



Progress with upgrading of SDN metadata formats for ISO/INSPIRE compliancy

Enrico Boldrini, Stefano Nativi

SeaDataNet-II

Technical Task Group Meeting

18th September 2012, Rhodes



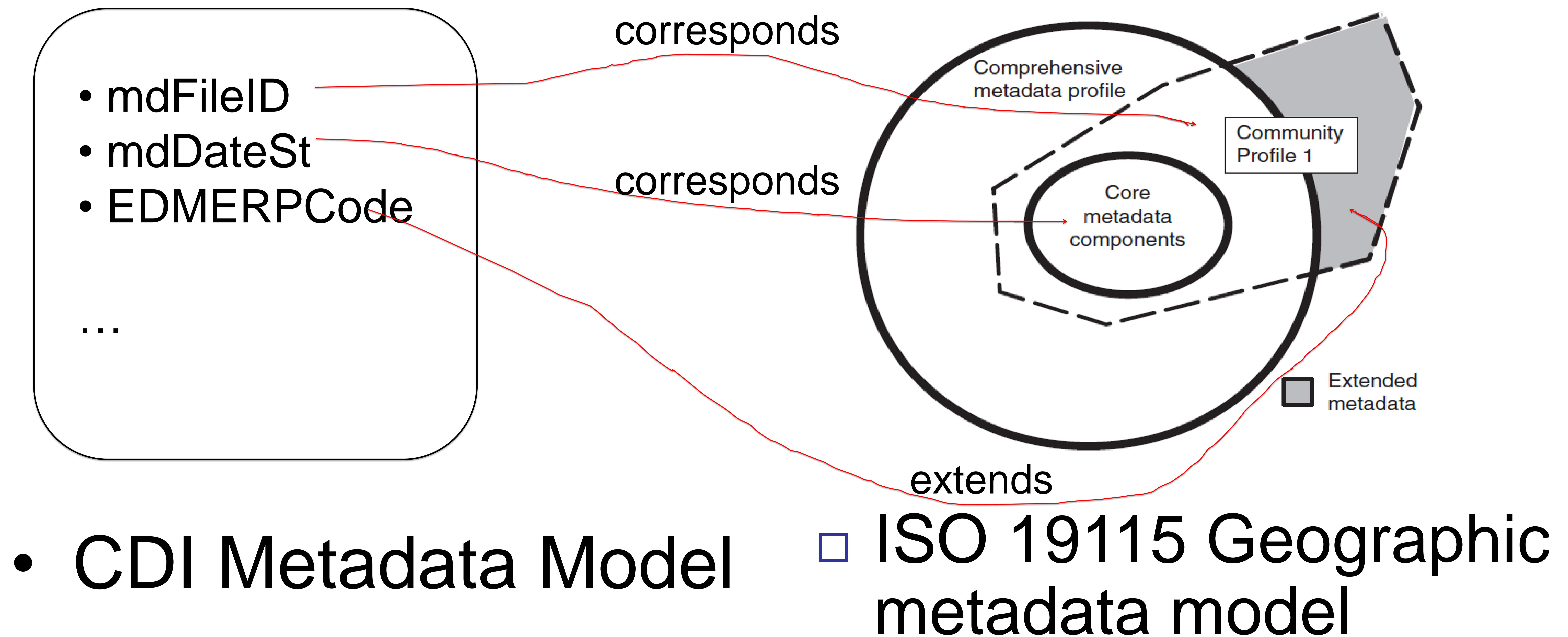
Context: SeaDataNet – second phase



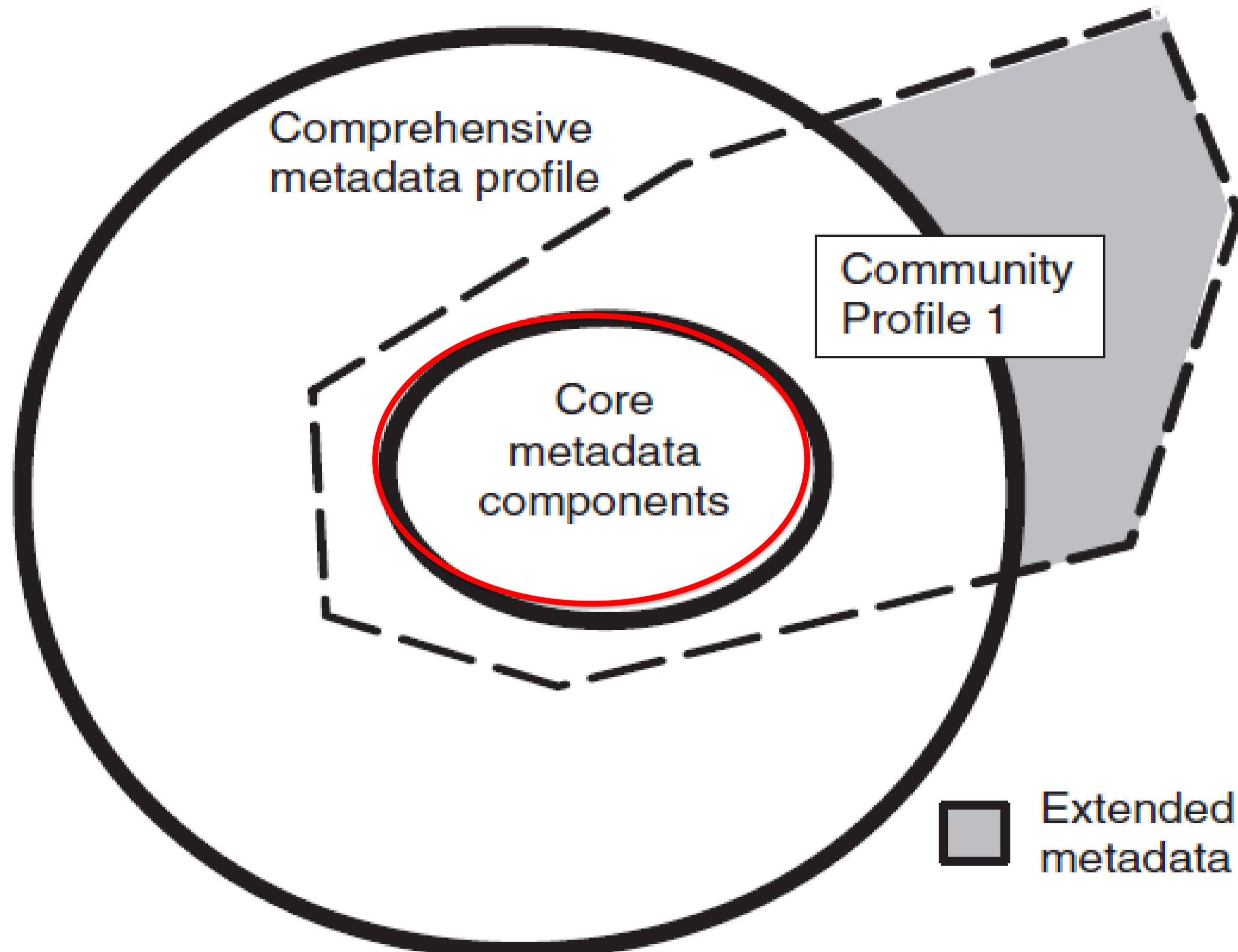
- Infrastructure upgrade
 - Operational robust
 - State of the art
- Methodology
 - Setting, adoption, promotion of international standards
 - Realize technical and semantic interoperability

