Project No. 249024

NETMAR

Open service network for marine environmental data

Instrument: *Please tick*	CA	✓ **STREP**	IP	NOE

ICT - Information and Communication Technologies Theme

D7.2.2 Project highlights for Year 1

Reference: D7.2.2_Project_Highlights_Year1_r1 _20110128

Due date of deliverable (as in Annex 1): M0 + 12
Actual submission date: 28 January 2011

Start date of project: 1 February 2010 Duration: 3 years

Nansen Environmental and Remote Sensing Center

Revision 1

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

NETMAR

Open service network for marine environmental data
Project Reference: 249024
Contract Type: Collaborative Project
Start/End Date: 01/03/2010 - 31/01/2013
Duration: 36 months

Coordinator: Prof. Stein Sandven
Nansen Environmental and Remote Sensing Center
Thormøhlensgate 47, Bergen, Norway
Tel.: +47 55 20 58 00
Fax. +47 55 20 58 01
E-mail: stein.sandven@nersc.no

Acknowledgements

The work described in this report has been partially funded by the European Commission under the Seventh Framework Programme, Theme ICT 2009.6.4 ICT for environmental services and climate change adaptation.

Consortium

The NETMAR Consortium is comprised of:

- Nansen Environmental and Remote Sensing Center (NERSC), Norway (coordinator).
  Project Coordinator: Prof. Stein Sandven (stein.sandven@nersc.no)
  Deputy Coordinator: Dr. Torill Hamre (torill.hamre@nersc.no)
  Quality Control Manager: Mr. Lasse H. Pettersson (lasse.pettersson@nersc.no)
- British Oceanographic Data Centre (BODC), National Environment Research Council, United Kingdom
  Contact: Dr. Roy Lowry (rkl@bodc.ac.uk)
- Centre de documentation de recherche et d'expérimentations sur les pollutions accidentelles des eaux (Cedre), France.
  Contact: Mr. François Parthiot (Francois.Parthiot@cedre.fr)
- Coastal and Marine Resources Centre (CMRC), University College Cork, National University of Ireland, Cork, Ireland.
  Contact: Mr. Declan Dunne (d.dunne@ucc.ie)
- Plymouth Marine Laboratory (PML), United Kingdom.
  Contact: Mr. Steve Groom (sbg@pml.ac.uk)
- Institut français de recherche pour l'exploitation de la mer (Ifremer), France.
  Contact: Mr. Mickael Treguer (mickael.treguer@ifremer.fr)
- Norwegian Meteorological Institute (METNO), Norway.
  Contact: Mr. Jan Ivar Pladsen (janip@met.no)

Author(s)

- Torill Hamre (torill.hamre@nersc.no)
- Roy Lowry (rkl@bodc.ac.uk
- Adam Leadbetter (alead@bodc.ac.uk)
- Yassine Lassoued (y.lassoued@ucc.ie)
- Anthony Patterson (A.Patterson@ucc.ie)
- Mike Grant (mggr@pml.ac.uk)
- Peter Walker (petwa@pml.ac.uk)
- Jorge Mendes de Jesus (jmdj@pml.ac.uk)
- Mickael Treguer (mickael.treguer@ifremer.fr)
- Declan Dunne (d.dunne@ucc.ie)
- Vincent Gouriou (vincent.gouriou@cedre.fr)

Document approval

- Document status: Revision 1
- WP leader approval: 2011-01-28
- Quality Manager approval: 2011-01-28
- Coordinator approval: 2011-01-28

Revision History

Issue	Date	Change records	Author(s)
Draft	2010-11-15	Topic layout.	T.Hamre
Draft	2010-12-15	Included descriptions of Pilot 1, reviews, ontology extension, semantic framework, system architecture and outline of services.	R.Lowry, A.Leadbetter, Y.Lassoued, A.Patterson, T.Hamre
Draft	2010-12-15	Included description of Pilot 3 and Pilot 4.	M. Grant, P. Walker
Draft	2010-12-20	Included description of PyWPS development, Pilot 6 and Pilot 5.	J. Mendes de Jesus, M. Treguer D. Dunne
Draft	2010-12-21	Included descriptions of Pilot 2 and Executive Summary.	V. Gouriou, T. Hamre
Draft	2011-01-17	Revised description of Pilot 2; update pilot 1 wiki screen shot; correct reference list.	V. Gouriou, T. Hamre
Draft	2011-01-27	Reviewed text, clarified products for Pilot 2.	T. Hamre, Ø. Torget
1	2011-01-28	First version of the report.	T. Hamre

Executive Summary

NETMAR aims to develop a pilot European Marine Information System (EUMIS) for searching, downloading and integrating satellite, in situ and model data from ocean and coastal areas. EUMIS will use a semantic framework coupled with ontologies for identifying and accessing distributed data, such as near-real time, forecast and historical data. This report describes major results from year 1 of the project.

Six pilots have been defined to clarify the needs for satellite, in situ and model based products and services in selected user communities. The pilots are:

- Pilot 1: Arctic Sea Ice Monitoring and Forecasting
- Pilot 2: Oil spill drift forecast and shoreline cleanup assessment services in France
- Pilot 3: Relationships between physical and biological variables
- Pilot 4: Ecosystem model validation
- Pilot 5: International Coastal Atlas Network (ICAN) for coastal zone management
- Pilot 6: Phytoplankton blooms in Gulf of Biscay and English Channel

Each pilot has defined a series of use cases that will be used to demonstrate the capabilities of the EUMIS pilot. The included use cases will demonstrate that EUMIS is capable of delivering decision-support to the targeted user groups in public sector, national authorities, industry and scientific communities.

A second major activity in year 1 has been to review Environmental Information System (EIS) architectures, initiatives and projects, as well as relevant frameworks, technologies and tools. This led to the recommendations that NETMAR's architectural design be driven by service-oriented architecture (SOA) design principles, and based on open standards from the Internet Engineering Task Force (IETF), World Wide Web Consortium (W3C), Organization for the Advancement of Structured Information Standards (OASIS), International Organization for Standardization/Technical Committee 211 (ISO/TC 211), and Open Geospatial Consortium, Inc. (OGC). Specific recommendations were made for tools for NETMAR protocols and subsystems (Table 2-1). An initial version of the NETMAR system architecture was also developed, using the RM-ODP (Reference Model for Open Distributed Processing) standard. This defined different aspects of the architecture in the form of five viewpoints. As an example, the information viewpoint describes the use of metadata in NETMAR, using an invariant schema to show constraints on metadata which must be observed throughout the project and by all partners and data providers.

