Title: A7.4-D2 Profiles for ocean and marine environment (Interim version)
Title: A7.4-D2 Profiles for ocean and marine environment (Interim version)
Author(s)/Organisation(s):
Marian de Vries (TUD), Peter Thijsse (MARIS), Patricia Moreno-Arancibia (SMHI), Sophie Baudel (CLS), Olivier Lauret (CLS), Eric Moussat (Ifremer), Mark Charlesworth (BODC), Vasilis Lakes (HCMR)
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 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

Short Description:
This document contains the interim 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 must support 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). The Ocean Scenario has in the first phase of the system specfication 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.
Keywords:
Ocean Scenario, information analysis, conceptual data model, harmonisation issues, oil spill monitoring and reporting, forecasting

History:			
Version	Author(s)	Status	Comment

000	Marian de Vries	new	Based on data specification components of WP9.x Reports
001	Marian de Vries	draft, rfc	
002	Marian de Vries	update	Incorporated comments of SMHI and Ifremer
003	Hugo Ledoux, Olivier Lauret, Eric Moussat, Marian de Vries	update	Data model adapted based on discussions with Ocean partners between Delft meeting (November 2008) and Milan meeting. Version date of data model: 23 March 2009.
004	Marian de Vries	final	Partial update of text reflecting discussions

Table of contents

1 Purpose and scope.....5

2 Information analysis.....6

2.1 Use cases of the Scenario.....6

2.2 Spatial information used and produced in the Scenario.....6
2.2.1 Data used to initialize and force the oil drift calculation.....6
2.2.2 Data produced by the system.....7
2.2.3 Data used to aid in the visualisation of the results.....8

3 Common data model.....9

3.1 Rationale for this data model.....13

4 Data sources.....13

4.1 Information that is published (output or reference layer).....13
4.1.1 Oil spill.....13
4.1.1.1 Observed oil spill.....13
4.1.1.2 Forecasted oil spill.....13
4.1.2 Ship trajectory.....13
4.1.3 Sensitive area, protected site.....14
4.1.4 Administrative unit (with its boundaries).....14
4.1.5 Coastline / shoreline.....14
4.1.6 Coast infrastructure: roads, waterways.....14
4.1.7 Geographic local names.....14
4.1.8 Reference / background layers for orientation.....14

4.2 Input data for forecasts.....14
4.2.1 Weather data (including Wind speed).....14
4.2.2 Ocean data (including Surface currents, Ice concentration).....15
4.2.3 Bathymetry.....15

5 Harmonisation issues.....16

5.1 Data format.....16

5.2 Spatial and temporal reference systems.....16
5.2.1 Spatial reference system.....16
5.2.2 Temporal reference system (time).....16

5.3 Conceptual data model.....17
5.3.1 Geometry types.....17
5.3.2 Identification and Versioning.....17
5.3.3 Importance of time / Dynamic data.....17

5.4 Classification.....17

5.5 Terminology.....18

5.6 Metadata.....18

5.7 Scale/resolution, level-of-detail, aggregation.....19

5.8 Portrayal.....19

5.9 Processing functions....19
5.10 Multi-linguality....19
5.11 Spatial, temporal and thematic consistency of data....19
5.11.1 Spatial consistency....19
5.12 Other....20
5.12.1 Maintenance....20
5.13 Priorities....20
6 To discuss....20
6.1 Scope of the common Scenario data model....20
6.2 Access to needed datasets....21
7 Related projects and initiatives....22
8 Glossary....23

1 Purpose and scope

This document contains the interim version of the data specification for the Ocean Scenario. It is based on a number of documents containing the WP9 system specification for the Scenario and on the information analysis discussions leading to the construction of the common data model for the Scenario. This common data model must support the use cases in the Scenario, that 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 will serve as the core of a somewhat broader application profile for the ocean and marine application domain. This broader profile is not the aim of this version of the document, but of the final version. Primary aim of this interim version is to establish a good fitting common data model for the spatial information processed and produced in the Ocean Scenario use cases.

