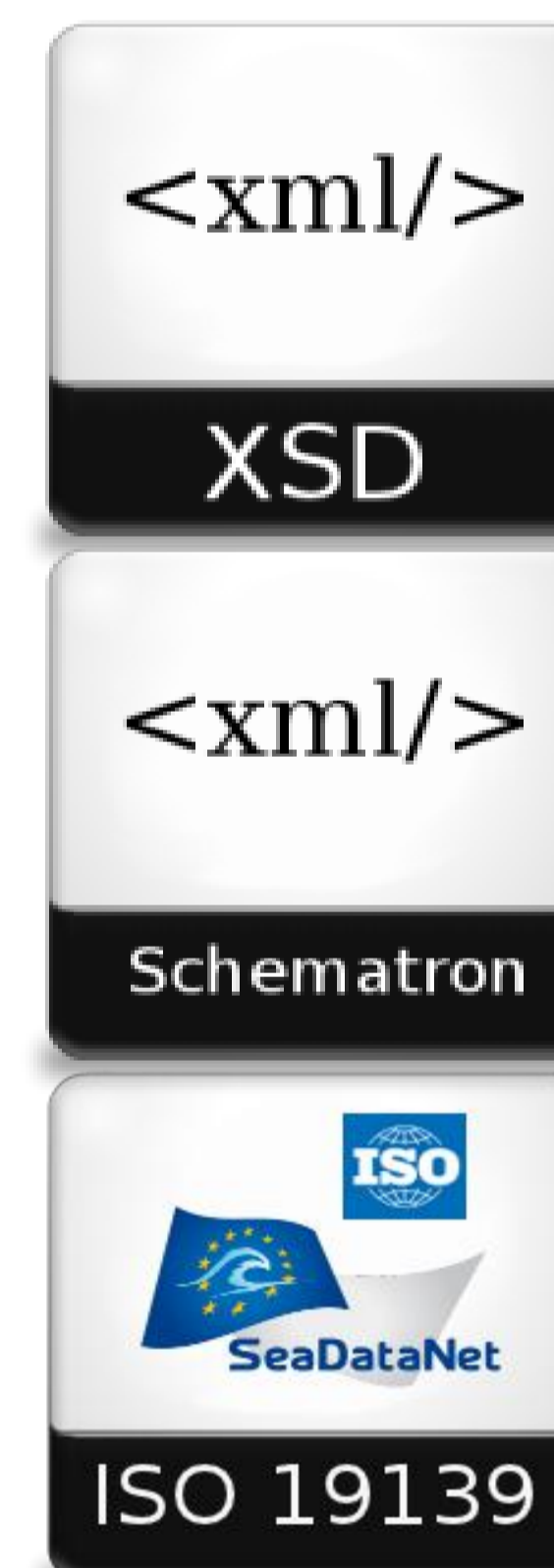
- XML **encoding** implementation (7.0):