The review of ontologies, ontology tooling and semantic frameworks also identified important semantic resources that can be used in the development of EUMIS, and recommendations for technologies and tools for its semantic components. Some of the identified resources include the GEMET concept thesaurus, NERC Vocabulary Server and the Marine Metadata Interoperability Ontology Registry and Repository. The development of semantic EUMIS components will be based on standards like RDF (Resource Description Framework), SPARQL and the Jena Framework. The NERC Vocabulary Server will be another important building block, which with ontology extensions mapping from its inherent semantic resources will map to external resources such as GEMET and MMI-ORR. A generic Semantic Web Service will be developed on top of this server, offering an OGC-like interface to the extended ontologies.

Partners have also set up a set of data delivery services for different types of satellite, in situ and model data, for the targeted user communities. NETMAR will also offer processing services, and during year 1 the NETMAR project has been a major contributor to the open source Python implementation of WPS, PyWPS. Part of this contribution has been a wiki containing information for basic users and system developers. Finally, the development of the NETMAR portal has started with a basic wiki and forum.

Contents

1 Definitions of pilots

During year 1 of NETMAR; six pilots have been defined to clarify the needs for satellite, in situ and model based products and services in selected user communities. Each pilot has defined a series of use cases that will be used to demonstrate the capabilities of the EUMIS pilot. The included use cases will demonstrate that EUMIS is capable of delivering decision-support to the targeted user groups in public sector, national authorities, industry and scientific communities. Below a short summary of each pilot is given, focusing on objectives, targeted users and examples of products that will be offered.

1.1 Pilot 1: Arctic Sea Ice Monitoring and Forecasting

The Arctic Ocean and the adjacent sub-Arctic Seas, including Europe's northern frontiers, have become more important as a result of increased economic activities related to energy exploration, marine resources and transportation. There is an increasing supply of sea ice information from satellites, in situ platforms and modelling systems that can be exploited by several user groups. Both operational and scientific users need to combine data from the available observing systems as well as from modelling systems to make the best possible assessment of the present and forthcoming ice situation, e.g. by assimilation of observations into forecasting models.

The Arctic Sea Ice and Metocean Observing System pilot is thus aiming to serve near real time data and forecasts as well as historical data to: (1) representatives from offshore and shipping companies, (2) ship and icebreaker captains/ice pilots, (3) national authorities and regional environmental agencies, (4) national ice services, and (5) scientists.

The pilot will offer a combination of scientific and operational products, including:

- Satellite Synthetic Aperture Radar (SAR) images (Figure 1-1(a))
- Sea ice type classification maps from SAR images
- Ice model forecasts from the Topaz model
- Time series of satellite and model products
- Sea Ice Charts from the operational Sea Ice Service at the Meteorological Institute of Norway (METNO) (Figure 1-1(b))
- Atmospheric parameters from the operational HIRLAM model

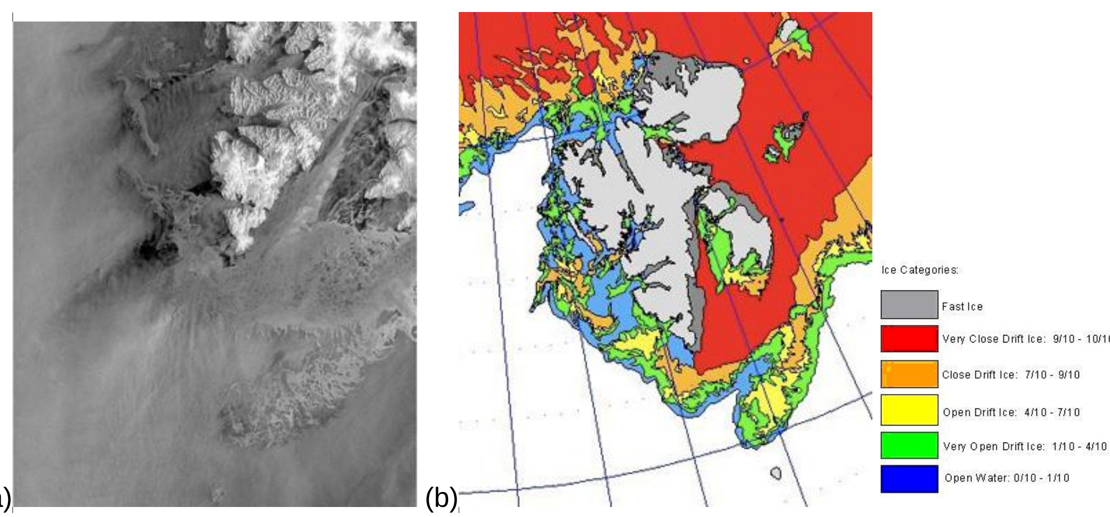
(a) (b)


Figure 1-1 (a) ENVISAT ASAR image 28 April 2010 showing a partly ice covered area close to Spitzbergen, Svalbard. (Data provider: NERSC). (b) Sea ice classification map generated from operator analysis of satellite SAR data from 28 April 2010. (Data provider: METNO).

1.2 Pilot 2: Oil spill drift forecast and shoreline cleanup assessment services in France

The two NETMAR oil spill use cases primarily concern tasks that need to deal with a large amount of data day after day. These data are needed to support emergency response services during long lasting oil spill crisis such as the ERIKA and PRESTIGE pollution providing fast and reliable information to allow for an efficient use of human resources and equipment both at sea and onshore.

The first use case is corresponding to an oil spill drift forecast service in France that will be based on new data processing tools with enhanced functionalities to better support the experts in making estimates of where oil slicks will drift. The oil spill drift forecast service will be used by the members of the French Slick Drift Monitoring and Prediction Committee, which include Cedre, Ifremer, MRCC, French Navy and Met offices. These organisations are in charge of obtaining and analysing all relevant data in order to make an assessment of how the pollution will spread. The need of a consensus among the experts may imply quite a lot of met-ocean simulation with different hypothesis and an easy way to retrieve, store and compare the results.