CDI Metadata profile of ISO 19115

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



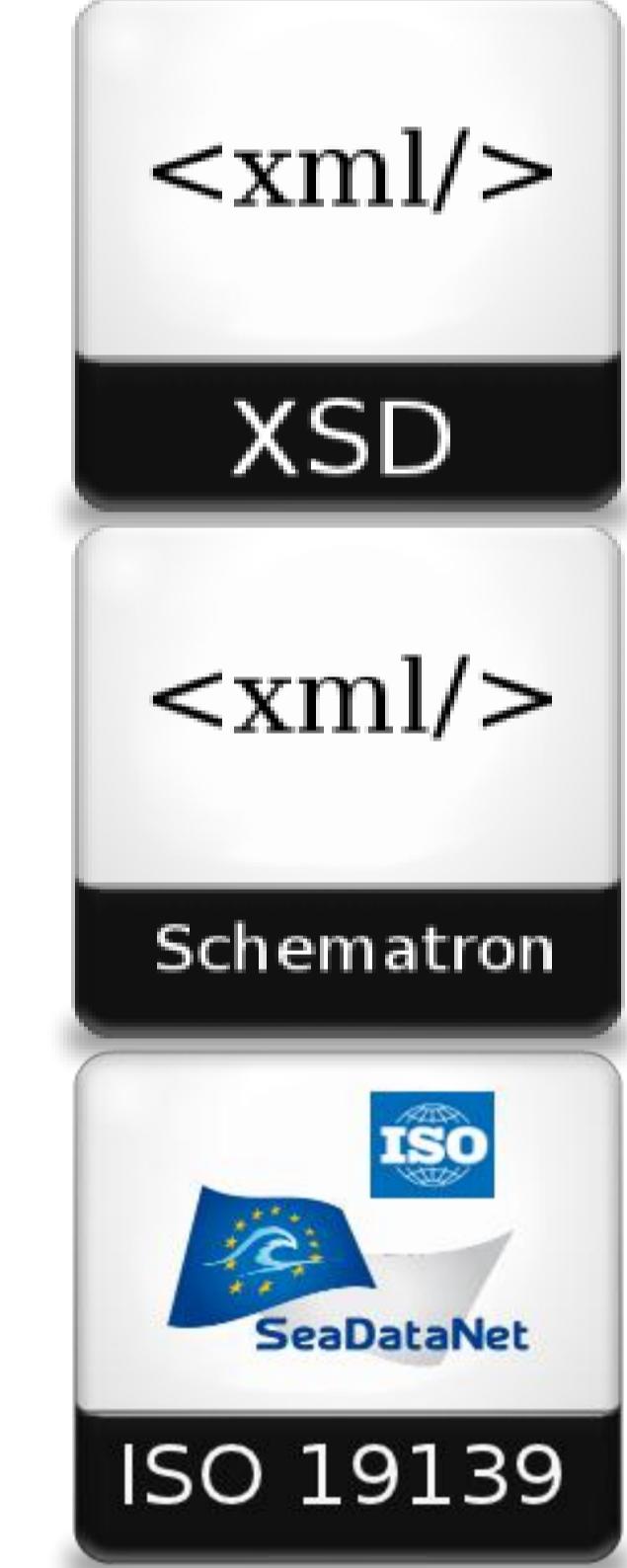
CDI Metadata profile of ISO 19115



- **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.

The result: 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

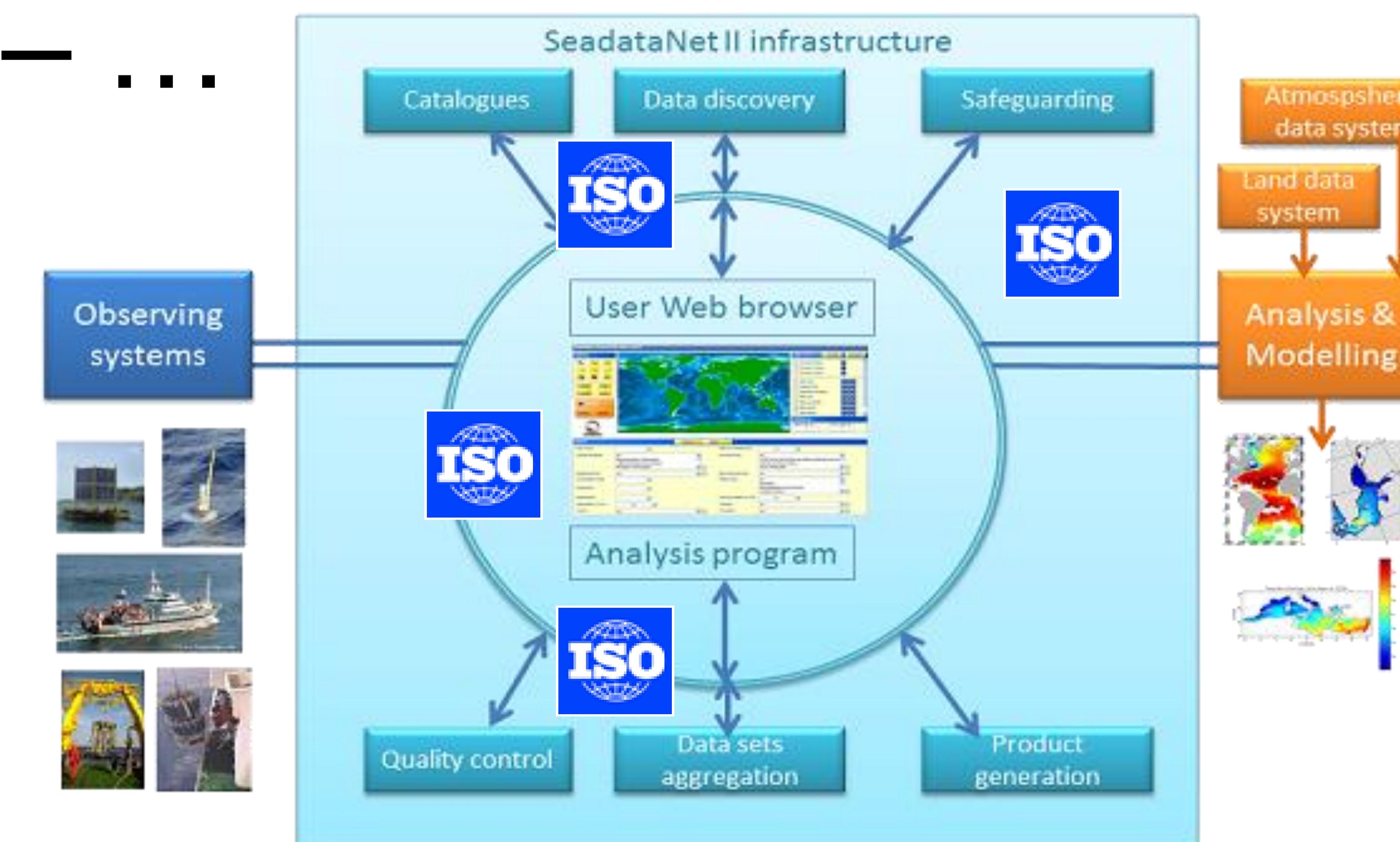
www.seadatanet.org/Standards-Software/Metadata-formats