The Ocean Scenario has in the first phase of the system specfication split-up the work into 4 sub-scenarios: Baltic Sea, French coast, English coast, Aegean Sea. In this WP7 document the system specification of the Baltic Sea sub-scenario was taken as basis, and the other documents were used to add to the information taken from the Baltic Sea sub-scenario. In the second phase of the system specification a combined cross-regional scenario has been developed, with several reporting uses cases (see document 904). In addition the French coast and Aegean Sea sub-scenarios have been extended. Not all of these extensions or changes have already been incorporated in this version of the text.

Chapter 2 gives an overview of the spatial information involved in the Scenario use cases: what are the information items, and for what specific purpose are they needed. 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.

In chapter 4 the object classes of the common data model are related to concrete data sources (data sets, dynamic observations, web services). These concrete (existing) data sources often differ from sub-scenario (country) to sub-scenario. When the sub-scenariios are merged, these existing data sources must be transformed somehow to the common data model, and also other differences must be resolved (data format, spatial reference system, codelists, language ...). The comparison between existing (source) data and target data model is necessary as part of the ‚gap analysis' between required data model, data format etc. on the one hand and current characteristics of the existing datasets in the different countries on the other hand. Based on this gap analysis the concrete data harmonisation issues in the Scenario can be established.

In chapter 5 follows the inventory of harmonisation issues for the Scenario, taken from the respective WP9 documents, with additional points added.

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 Information analysis

2.1 Use cases of the Scenario

The overal Scenario application functionality is: 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:

01 – Calculate the drift of an oil spill

02 – Forecast when and where an oil spill reaches a sensitive area

03 – Find ship responsible for an oil spill

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

UC03 – Initializing the processes of the (monitoring and forecasting) system

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

UC04 – Make available additional data

This additionnal use case was introduced because at the moment the news of an oil spill comes in, and the monitoring and forecasting system is initialized, depending on the probable location of the oil spill, 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.) For the dynamic data pre-loading is never an option, but also for the ‚static' reference data it is better to search for and upload these data sets at the last moment, from the original sources, 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.

So all these 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 will have to be carried out on-the-fly, because the data cannot be prepared beforehand.

2.2 Spatial information used and produced in the Scenario

The description of the information used in the Scenario is divided into three sections. In the first one (3.1) the data needed to initialize and force the oil drift calculation is described. In the second section (3.2) the data that the system produces is described. In the last section (3.3) the data used for the visualisation of the results are discussed.

The use case template of in the WP9 Report distinguishes between input and output data of the use cases. This distinction was not explicitely made in the information analysis tables, therefore this distinction is now added to the tables below (third column).

2.2.1 Data used to initialize and force the oil drift calculation

Information item	Purpose	Input/output/internal
Weather data • Wind data	Wind speed and wind direction are used as parameters to compute the oil spill drift, and to predict the weathering	input

	processes that influence the oil spill.	
Ocean data • Current velocity • Temperature • Salinity • Turbulence parameters • Ice concentration • Ice drift velocity	Oceanographic data is used to 1. force the transport and spreading of the oil in the oil drift model 2. to calculate the chemical state of the oil (evaporation, emulsification, etc)	input
Coastline	To determine the geographical location; To determine the boundary of the computational grid.	
Bathymetry	To determine the geographical 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
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
Type of oil	To set density, viscosity, and empirical parameters that determine the weathering process of the oil. There parameters determine how the oil is transported in the ocean.	input, stored for reporting

2.2.2 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	Of interest to end-user	output

has evaporated, dispersed, and surface or stranded		

2.2.3 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 • Wind speed • Surface currents • Ice concentration	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 currents and ice conditions at the time of the oil spill, should be stored for later analysis.)
AIS (Automatic Information System) data • Ship ID and name • Position every 15 minutes • Destination	To determine which ships are (or have been) in the area where the pollution occurs; If needed to help identify which ship is responsible for the oil spill.	input, stored for reporting Depending on the further development of the use cases in the Scenario, this information item can become optional.

