
SeaDataNet Common Data Index (CDI) metadata model for Marine and Oceanographic Datasets – XML encoding

November 2014

Document type: Technical Specification

Current status: Proposal

Submitted by:

Dick M.A. Schaap
Technical Coordinator SeaDataNet
MARIS
The Netherlands

Enrico Boldrini
CNR-IIA
Italy

Stefano Nativi
CNR-IIA
Italy

Michele Fichaut
Coordinator SeaDataNet
IFREMER
France

Title: SeaDataNet Common Data Index (CDI) metadata model for Marine and Oceanographic Datasets – XML encoding

Scope: Proposal to adopt SeaDataNet CDI XML encoding as the reference XML implementation of SeaDataNet CDI metadata model.

SeaDataNet CDI XML encoding has been drafted, published and firstly implemented in the context of SeaDataNet, the leading infrastructure in Europe for marine & ocean data management. Its wide implementation, both by data centres within SeaDataNet and by external organizations makes it also a de-facto standard in the Europe region.

The acknowledgement of SeaDataNet CDI XML encoding as a recommended standard by IODE/JCOMM will further favour interoperability between European data centres and outreach to the broader marine community.

Envisaged publication type: The proposal target audience includes all the European bodies, programs, and projects that manage and exchange marine and oceanographic data. Besides, the proposed document informs all the international community dealing with marine and oceanographic data about the reference XML implementation of the SeaDataNet CDI metadata model.

Purpose and Justification: Provide details based wherever practicable.

1. Describe the specific aims and reason for this Proposal, with particular emphasis on the aspects of standardization covered, the problems it is expected to solve or the difficulties it is intended to overcome.

By acknowledging SeaDataNet CDI XML encoding as the reference XML implementation for the SeaDataNet CDI metadata model, multiple objectives are sought:

- Wider adoption of SeaDataNet CDI XML encoding by additional marine data centres around European waters. The process will favour further harmonisation and standardisation of European ocean and marine metadata as well as interoperability by reducing the existing metadata formats heterogeneity. Organizations adopting this standard will be able to encode their marine metadata according to a well-known and well specified structure, therefore the data management and exchange of marine and oceanographic information will be eased in many ways (see following point 2).
 - Ease interoperability and outreach towards international communities and initiatives. The existence of a recognized standard encoding for SeaDataNet CDI at European level will favour its understanding also at a broader level. Example given, international marine and oceanographic communities will be able to correctly understand (both syntactically and semantically) the information contained in SeaDataNet CDI documents (e.g. also with the help of appropriate software).
2. Describe how this proposed standard supports data management, exchange or interoperability. When applicable include mention of what data management functions (e.g. date transport, quality control, archive) the proposal supports.

SeaDataNet CDI XML encoding supports data management by providing an implementation based on ISO 19139 XML and containing as well specific SeaDataNet schema extensions (established through the ISO extension methodology). This XML implementation fully enable encoding and validation of the information described by SeaDataNet CDI metadata model.

A set of Schematron rules further enacts validation, assuring compliancy of encoded XML documents to SeaDataNet CDI metadata model, ISO 19115 and European directive INSPIRE (in particular INSPIRE implementing rules have been followed). Interoperability with software tools supporting such international standards and directives (such as catalogs or metadata editors) is thus favoured.

The Schematron rules enable as well validation of vocabulary restricted elements against online code lists published by SeaDataNet following ISO guidelines.

3. Describe the main interests benefitting from or affected by the proposed standard, such as industry, consumers, governments, distributors. Identify any relationships and/or dependencies.

Adoption by IODE/JCOMM of SeaDataNet CDI XML encoding as a reference implementation will give extra momentum to European marine and ocean data centres adopting SeaDataNet CDI. This will also benefit users from all over the world from various sectors. Moreover, it will benefit efforts for global interoperability because that process can focus on a limited set of marine metadata profiles (and encodings), whereby SeaDataNet CDI represents European input.

4. Describe the feasibility of implementing the proposed standard. Include any factors that could hinder the successful establishment or global application of the Proposed standard. Are there any associated issues? Identify resource implications resulting from the recommendations.