Latest features (v. 7):

- Bibliographic information (citations)
- Main namespace change
- (Optional) parentIdentifier element
- EDMED codelist
- Updated KeywordTypeCode,
GeometricObjectTypeCode
- Aggregation information cardinality
- Update after official lists publication

CDI to ISO 19139 tools migration

- SeaDataNet tools update:
 - Metadata editors (e.g. MIKADO)
 - Discovery services (e.g. CSW ISO)
 - SeaDataNet portal



CDI to ISO 19139 datasets migration

- About 1 Million **CDI** entries currently in use
- Conversion to **ISO 19139** through **GI-cat XSLT**



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.

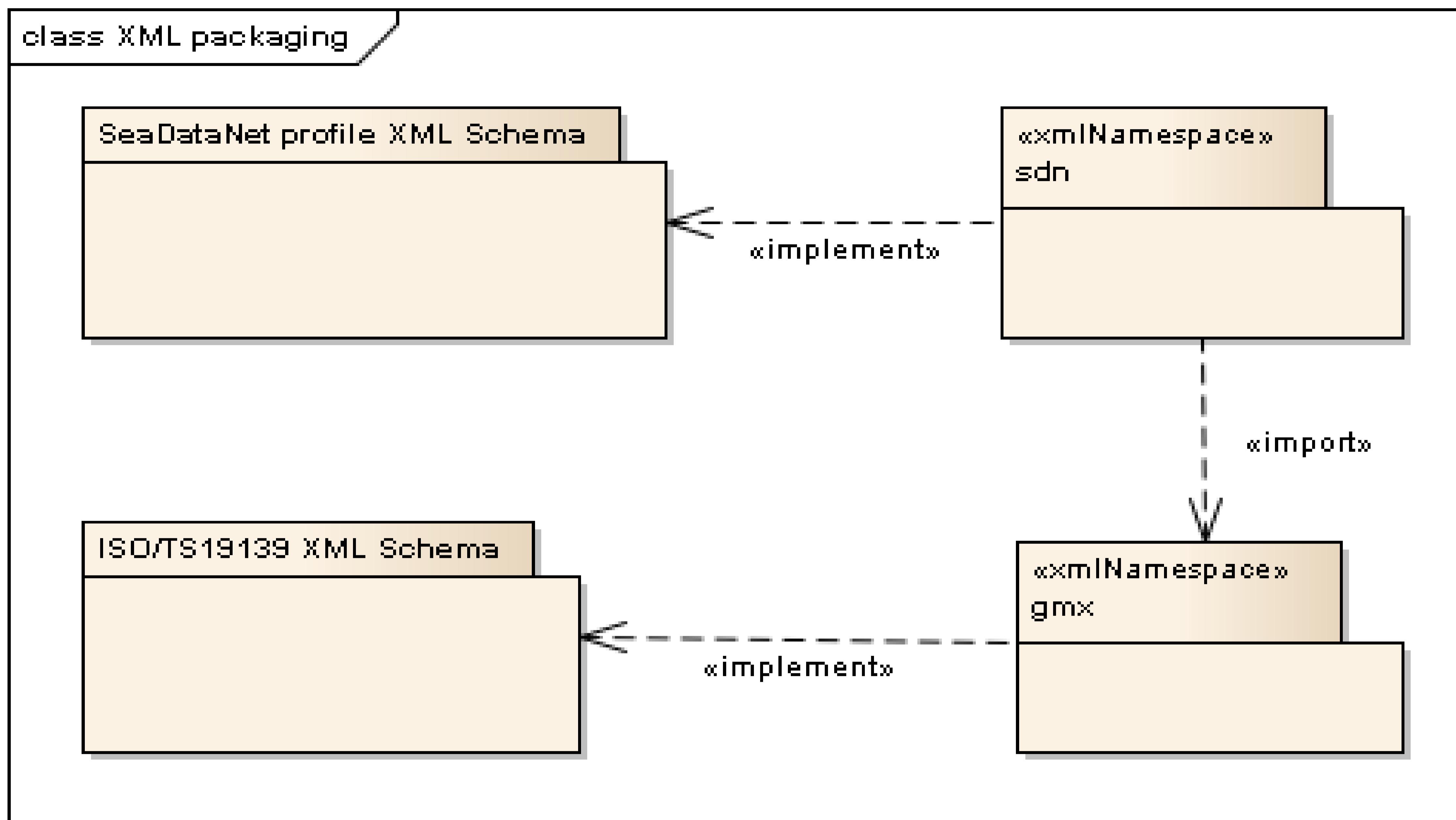
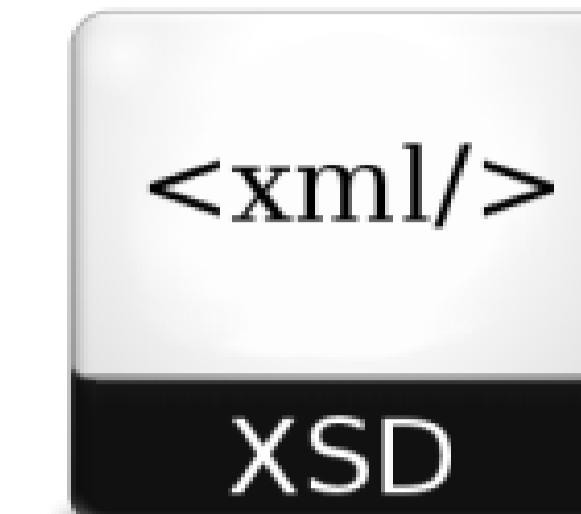
- all the profile elements are listed in the data dictionary
- Changes from ISO 19115 are highlighted in red
 - Changes in obligation (e.g. ***identifier*** from optional to mandatory)
 - Changes in cardinality (e.g. ***contact*** from 1..N to 1..1)
 - Changes in domain (e.g. ***format*** from free text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

