

Title:

**Title: A7.4-D3 Profiles for ocean and marine environment (Interim version)**

Title: A7.4-D3 Profiles for ocean and marine environment (Interim version)

**Author(s)/Organisation(s):**

Marian de Vries (TUD), Peter Thijssse (MARIS), Patricia Moreno-Arancibia (SMHI), Sophie Baudel (CLS), Olivier Lauret (CLS), Eric Moussat (Ifremer), Mark Charlesworth (BODC), Vasilis Lakes (HCMR), Sisi Zlatanova (TUD)

**Working Group:**

WP7

**References:**

758-hs-ocean\_-\_aegan\_sea\_use\_case\_document-fhg-igd-001-new

759-a9\_8d1-baltic-smhi-001-new

760-a9\_8-d1\_france-ifremer-002-new

761-a9\_8-d1\_uk-nerc-bodc-001-new

837-a7\_4-d2\_profiles\_for\_ocean\_and\_marine\_environment-tud-002-final

843-d9\_8d1\_hcmr\_-\_updated\_version\_of\_scenario-hcmr-002-new

904-a9\_8d1system\_specification\_ocean\_umbrella\_scenari-maris-001-new

921-ocean\_data\_sets\_description-hcmr-001-new.xls

**Quality Assurance:**

<input checked="" type="checkbox"/> Review WP Leader	(WP Leader: )
<input type="checkbox"/> Review dependend WP Leaders	(Depended WP Leaders: )
<input type="checkbox"/> Review Executive Board	(Executive Board Members: )
<input type="checkbox"/> Review others	(Other Reviewers: )

**Delivery Date:**

**Short Description:**

This document contains the final version of the data specification for the Ocean Scenario. It is based on the WP9 system specification for the Scenario, and on the information analysis discussions leading to the construction of the common Scenario data model. This common data model supports the use cases in the Scenario, that are centred around the monitoring and reporting of oil spills, with special attention to the risk of oil spills to ,sensitive' areas (economically and environmentally sensitive areas).

Title:

The Ocean Scenario has in the first phase of the system specification split-up the work into 4 sub-scenarios: Baltic Sea, French coast, English coast, Aegean Sea. In the course of 2009 also a combined cross-regional scenario has been developed, with focus on gathering statistics about past oil spills. The common data model of the Scenario will for the final version be extended to cover this new use case.

This final version provides a broader scope making a link to similar or related models, projects and standards.

**Keywords:**

Ocean Scenario, information analysis, conceptual data model, harmonisation issues, oil spill monitoring and reporting, forecasting

<b>History:</b>			
Version	Author(s)	Status	Comment
000	Marian de Vries	new, rfc	Initial set up
001	Peter Thijsse	Rfc	Scope extension
002	Vassilis Lakes	Rfc	Scope extension
003	Peter Thijsse	Rfc	Model adaptation
003	Sisi Zlatanova	Final	Check content

Title:

## Table of contents

1	Purpose and scope .....	4
2	Profile of the Scenario .....	4
2.1	Data used to initialize and force the oil drift calculation .....	5
2.2	Data produced by the system .....	6
2.3	Data used to aid in the visualisation of the results .....	6
3	Common data model .....	7
4	Relevant data models .....	11
4.1	OGC Observations & Measurements .....	12
4.2	SensorML .....	13
4.3	CSML (Climate Science Modelling Language) .....	14
5	Related projects .....	15
6	Glossary .....	16
Annex A: Data harmonisation issues .....		17
1.1	Data format .....	17
1.2	Spatial and temporal reference systems .....	17
1.2.1	Spatial reference system .....	17
1.2.2	Temporal reference system (time) .....	17
1.3	Conceptual data model .....	18
1.3.1	Geometry types .....	18
1.3.2	Identification and Versioning .....	18
1.3.3	Importance of time / Dynamic data .....	18
1.4	Classification .....	18
1.5	Terminology .....	19
1.6	Metadata .....	19
1.7	Scale/resolution, level-of-detail, aggregation .....	20
1.8	Portrayal .....	20
1.9	Multi-linguality .....	20
1.10	Spatial consistency .....	20
1.11	Maintenance .....	20

Title:

## 1 Purpose and scope

The Ocean Scenario specifications are based on a number of documents and on the information analysis discussions leading to the construction of the common data model for the Scenario. This common data model supports the use cases in the Scenario, which are centered around the monitoring and reporting of oil spills, with special attention to the risk of oil spills to ,sensitive' areas (economically and environmentally sensitive areas). At the same time the common data model for the Scenario serves as the core of a somewhat broader application profile for the ocean and marine application domain.

The Ocean Scenario has in the first phase of the system specification split-up the work into 4 sub-scenarios: Baltic Sea, French coast, English coast, Aegean Sea. The system specification of the Baltic Sea sub-scenario is taken as basis, and the other scenarios are used to add to the information. In the second phase of the system specification a combined cross-regional scenario has been developed, with several reporting uses cases. The French coast and Aegean Sea sub-scenarios have been extended.

The structure of this document is as follows:

Chapter 2 gives an overview of the spatial information involved in the Scenario use cases. A distinction is made between input and output information, but in some cases the input data (used for forecasting especially) is in a later stage also needed for analysis and reporting, for example to display the drift and severity of the oil spill event in combination with weather and ocean conditions at that time.

Chapter 3 presents the common data model for the Scenario (current version) in the form of UML class diagrams.

Chapter 4 discusses related models and standards in the broader scope.

Chapter 5 presents the projects dealing with similar topics.

Chapter 6 lists two open issues, of which especially the first is relevant: not all data sources needed in the Scenario are at the moment ,freely' available. Nevertheless it was decided not to adapt the common data model too much to the availability of data.

## 2 Profile of the Scenario

The overall Scenario application functionality is as follows: Oil spill detection, monitoring and reporting, with special attention to the risk to ,sensitive areas' along the sea or ocean coast.

In the Baltic-Sea sub-scenario the following use cases have been specified:

- Use Case 1 (UC1): Calculate the drift of an oil spill
- Use Case 2 (UC2): Forecast when and where an oil spill reaches a sensitive area
- Use Case 3 (UC3): Find ship responsible for an oil spill

In the Aegean Sea and the French sub-scenarios an extra use case is specified:

- Use Case 3 (UC3): Initializing the processes of the (monitoring and forecasting) system

The French sub-scenario has introduced in addition the use case:

Title:

- Use Case 4 (UC4): Make available additional data

This fourth use case was introduced because at the moment the news of an oil spill comes in, and the monitoring and forecasting system is initialized, a large number of data sets are needed as reference data (coastline, administrative units) or as input to the computational models (weather and ocean data for that region, etc.). Preparing data in advance is never an option since the probability nature of the event. This is true for dynamic but also for static data. Static reference data has to search for and upload at the last moment to avoid out-of-date or not useful data sets and maps. Furthermore, during the monitoring of the oil spill when the location of the oil spill changes, new reference maps or additional input data must again be added to the system.