The second use case concerns the collection of all available information about both the observation of pollution landings and the related onshore mitigation actions during the response phase. This second use case is intended to collect all of these pieces of information and make them readily available to regional and national authorities during the crisis. The main objective is to provide a detailed view of the evolution of the pollution response on shore in near real time, exploiting all the cleanup data at different levels (commune, department, region) of organization (count of stranded oil slick, on going cleanup operations, actual manpower resources, special equipments being used, collected and disposed waste, etc.).

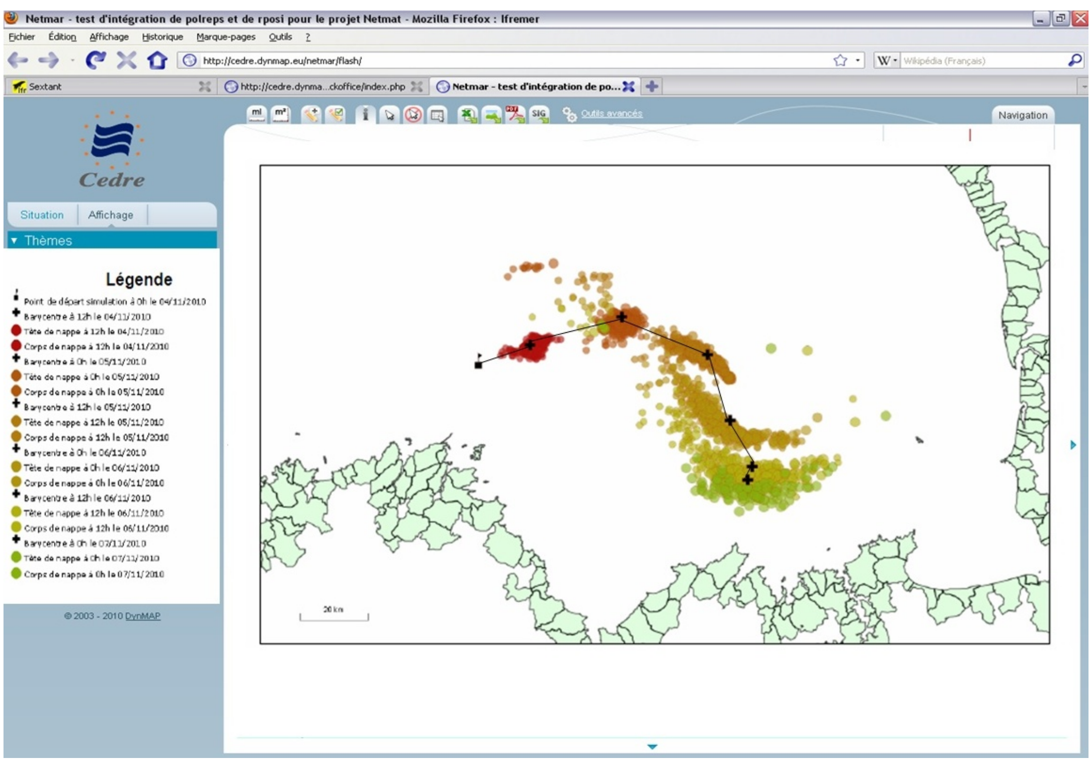


Figure 1-2 Screenshot of OILMAP (oil spill drift) service.

1.3 Pilot 3: Relationships between physical and biological variables

Researchers in ecosystems need to be able to identify and use long term time series in order to quantify ecosystem responses to natural variability, climate change or the impact of anthropogenic activities. Examples may include comparing long term change in zooplankton concentration to water temperature or relating optical properties to chlorophyll concentration. Operational users may find it useful to compare, in near real time, contemporary satellite and in situ data in order to provide input to water quality monitoring systems, for example, on phytoplankton chlorophyll-a concentration.

At present these data are usually held in multiple locations by different providers in different formats with different naming conventions. The pilot aims to provide discovery services to enable the users to find the data; standards based retrieval tools and a semantic framework for managing the parameters. For instance the semantic framework could be used to enhance discovery by allowing users to select datasets with specific parameters.

The pilot will use the Western English Channel Observatory as a test case for combining in situ and satellite data (using a WPS) showing how the tools developed make it easier to find, extract and compare data.

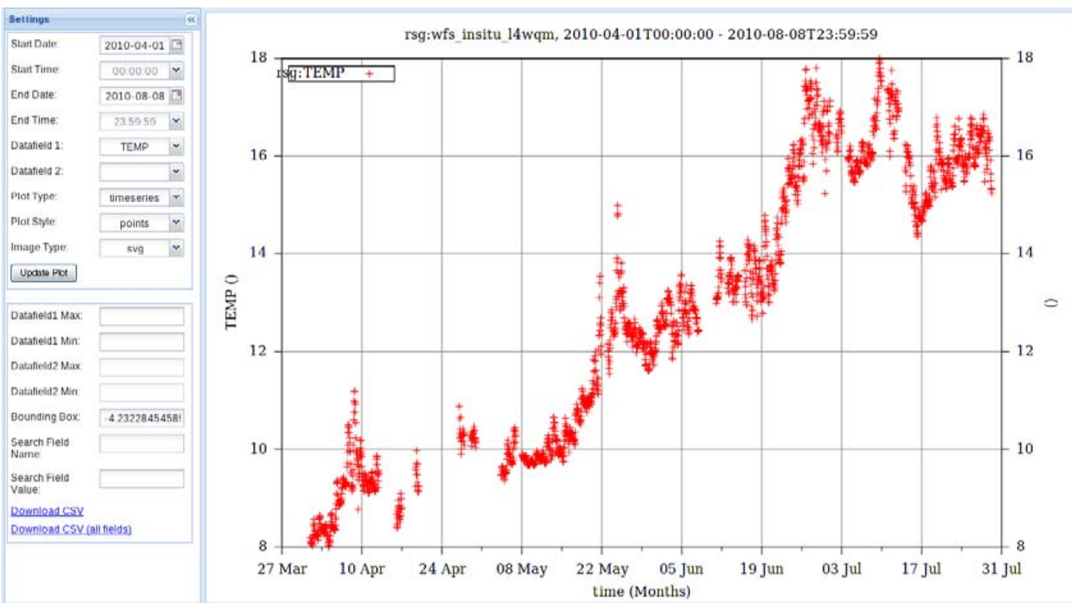


Figure 1-3 Sea surface temperature measured at station L4.

