ODANext



|  |  |
| --- | --- |
| Oceanographic Data Acquisition: the Next Age | |
| Deliverable title | Investigating interoperability of O&M with common dissemination formats (INSPIRE, SeaDataNet ODV/NetCDF and GOSUD) |
| Work Package Title | Data management architecture specifications |
| Deliverable number | 2.1.3 |
| Lead group | BMDC |
| Lead authors | de Ville de Goyet Nicolas, Thomas Vandenberghe |
| Contributors |  |
| Revision number | 1.2 |
| Revision Date |  |
| Security | Public |

[*I.* *Introduction and objectives* 3](#_Toc76442264)

[*II.* *OGC SWE standards* 3](#_Toc76442265)

[*1.* *Observations and measurements* 3](#_Toc76442266)

[*2.* *SensorML for procedure description* 5](#_Toc76442267)

[*III.* *O&M vs INSPIRE* 9](#_Toc76442268)

[1. Environmental monitoring facilities 9](#_Toc76442269)

[2. Oceanographic geographical features 9](#_Toc76442270)

[*IV.* *O&M vs SeaDataNet CDI* 10](#_Toc76442271)

[1. Technical implementation 10](#_Toc76442272)

[2. Mapping 11](#_Toc76442273)

[O&M to CDI metadata mapping 11](#_Toc76442274)

[O&M to ODV data mapping 13](#_Toc76442275)

[*V.* *O&M vs GOSUD* 14](#_Toc76442276)

[1. Technical implementation 14](#_Toc76442277)

[2. Mapping 16](#_Toc76442278)

[O&M to GOSUD Colcor format 16](#_Toc76442279)

**List of Accronyms**

API – Application Programming Interface

CDI – Common Data Index

ODV – Ocean Data View

O&M – Observations and Measurements

RV - Research Vessel

SDN - SeaDataNet

SML - SensorML

SOS – Sensor Observation Service

SWE – Sensor Web Enablement

XML – Extensible Markup Language

# *Introduction and objectives*

The new research vessel (RV) Belgica, built to replace the aging old Belgica, will be launched at sea during the year 2021. This represents an opportunity to renew the entire data flow from the data acquisition on the RV with the main idea of increasing the FAIRness of the acquisition system. A key element in the data flow is the choice of the schema used to store the various observations recorded by the sensors installed on the Belgica. The database structure must allow the storage of all type of observations and the associated metadata (e.g., sensor type, calibration, observations location...). The structure must also ease the dissemination of the observations in various data repositories.

Based on the experience of the project coordinators and the objectives of the ODANext project, the Open Geospatial Consortium Sensor Web Enablement standards (OGC SWE[[1]](#footnote-1)) have been selected as the core structure of the data and metadata storage. This suite of standards includes, amongst others, the Observations and Measurements (O&M) standard for the description of the observations, SensorML (SML) for the description of the measurement processes and Sensor Observation Service (SOS) for the dissemination of the observations. The choice of a recognized and validated international standard in this project is the first step towards the interoperability as it will be shown in this deliverable.

The objective of this deliverable is first to briefly introduce the OGC SWE standards used to store the information (O&M and SML). Then the ease of mapping SWE with the targeted international repositories is assessed, namely INSPIRE[[2]](#footnote-2), SeaDataNet[[3]](#footnote-3) and GOSUD[[4]](#footnote-4). The choice of those three repositories is guided by our mandatory legal obligations (e.g., INSPIRE – Directive 2007/2/EC, NODC) and our historical involvement in various projects (e.g., Seadatanet and GOSUD). Aside from mappings, we also discuss possible technical implementations.

# *OGC SWE standards*

The OGC SWE is a group of standards that is initially designed for the storage and machine-to-machine exchange of sensors data on the web. It includes the necessary standards for the encoding of the observations and sensors metadata with Observations and Measurements (O&M) and SensorML and also the necessary protocols for the dissemination of the data on the web (Sensor Observation Service).

In the following chapters, we compare the O&M and SensorML schemas to some common standards. The idea is to investigate whether and how the O&M and SensorML models can be expressed in those standards, not the XML expressions of the model. Translating the XML itself would be inefficient. In practice, the models are expressed in database tables compatible with resp. O&M and SensorML.