All the use cases are 'operational' use cases, with mostly dynamic, real-time data. Only a small part of the data can be pre-loaded before an event, and many data transformation steps need to be carried out on-the-fly, because the data cannot be prepared beforehand.

The description of the information used in the Scenario is divided into three sections: data needed to initialize and force the oil drift calculation, data that the system produces and data used for the visualisation of the results. The data are presented in the tables below

## 2.1 Data used to initialize and force the oil drift calculation

<b>Information item</b>	<b>Purpose</b>	<b>Input/output/internal</b>
Weather data <ul style="list-style-type: none"> <li>• Wind data</li> </ul>	Wind speed and wind direction are used as parameters to compute the oil spill drift, and to predict the weathering processes that influence the oil spill.	Input
Ocean data <ul style="list-style-type: none"> <li>• Current velocity</li> <li>• Temperature</li> <li>• Salinity</li> <li>• Turbulence parameters</li> <li>• Ice concentration</li> <li>• Ice drift velocity</li> </ul>	Oceanographic data is used to <ol style="list-style-type: none"> <li>1. force the transport and spreading of the oil in the oil drift model</li> <li>2. to calculate the chemical state of the oil (evaporation, emulsification, etc)</li> </ol>	Input
Coastline	To determine the geographical location;  To determine the boundary of the computational grid.	
Bathymetry	To determine the geographical	

Title:

	location;  To determine the depth of the computational grid.	
Initial position of the oil spill	To initialize the oil drift calculation	input, stored for reporting (PolRep = Pollution report)
Amount of oil	To initialize the oil drift calculation  Note: in the French case maybe not relevant as input (depends on MOTHY model)	input, stored for reporting (PolRep = Pollution report)
Type of oil	To set density, viscosity, and empirical parameters that determine the weathering process of the oil. These parameters determine how the oil is transported in the ocean.	input, stored for reporting (PolRep = Pollution report)

## 2.2 Data produced by the system

Information item	Purpose	Input/output/internal
Location of oil spill	Of interest to end-user	Output
Mass fraction of the oil that has evaporated, dispersed, and surface or stranded	Of interest to end-user	Output

## 2.3 Data used to aid in the visualisation of the results

Information item	Purpose	Input/output/internal
General GIS layers: coastline, national borders, territorial water boundaries, etc	Orientation purposes	Reference layers, needed for reporting
GIS layer of sensitive areas: Baltic Sea Protected Areas, important bird areas, Maris (Maritime Accident Response Information System) zone	To determine when and where an oil spill reaches sensitive areas	(static) input, also needed for reporting
Forcing <ul style="list-style-type: none"> <li>Wind speed</li> </ul>	To aid in the calculation (during the event) and analysis (after the event) of the drift of the oil spill	input, also stored for reporting  (Some of the data used to produce the oil drift calculation i.e. wind speed, surface

Title:

<ul style="list-style-type: none"> <li>• Surface currents</li> <li>• Ice concentration</li> </ul>		currents and ice conditions at the time of the oil spill, should be stored for later analysis.)
<p>AIS (Automatic Information System) data</p> <ul style="list-style-type: none"> <li>• Ship ID and name</li> <li>• Position every 15 minutes</li> <li>• Destination</li> </ul>	<p>To determine which ships are (or have been) in the area where the pollution occurs;</p> <p>If needed to help identify which ship is responsible for the oil spill.</p>	<p>input, stored for reporting</p> <p>Depending on the further development of the use cases in the Scenario, this information item can become optional.</p>

For the broader application domain Profile (Ocean and marine environment) the following information is of importance and should be added:

- Real time and near real time data
- Marine biodiversity
- Model output: climate change information (time series), other weather and ocean dataproducts (not used for oil spill forecasting but for other research or operational situations)
- How to describe (with metadata) and manage in-situ sensors
- Controlled vocabularies with marine/ocean terms (SeaDataNet vocabularies)
- Ecological protection of sea and coasts in general
- Production of synthetic Data Sets (outputs) with strategic information regarding environmental and security issues. These Data Sets may be valuable for decision makers in case of events with environmental hazards.

### 3 Common data model

As discussed previously may data within the Scenario needs to be harmonised (and based on that, many harmonization issues can be identifies). The most important aspect in the Scenario is the ,output' data, because this indicates what information content the Scenario will produce/publish to the outside world. This output information needs to be homogeneous (same data model, same spatial reference system, when necessary same scale/resolution, same portrayal) in all 4 sub-scenarios. This output information can be published on a website that shows the current and forecasted position of the oil spill during the incident, or can be used for reporting afterwards in studies that give a risk/damage calculation of oil spills in Europe in a certain year.

The reason for this present scope of the data model is that what is designated as ,input data' in the use case descriptions (e.g. weather and ocean conditions) does not necessarily have to be homogeneous. Which of the mentioned input data needs to be harmonised depends on the further detailing of the system specification (see also A7.4 D2).

Title:

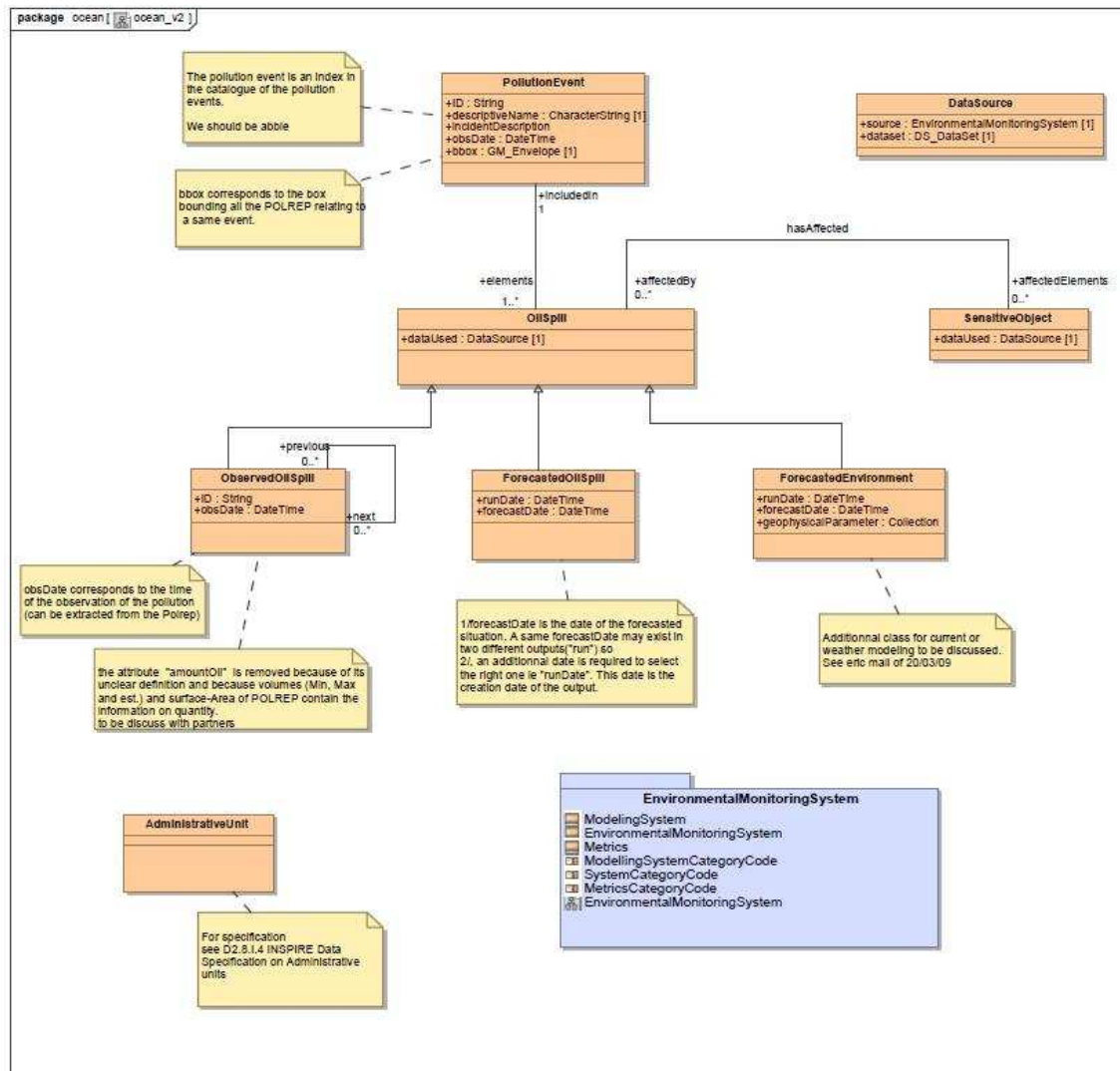


Figure 1 Data model for oil spill monitoring and reporting (output and reference data)

Therefore the UML class diagram in Figure 1 contains only the core output information (as this was understood from the WP9 system specification) and not the input data that is specified in the WP9 tables, like weather and ocean conditions, bathymetry and tide data. Nevertheless, while not a high priority issue, a harmonisation of input data can also be helpful to the end-user (e.g the observed height of waves at a specific place to make a comparison with an hydrodynamic model forecast). At the moment there is a collection of attributes called 'geophysicalParameters' in the class ForecastedEnvironment, as a placeholder. There will probably also be an ObservedEnvironment class in the next version of the model, with a comparable solution for the parameters that are input for the computational models.

As it can be seen the model is built around OilSpill class. The main classes of interest for the scenario are ObservedOilSpill (coming from a Pollution Report, POLREP, being a manmade report of an observation), ForecastedOilSpill, ForecastEnvironment, SensitiveObject and PollutionEvent (multiple Pollution Reports form an event). The data model contains the main object classes and their attributes and associations with other object classes. For some object classes also one or more operations are

Title:

indicated (last block in the class, after the attributes), to denote that calculations are needed to derive the value of some of the attributes from other (static or dynamic) data.

This is a high-level conceptual data model, meaning that details like 'is this attribute a string, or a number' are still left out. There are some exceptions: for example in case of the spatial attributes we indicated whether it is a line, point, or polygon, or something else; and in case of a timestamp (time/date) attribute this is also indicated.

Two kinds of information are defined according to the INSPIRE Annex I data specifications: AdministrativeUnit and ProtectedSite. These two classes are commented in the models.