3 Common data model

As a starting point for the question what data in this Scenario needs to be harmonised (and based on that, what exactly the data harmonisation issues are) we focus in this chapter on the ‚output' data of the operational use cases, because this indicates what information content the Scenario will produce/publish to the outside world. This output information needs to be homonegeous (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 (also see chapter 6).

Therefore In the UML class diagram in Figure 1 only the core output information (as this was understood from the WP9 system specification) is incorporated in the common data model for the Scenario, 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). How to express this in the data model will be further discussed. At the moment there is a collection of attributes called ‚geophysicalParameters' in the class ForecastedEnvironment, as a placeholder. There will probably also be a ObservedEnvironment class in the next version of the model, with a comparable solution for the parameters that are input for the computational models.

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 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.

For two kinds of information the idea is to follow the definition of spatial object types (feature types) as defined in the INSPIRE Annex I data specifications: AdministrativeUnit and ProtectedSite. These two classes in the model are depicted in another colour, and are not detailed further except with their name.

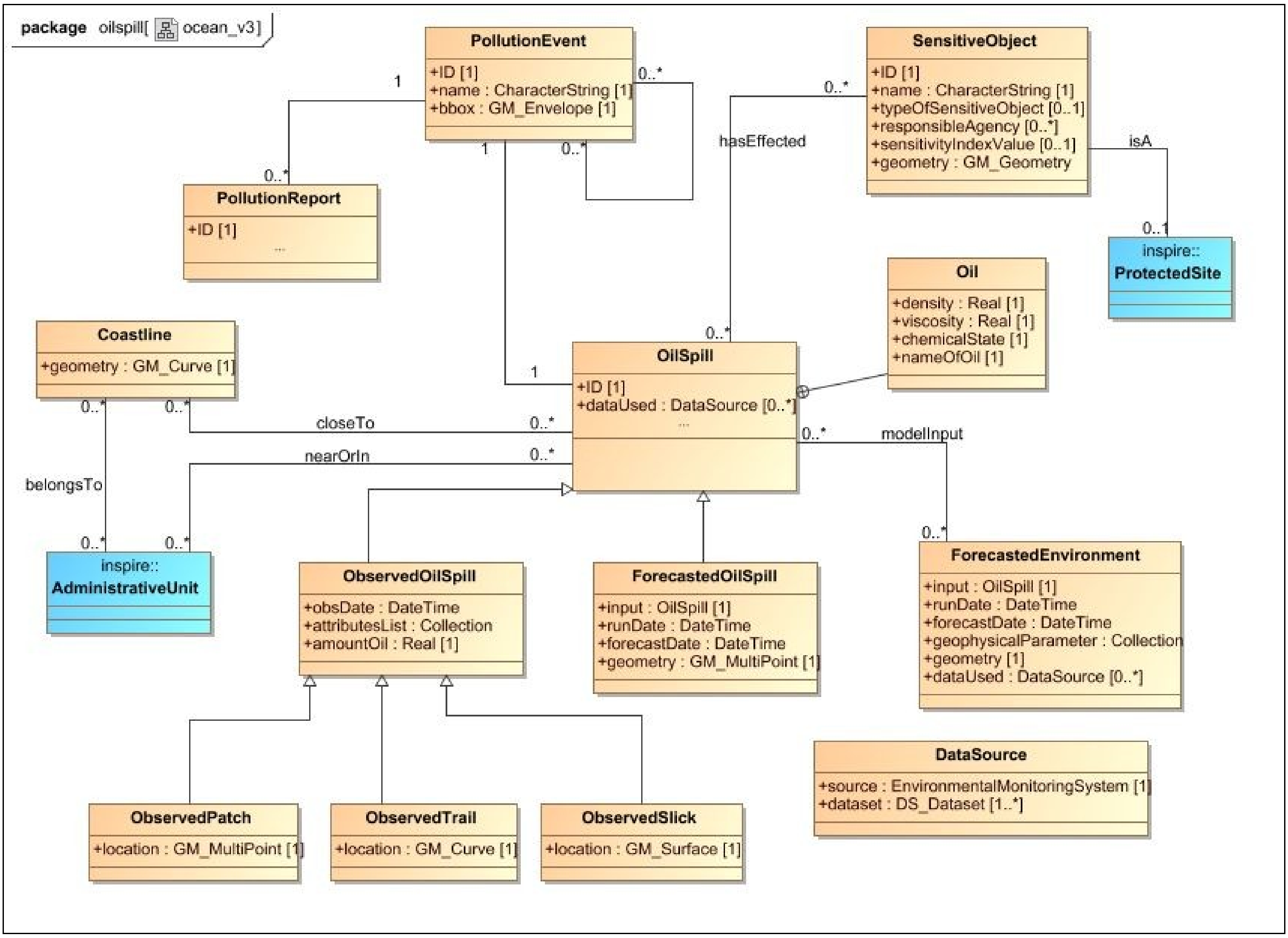

Figure 1 Data model for oil spill monitoring and reporting (output and reference data)

Note: The prefix GM_ means the official ISO 19107 narmes 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 usefull 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.

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 2 and 3.