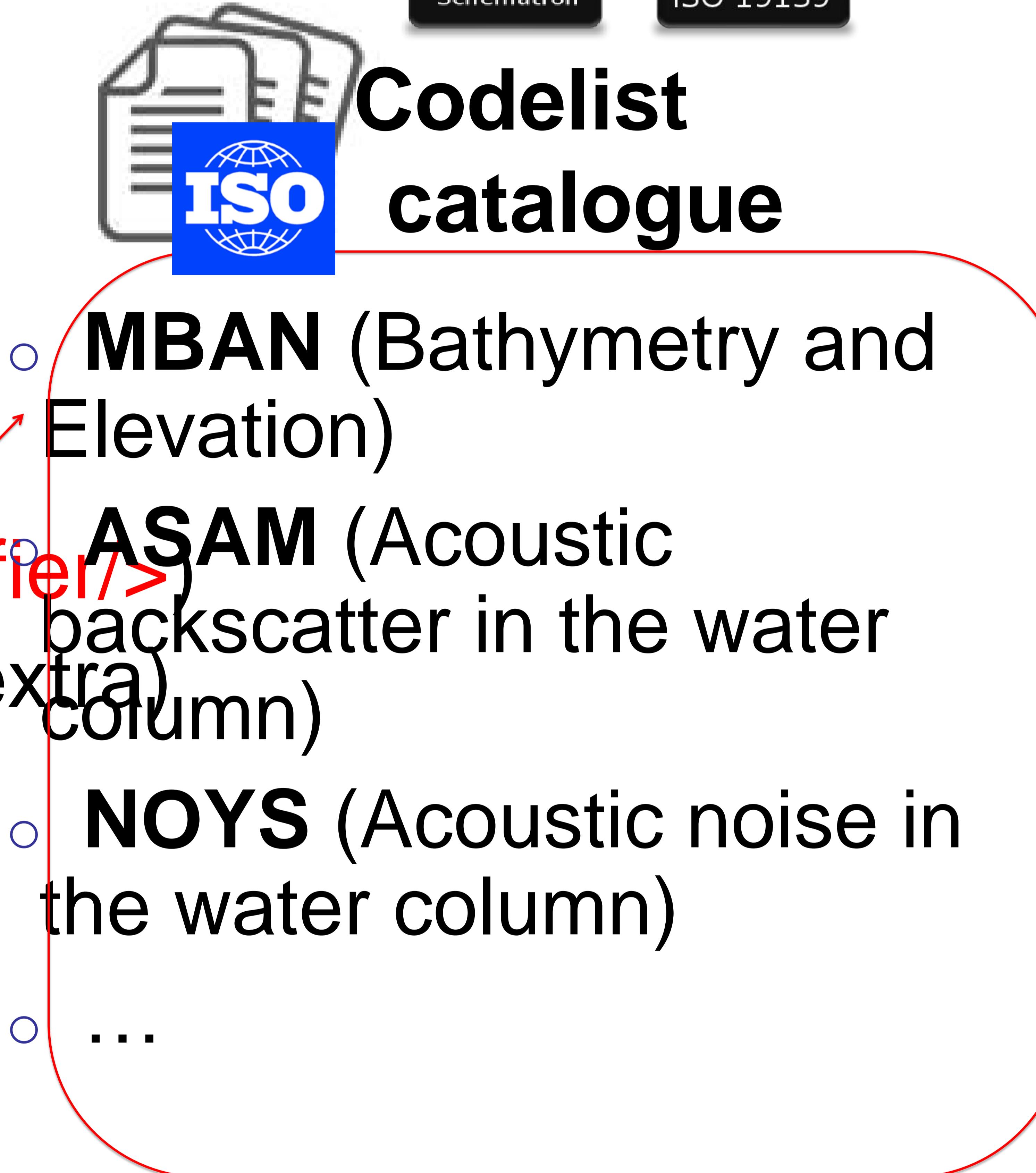


- Imports:
 - **ISO 19139 XML schema (CSW AP ISO version)**
- Includes:
 - Definitions of **new SDN codelists**:
 - SDN_FormatNameCode
 - SDN_HierarchyLevelNameCode
 - SDN_DeviceCategoryCode
 - SDN_PlatformCategoryCode
 - SDN_ParameterDiscoveryCode
 - SDN_EDMERPCode
 - SDN_EDMOCode
 - SDN_CRSCode
 -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