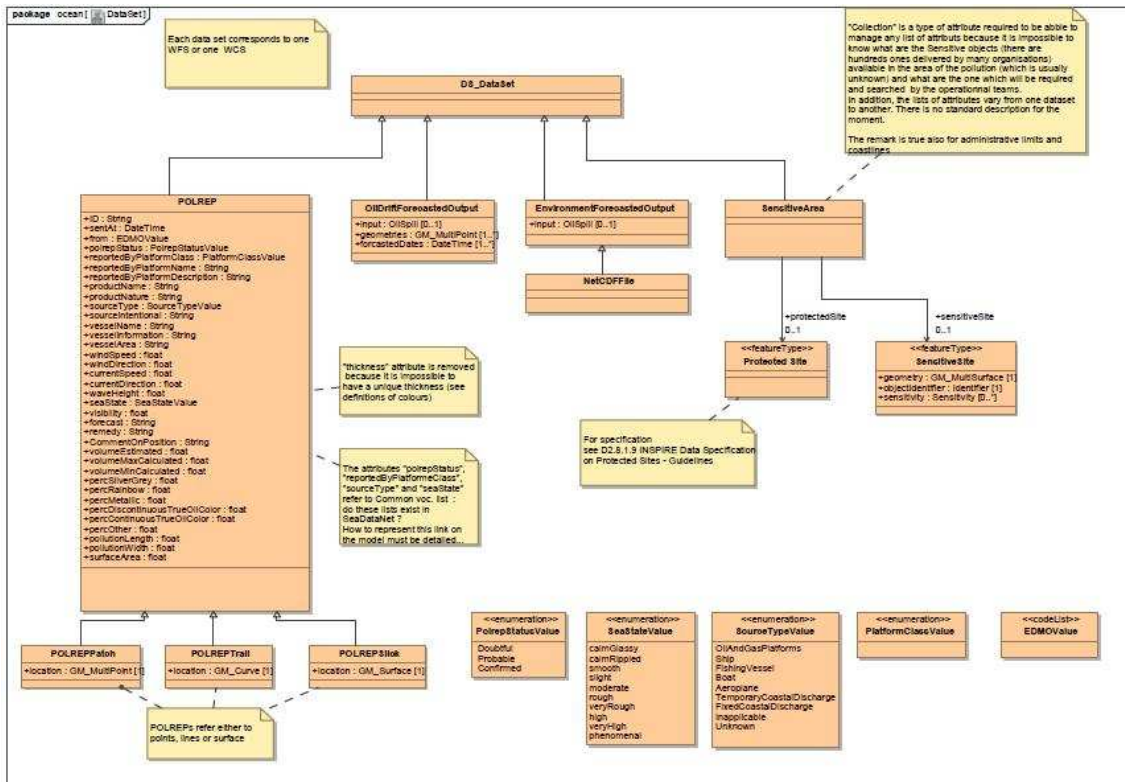


Figure 2: Data model for the data sets to be used with WFS and WCS

Figure 2 represent the data that used in the web environment. All the data sets are to be requested by WFS or WCS

In both UML schemas, the prefix GM\_ means the official ISO 19107 names for the geometry types are used. For oil spills that are represented as lines (trails) or polygons (slick) it is not known what common practice is, therefore this question arises: if an oil spill trail or slick consists of several parts that are not connected: is it useful to still consider this as one oil spill? In that case the ,multi'geometry types must be used: GM\_MultiCurve or GM\_MultiSurface. Otherwise the geometry type should be as it is now (GM\_Curve or GM\_Surface). As can be seen in the data model in Figure 1 several types of geometry of an oil spill (observed and forecasted) are possible depending on the situation.

Title:

For France the situation is as follows (also making a distinction between observations of oil spills and forecasts):

**Observations:** On the SAR images, slicks are represented by polygons. However in the POLREP, the geometry of an observed oil slick is either a point, or a line (2 points) for discharge of hydrocarbons at sea, or a polygon (usually there are 4 points but it may be more). Each large observed slick, separated from others by several km, is considered as a separate 'object' of which the drift should be surveyed independently.

**Forecasts:** In France, the MOTHY forecast output of the oil spill drift has a point geometry. The output of a run of the forecast model gives the most likely positions of 500 elements starting from a same given point. According to Cedre (Centre de Documentation, de Recherche et d'Expérimentations sur les pollutions accidentelles des eaux), in the future the initial position of a slick will not be any more given by a point or by a set of points but by a polygon (surface).

While Figure 1 gives the overview model, some parts are further detailed in the Figures 3 and 4.

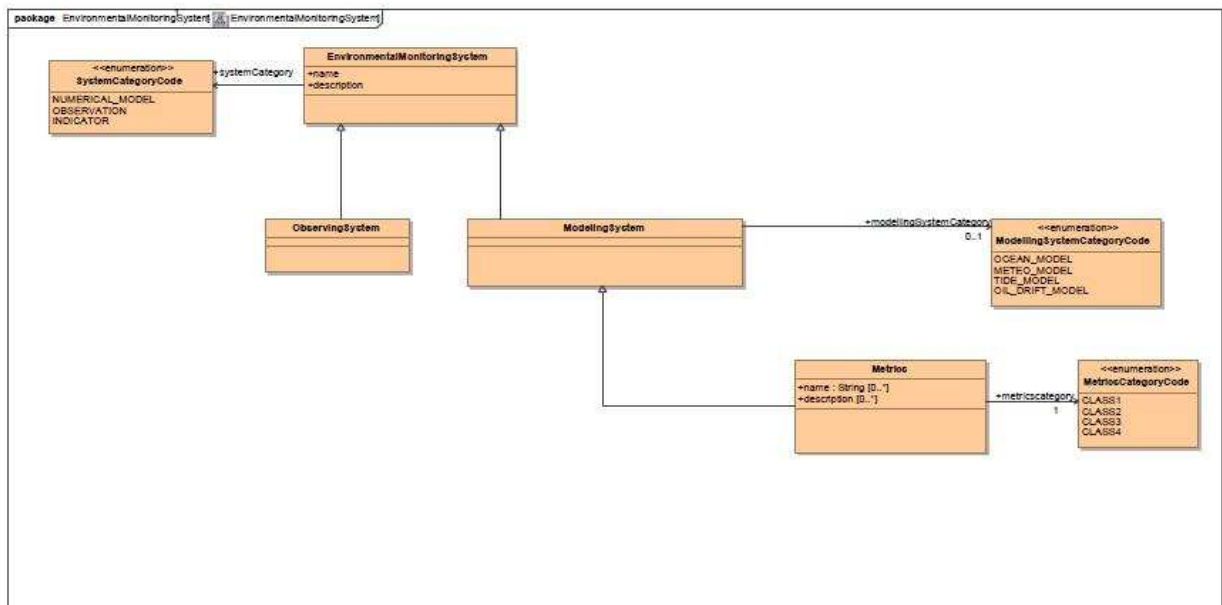


Figure 3 Data model part: Environmental Monitoring System

Figure 3 gives a sketch of the EnvironmentalMonitoringSystem class and its subclasses. Figure 4 shows the ObservingSystem class, its subclasses and associations.

Part of the defined classes have to do with metadata about the observing and monitoring systems and instruments used, and also can be considered belonging to the metadata collected and published in the Scenario. In Figure 4 a hint to this is the fact that for the organisation responsible for the observing platform and the mission, the ISO 19115 type 'ResponsibleParty' is re-used.