## *Observations and measurements*

ISO1956 O&M is a widespread standard used in the exchange of environmental observation. The basic elements of the standard are defined hereunder. As the schema is already extensively described in various online resources, this deliverable won’t go too much into the details of the schema.

The O&M schema can be summarized as:

An Observation is an action whose result is an estimate of the value of some property of the feature-of-interest, obtained using a specific procedure (illustration 1).

|  |
| --- |
|  |
| Illustration 1: O&M standard UML model. |

The important elements are defined in the table below:

|  |  |
| --- | --- |
| element | Definition |
| type | Type of O&M observation describe (e.g., point observation, time series observation...) |
| Phenomenon time | Time when the measurement is performed |
| Result time | Time when the result of the measurement is known |
| procedure | sensor used for the measurement, defined in the SensorML standard. |
| observed property | property that is being measured (e.g. temperature). |
| feature of interest | geographical object/location that is being observed (e.g. North Sea). |
| result | numerical value of the observation. |

As an example, the Table 1 below shows the encoding of a simple observation in the O&M schema.

|  |
| --- |
| <om:OM\_Observation gml:id="o\_6393366">  <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM\_Measurement"/>  <om:phenomenonTime>  <gml:TimeInstant gml:id="phenomenonTime\_6393366">  <gml:timePosition>2012-11-19T13:00:01.000Z</gml:timePosition>  </gml:TimeInstant>  </om:phenomenonTime>  <om:resultTime>  <gml:TimeInstant gml:id="ti\_6960D05"> <gml:timePosition>2012-12-30T15:45:01.000Z</gml:timePosition>  </gml:TimeInstant>  </om:resultTime>  <om:procedure xlink:href="LISST" xlink:title="LISST"/>  <om:observedProperty xlink:href="http://vocab.nerc.ac.uk/collection/P01/current/MDGSISXX/"/>  <om:featureOfInterest xlink:href="http://spatial.naturalsciences.be/ef/marine/BEFED\_MOW1"/>  <om:result uom="micrometres" xsi:type="ns:MeasureType">64.74</om:result>  </om:OM\_Observation> |
| Table 1 : O&M xml encoding of a single observation |

For more detailed information on the O&M standard, please read the following documentation:

- OGC 10-004r3 (<https://www.ogc.org/standards/om>)

## *SensorML for procedure description*

The primary focus of SensorML is to provide a robust and semantically-tied means of defining processes and processing components with the measurement and post-measurement transformation observations[[5]](#footnote-5). SensorML can be used to describe process independently of the environment in which they might be executed, this includes both real-world physical sensors and model/computation processes.

The standard is rich and allow for the complete description of the sensor's specifications, including keywords and classifiers for the discovery of the sensors and events history (calibration, cleaning…). The core elements of the standard are defined in Table 2.

|  |  |  |
| --- | --- | --- |
| Element | multiplicity | Definition |
| description | 0...1 | Textual description for the feature. |
| name | 0...n | Common name for the feature. |
| Unique identifier | 1 | Unique identifier for the described process |
| keyword | 0...n | Short keywords describing the context of this document to aid in discovery. |
| identifiers | 0...n | Identifiers useful for discovery of the process (e.g., short name, mission id, wing id, serialnumber, etc.) |
| classifiers | 0...n | Classifiers useful for discovery of the process (e.g., process type, sensor type, intendedapplication, etc.) |
| Security constraints | 0...n | Overall security tagging of process description; individual tagging of properties can bedone using extension element. |
| Valid time | 0...n | The time instance or time range during which this instance description is valid. |
| Legal constraints | 0...n | Legal constraints applied to this description (e.g., copyrights, legal use, etc.) |
| capabilities | 0...n | Properties that further clarify or quantify the output of the process (e.g., dynamic range,sensitivity, threshold, etc.). These can assist in the discovery of processes that meetparticular requirements. |
| characteristics | 0...n | Useful properties of this process that do not further qualify the output values (e.g.,component dimensions, battery life, operational limits, etc). |
| contacts | 0...n | Persons or responsible parties that are relevant to this process (e.g., designer,manufacturer, expert, etc.) |
| documentation | 0...n | Additional external online documentation of relevance to this process (e.g., user’s guides,product manuals, specification sheets, images, technical papers, etc.) |
| history | 0...n | A collection of time-tagged events relevant to this process. |
| Table 2 : SensorML core elements as defined in the OGC SensorML model and xml encoding Standard[[6]](#footnote-6) | | |