– Online validation of codelists

- E.g: SDM_ParameterDiscoveryCode codelist
- `<sdn:SDN_ParameterDiscoveryCode`

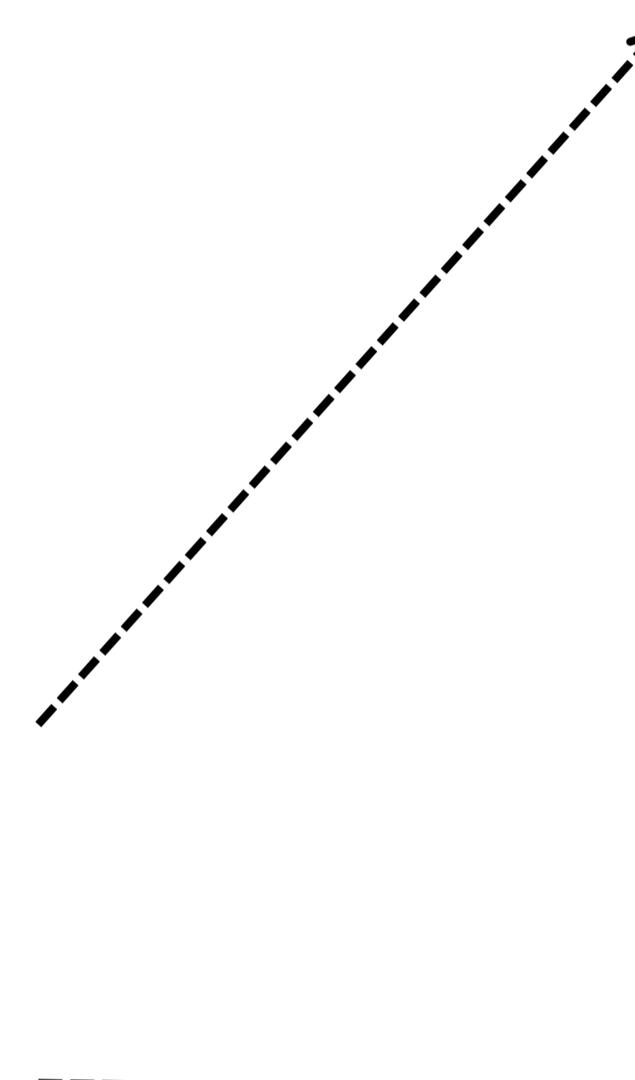
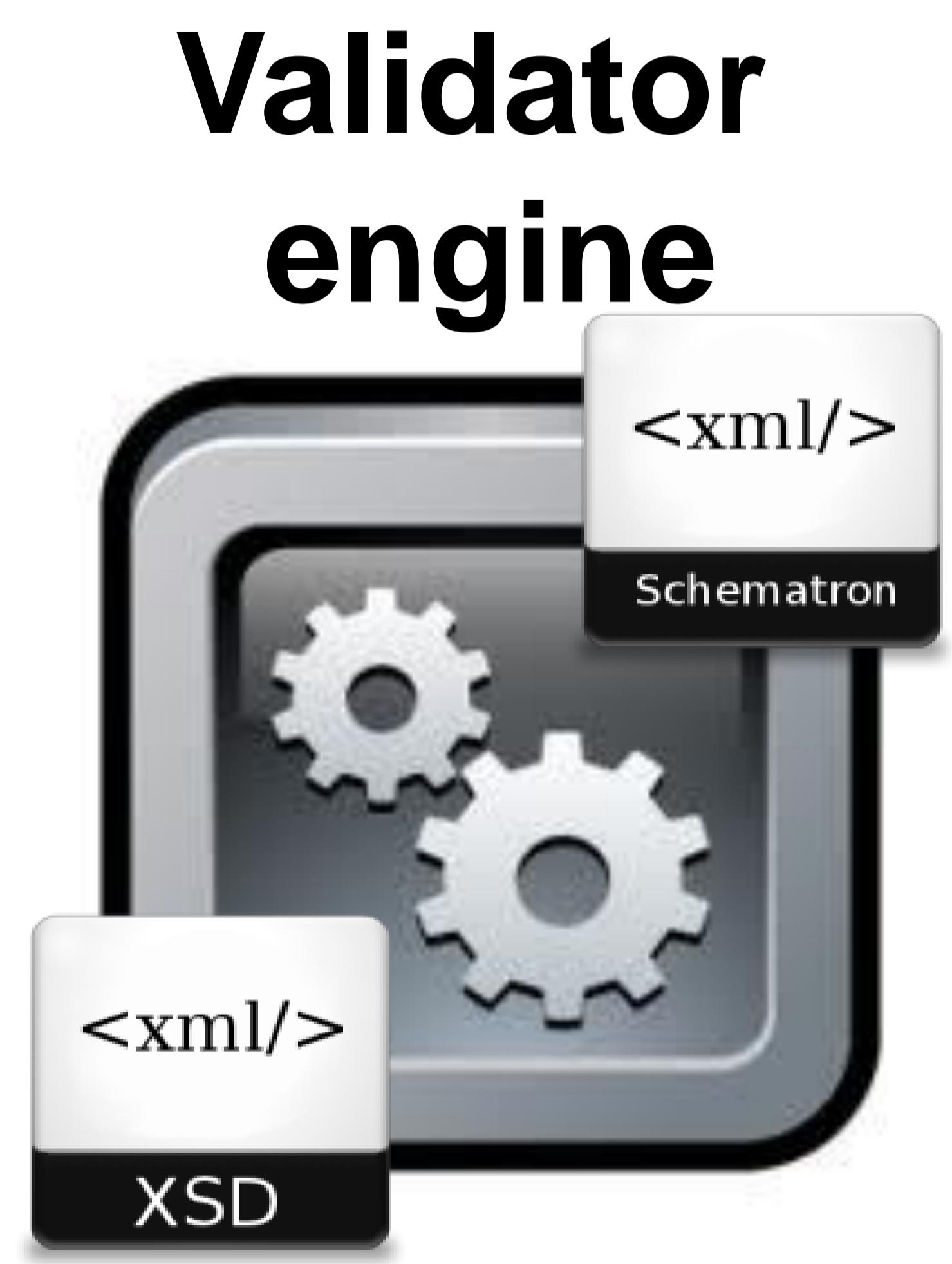
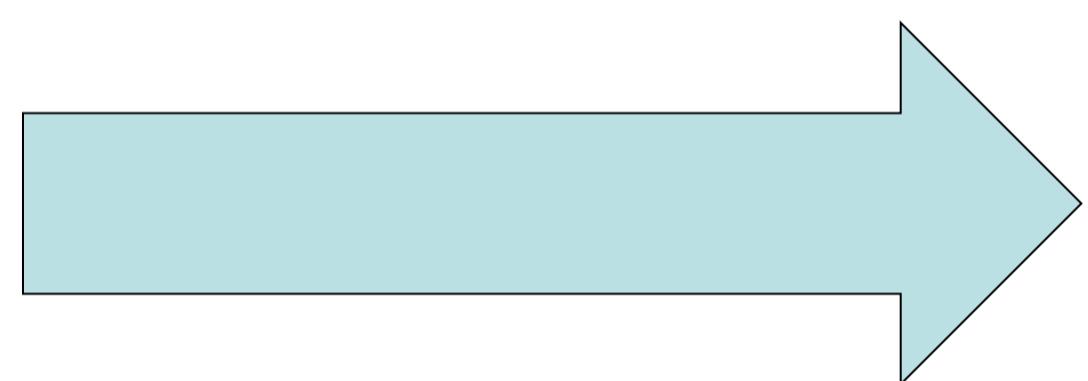
`codeList=..../schema/resources/Codelist/sdnCodelists.xml#SDN_ParameterDiscoveryCode"`
`codeListValue="MBAN" codeSpace="SeaDataNet">Bathymetry and`