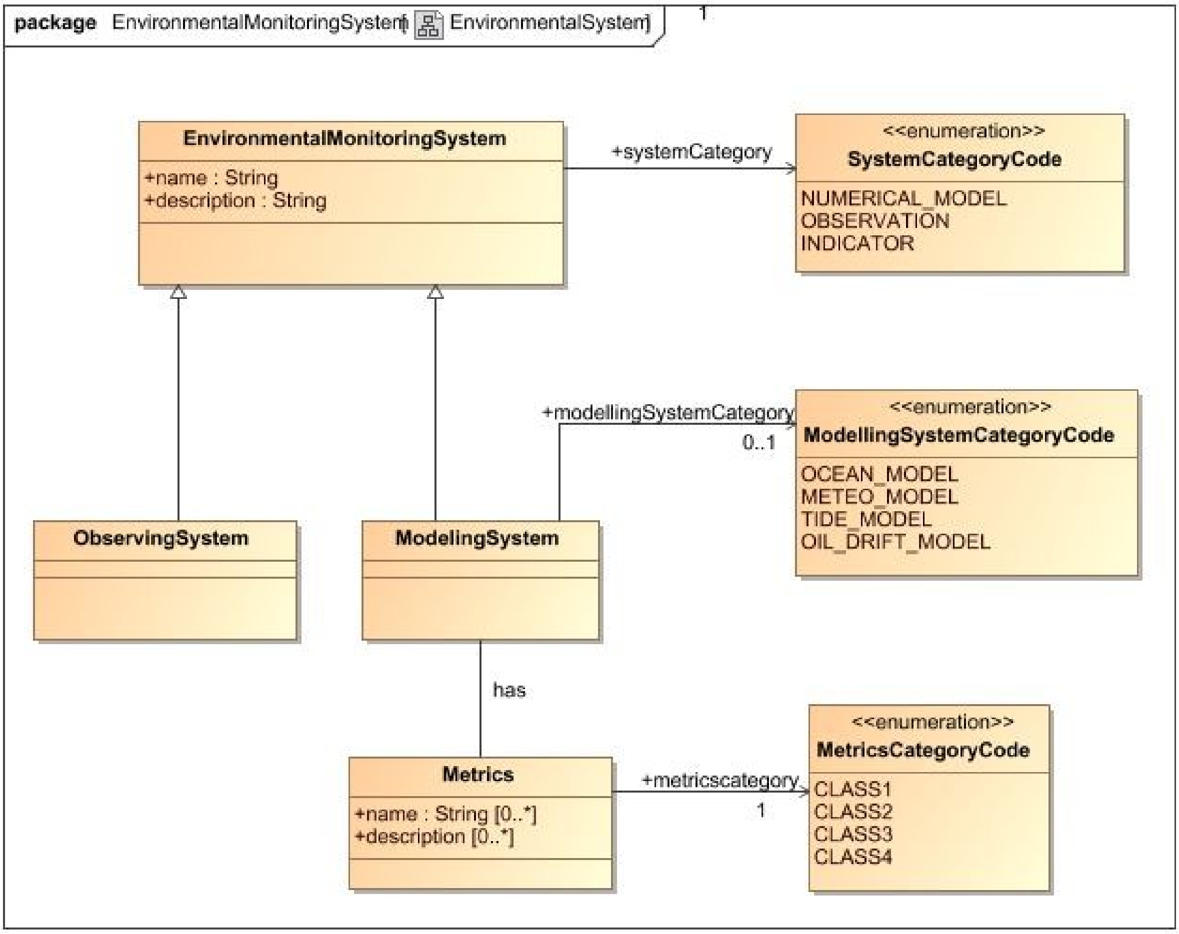

Figure 2 Data model part: Environmental Monitoring System

Figure 2 gives a sketch of the EnvironmentalMonitoringSystem class and its subclasses. Figure 3 shows the ObservingSystem class, its subclasses and associations.

At the moment it is still under discussion how these parts of the data model will be implemented. 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 3 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.