1.4 Pilot 4: Ecosystem model validation

Projects such as MEECE (Marine Ecosystem Evolution in a Changing Environment) aim to use a combination of data synthesis, numerical simulation and targeted experiments to boost knowledge and develop the predictive capabilities needed to learn about the response of marine ecosystems. In order to validate and improve models developed by projects such as this there is a need to compare the coupled physical and biological models run in hindcast mode with historical EO data for validation.

This can be achieved using in situ data sets such as the Continuous Plankton Recorder (CPR); long term sampling stations, such as in the Western English Channel, in operation for over 100 years, and satellite remote sensing for which SST data have been continuously available for approximately 30 years and ocean colour for approximately 12 years. These time series are therefore valuable “ground truthing” to test hindcast runs of numerical models.

The pilot will initially provide a framework allowing the hindcast output from the POLCOMS ERSEM model (provided by MEECE) to be compared with the EO data for the North West Shelf. This framework will then be extended to cover additional model and EO data which adheres to the format required.

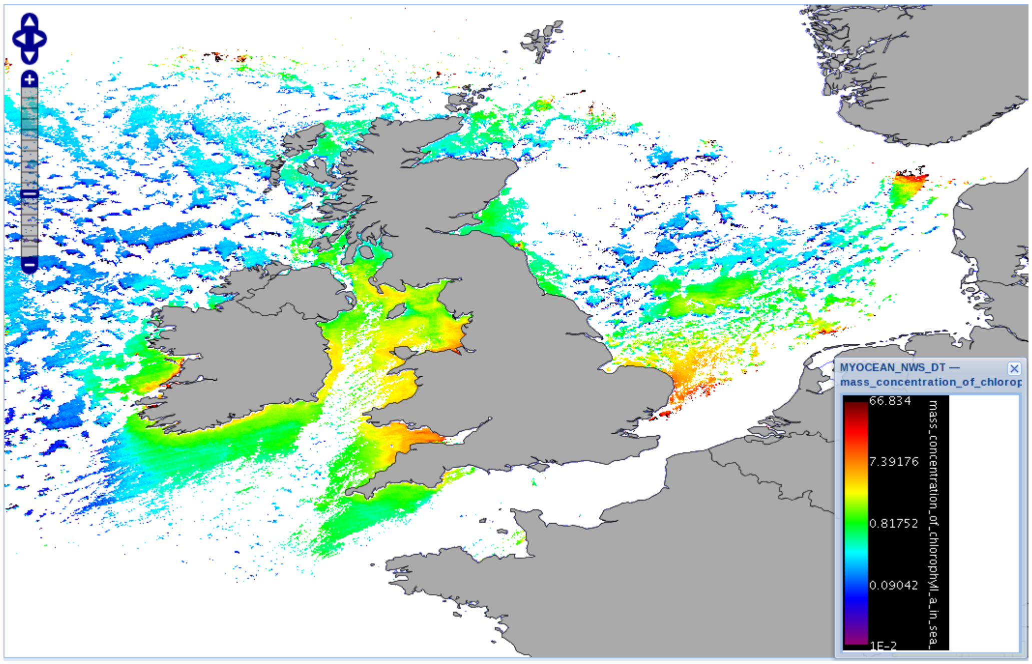


Figure 1-4 MODIS Chl-a viewed through WCO Data Portal.

1.5 Pilot 5: International Coastal Atlas Network (ICAN) for coastal zone management

In recent years significant momentum has occurred in the development of Internet resources for decision makers, scientists and the general public who are interested in the coast. A key aspect of this trend has been the development of coastal web atlases (CWA), based on web enabled geographic information systems (GIS). In this respect, ICAN, an informal group of organisations, has been meeting since 2006 to scope and implement data interoperability approaches to CWAs.

The mission/strategic aim of ICAN is to share experiences and to find common solutions to CWA development (e.g., user and developer guides, handbooks and articles on best practices, information on standards and web services, expertise and technical support directories, education, outreach, and funding opportunities, etc.), while ensuring maximum relevance and added value for the end users. Coastal web atlases in the ICAN network deal with a variety of thematic priorities and can be tailored to address the needs of a particular user group. Coastal mapping plays an important role in informing decision makers on issues such as national sovereignty, marine spatial planning, resource management, maritime safety, hazard assessment, etc.

While coastal web atlases contain a diverse range of datasets, the inclusion of both near real-time and historical data products from the operational oceanography and remote sensing communities has been more limited, often because such data has been difficult to access in terms of both data policy and data interoperability. This interoperability issue will be addressed by NETMAR. In addition, the ability to generate new value-added data products on the fly will be demonstrated by NETMAR.

It is also planned that the ontology structure and semantic framework specification proposed as part of the NETMAR project will be adopted by ICAN to replace the current ad-hoc ICAN ontology structure. It is hoped that this will help increase the number of participating coastal web atlases as it will provide a more robust and standardised model and infrastructure. The NETMAR catalogue service mediator will be exploited as it will replace the existing ICAN mediator. The new mediator will have a more standardised service interface and will address existing weaknesses in terms of scalability, performance, and technical bugs.

1.6 Pilot 6: Phytoplankton blooms in Gulf of Biscay and English Channel

Permanent monitoring of micro-algae blooms is performed by the Phytoplankton and Phytotoxins networks (REPHY) operated under the control of Ifremer. Other sources of information are also taken in account such as satellite sea-colour imagery and hydrodynamical models outputs.

One of Ifremer's objectives is to use systematically joint time series of in situ and satellite data for these parameters at some selected stations for the surveillance of the eutrophication risk (Chlorophyll).