Elevation</sdn:SDN_ParameterDiscoveryCode>

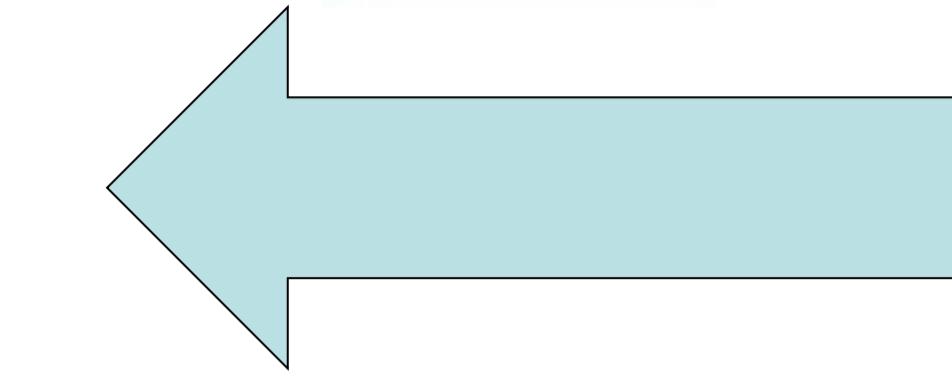
Online Validation tool



Upgraded
metadata
record



SeaDataNet ISO
Codelist catalogue
(EDMERP, CRS,
...)



SDN
EDMERP



ISO 19115 Codelist catalogue
(contains CI_DateTypeCode,
CI_OnlineFunctionCode, ...)

CDI encoding

New profile encoding

```
<?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:EDMO::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>
```

```
<?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>
```

The next challenge: Cruise Summary Reports

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

SeaDataNet

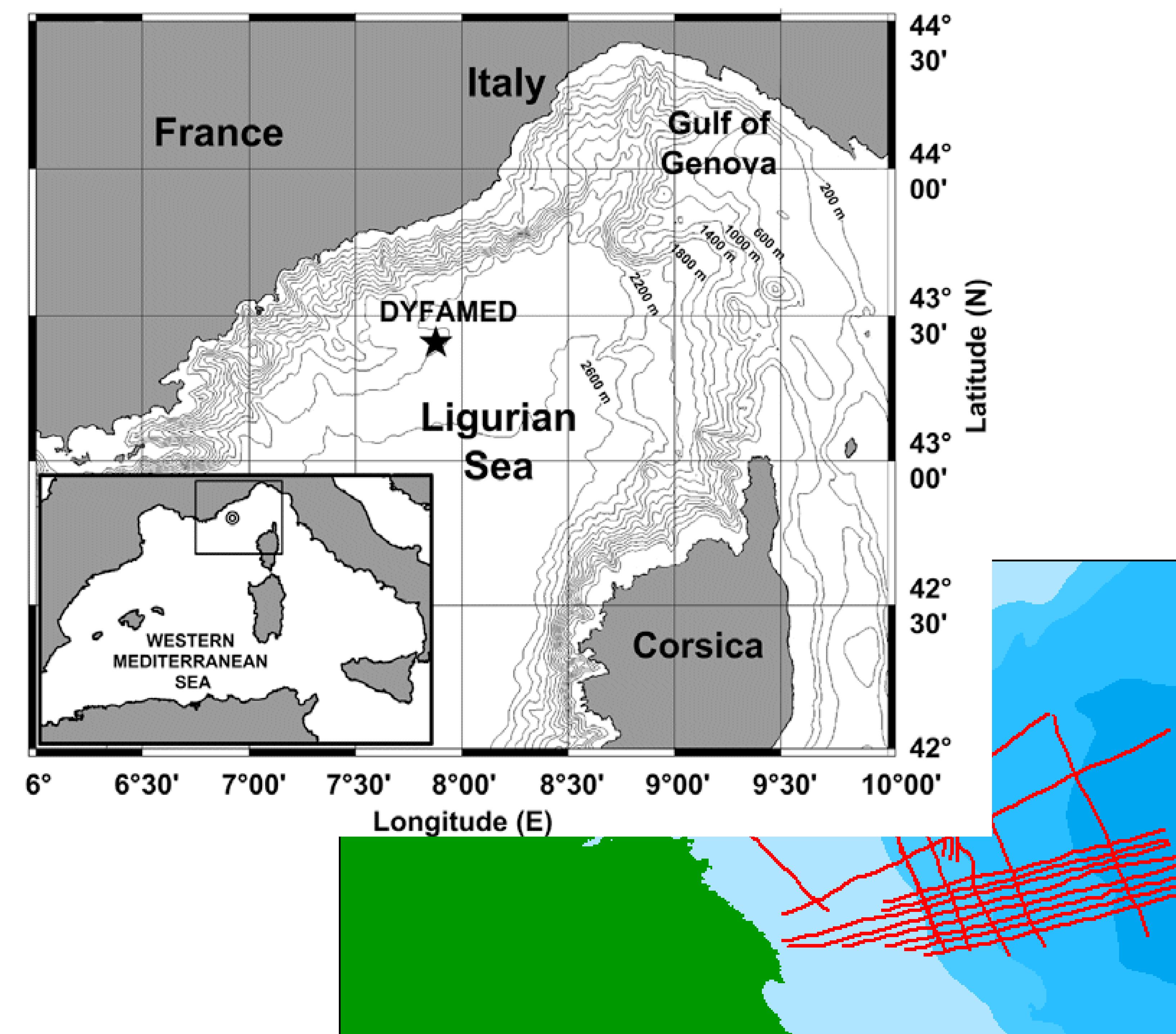
Cruise Summary Report Inventory (CSR)

| XML V0 | XML V1 | print | New Query | Results | Found 59 | Show 2 | Previous | Next |

Details of Sarmiento de Gamboa Cruise TOPOMED_LEG1 (BSH Ref-No.: 20116356)