As an example, Table 3 shows a simplified example of a LISST sensor described with SML.

|  |
| --- |
| <soap:Body>  <swes:DescribeSensorResponse>  <swes:procedureDescriptionFormat>http://www.opengis.net/sensorml/2.0</swes:procedureDescriptionFormat>  <swes:description>  <swes:SensorDescription>  <swes:validTime>  <gml:TimePeriod gml:id="tp\_06751FF2">  <gml:beginPosition>2020-03-26T09:06:28.025Z</gml:beginPosition>  <gml:endPosition indeterminatePosition="unknown"/>  </gml:TimePeriod>  </swes:validTime>  <swes:data>  <sml:PhysicalSystem xmlns:sml="http://www.opengis.net/sensorml/2.0">  <gml:description>The workhorse instrument of Sequoia, the LISST-100X, is a multi-parameter system for in-situ observations of particle size distribution and sediment concentration. It also records the optical Volume Scattering Function, the optical transmission, pressure and temperature. It features an automatic gain switch for clear water use and is compatible with the BioBlock anti-fouling copper shutter , also developed by Sequoia Scientific.</gml:description>  <gml:identifier codeSpace="uniqueID">LISST</gml:identifier>  <gml:name>LISST</gml:name>  <sml:keywords>  <sml:KeywordList>  <sml:keyword>water</sml:keyword>  <sml:keyword>backscatter</sml:keyword>  </sml:KeywordList>  </sml:keywords>  <sml:identification>  <sml:IdentifierList>  <sml:identifier>  <sml:Term definition="urn:ogc:def:identifier:OGC:uniqueID">  <sml:label>uniqueID</sml:label>  <sml:value>LISST</sml:value>  </sml:Term>  </sml:identifier>  </sml:IdentifierList>  </sml:identification>  <sml:validTime>  <gml:TimePeriod gml:id="deploymentDates">  <gml:beginPosition>2013-01-01</gml:beginPosition>  <gml:endPosition>2014-12-31</gml:endPosition>  </gml:TimePeriod>  </sml:validTime>  <sml:featuresOfInterest>  <sml:FeatureList definition="http://www.opengis.net/def/featureOfInterest/identifier">  <swe:label>BEFED\_MOW1</swe:label>  <sml:feature xlink:href="http://spatial.naturalsciences.be/ef/environmentalmonitoringfacility/marine/BEFED\_MOW1"/>  </sml:FeatureList>  </sml:featuresOfInterest>  <sml:outputs>  <sml:OutputList>  <sml:output name="phenomenonTime">  <swe:Time definition="http://www.opengis.net/def/property/OGC/0/PhenomenonTime">  <swe:uom xlink:href="http://www.opengis.net/def/uom/ISO-8601/0/Gregorian"/>  </swe:Time>  </sml:output>  <sml:output name="Median\_Distribution">  <swe:Quantity definition="http://vocab.nerc.ac.uk/collection/P01/current/MDGSISXX/">  <swe:uom code="micrometres"/>  </swe:Quantity>  </sml:output>  </sml:OutputList>  </sml:outputs>  <sml:position>  <swe:Vector id="SYSTEM\_LOCATION" referenceFrame="http://www.opengis.net/def/crs/EPSG/0/4326">  <swe:coordinate name="easting">  <swe:Quantity axisID="Lon" definition="longitude">  <swe:uom code="degree" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UAAA/"/>  <swe:value>3.113889</swe:value>  </swe:Quantity>  </swe:coordinate>  <swe:coordinate name="northing">  <swe:Quantity axisID="Lat" definition="latitude">  <swe:uom code="degree" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UAAA/"/>  <swe:value>51.358333</swe:value>  </swe:Quantity>  </swe:coordinate>  <swe:coordinate name="altitude">  <swe:Quantity axisID="Alt" definition="altitude">  <swe:uom code="m" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/ULAA/"/>  </swe:Quantity>  </swe:coordinate>  </swe:Vector>  </sml:position>  </sml:PhysicalSystem>  </swes:data>  </swes:SensorDescription>  </swes:description>  </swes:DescribeSensorResponse>  </soap:Body> |
| Table 3 : Simplified XML example of a process described with the SensorML standard. |