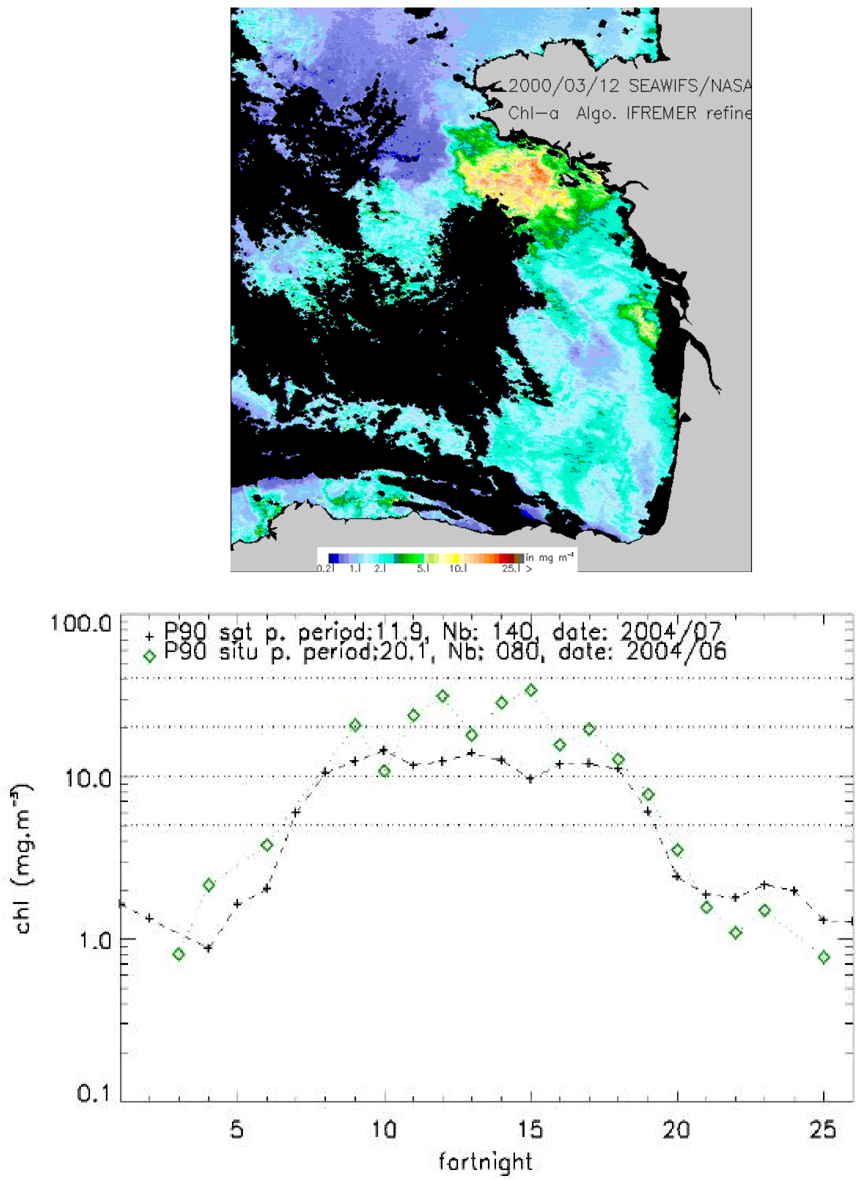


Figure 1-5 Regional Chl-a algorithm validation at the network stations (Cabourg, MODIS OC5 vs In situ 2003-2005).

2 Reviews of standards, technologies and tools

In the beginning of the project, we reviewed Environemental Information Systems (EISs) architectures, ontologies and ontology tooling, and semantic frameworks for EIS. The main results are presented below. The complete review reports are available online at http://netmar.nersc.no/?q=node/39.

2.1 Review of EIS architectures

We reviewed initiatives, projects, and technologies addressing the area of system architectures for distributed EISs. The review focused on identifying standard notations and methodologies for architectural design, as well as technologies and tools for web service development in the field of EIS [Dun10].

The majority of the reviewed initiatives and projects have developed a service-oriented architecture (SOA) and several follow the ISO/IEC 10746 RM-ODP (Reference Model for Open Distributed Processing) [Val01] standard to define their architecture. We therefore recommended to use RM-ODP to describe the NETMAR architecture using an incremental and iterative approach for analysis and design.

It is further recommended that NETMAR's architectural design be driven by service-oriented architecture (SOA) design principles, and based on open standards from the Internet Engineering Task Force (IETF), World Wide Web Consortium (W3C), Organization for the Advancement of Structured Information Standards (OASIS), International Organization for Standardization/Technical Committee 211 (ISO/TC 211), and Open Geospatial Consortium, Inc. (OGC).

OGC standards are particularly important for implementing the NETMAR service network. NETMAR will deploy Catalogue Service for the Web (CSW), Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Sensor Web Enablement (SWE), Open-source Project for a Network Data Access Protocol (OPeNDAP), and Web Processing Service (WPS) services. A number of tools supporting these protocols have been reviewed. Table 2-1 lists the recommended tools for NETMAR.

Table 2-1 Recommended tools for NETMAR protocols and subsystems.

Protocol/Subsystem	Recommended tool
CSW	GeoNetwork
WMS	GeoServer or MapServer
WFS	GeoServer
WCS	GeoServer, MapServer of THREDDS Data Server
SWE	Constellation
OPeNDAP	THREDDS Data Server/Dapper (server); DChart/ncWMS (client)
WPS	PyWPS
Workflow engine	None (further investigation and trials are needed)
Portal framework	Liferay or Apache Jetspeed 2

The NETMAR service chaining subsystem will make it easier to construct complex services on demand. A graphical chaining editor will be created which will allow users to define their own workflows. Although some tools were identified that could be useful for the development of the service chaining editor, these are not yet as mature as the OGC data delivery protocols listed above. It is therefore needed to further investigate such tools before choosing the most promising chaining editor system.

NETMAR will use a portal framework to integrate all client-side subsystems of NETMAR (i.e. service discovery and access, service composition, and GIS viewer with decision-support) into a coherent portal. Several portal frameworks and two Java-based standards for developing portal components have been reviewed. The most promising portal frameworks found were Liferay and Apache Jetspeed 2. Subsequent testing led to Liferay being chosen.