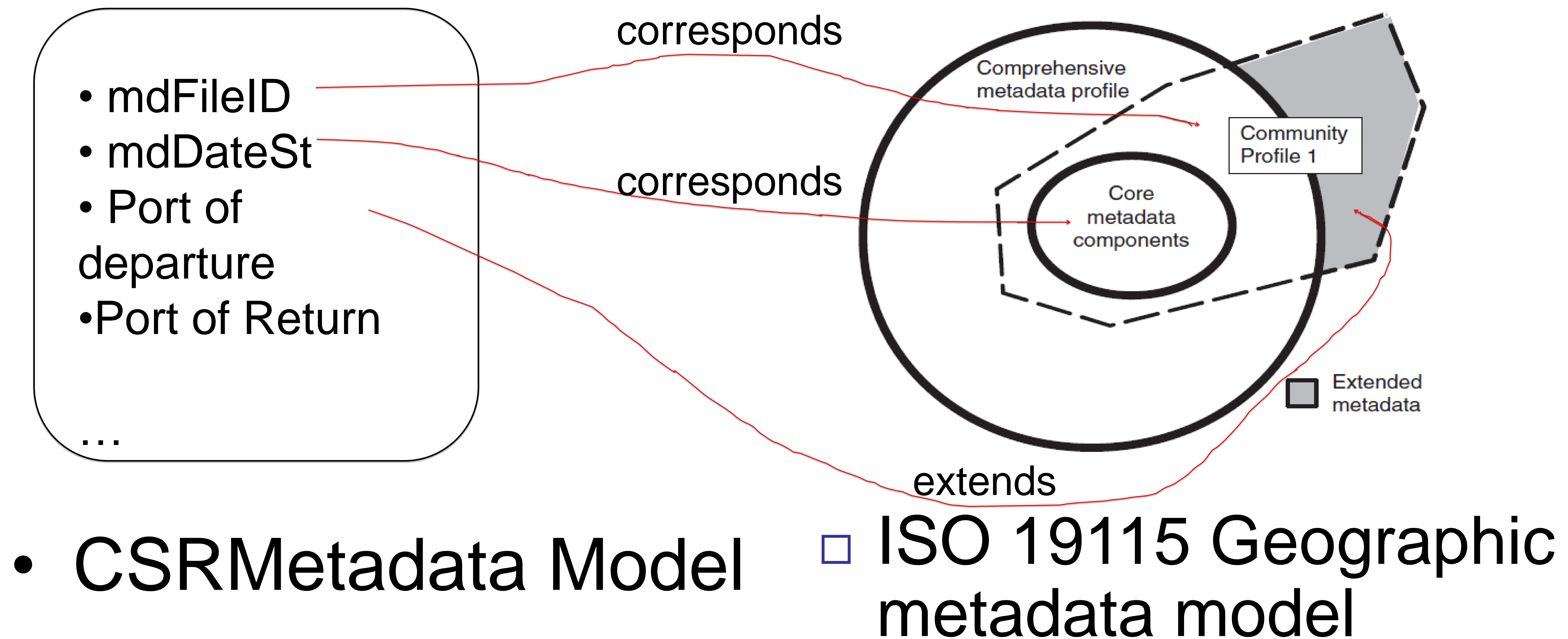
GENERAL INFORMATION	
Platform/Ship	Sarmiento de Gamboa
Cruise begin	03.10.2011
Cruise end	17.10.2011
Port of Departure	Vigo, Spain
Port of Return	Cadiz, Philippines
Chief Scientist(s)	EULALIA GRACIA - Marine Technology Unit. Mediterranean Marine and Environmental Research Centre CESAR RANERO - Marine Technology Unit. Mediterranean Marine and Environmental Research Centre BARTOLOMÉ, RAFAEL - Marine Technology Unit. Mediterranean Marine and Environmental Research Centre
Responsible Laboratory	Marine Technology Unit. Mediterranean Marine and Environmental Research Centre
LOCATION	
General Ocean Areas	Alboran Sea (Mediterranean Sea)
Marsden Squares(S, N, E, W)	
Bounding Box(es)	WEST EAST SOUTH NORTH -6.5154 -.975 35.3512 37.5932
Specific Geographic Areas	
Link to Track Charts	» Trackchart
PROJECT	
Project Title / Coordinating Body	
OBJECTIVES	
Description	Investigate the deep crustal structure of the Alboran sea
ADDITIONAL INFORMATION	
Parameters measured	Gravity Air temperature

- Cruise Summary Reports (CSR = former ROSCOPs) are the usual means for reporting on cruises or field experiments at sea
- Based on ISO 19115 DTD