Title:

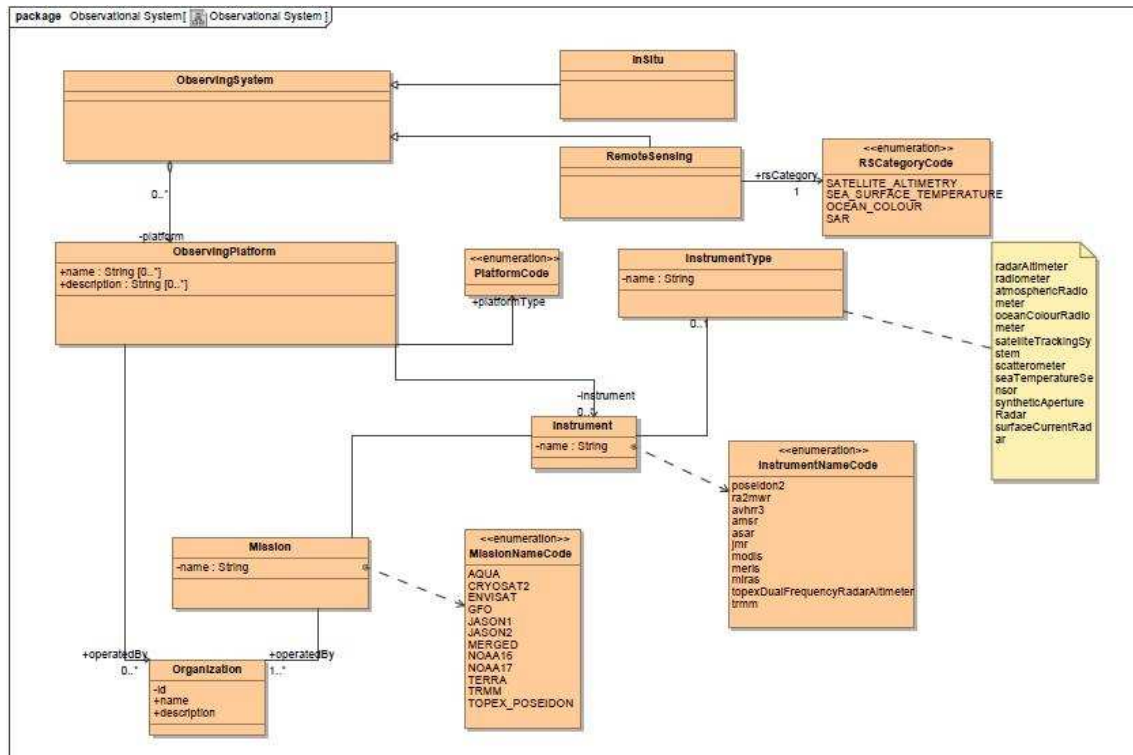


Figure 4: Data model part: Observing System

## 4 Relevant data models

Within the enlarged scope of the scenario, several new kinds of information is needed. This information can either be modeled or re-used from existing models. In this respect we have studies and identified several data models and standards that are relevant for this Scenario. Most important of them are listed below:

- CSML (Climate Science Modelling Language) <http://csml.badc.rl.ac.uk/>
- NERC/BODC thesaurus / vocabulary [http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)
- IHO S-100 „The New Hydrographic Geospatial Standard for Marine Data and Information“ [http://seabass.ieee.org/groups/geoss/index.php?option=com\\_sir\\_200&Itemid=157&ID=128](http://seabass.ieee.org/groups/geoss/index.php?option=com_sir_200&Itemid=157&ID=128)
- IHO S-57 for Electronic Navigational Charts (ENC) <http://marinemetadata.org/references/ihos57>
- NetCDF CF datamodel [http://www.seadatanet.org/standards\\_software/data\\_transport\\_formats](http://www.seadatanet.org/standards_software/data_transport_formats) or <http://www.unidata.ucar.edu/software/libcf/>
- ODV ASCII datamodel [http://www.seadatanet.org/standards\\_software/data\\_transport\\_formats](http://www.seadatanet.org/standards_software/data_transport_formats)

Since the Scenario uses different sensors to collect information about the movement of the oil spill, the weather conditions and environmental conditions, two of the OGC standards are also of great importance:

Title:

- SensorML, and the OGC specifications related to sensor networks  
<http://www.opengeospatial.org/standards/sensorml>
- OGC Observations & Measurements <http://www.opengeospatial.org/standards/om>

The interdependencies between OGC standards is shown in Figure 5.

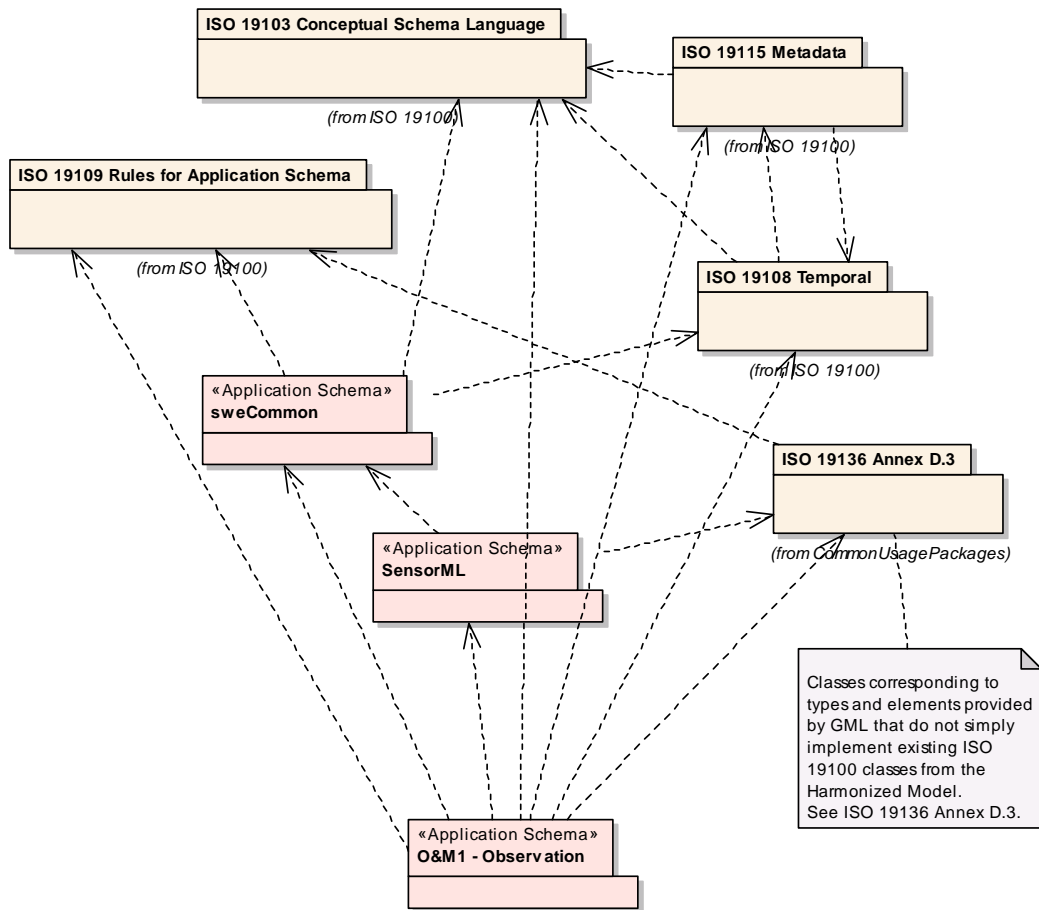


Figure 5: Dependencies between application schemas in Sensor Web context, source: OGC 2007b

As can be seen in this figure, both SensorML and Observations & Measurements are based on the ISO 19xxx set of standards.