2.2 Review of ontologies and ontology tooling

The strategy of NETMAR is to build on existing ontologies and tools. Several ontologies that can be reused as building blocks were identified. The most relevant ones are listed below; for a full list and description see [LLC10]:

- The GEMET concept thesaurus is a rich multilingual resource with a suite of simple to use access methods that should be included in part if not in its entirety in the NETMAR semantic infrastructure.
- The NERC Vocabulary Server holds a very large amount of English language content (for example a sea area gazetteer; a hierarchy of sensors and sampling devices; and a hierarchy of parameters for both discovery and usage metadata) and resource bridges that could be subsumed developed and extended through ontology extension to form the core of the NETMAR semantic resource.
- The Marine Metadata Interoperability Ontology Registry and Repository is a resource that is of significant interest to NETMAR as it is the preferred repository for the US International Coastal Atlas Network community, which is one of the targeted user communities in NETMAR.
- The Global Change Master Directory (GCMD) keyword lists, particularly the Science Keywords, have significant potential for NETMAR. Further investigation is however needed to verify its applicability, among others if a planned GCMD keyword server becomes available during the NETMAR project.
- The NASA SWEET ontologies represent a rich resource of knowledge of relevance to the NETMAR semantic framework, designed as an upper-level ontology for Earth system science.
- The InterRisk ontologies [LLH+08] were developed for a NETMAR forerunner project and cover two of the NETMAR example scenarios for oil spill and harmful algal bloom monitoring and forecasting.
- NETMAR has a requirement to be able to document service chain interfaces. Some of the required semantic information is contained within Quantities, Units, Dimensions and Types ontology.

We also reviewed technologies and tools for creating, manipulating and storing ontologies. This review led to a recommendation that a Resource Description Framework (RDF) based ontology be developed for use by the NETMAR project. This ontology should be queried by SPARQL and served using the Mulgara server or the Jena framework, unless a cluster server is available, in which case 4store is recommended. Editor tools such as Semantic Turkey or a query language web interface should be used and to bridge or extend existing ontologies the CMAPTools Ontology Editor and MMI VINE software should be used.

2.3 Review of semantic frameworks

The main semantic frameworks for environmental information systems (EIS) reviewed were [Las10]: (1) the International Coastal Atlas Network (ICAN) semantic mediation prototype, (2) the Open Access Ontology/Terminology for the GMES Space Component (OTEG) semantic framework, (3) the ORCHESTRA Semantic Catalogue Architecture, (4) the NERC Data Grid Vocabulary Server, and (5) the NASA Global Change Master Directory (GCMD).

These semantic frameworks offer different architectures, capabilities and functionalities of interest to the NETMAR project.

The main requirements for the NETMAR semantic framework as identified through the defined NETMAR case studies are the following:

- *Data Discovery* – ability to perform semantic data discovery,
- *Service Discovery* – ability to perform semantic service discovery,
- *Interoperability* – ability to perform semantic data, metadata or service interoperability,
- *Service Chaining* – ability to perform semantics-based service chaining,
- *Multi-Domain Support* – support for multi-domain ontologies,
- *Multilingual Support* – support for multi-lingual vocabularies or ontologies,
- *Multi-Facet browsing/search* – support for multi-facet data/service search or browsing,
- *Ontology Browsing* – support for ontology or vocabulary browsing by users,
- *Smart Search* – support for meaning-based free text search by users,
- *Semantic Queries* – support for semantics-based queries by users.

None of the reviewed semantic frameworks offer all of these capabilities and functionalities on its own. Therefore, we recommend that the NETMAR semantic framework should build on existing semantic framework architectures and adapt these as needed to fulfil the NETMAR use case requirements. This allows the NETMAR project to draw upon best practices and proven solutions for semantic framework development in order to obtain a flexible and maintainable semantic framework for EIS.

3 Initial system architecture

3.1 Design methodology

The RM-ODP (Reference Model for Open Distributed Computing) is a standard for describing system architectures through five complementary but consistent viewpoints:

- The **enterprise viewpoint**, which has focus on the overall purpose, scope and policies for the system, and describes its business requirements and how these can be met.
- The **information viewpoint**, which defines the semantics of the information that the system will handle and the needed information processing. This viewpoint describes the information handled by the system as well as the structure and type of supporting data.
- The **computational viewpoint**, which specifies the system as a set of objects (or services) interacting through interfaces. This view describes the functionality of the system and the functional decomposition.
- The **engineering viewpoint**, which focuses on the mechanisms and functions required to support distributed interaction among objects in the system.
- The **technology viewpoint**, which describes the technologies chosen to implement the system.

In the initial system architecture we have concentrated on giving an example of each viewpoint, describing how it can be used to increase communication and collaboration among stakeholders. This allows us to get early feedback on how well each viewpoint can be tailored to meet the needs of NETMAR.

3.2 Information Viewpoint

Figure 3-1 illustrates the use of the RM-ODP information viewpoint to describe the use of metadata in NETMAR. It uses an invariant schema to show constraints on metadata which must be observed throughout the project and by all partners and data providers. The constraints are expressed as typed relations between elements, cardinality restrictions on the roles in those relations, and additional constraints expressed in the Object Constraint Language (OCL).

One constraint expressed in this diagram is that each dataset must have an associated metadata record. Another is that the name for each metadata element must be associated with a unique type within the EUMIS. In this way, we describe a set of rules which will ensure that the data produced by NETMAR is usable and that the system is buildable.