# *O&M vs INSPIRE*

The INSPIRE Directive aims to create a European Union spatial data infrastructure for the purposes of EU environmental policies and activities[[7]](#footnote-7). The Directive came into force in 2007 and the full implementation is required for 2021. The scope of the Directive that is of interest for the ODANext project is related to marine monitoring programs for which the new RV Belgica will be used.

The Directive addresses 34 spatial data themes needed for environmental applications and two are of relevance for the RV Belgica: Environmental monitoring facilities and Oceanographic geographical features.

Environmental monitoring facilities (EF) in the context of INSPIRE is defined as: ‘Location and operation of environmental monitoring facilities includes observation and measurement of emissions, of the state of environmental media and of other ecosystem parameters (biodiversity, ecological conditions of vegetation, etc.) by or on behalf of public authorities[[8]](#footnote-8). In this case, the EF is the RV Belgica itself and should be described according to the EF encoding rules defined in the INSPIRE data specification.

Oceanographic geographical features (OF) in the context of INSPIRE is defined as:’ Physical conditions of oceans (currents, salinity, wave heights, etc.)[[9]](#footnote-9). In this case, it includes all the physico-chemical observations measured by sensors installed on-board of the RV Belgica.

To correctly implement the INSPIRE specification, we have to represent the data using those two themes.

## Environmental monitoring facilities

An efficient mapping of OGC SWE with the INSPIRE EF data model is not a critical part in this project as the RV Belgica is the only EF that needs to be described. The description of the RV Belgica as an INSPIRE EF xml and SensorML document will be performed independently.

INSPIRE EF xml file will be generated using the open-source Hale studio software[[10]](#footnote-10). All specifications regarding the data model can be found on the INSPIRE website and in the data specifications[[11]](#footnote-11).

## Oceanographic geographical features

At the time of defining the Oceanographic geographical features (OF) data model in the INSPIRE data specifications, the O&M standard has been identified as the most complete, generic and recognized standard for the representation of environmental data and its associated metadata. Therefore, the OF theme was built directly upon the O&M specifications.

The implementation of the OGC SWE standards suite therefore ensure that the data provided by the Sensor Observation Service will reach the INSPIRE-compliance without any action needed. A simple SOS request such as ‘GetObservation’ with ‘ResponseFormat=[*http://inspire.ec.europa.eu/schemas/omso/3.0*](http://inspire.ec.europa.eu/schemas/omso/3.0) will provide O&M representation of the data (e.g. <https://sensors.naturalsciences.be/sos/service?service=SOS&version=2.0.0&request=GetObservation&procedure=SEABIRD_SBE21&temporalFilter=phenomenonTime,2019-08-29T06:00:00Z/2019-08-30T06:00:00Z&responseFormat=http://inspire.ec.europa.eu/schemas/omso/3.0>).

# *O&M vs SeaDataNet CDI*

## Technical implementation

SeaDataNet (SDN) is a distributed Marine Data infrastructure for the management of large and diverse sets of data deriving from in situ of the seas and oceans[[12]](#footnote-12). Amongst other, its objective is to develop a Pan-European integrated database populated by numerous international data centres in a standardized way. Metadata are described using the Common Data Index (CDI) standard, which is a community profile on the ISO19115 metadata standard. Vanilla ISO 19115 is also used for INSPIRE metadata descriptions. Data can be encoded in two different formats, namely NetCDF and Ocean Data View (ODV) files.