There two standards and CSML are given with more details in the sections below.

## 4.1 OGC Observations & Measurements

Observations & Measurements (OM) provides general models and schema for supporting the packaging of observations from sensor system and sensor-related processing. The O&M Observation XML encoding is very general in the sense that the result can be packaged in any structure specified in XML. The Observation and Measurement (O&M) standard defines an abstract model and an XML schema encoding for observations and measurements. This framework is required for use by other

Title:

OGC Sensor Web Enablement (SWE) standards as well as for general support for OGC compliant systems dealing in technical measurements in science and engineering.

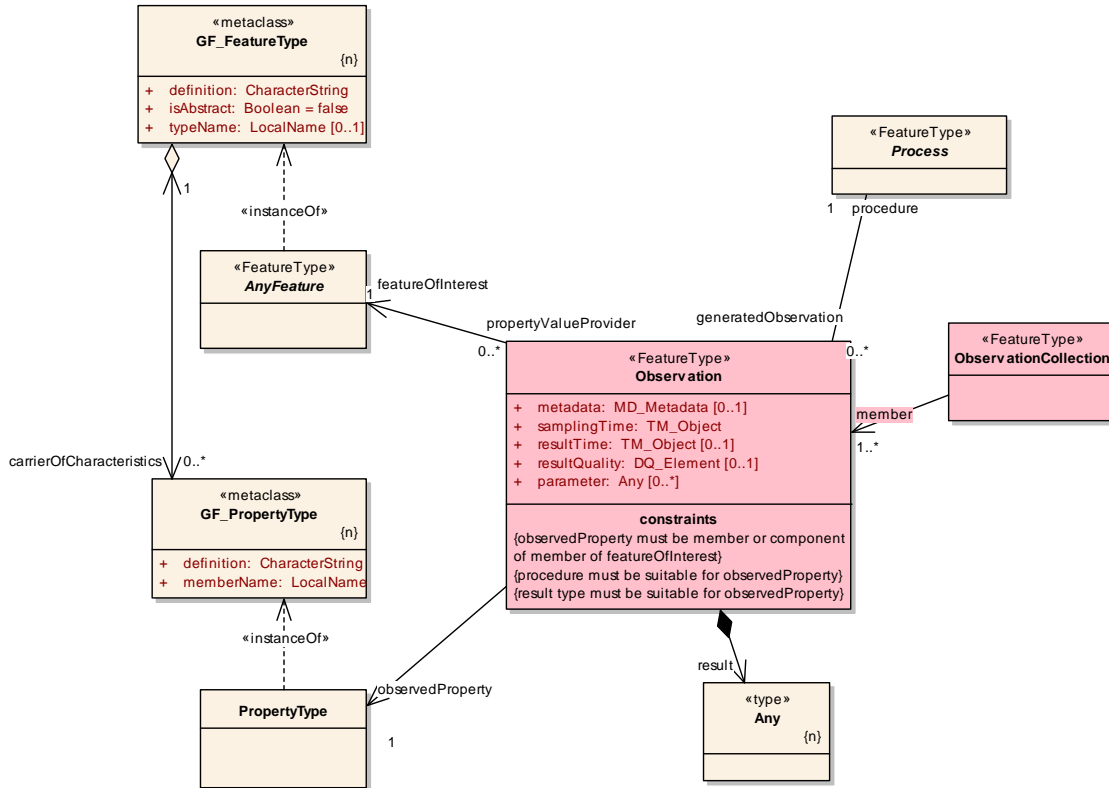


Figure 6: Core of the O&M data model: the Observation feature class (source: OGC 2007b)

Figure 3 shows the core class of the O&M data model: the 'Observation' feature type.

## 4.2 SensorML

SensorML (Sensor Model Language) is an OGC implementation standard that provides data and process models and an XML encoding for sensor measurements and processes, including methods for deriving higher-level information from sensor observations.

Processes described in SensorML define their inputs, outputs, parameters, and methods, as well as provide relevant other metadata. With SensorML detectors and sensors can be modeled as processes that convert measured phenomena into (geo)data.

Applications that use sensor data as input have to be able to process series of data items that somehow have to be kept together in aggregate data types (Figure 7).

Title:

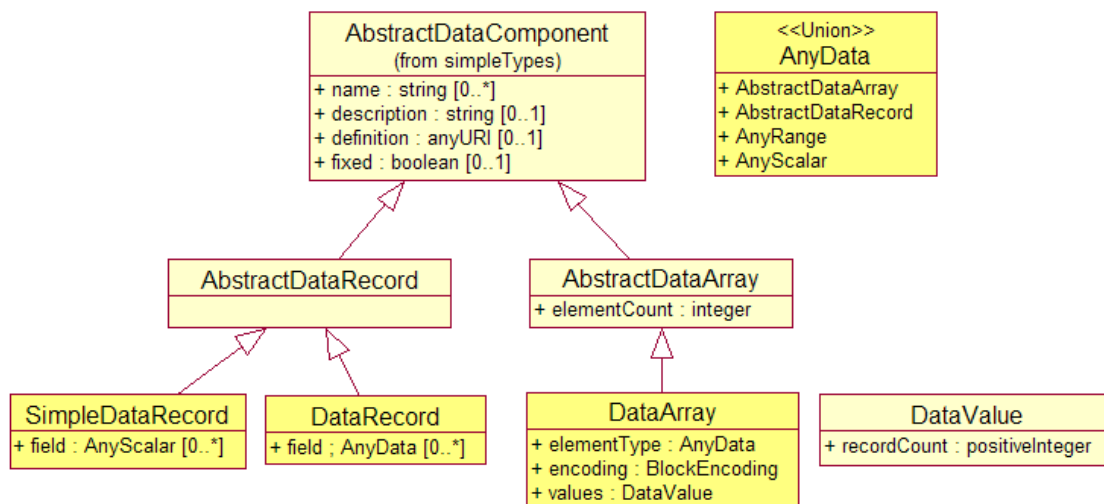


Figure 7: SensorML: model elements for data aggregates (source: OGC 2007a)