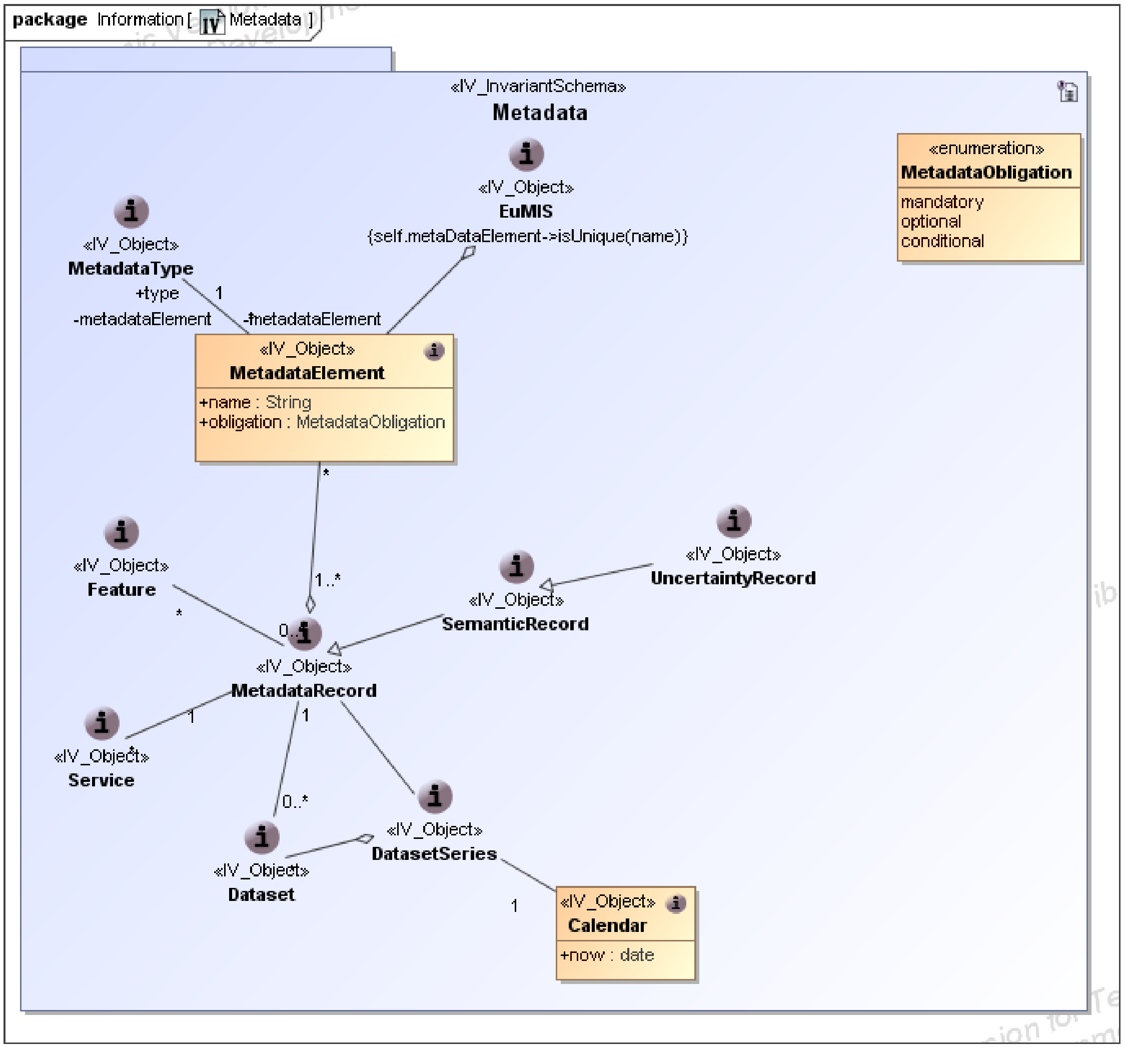


Figure 3-1 Information Viewpoint: Metadata Invariant Schema

4 Ontology extension and semantic framework design

4.1 Ontology extension methodology

There are two issues to address for ontology extension across distributed semantic resources:

- the addressing (URLs) of and access (API methods) to concepts and concept schemes in a uniform manner
- the harmonisation of their representation as XML documents.

Ideally this should be achieved through standards conformance. However, there are no standards for concept addressing and access. Whilst there are W3C outline standards for concept representation such as OWL and SKOS there are no standard usage profiles for use in operational interoperability.

A two-part methodology strategy is therefore required. First, there needs to be NETMAR involvement in the development process of the necessary standards. Secondly, there is a need to develop pilot tools for interoperability based on proto-standards developed within NETMAR. The latter should obviously be fed into the former.

Standards development involvement is focussed on contact, and future involvement, with a W3C incubator group on vocabulary standards that has been proposed by Peter Fox (Rensselaer Polytechnic Institute, USA) and Rob Atkinson (CSIRO, Australia). Contact expressing interest in this work has been made with Peter Fox. The pilot development will be based on bringing the NERC Vocabulary Server (NVS) into line with the standards requirements of the NETMAR Semantic Web Service (SWS) [LL10].

The NVS is a semantic resource developed over the past 20 years by a NETMAR partner (BODC) that is populated by over 140,000 concepts in over 100 schemes interlinked by over 100,000 semantic relationships. Concepts and schemes are addressed through pseudo-RESTful URLs, accessed through a SOAP API and represented in an early form of the W3C SKOS standard. There are three developments required to bring the current system up to the level required:

- implementation of semantic linkage infrastructure to concepts outside the NVS namespace
- development of truly RESTful URL addressing and developing the SOAP API to the level required by other potential users of vocabulary standards
- development of the concept and scheme representation documents to the SKOS version adopted by W3C and specified for the SWS.

A bridge, even a pilot, obviously needs to have two ends. The basic requirement for the ‘other end’ is a semantic resource offering a Web Service delivering XML documents in response to a concept or scheme URL. The Marine Metadata Interoperability (MMI) Ontology Register and Repository (ORR) has been identified as the most promising prospect for initial development of a pilot bridge with NVS due to the high level of external empowerment over content and the long history of collaboration between BODC and MMI. Other resources offering suitable services have been catalogued in [LLC10].

Should the development of a bridge to a semantic resource that doesn’t offer the necessary service interface be deemed essential to NETMAR service development then that may be accomplished by ingestion of that resource into NVS and implementation of mappings within the NVS namespace.

4.2 Semantic framework design

The semantic framework design involves the design of an architecture for the NETMAR semantic framework, and the design of the web interface and protocols for accessing the semantic resources.