- ✓ Schema definition

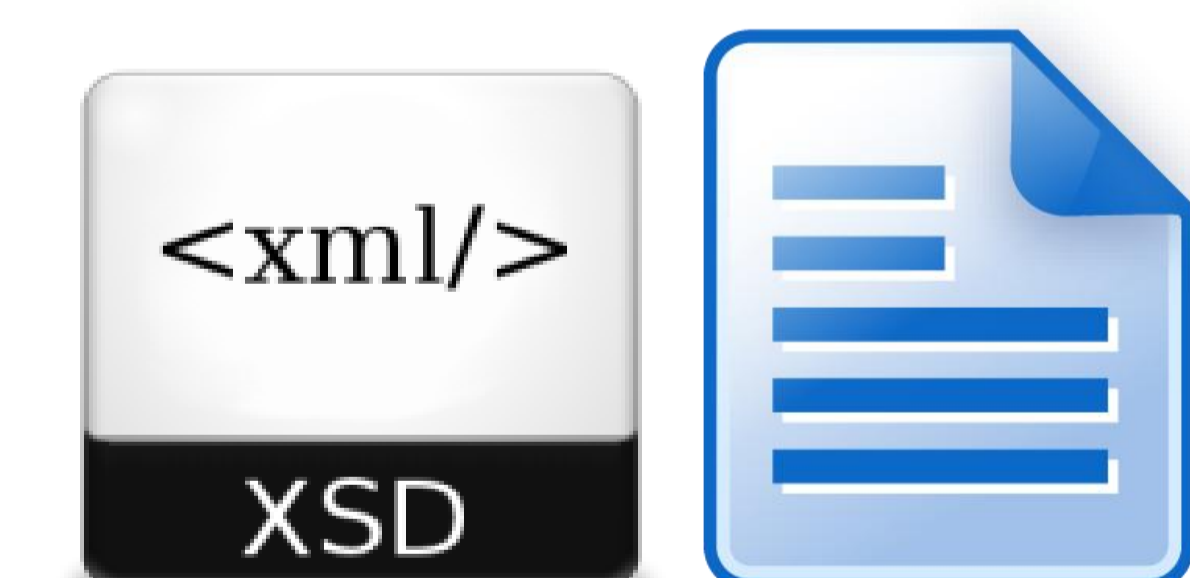
- ✓ Schematron rules

- ✓ Sample metadata

- ✓ XML implementation documentation



- ✓ **ISO 19139** compliant



www.seadatanet.org/Standards-Software/Metadata-formats

Metadata profile documentation



Data dictionary excerpt

	Name / Role name	Definition	Obligation / Condition	Maximum occurrence	Data type	Domain	Comment
1.	MD_Metadata	root entity which defines metadata about a resource or resources	M	1	Class	Lines 18-49	
2.	fileIdentifier	unique identifier for this metadata file	O -> M	1	CharacterString	Free text -> urn as defined in RFC 1737 and starting with the string "urn:SDN:CDI:"	SeaDataNet requires one file identifier urn, starting with the default string urn:SDN:CDI:. ISO has Optional fileIdentifier.

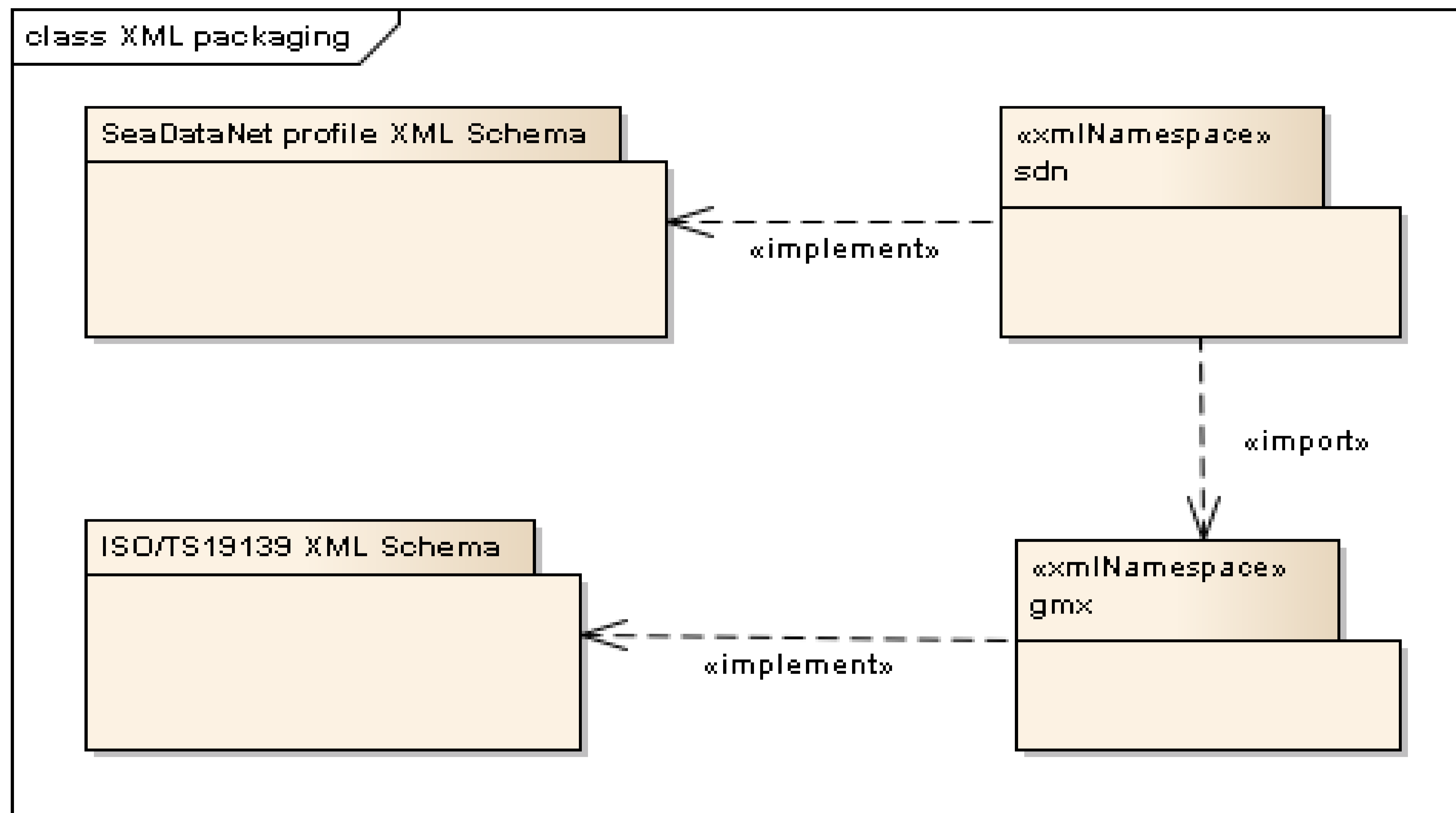
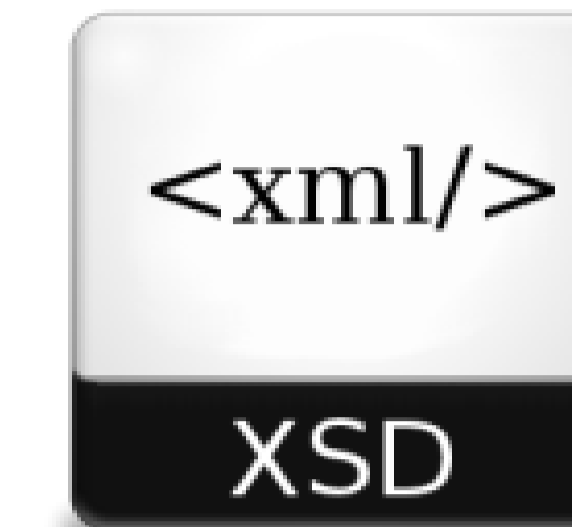
- **all the profile elements** are listed in the data dictionary
- Changes from ISO 19115 **are highlighted in red**
 - Changes in obligation (e.g. **identifier** from optional to mandatory)
 - Changes in cardinality (e.g. **contact** from 1..N to 1..1)
 - Changes in domain (e.g. **format** from free text to restricted to list)