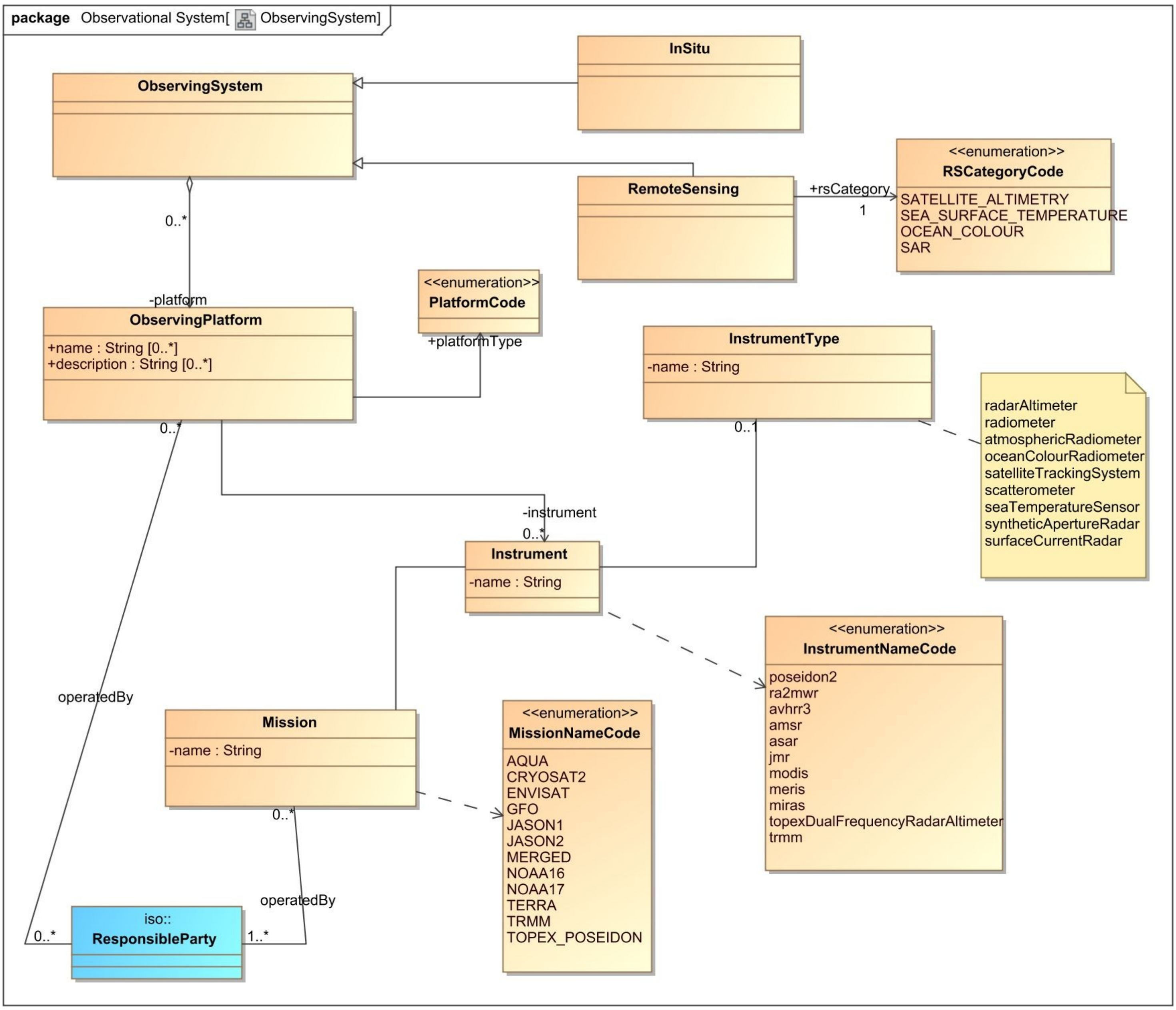
package Observational System[ObservingSystem]
InSitu
ObservingSystem
<<enumeration>>
RSCategoryCode
SATELLITE_ALTIMETRY
SEA_SURFACE_TEMPERATURE
OCEAN_COLOUR
SAR
RemoteSensing
+rsCategory
1
0..*
-platform
ObservingPlatform
+name : String [0..*]
+description : String [0..*]
<<enumeration>>
PlatformCode
+platformType
InstrumentType
-name : String
0..1
radarAltimeter
radiometer
atmosphericRadiometer
oceanColourRadiometer
satelliteTrackingSystem
scatterometer
seaTemperatureSensor
syntheticApertureRadar
surfaceCurrentRadar
0..*
-instrument
0..*
Instrument
-name : String
<<enumeration>>
InstrumentNameCode
poseidon2
ra2mwr
avhrr3
amsr
asar
jmr
modis
meris
miras
topexDualFrequencyRadarAltimeter
trmm
operatedBy
Mission
-name : String
<<enumeration>>
MissionNameCode
AQUA
CRYOSAT2
ENVISAT
GFO
JASON1
JASON2
MERGED
NOAA16
NOAA17
TERRA
TRMM
TOPEX_POSEIDON
0..*
operatedBy
0..*
iso::
ResponsibleParty
1..*

Figure 3 Data model part: Observing System

3.1 Rationale for this data model

This section is meant to give some explanation and reasons for defining this data model. This is also a check of the data model against the use cases as a first evaluation of its suitability.

The common, harmonised data model is meant to describe and formalize the main information content of this Scenario. Depending on the further detailing of the system specification this data model can still change as a result of the work in WP9. Therefore a more complete explanation of and justification for the data model will be given after the system specification is frozen.

We have looked at comparable projects and initiatives to see if we could adapt existing (conceptual) data models, classifications, portrayal standards, see references in chapter 7. The EMSA CleanSeaNet satellite service is an interesting comparable project. Whether this project has a data model for oil spill monitoring of which parts can be re-used in the Ocean Scenario will be evaluated for the final version of the Ocean data model.