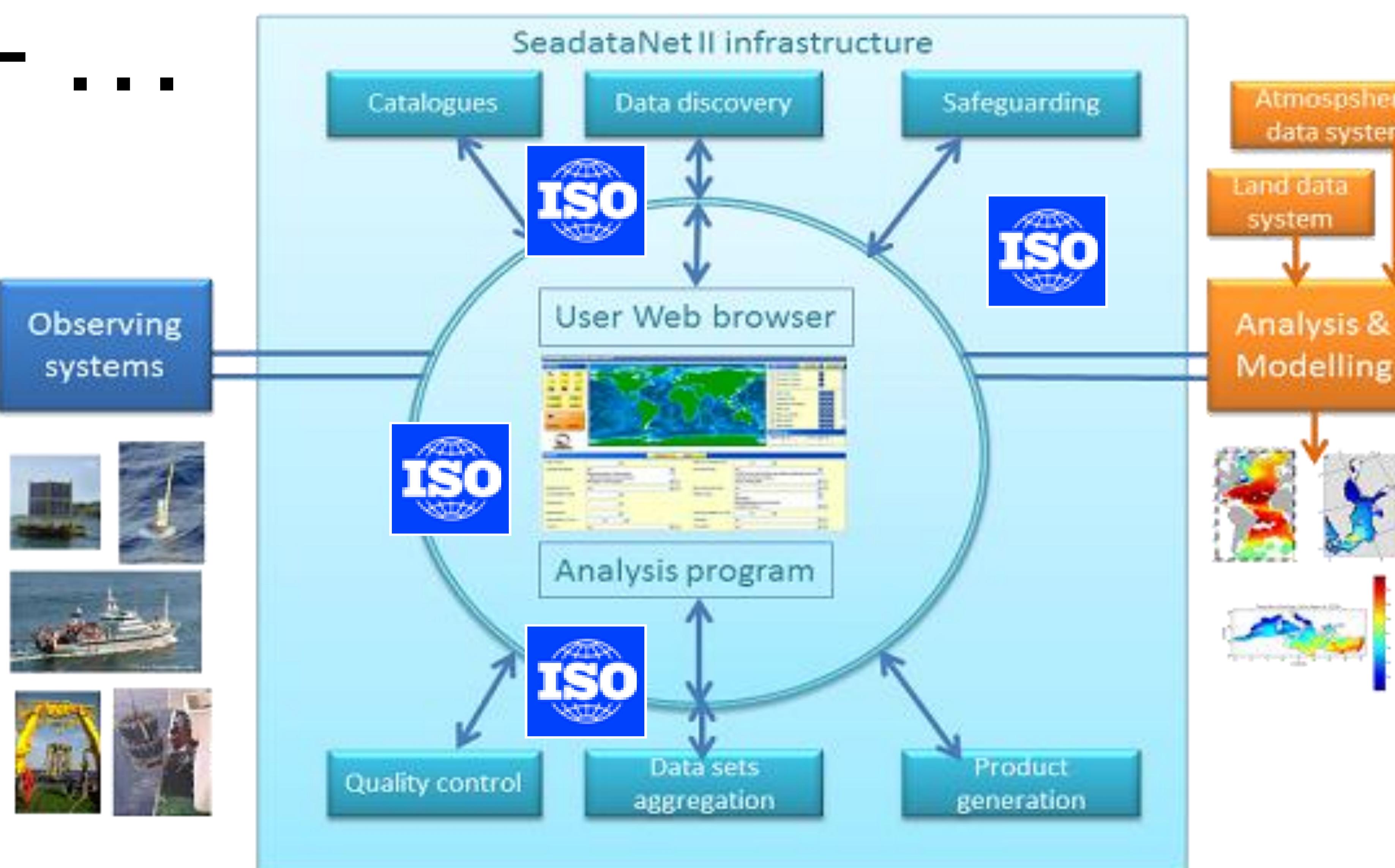
CSR Metadata profile of ISO 19115

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



Impacts on tools

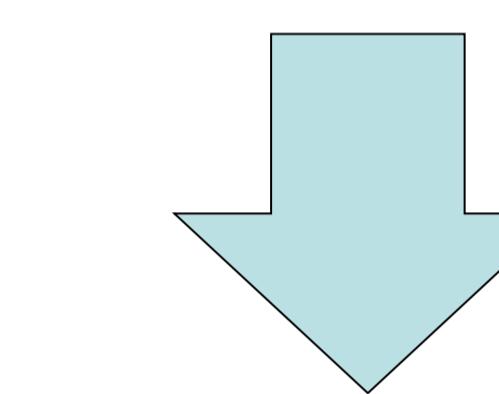
- SeaDataNet tools update:
 - CSR Inventory & CMS
 - Mikado
 - GI-cat broker
 - ...



Steps to be undertaken

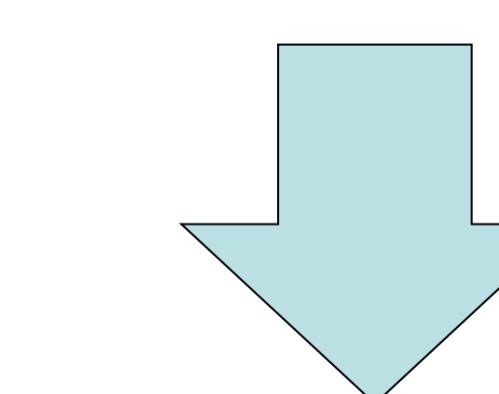
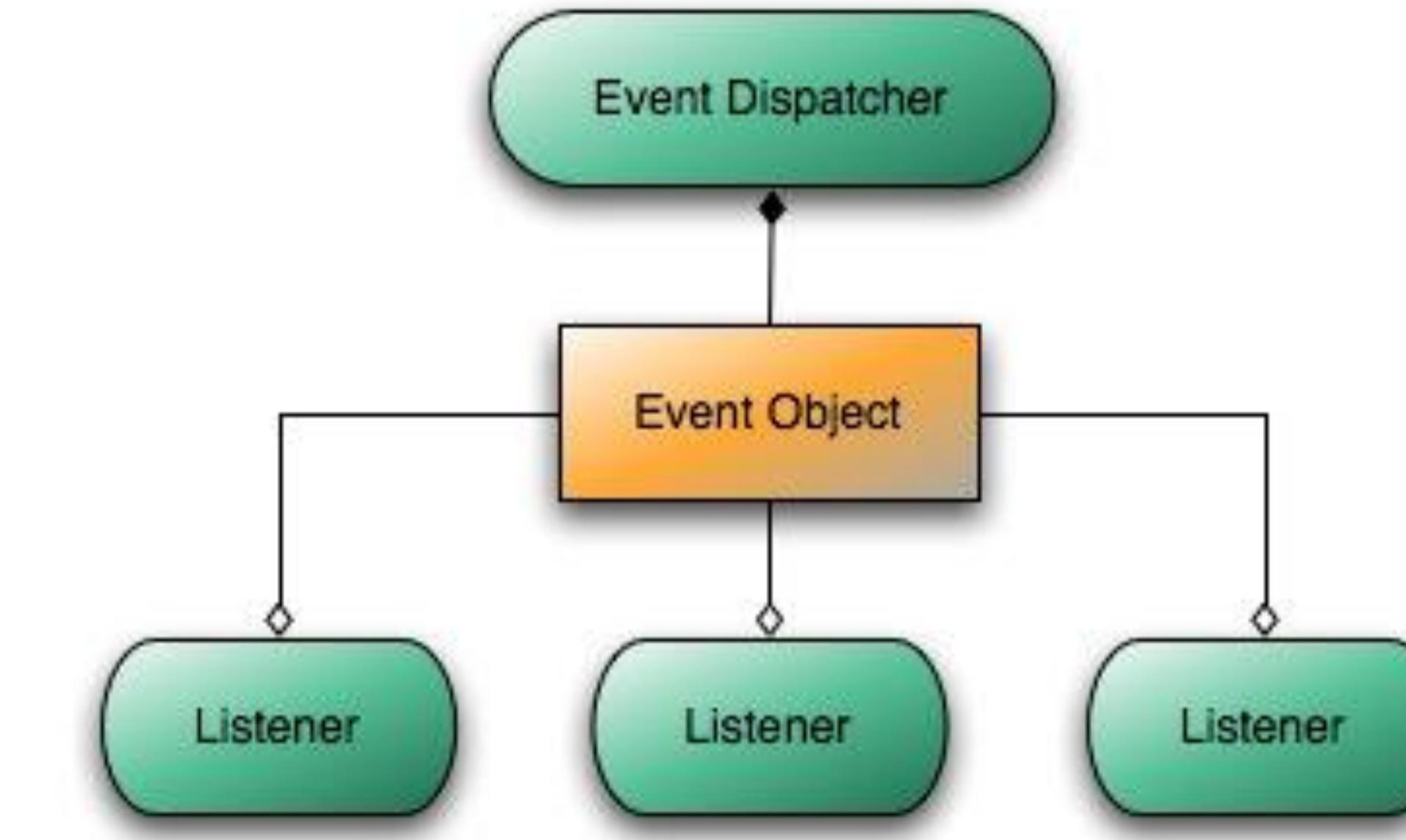
1. Collect all the existing documentation:

- XSD schema - DONE
- Excel documentation?
- Word documentation?



2. Identify the CSR elements :

- their definition
- Cardinality
- Additional constraints:
 - ISO, INSPIRE, SDN community



3. Implementation as for the CDI case ... (estimated first version: end of the year)

```
<?xml version="1.0"?
<quiz>
  <question>
    Who was the forty-second
    president of the U.S.A.?
  </question>
  <answer>
    William Jefferson Clinton
  </answer>
  <!-- Note: We need to add
       more questions later.-->
</quiz>
```

XML

Thank you for your attention!