Metadata profile documentation



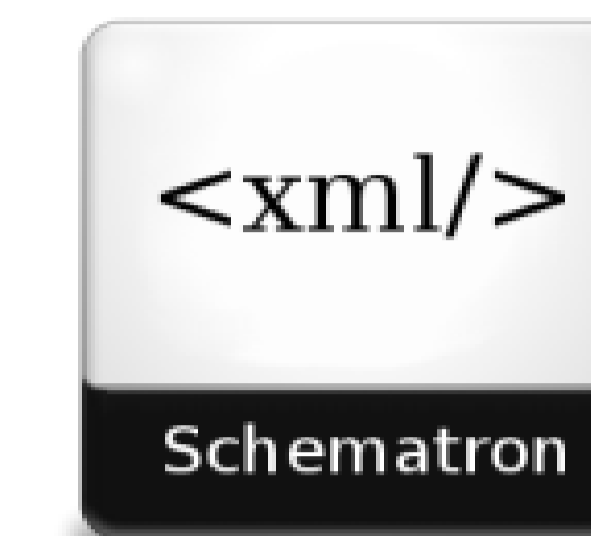
- **New elements definitions (e.g. codelists)**
 - New codelists (e.g. *format*)
 - Expanded codelists (e.g. *keyword type*)
 - Restricted codelists (e.g. *character set encoding*)
- **Additional constraints definitions**
 - **INSPIRE constraints** (e.g. at maximum one “creation” date, conformance result report, ...)
 - **SeaDataNet constraints** (e.g. at least one keyword with type ‘parameter’ should be documented)

XML encoding: Schema definition



- Imports:
 - **ISO 19139 XML schema (CSW AP ISO version)**
- Includes:
 - Definitions of **new SDN codelists**:
 - SDN_FormatNameCode
 - SDN_HierarchyLevelNameCode
 - SDN_DeviceCategoryCode
 - SDN_PlatformCategoryCode
 - SDN_ParameterDiscoveryCode
 - SDN_EDMERPCode
 - SDN_EDMOCCode
 - SDN_CRSCCode
 - SDN_CountryCode

XML encoding: Schematron rules



Used to enforce:

– Restrictions on elements:

- Obligation
- Cardinality
- Domain

– Correct use of null elements (e.g. `<gmd:fileIdentifier/>`)