4 Data sources

In this chapter the object classes identified in the data model are related to concrete data sources (in the form of static data in files or databases, dynamic data produced in real-time, or web services that serve the data on request). Note: this is the interim version of the list of data sources (March 2009). Some data sources must be added for the non-ocean information items. And there are some data sources of which it is not sure whether they will be available for demonstration purposes.

4.1 Information that is published (output or reference layer)

Per object class identified in the common data model

4.1.1 Oil spill

4.1.1.1 Observed oil spill

Aegean Sea: ENVISAT ASAR Wide Swath data, detection reports

France: POLREPs; SAR images overlapped with uploaded vector information delimiting the slick area as reported by the SAR sensor; drift of monitoring buoys.

SAR images have data format GeoTIFF, CEOS or PDS

4.1.1.2 Forecasted oil spill

Calculated based on observations and estimates

4.1.2 Ship trajectory

Baltic Sea: AIS data (‘Automatic Identification System’, all ‘sizable’ ships have a transponder that is the source for this data)

4.1.3 Sensitive area, protected site

Baltic: Baltic Sea Protected Areas, important bird areas, Maris (Maritime Accident Response Information System) zone

France: all kinds of areas that are negatively impacted (environmentally or economically) by the oil spill, including fish farms, barriers, water inlets, ...

UK: Environmental data provided by web service 'MAGIC': environmentally sensitive areas, designated sites

4.1.4 Administrative unit (with its boundaries)

France: administrative limits (national borders, provinces, municipalities etc.); French surveillance areas

4.1.5 Coastline / shoreline

France: aerial orthophotoplan

4.1.6 Coast infrastructure: roads, waterways

Most probably having vector data is not needed here, and a raster map as reference layer is sufficient (see 4.1.8). The main purpose is to easily be able to visualize the coastal infrastructure in an oil drift visualisation tool. This information is important when planning the clean-up of an oil spill that has reached the coast. So raster maps are sufficient at the moment. Nevertheless having access to vector data with height information would be really useful to get specific characteristics of a location e.g. the % of the slope of a ramp for hauling, or in general of the coastal area.

4.1.7 Geographic local names

France: point datasets with local names (French names for France, etc.)

4.1.8 Reference / background layers for orientation

France: B&W aerial photo of 2kmX2km needed for orientation purposes. Providing information on types of land cover (forest, settlement, grassland, water, …)

4.2 Input data for forecasts

Per information item (for the moment kept out of the common data model, see To discuss, chapter 6).

Note: Apart from being input data for the forecasts, some of this data will also be used for: visualisation (for example bathymetry, as background map layer) or for later analysis and reporting (weather and ocean conditions at the time of the event).

4.2.1 Weather data (including Wind speed)

UK: weather data from UK Met Office (wind velocity and strength)

France: sea level atmospheric pressure in Pa, u component of wind at 10 m in m/s, v component of wind at 10 m in m/s.

Data format: GRIB or netCDF (Opendap); for animation, the approach is to use WMS with recurrent call to data.

4.2.2 Ocean data (including Surface currents, Ice concentration)

UK: data provided on request by BODC: current velocity and direction (?), temperature, salinity.

France:

A. Absolute Dynamic Topography = Sea Level Anomalies, Significant Wave Height (h1/3), Wind speed modulus, Sea surface temperature

B. Ocean forecast data = 3D Temperature (°C), 3D Salinity (psu), Zonal velocity (m/s), 3D Meridian velocity (m/s) etc.

1. open ocean model forecast data: netCDF (output of Opendap web service). Projection type: Mercator or geographic or stereo-polar (in arctic area)
2. coastal ocean model forecast data: netCDF (output of Opendap web service), WGS84

4.2.3 Bathymetry

UK: bathymetry and tide data: built into POL model

France:

1. Isolines (Hydro. std S57): ESRI shapefiles or Oracle Spatial, WGS84.
2. Nautic charts: ESRI shapefiles. The projection can be Mercator or UTM and the geodetic system may be different from WGS84.
3. Gridded data set: NetCDF ESRI compatible grid format, WGS84.