SensorML offers several options for data format (the encoding). For singular values there is a 'value' element:

```

<swe:DataRecord definition="urn:ogc:def:property:OGC:atmosphericConditions">
  <swe:field name="AirTemperature">
    <swe:Quantity definition="urn:ogc:def:property:OGC:AirTemperature">
      <swe:uom code="Cel"/>
      <swe:value> 35.1 </swe:value>
    </swe:Quantity>
  </swe:field>
  <swe:field name="WindSpeed">
    <swe:Quantity definition="urn:ogc:def:property:OGC:WindSpeed">
      <swe:uom code="m/s"/>
      <swe:value> 6.5 </swe:value>
    </swe:Quantity>
  </swe:field>
</swe:DataRecord>

```

Series of measurements can however be encoded as text block, with the 'header' information included in the DataArray element (see Annex C for the complete fragment):

```

<swe:values>
  2006-10-05T12:30:00Z 35.1 950.0 32.0 clear,
  2006-10-05T13:00:00Z 35.8 940.0 331 clear,
  2006-10-05T13:30:00Z 36.5 938.0 35.8 hazy,
  2006-10-05T14:00:00Z 38.0 935.0 37.0 cloudy
</swe:values>
</swe:DataArray>

```

### 4.3 CSML (Climate Science Modelling Language)

CSML is a standards-based data model and GML (Geography Markup Language) application schema for atmospheric and oceanographic data. The key subcomponents of CSML are:

Title:

- Feature Type Definitions: A set of UML conceptual models ('feature types') for a range of atmospheric and oceanographic data types, based on the framework and components provided by ISO standards for geospatial information modelling.
- CSML itself (an application schema of GML), built around these feature types providing a standards-based reference encoding for atmospheric and oceanographic datasets.

See the CSML user guide for further information about the CSML data model and XML encoding (Woolf and Lowe, 2007).

## 5 Related projects

This is based on the table in chapter 5 of the WP9 system specification report that was originally filled in per sub-scenario.

CDI (Common Data Index), [http://www.nodc.nl/v\\_cdi\\_v1/search.asp](http://www.nodc.nl/v_cdi_v1/search.asp)

EMSA CleanSeaNet satellite service, <http://cleanseanet.emsa.europa.eu/>

Global Marine Oil Pollution Information Gateway, <http://oils.gpa.unep.org/facts/glossary.htm>

Huijjer, K. (2005). Trends in Oil Spills from Tanker Ships 1995-2004. 28th Arctic and Marine Oilspill Program (AMOP) Technical Seminar. Calgary, Canada.

International Maritime Organization (IMO), <http://www.imo.org/>

International Tanker Owners Pollution Federation Limited (ITOPF), <http://www.itopf.com/>

InterRisk project (GMES), <http://interrisk.nersc.no/>

MAGIC: Coastal and Marine Resource Atlas (UK)

MarCoast, <http://213.236.12.71:8010/marcoast/htdocs/index.html>

MARitime Security Service (MARISS)

MERSEA (Marine EnviRonment and Security for the European Area - Integrated Project), [http://www.mersea.eu.org/Information/product\\_template\\_help.html](http://www.mersea.eu.org/Information/product_template_help.html) (metadata profile) and <http://behemoth.nerc-essc.ac.uk/ncWMS/mersea.html> (WMS service(s) and viewer)

MOTIIVE, <http://www.motiive.net/>

NERC DataGrid vocabulary server, [http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)

Oceanides project (GMES), <http://oceanides.jrc.it/>

SeaDataNet, <http://www.seadatanet.org/>

U.S. National Oceanic and Atmospheric Administration (NOAA), <http://response.restoration.noaa.gov/>

Title:

## 6 Glossary

MOTHY model = **M**odèle **O**céanique de **T**ransport d'**HY**drocarbures, a model developed by Météo-France to predict the drift of pollutants on the ocean surface

POL model = model of the **P**roudman **O**ceanographic **L**aboratory on which the UK oil spill model is based

POLREP = Marine Pollution Incident Report, used by France

Title:

## **Annex A: Data harmonisation issues**

This is based on the A7.4-D2 (interim version).

### **1.1 Data format**

#### **France**

SAR Radar data may be in several formats (CEOS, PDS, GeoTIFF). The harmonization process must be able to convert the SAR data to the chosen raster format for HUMBOLDT (GeoTIFF). Synoptic visualisation of raster data (netCDF format for ocean model, GRIB format for weather model, vector format for oil drift forecast) is an important issue for the Ocean Scenario. For vector data the goal is to publish the data preferably via a WFS (output GML).

#### **UK**

The MAGIC data sources are ESRI shapefiles, the BODC data are in netCDF (QXF). A consistent format between data provided by MAGIC, BODC and the POL model would aid quicker interpretation by the Synthesiser and Decision Maker in the use cases.

### **1.2 Spatial and temporal reference systems**

#### **1.2.1 Spatial reference system**

The spatial reference system for the generated data in this Scenario will be WGS84. Coordinate transformation is necessary when data from different countries is combined, for those (land, coastline) datasets that are in another (national) spatial reference system.

For the static data this can be done beforehand (offline), but if there are dynamic data that are not in WGS84 the coordinate transformation must be carried out on-the-fly.

#### **France**

The spatial reference systems used to be different between land and sea as the user needs were different. At present, there is a single – legal -Geodetic Reference System compatible with the global WGS84. However, many documents are still in the old legal system (NTF on land and ED50 for old nautic documents) and projections differ between land and sea. In addition, recently new legal local conic projections have been published for land areas. Combining land and sea data is therefore often a challenge in France.

#### **UK**

Input data sources (MAGIC and BODC) are already in WGS84 (decimal lat/long).

#### **1.2.2 Temporal reference system (time)**

In the data sets in use, dates refer usually to the Gregorian calendar and time is expressed in UTC time in data sets. When observations refer to local time (or should be given in local time), time is easy to convert.

Title:

Recommendation: UTC will be the reference time.

## 1.3 Conceptual data model

**UK:** The conceptual models for MAGIC, BODC and the POL model differ which is acceptable. It is important to be able to harmonise the outputs from these 3 sources of information, so for the output data of the Scenario a common data model is needed (see chapter 3)

### 1.3.1 Geometry types

**France:** Harmonisation of the geometry types is not an issue at present as these are standard ones: point, line, polygon without topology and tables.