In the frame of the SeaDataCloud project, SeaDataNet has been studying and implementing the use of O&M and Sensor Observation Service as a means to display the different data sources in an interoperable way[[13]](#footnote-13). As a result, a 52north SOS with a browser data viewer Helgoland has been deployed and is accessible online[[14]](#footnote-14). In the context of the SDC2 project (not funded so far), the integration of the SOS standard with the existing SDN infrastructure would be pushed even more. The adoption of O&M and the SOS, as planned in the context of the ODANext project, already meets some of the goals of SDC2 (and the Belgica SOS could be showed there as one of the use cases). These are the key points:

* En route data is ***not*** integrated into the CDI data portal via the NODC’s SOS, which is only used to provide viewing services in the Helgoland viewer of the SeaDataNet Sensor Web Viewer Services
* Data should still be passed to SeaDataNet via the classical CDI mechanism, ie. with CDI metadata, coupling table and mapping file. For station and trajectory data, SDN uses these modi:
  + Modus=1. Pre-processed static mono-station files
  + **Modus=2. Database connection, Replication Manager generates ODV files**
  + Modus= 3. Pre-processed static multi-station files (Trajectory)
    - Modus 1 and 3 are more complicated to achieve as first static files would need to be generated from the underlying data

RBINS is involved in a thematic ENVRI-FAIR working group (WP7), where these recommendations for CDI metadata are made:

* + link to the GetObservation matching the granularity of the CDI (=out-of-band) encoding. From the O&M, link to the Sensor ML using om:procedure.
  + Add L22 model as a keyword (upcoming as it requires an update to the CDI schema, keyword type list and MIKADO). Will be ready at SDN autumn/winter 2021.

The integration of the Belgica data into the SDN repository will be performed in two complementary steps:

First, the RBINS SOS (possibly subsetted for only Belgica data) should be added to the SDN Sensor Web Viewer Services. This is a basic and simple step as the SOS protocol is designed for seamless integration of different SOS sources into one data viewer. There have already been contacts with SDN representatives about this.

The second step is the creation of CDI, mapping and ODV files and the publication of those files via the replication manager developed by SDN. A prototype implementation using Mikado and a view inside the SOS database has been made for ENVRI-FAIR using the RBINS tripod data (June 2021). This prototype will be further developed in the framework of EMODnet DIP (summer-autumn 2021) and finally made active for ODANext in 2022. Crucial is the use of Mikado’s command-line batch generation of CDIs and associated files.

## Mapping

### O&M to CDI metadata mapping

To prepare the mapping, the following resources were investigated:

* <https://www.seadatanet.org/content/download/2746/file/2018_11_IMDIS2018_Abstract_Partescano.pdf> (a mapping of the XML elements from O&M to CDI)
* <https://www.seadatanet.org/content/download/636/file/SDN2_D85_WP8_Datafile_formats.pdf>
* <https://www.seadatanet.org/content/download/651/file/sdn_Mikado_UserManual_V3.6.3.pdf>

The mapping was done based on the SOS database, which can be translated into O&M. It was decided to follow the granularity of an offering, which corresponds to the full deployment period of a measurement device that can potentially measure multiple parameters. All mandatory fields have been mapped and the CDI validates. It will be investigated how additional variables can be added. Some values have been fixed for the RBINS use case. Natively, the SOS has no idea about observedproperty categories (P02) and device categories (L05). These concepts should be added and mapped to the P01 and L22 codes in the SOS. This mapping is based on a fixed benthic platform that does not move, where the location of the featureofinterest is sufficient. On the Belgica, in order to get the sample location and calculating the CDI bounding box, the observation.samplinggeometry needs to be read for each observation.