5 Harmonisation issues

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

5.1 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.

Approach: to be discussed and agreed.

5.2 Spatial and temporal reference systems

5.2.1 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).

5.2.2 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.

Recommendation: UTC will be the reference time.

5.3 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)

5.3.1 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.

5.3.2 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.

5.3.3 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.

5.4 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.

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.

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).

5.5 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/

Based on these existing initiatives the point is not that there are no common vocabularies, but rather: how to deploy them in the Scenario. Ifremer will ask if 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.

5.6 Metadata

Metadata published in a metadata registry is very important for especially use case “UC04 - Make available additional data”. For this use cse 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.

5.7 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, also see 5.8.)

5.8 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 (M4), but are not adapted as a whole, as they are mainly designed for the safety of navigation.

It will be investigated whether there are (other) portrayal standards for the sea and marine community that can be used.

5.9 Processing functions

The initial answer is: no

To be discussed is whether this is an issue in the Scenario if the 4 sub-scenarios are merged. Should there be a comparison between the forecast models, parameters etc. now in use in the Ocean partners' organisations for estimates, forecasts, calculations?

5.10 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).

5.11 Spatial, temporal and thematic consistency of data

5.11.1 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.

5.12 Other

5.12.1 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. This calls for a common data model (see chapter 3).

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 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.

5.13 Priorities

France: The main problem relates to the lack of harmonisation of the codifications in use (= classifications). This is the first priority, and it effects both the conceptual data model (enumerations, code lists) and the map portrayal (legend classifications).

The second priority concerns the harmonisation of data format.

The third, but still high priority is the possibility to have a user context which allows to take into account different important aspects of the user profile, such as: the national language, the level of knowledge of the end-users of the system (basic user, skilled user…), the user rights to access data (confidential data are used in the Scenario, but existing OGC web service interface protocols do not offer protection up to now, which is an obstacle for data sharing).

UK: Harmonisation of data format has a high priority. Possible complication: Changing outputs from the POL model and the MAGIC Coastal and Marine Resource Atlas would be by agreement with the organisations who manage these applications.

6 To discuss

In this chapter some general issues are collected that must be discussed and dealt with in WP7 and WP9.

6.1 Scope of the common Scenario data model

What information elements should be incorporated in the common Scenario data model (and consequently: which data sources must be harmonised towards this common data model)?

The input data for the use cases (e.g. weather and ocean conditions) does not necessarily have to be homogeneous. This depends on the further detailing of the system specification, and the answer on questions like:

- will the weather and ocean conditions also be stored for later reference, e.g. for later scientific analysis in the reports on oil spill incidents, maybe their severity depending on weather and ocean conditions, or the speed with which the oil spill drifts to shore or is dispersed without doing much harm?
- do the Scenario partners want to use the same software for the calculations, forecasts, monitoring etc. In that case also the input data should be harmonised as much as possible, so that no (syntactic) interoperability problems will arise.

The answer to these questions is part of the further system specification work in WP9. Depending on these decisions for the Scenario use cases, the common data model in Figure 1 can be extended in the next version of this Report.

6.2 Access to needed datasets

Part of the information processed in the Ocean Scenario is not owned by the partners, therefore a separate activity in the Scenario must be to acquire this data (in case of files) or negotiate access to this data (in case of web services). Because this is not only the case for the Ocean Scenario, it would be good to have a common approach for this.

In addition, an important issue, at least in France, is the management of the rights of the end-user to access the data (output as well as input data) depending on the user category he or she belongs to: some information cannot be accessed publicly while other information is freely accessible; some data can only be viewed by some user category while it can be processed only by another. For the Ocean Scenario a way must be found to use high quality, real data, without being hindered by these restrictions.

7 Related projects and initiatives

CDI (Common Data Index), 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

Huijer, 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 (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/

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/

8 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