### 1.3.2 Identification and Versioning

Most input data sources will have id's for the objects, and the id's will be unique within that data set. When different data sources are combined a convention to make the id's also unique in the merged result must be adopted. INSPIRE provides some recommendations (options) how to do this. We will make a choice in a later stage.

### 1.3.3 Importance of time / Dynamic data

Very important due to the oil spill monitoring and forecasts aspect of the use cases. Also for post-event reporting and analysis the storage of time and date of observations and forecasts are essential. There are two time variables in the forecasts: the time of the run, and within the output of the run, the time of the forecasted phenomenon as calculated in that specific run.

## 1.4 Classification

'Classification' and 'data model' are closely related. 'Classification' can influence class and sub-class hierarchies in the data model, but that is not the case in this Scenario (in the draft version of the data model). Rather the attribute values will be influenced by the classifications chosen.

For all attributes that have a fixed list of possible values (an enumeration attribute) the Scenario will try to use already existing standard (official or not official) classifications, e.g. for the type of administrative region, or for the type of oil.

### France

Classification is an important issue, especially for sensitivity atlases, which require synthetic indexes, but no standard has been approved in France up to now. Only attempts to define them have been carried out. In France, 3 types of sensitivity indexes have been considered: morphological, protected ecological area, and economical ones, but no standard has been adopted up to now.

Morphological indexes: An attempt to derive more appropriate morphological indexes from existing standard (NOAA's Environmental Sensitivity Index (ESI), European Impact Reference System) has been studied at a local level. The first reference has been considered more appropriate but it has not been adopted unanimously.

Title:

Protected areas: There are 7 kinds of (legal) protected areas exist in France among which the Natura 2000. Attempts at local level have been carried out to define synthetic ecological indexes but there are still no consensus on the definition of these indexes.

Economical index: An attempt to define an economical index based on the duration of the interruption of the activity and the weight of each economical sector (fisheries, transportation, tourism ...) has been carried out.

### **Current approach**

For this we will look at other projects or initiatives that deal with oil spill incidents. Two projects are especially relevant: the EMSA CleanSeaNet project, and the database and web site maintained by the International Tanker Owners Pollution Federation Limited (ITOPF) (see chapter 7).

We consider to use for the common Scenario data model the ITOPF classifications for 'type of oil' and 'cause' of oil spill (to discuss, this also depends on the forecast data providers and other stakeholders, e.g. Cedre in France).

## **1.5 Terminology**

Harmonisation of terminology in the Ocean Scenario is important, because having the same terminology in case of oil spills at borders between countries is essential for crisis response and later reporting. There are a number of initiatives to harmonise and publish expert terminology used in the ocean and marine community, see for example <http://marinemetadata.org/references/cdivocabs>, or the NERC DataGrid Vocabulary server web site: [http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)

Based on these existing initiatives the point is not that there are no common vocabularies, but rather: how to deploy them in the Scenario. The terminology proposed in the study on the indices of sensibility of the coasts carried out for the Atlas of Finistere on request of the Departement du Finistere can be supplied to the Ocean partners for use in HUMBOLDT. For the other information used in the Scenario (land data, sensitive areas) it must be decided whether controlled vocabularies for these subjects are needed, and if so, which organisation could supply them.

## **1.6 Metadata**

Metadata published in a metadata registry is very important for especially use case "UC04 - Make available additional data". For this use case an efficient tool is needed to search, discover and access data with the use of catalogue Web services and with an interface which allows to collect all but only all appropriate layers requested by the end-users. To accomplish this a common definition of metadata in the Scenario including the way to access data via the metadata registry (see A9\_8.D1) is required.

There already have been initiatives in the Ocean community to harmonise the metadata needed to describe ocean and marine data sources, see e.g. the Common Data Index (CDI) hosted at the Sea-Search website. In the context of the discussion about the core HUMBOLDT metadata profile, the Scenario will list the metadata elements that should be mandatory from the perspective of the Ocean Scenario.

**UK and France:** Further investigation of metadata schemas is required, also regarding the question if the core HUMBOLDT metadata profile is sufficient for ocean and marine data, or whether extensions are needed.

Title:

## 1.7 Scale/resolution, level-of-detail, aggregation

This aspect does not have priority for the use cases in this Scenario. (Note: with this remark, that in a visualisation tool the geographical names (labels), and the small spatial objects represented as point symbols, should stay readable when zooming in and out)

## 1.8 Portrayal

For the visualisation of forecasted oil spill trajectory, amount and type of oil, and the sensitive areas that are potentially at risk, harmonised map portrayal is important, but it has a low priority at the moment.

### France:

As a consequence of the absence of common symbolism of some basic data, including rules for the representation on synthetic maps, portrayal is still improvised for a large part. Note: some standards exist such as for hydrographic maps, but are not adapted as a whole, as they are mainly designed for the safety of navigation.

## 1.9 Multi-linguality

For the monitoring during the oil spill: yes, this can be an issue when an oil spill incident involves authorities from more than one country (non-expert end-users). For the visualisation on maps and other reports after the event: this will probably be in English, so multi-linguality is not an issue in that case.

However, documents can be requested by national authorities, local administrations and courts after the crisis, so reports and maps must also be produced in the national language of the end-users for legal reasons, at least in France (according to French Law N° 94-665 08/04/1994 on the use of the French language).

## 1.10 Spatial consistency

In case of spatially overlapping datasets (same area, several datasets): Yes, this can be an issue. For example if the feature is a national response zone boundary (in case of an oil spill disaster). The boundary needs to be the same regardless of the dataset used. In case of spatially adjoining datasets (cross-border situations): Yes, this is an issue. For example when wanting to visualise different features which cross or connect at a sea-land boundary.

## 1.11 Maintenance

Maintenance of oil spill data after the incident is an harmonisation issue. In order to be able to produce reports later on the data should be stored in a uniform way. Maintenance of the static datasets (coastline, sensitive areas, administrative boundaries) depends on the approach taken: when this data is acquired and stored on the data servers of the Scenario partners, then there is an issue how to keep this data up-to-date. When however this data is not acquired and stored on the Scenario servers, because it is possible to access the coastline, sensitive areas, and administrative boundary data online via a web service of the respective (outside HUMBOLDT) data providers, then it is less a maintenance issue of data, but a maintenance issue of connections and securing online access to

Title:

these remote data sources. This is then more an architectural issue (how to cope with decentralised data storage, relying on web services of other data providers) than a data harmonisation issue.