|  |  |  |  |
| --- | --- | --- | --- |
| table | Field | CDI variable | CDI variable name |
| offering | identifier | cdi\_identifier | CDI identifier |
|  | '3327' | var01 | METADATA CREATING ORGANISATION |
|  | 'point' | var02 | MEASURING AREA TYPE |
|  | '4326' | var03 | DATUM OF COORDINATE SYSTEM |
| offering | name | var04 | NAME/ALTERNATIVE NAME OF THE DATASET |
| offering | identifier | var05 | DATASET-ID |
|  | now() | var06 | REVISION-DATE OF DATASET |
|  | '3327' | var07 | ORIGINATORS OF THE DATASET |
|  | abstract based on the offering's name and measured parameters | var08 | ABSTRACT ON DATASET |
|  | '1578' | var09 | ORGANISATION MANAGING THE DATASET |
| *extended\_procedure* | p02\_identifier | var10 | PARAMETERS |
| *extended\_observableproperty* | l05\_identifier | var11 | INSTRUMENT and POSITIONING SYSTEM |
|  | '11' | var12 | PLATFORM TYPE |
|  | 'CB' | var14 | DATASET ACCESS RESTRICTIONS |
| featureofinterest | name | var18 | STATION NAME |
| featureofinterest | name | var19 | CRUISE NAME |
|  | now() | var20 | STATION NAME and/or CRUISE NAME |
| featureofinterest | xmin(geom) | var24 | GEOGRAPHICAL COVERAGE  WEST |
| featureofinterest | xmax(geom) | var25 | EAST |
| featureofinterest | ymin(geom) | var26 | SOUTH |
| featureofinterest | ymax(geom) | var27 | NORTH |
| series | min(firsttimestamp) | var28 | START DATE (AND TIME) |
| series | max(lasttimestamp) | var29 | END DATE (AND TIME) |
|  | '0.2' | var30 | MINIMUM DEPTH OF OBSERVATION |
|  | '2.5' | var31 | MAXIMUM DEPTH OF OBSERVATION |
|  | 'D09' | var34 | VERTICAL DATUM |
|  | '1578' | var36 | ORGANISATION DISTRIBUTING THE DATASET |
|  | 'ODV' | var37 | Dataformat |
|  | '0.4' | var38 | Version |
| procedure | A link towards the sensor ML with the procedure url argument set to the identifier of the procedure table | var40 | Distribution website |
|  | 'OGC:SOS' | var42 | Distribution protocol |
|  | 'CDIMTH05' | var43 | Distribution Method |

### O&M to ODV data mapping

For valid ODV, values for all observedproperties of an offering should be repeated as value columns for each point in time. In this prototype, all QC has been set to 2, ‘probably good value’. It will need to be investigated how to add quality information in the SOS, for each value. This is possible by using value parameters.

|  |  |  |
| --- | --- | --- |
| Table | column | ODV column |
|  | 'MOMO tripod' | Cruise |
| featureofinterest | identifier | Station |
|  | 'C' | Type |
| observation | phenomenontimestart | Event |
| featureofinterest | xmin(geom) | Longitude |
| featureofinterest | ymin(geom) | Latitude |
| offering | identifier | cdi\_identifier |
| numericvalue | value | valuecolumn |
|  | … | … |
|  | '2' | Qc |

# *O&M vs GOSUD*

Global Ocean Surface Underway Data (GOSUD) is an initiative of the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) programme.

GOSUD aims at assembling in-situ observations of the world ocean surface collected by a variety of ships and at distributing quality-controlled datasets. At present time the variables considered by GOSUD are temperature, conductivity and salinity.

|  |
| --- |
|  |
| Figure 1: Thermosalinometry dataflow from the research vessel to the Coriolis database and international data repositories. Credit to L. Drouineau |

## Technical implementation

The format used for near-real time data sharing to GOSUD is name ‘Colcor’[[15]](#footnote-15). It comprises a header that defines sensors metadata and rows that gives the measured values. This format is designed by IFREMER and a limited number of RVs make use of it for data sharing. Aware of this limitation, IFREMER showed some flexibility in the accepted format in the past.

The publication of data to the GOSUD repository is done in two phases:

1. Simplified near-real time data
2. Complete quality-controlled data with metadata (add calibration info, perform QC …)

In the context of the ODANext project, the objective is first to restart the near-real time data flow from the new Belgica to the GOSUD repository using a modern interoperable protocol such as O&M and SensorML. Data from the RV Belgica were shared until 2016 via email so there is a gap in the data between 2016-2020. Those data will be used to test the use of Sensor Observation Service protocol for near-real time data transfer.

The mappings show that both phases (simple data vs. QCed data including calibration information) can be resolved by the SOS+SML approach. The mapping also shows that a second approach is also possible, namely to program a colcor output library on top of the SOS DB+SML DB.