– Conformance rules from ISO 19139 (Table A1+ extra)

– Additional Constraints:

- **INSPIRE**
- SeaDataNet



Codelist catalogue

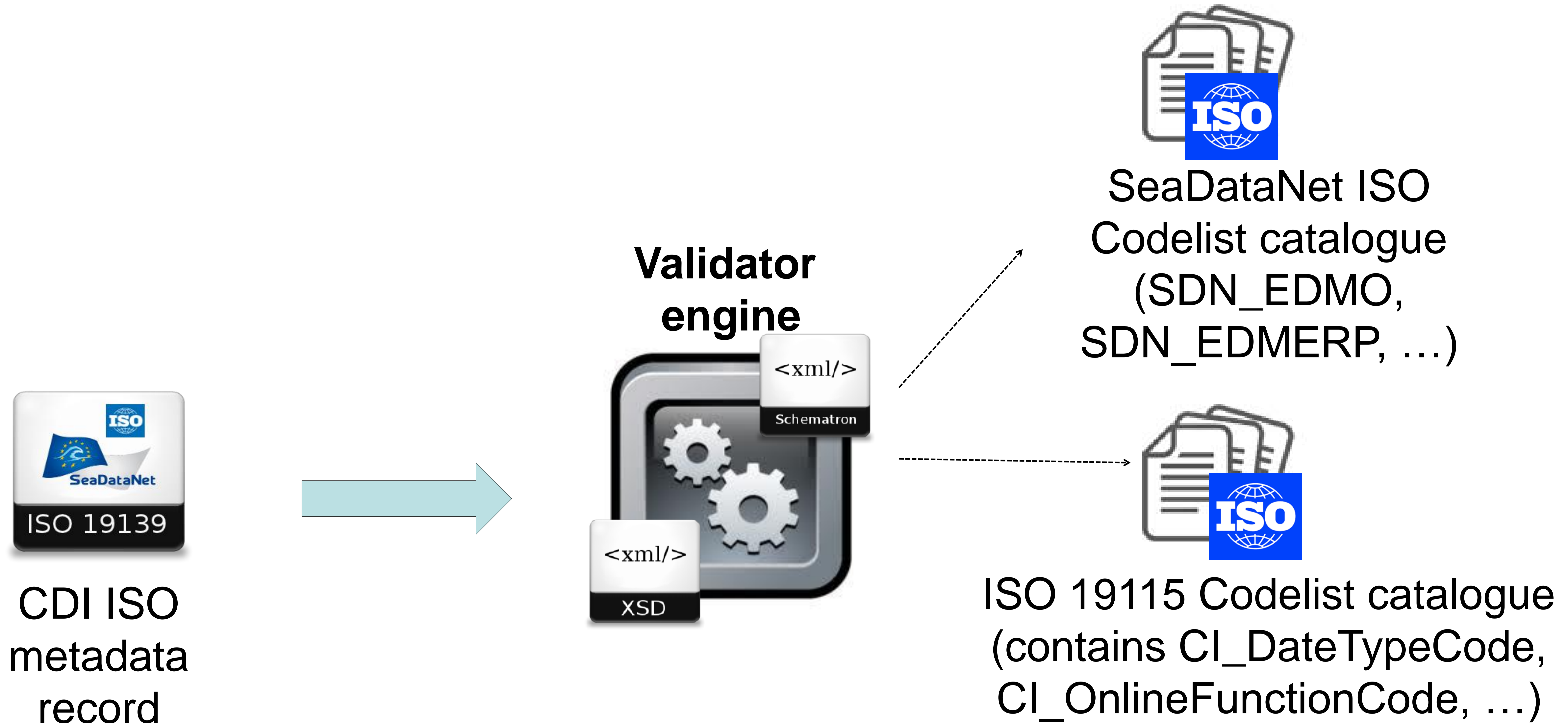
- **MBAN** (Bathymetry and Elevation)
- **ASAM** (Acoustic backscatter in the water column)
- **NOYS** (Acoustic noise in the water column)
- ...

– **Online** validation of **codelists**

- E.g: SDM_ParameterDiscoveryCode codelist

- `<sdn:SDN_ParameterDiscoveryCode codeList="../schema/resources/Codelist/sdnCodelists.xml#SDN_ParameterDiscoveryCode" codeListValue="MBAN" codeSpace="SeaDataNet">Bathymetry and Elevation</sdn:SDN_ParameterDiscoveryCode>`

Online Validation tool



Mapping CSR to ISO 19115-2