The feasibility and practicality of implementing the SeaDataNet CDI XML encoding can be, and has already been successfully accomplished at 57 data centres within the SeaDataNet partnership. Moreover another 47 data centres in Europe at present have realized the CDI XML encoding implementation for their managed data sets giving in total access to already more than 1.6 million data sets for physical oceanography, chemistry, geology, geophysics, bathymetry and biology. The results of these activities can be followed at the operational CDI data access service, that is part of the SeaDataNet portal (<http://www.seadatanet.org>). The implementation is supported by dedicated Training Workshops which deal with explaining the standards and the associated tools and which provide hands-on training activities to get fully acquainted with the standards and tools. The training material is also documented in Vademecums for study and consultation. The time needed for full implementation at a data centre is approximately 12 months considering the mapping of legacy datasets to SeaDataNet CDI and deployment of the associated SeaDataNet CDI tools.

5. Considering the needs of other fields or organizations, indicate the timeliness, target date(s), or if proposing a series of standards, suggest priorities. List any statutory requirement or other driving factors.

There are no statutory requirements for adoption of the SeaDataNet CDI XML encoding standard. The National Oceanographic Data Centres in Europe are bound to

implement the standard within their contractual obligations of several EU projects. The NODCs also motivate other data centres in their countries to adopt it. The IOC recommendation will add to this process.

6. Describe the possible benefits gained by the implementation of the proposed standard. Alternatively, describe the loss or disadvantage(s) if no standard is established within a reasonable time.

The advantage of using the SeaDataNet CDI standard in Europe is described in (2) and (3). There are no anticipated disadvantages to adopting it.

7. Indicate whether the proposed standard is or may become the subject of regulations or may require the harmonization of existing regulations. Describe any impacts of this activity.

The SeaDataNet CDI XML encoding standard is a de-facto standard in Europe and increasingly prescribed in calls for proposal and contracts by the European Commission for framework programmes and the EMODNET implementation as part of the EU Marine Directive.

Current Operational Implementations: At present already 57 National Ocean Data Centers (NODC's) and marine data centres within the SeaDataNet partnership have successfully implemented the SeaDataNet CDI XML encoding standard and are leveraging it at their local centre for giving overview and access to their managed data sets as part of the Data Discovery and Access Services of the SeaDataNet infrastructure (see <http://www.seadatanet.org>). Another 47 data centres in Europe at present have implemented the SeaDataNet CDI XML encoding standard as part of related EU funded projects (FP6-Upgrade Black Sea SCENE, FP6-CASPINFO, FP7/Geo-Seas, EMODNET Projects, FP7-EuroFleets, FP7-JERICO, FP7-CitClops, FP7-Micro B3, ...). The results of these activities can be followed at the SeaDataNet portal, where at present search can be distributed against 104 data centres, giving access to already more than 1.6 million data sets for physical oceanography, chemistry, geology, geophysics, bathymetry and biology.

A CSW ISO service is also available for querying the SeaDataNet content from automatic tools (e.g this is the interface leveraged by GEOSS), as well as an OAI-PMH service (e.g. leveraged to automatize harvesting by the ODP portal).

In addition, SeaDataNet has been adopted as the leading component for data management in the development of the European Marine Observation and Data Network (EMODNet) which was initiated in the framework of the MSFD. This contributes to SeaDataNet perspective towards long term sustainability.

Different software tools are implementing SeaDataNet CDI, such as the MIKADO metadata editor and the GI-cat discovery broker.

Relevant Documents:

The following documents (attached to the proposal) is the normative specification for the SeaDataNet CDI XML encoding:

- E.Boldrini, S.Nativi. SeaDataNet CDI metadata profile of ISO 19115 – XML encoding, Version 10.0.0, September 2013, published at
<http://www.seadatanet.org/Standards-Software/Metadata-formats/CDI>

The following online schemas are normative references for the CDI XML encoding:

- CDI profile XML schema definition, Version 10.0.0, September 2013, published at http://schemas.seadatanet.org/Standards-Software/Metadata-formats/SDN2_CDI_ISO19139_10.0.0.xsd
- CDI profile Schematron rules definition, Version 10.0.0, September 2013, published at http://schemas.seadatanet.org/Standards-Software/Metadata-formats/SDN2_CDI_ISO19139_10.0.0.sch
- CDI Extension information, Version 10.0.0, September 2013, published at <http://schemas.seadatanet.org/Standards-Software/Metadata-formats/cdiExtensionInformation.xml>

The following document is provided for informative purpose:

- Sample CDI metadata XML document, published at:
<http://www.seadatanet.org/Standards-Software/Metadata-formats/CDI>

The SeaDataNet CDI homepage represents as well an informative reference for SeaDataNet CDI, containing related standards (e.g. the SeaDataNet CDI metadata model) and useful documentation:

- SeaDataNet CDI metadata profile Homepage, at
<http://www.seadatanet.org/Standards-Software/Metadata-formats/CDI>

Cooperation and liaison:

1. **Existing Community:** All the organizations listed in the 'Current Operational Implementations' section are using SeaDataNet CDI in an operational environment and represent the SeaDataNet CDI community. In particular MARIS, CNR-IIA and IFREMER have been involved in the drafting and publication of the SeaDataNet CDI XML encoding standard (together with the rest of the SeaDataNet Technical Task Team) and are responsible for the current proposal submission.
2. **Expanded Community:** Firstly, other relevant marine and oceanographic data centres in Europe that are not yet engaged in the NODC national networks and/or any of the EU projects and would like to adopt SeaDataNet CDI as the metadata model for their datasets and CDI XML encoding as its XML implementation.

Moreover, other marine and oceanographic data centres worldwide eager to discover, evaluate and access SeaDataNet CDI datasets at full. In this regard, SeaDataNet is establishing exchanges from its infrastructure and portal to GEOSS, Ocean Data Portal (ODP) of IOC-IODE, EurOBIS and the European Nucleotide Archive (ENA) of EMBL-EBI. These activities have been extended with the active participation of SeaDataNet in the Ocean Data Interoperability Platform (ODIP) project where cooperation takes place with

leading oceanographic data infrastructures from the USA (US NODC, IOOS, R2R), Australia (IMOS) as well as IOC-IODE to explore common standards and interoperability solutions.

Contact information:

Name: Dick M.A. Schaap
Organization: Mariene Informatie Service ‘MARIS’
Email Address: dick@maris.nl
Telephone Number: +31 70 3004710

Proposal version: Version 1.0, November 2014

List of Acronyms:

- CDI – Common Data Index
- CNR-IIA – National Research Council of Italy – Institute of Atmospheric Pollution Research
- EDMED - SeaDataNet European Directory of Marine Environmental Data sets
- EDMERP - SeaDataNet European Directory of Marine Environmental Research Projects
- EDMO – SeaDataNet European Directory of Marine Organisations
- EMODNET – European Marine Observation and Data Network
- EU – European Union
- EuroFleets – EU FP7 project Towards an Alliance of European Research Fleets
- GEOSS – Group on Earth Observation System of Systems
- Geo-Seas - EU FP7 project for a Pan-European Infrastructure for Marine Geological and Geophysical Data Management
- IFREMER – Institut Francais de recherche pour l’exploitation de la mer
- IOC – Intergovernmental Oceanographic Commission
- IODE – International Oceanographic Data and Information Exchange
- ISO – International Organization for Standardization
- MARIS – Mariene Informatie Service
- MIKADO – SeaDataNet metadata editor software tool
- MMI – Marine Metadata Initiative
- MSFD - Marine Strategy Framework Directive
- NEMO – SeaDataNet ASCII data formats conversion software tool
- NODC – National Oceanographic Data Center
- ODIP - Ocean Data Interoperability Platform
- ODP - Ocean Data Portal
- SeaDataNet – EU FP6 project for a Pan-European Infrastructure for Marine and Oceanographic Data Management
- SeaVoX – mailing list governing the SeaDataNet Common Vocabularies
- Upgrade Black Sea SCENE - EU FP7 project for an Upgrade Black Sea Scientific Network
- URL – Uniform Resource Locator

Other Attachments: No other attachments.