Data imported in the SOS are accessible via web-service. IFREMER could therefore harvest thermosalinity data of the RV Belgica data with requests such as:

|  |  |
| --- | --- |
| Request | <https://sensors.naturalsciences.be/sos/service?service=SOS&version=2.0.0&request=GetObservation&procedure=SEABIRD_SBE21&temporalFilter=phenomenonTime,2019-08-29T06:00:00Z/2019-08-30T06:00:00Z&responseFormat=http://inspire.ec.europa.eu/schemas/omso/3.0> |
| Timeserie metadata | <om:OM\_Observation gml:id="o\_6548855" xsi:type="ns:TrajectoryObservationType">  <gml:description>campaign</gml:description>  <gml:identifier codeSpace="unknown">2019/21a</gml:identifier>  <om:type xlink:href="http://inspire.ec.europa.eu/featureconcept/TrajectoryObservation"/>  <om:phenomenonTime>  <gml:TimePeriod gml:id="phenomenonTime\_6548855">  <gml:beginPosition>2019-08-29T06:10:00.000Z</gml:beginPosition>  <gml:endPosition>2019-08-30T06:00:00.000Z</gml:endPosition>  </gml:TimePeriod>  </om:phenomenonTime>  <om:resultTime>  <gml:TimeInstant gml:id="ti\_8C13B6ED7F9DC29CEBC3062CD64C7ADEE54D257B">  <gml:timePosition>2019-08-30T06:00:00.000Z</gml:timePosition>  </gml:TimeInstant>  </om:resultTime>  <om:procedure xlink:href="SEABIRD\_SBE21" xlink:title="SEABIRD\_SBE21"/>  <om:observedProperty xlink:href="http://bmdc.be/property/water\_temperature/" xlink:title="water temperature"/>  <om:featureOfInterest xlink:href="https://ears.bmdc.be/geoserver/ows?version=2.0.0&SERVICE=WFS&REQUEST=getfeature&typename=odas:ship\_track&SRSNAME=EPSG:4326&viewparams=campaign:2019/21a&outputformat=gml3"/> |
| Data | <wml2:point>  <wml2:MeasurementTVP xsi:type="ns:MeasurementTimeLocationValueTripleType">  <wml2:time>2019-08-29T06:30:00.000Z</wml2:time>  <wml2:value>18.41</wml2:value>  <ns:location>  <gml:Point gml:id="Point\_6548855\_2">  <gml:pos srsName="http://www.opengis.net/def/crs/EPSG/0/4326">53.260 0.518</gml:pos>  </gml:Point>  </ns:location>  </wml2:MeasurementTVP>  </wml2:point> |

With the ResponseFormat <http://inspire.ec.europa.eu/schemas/omso/3.0>, data is sorted per time-series (e.g. water temperature) as MeasurementTimeLocationValueTripleType. The first section of the answer provides all the time-serie metadata (e.g., procedure, feature of interest, campaign…). The second section is a list of all the points with the time, the geographical location and the actual measured value.

The second approach, programming the Colcor format based on the databases underlying the SOS and the SML, ie. the 52°N SOS DB and a (still to be designed) sensor database is possible as well, if IFREMER is not interested in reading O&M and SML. Most straightforward is offering this data as an http endpoint; harvesting is then IFREMER’s responsibility. The response from the EARS webservices (ears2Nav, only data) is very similar to the Colcor format, but misses sensor and calibration information. The challenge will be to query both databases at the same time for each data request.

## Mapping

### O&M to GOSUD Colcor format

The Colcor format has been analysed and compared with the fields and variables available in the Sensor Observation Service database as developed by 52° North. It is an open question what the mandatory elements are. Header information can be delivered without a problem; however the calibration curve coefficients are requested in high detail. It needs to be asked to IFREMER what exactly the calibration coefficients are.

**Colcor metadata header**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| vessel\_code | first\_data\_date | instrument | sensor | serial\_number | install\_date | calib\_date | g | coef |  |
| #$xxCOR | first\_data\_date | SBE21 | CONDUCTIVITY | serial\_number | install\_date | calib\_date | g | coef | ... |
| #$xxCOR | first\_data\_date | SBE21 | TEMPERATURE | serial\_number | install\_date | calib\_date | g | coef | ... |
| #$xxCOR | first\_data\_date | SBE3 | TEMPERATURE | serial\_number | install\_date | calib\_date | g | coef | ... |
| #$xxCOR | first\_data\_date | SBE3 | IMMERSION | immersion |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| #$ATCOR | 03/08/13 | SBE21 | CONDUCTIVITY | 3237 | 24/10/12 | 09/09/11 | g | -4.22398567E+00 | ... |
| #$ATCOR | 03/08/13 | SBE21 | TEMPERATURE | 3237 | 24/10/12 | 02/02/07 | g | 4.23084440E-03 | ... |
| #$ATCOR | 03/08/13 | SBE3 | TEMPERATURE | 217 | 24/10/12 | 05/09/11 | g | 4.00000000E-03 | ... |
| ‘#$BECOR’ | min(observation.phenomenontimestart) | procedure.name | Fixed | Sensor database: OK: probe.serial\_number | Sensor database: OK, probe.installation\_datetime | Sensor database: OK, probe.last\_calibration\_date |  | ? | ? |
| #$ATCOR | 03/08/13 | SBE3 | IMMERSION | 3.50000000E+00 |  |  |  |  |  |
|  |  |  |  | Sensor database: OK, instrument.inlet\_depth |  |  |  |  |  |

All metadata can be expressed in the SOS database (blue) and the database foreseen for sensor information (orange). All metadata can also be expressed in respectively O&M and SensorML, although some fields get distributed quite far away. It is clear in O&M the responses become much more verbose. some observations:

* Start date of observations in file cannot be accommodated literally, but is of course implicit in the earliest phenomenon time. The problem is that the start date will never be the actual start date of the whole series/offering, because the earliest date of the response is dependent on the date range given in the url arguments. IFREMER will harvest the data incrementally, for instance once daily.
* Installation and calibration dates are part of the SML history and need to be requested from where they are stored.
* Intake depth is also in the SML and one of the sensor capabilities. Needs to be requested in <http://vocab.nerc.ac.uk/collection/W05/current/>

**Colcor data row**

A mapping has been made in Excel, which is impractical to show here. Fields/variables that have some difficulty to be mapped are:

* The reduction masks, we will check with IFREMER what they do and if they are mandatory
* Temperature at water intake
* “temperature of sea water” Intake T be taken for this? Part of meteo package?
* Pluviometer: presumed not available
* Dew point: presumed not available
* Apparent vs true wind speed and direction (both are available in ODAS)
* Flow speed in l/minute (where: at the water intake, over the whole pipe, at the entry of the instrument?)

1. <https://www.ogc.org/node/698> [↑](#footnote-ref-1)
2. <https://inspire.ec.europa.eu/> [↑](#footnote-ref-2)
3. <https://www.seadatanet.org/> [↑](#footnote-ref-3)
4. <http://www.gosud.org/> [↑](#footnote-ref-4)
5. https://www.ogc.org/standards/sensorml [↑](#footnote-ref-5)
6. <http://docs.ogc.org/is/12-000r2/12-000r2.html#65> [↑](#footnote-ref-6)
7. <https://inspire.ec.europa.eu/about-inspire/563> [↑](#footnote-ref-7)
8. <https://inspire.ec.europa.eu/theme/ef> [↑](#footnote-ref-8)
9. <https://inspire.ec.europa.eu/theme/of> [↑](#footnote-ref-9)
10. https://www.wetransform.to/ [↑](#footnote-ref-10)
11. https://inspire.ec.europa.eu/id/document/tg/ef [↑](#footnote-ref-11)
12. https://www.seadatanet.org/About-us [↑](#footnote-ref-12)
13. <https://www.seadatanet.org/content/download/3738/file/SDC_WP10_D10.17_SpecificationSOSviewingServices.pdf>

    <https://www.seadatanet.org/Software/Sensor-Web-Viewer/How-does-it-work> [↑](#footnote-ref-13)
14. <https://oceans.dev.52north.org/helgoland/timeseries/diagram> [↑](#footnote-ref-14)
15. <http://www.gosud.org/content/download/24640/169428/file/Colcor_format_2015_09_24.pdf> [↑](#footnote-ref-15)