An initial architecture has been proposed for the NETMAR semantic framework (Figure 4-1). The proposed architecture builds on the existing, well populated and governed NVS (c.f. section 4.1). On top of the NVS, a web interface, called the NETMAR Semantic Web Service (SWS), provides high-level SKOS-like operations required by most SWS clients (e.g. catalogue service mediator, service chaining engine, ontology browser, data and service discovery interface, and external web applications). In contrast to a SPARQL-based web interface (the design and implementation of which will be investigated as part of NETMAR), the SKOS-based SWS trades expressiveness for ease of use. It provides a very easy-to-use query interface tailored for SKOS-based thesauri, while retaining maximum usefulness for most common applications requiring such semantic resources.

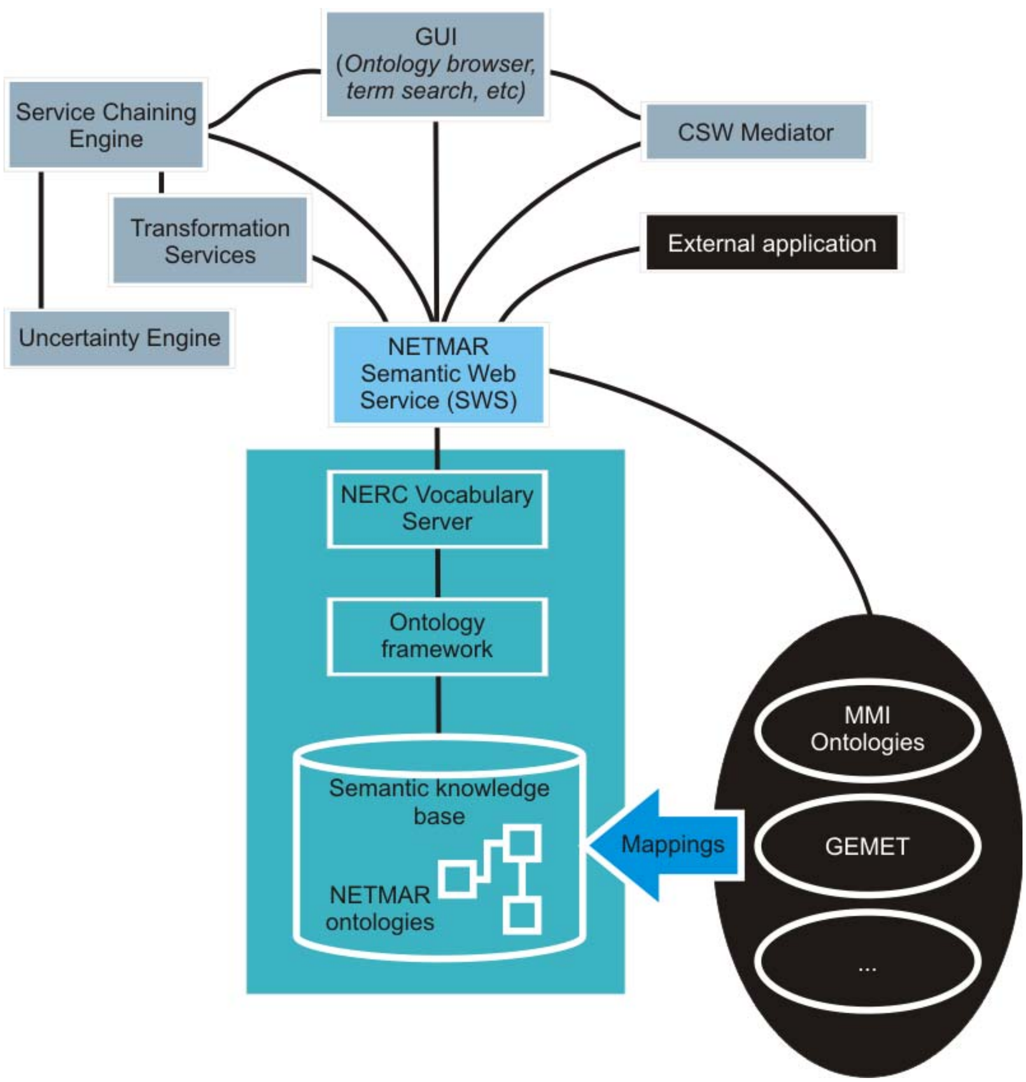


Figure 4-1 NETMAR High-Level Semantic Framework Architecture.

Examples of the SKOS-based operations supported by the designed NETMAR SWS (version 1.0) are:

- ***GetConceptScheme***, for retrieving the content of a specified concept scheme;
- ***GetConcept***, for retrieving the definition of a specified concept;
- ***GetRelatedConcepts***, for retrieving all concepts related to a specified concept with a specified SKOS relationship.

In total, we have proposed a set of eight operations to be supported by the first version of the NETMAR SWS [LL10]. As part of the SWS design [LL10], we specified the request, response and exception formats for each of these operations. We specified requests in both XML and Keyword Value Pair (KVP) encodings. For XML encoding of requests, responses and exceptions, an XML schema will become publicly available. We also specified the way XML requests, responses, and exceptions shall be embedded in SOAP messages, as a W3C-compliant SOAP interface turns out to be essential.

The SWS interface supports multilingual ontologies and requests. For instance, SWS clients can search concepts in any language supported by the semantic resources. The interface may also specify the language used for resource description in SWS responses.

As part of the SWS specification, we defined the type and format of semantic information to be returned by each of the SWS operations. This needs further refinement by specifying the SKOS profile to use within the response. For this purpose, we will develop an XML schema that will define the exact XML structure and content of the SKOS resources returned within each SWS response type.

5 Other major achievements

5.1 Data delivery services

For the six pilots, NETMAR partners have established a set of data delivery services, as listed in Table 5-1. These products constitute a combination of observations from satellites and in situ instruments, parameters derived from such measurements, and forecasts of future met-ocean, ecosystem or pollution conditions.

Table 5-1 NETMAR data delivery services.

Pilot	Product	Provider
1. Arctic sea ice monitoring and forecasting	Sea ice concentration from satellite	NERSC
	Sea ice concentration forecast	NERSC
	SAR quick-looks and SAR-wind	NERSC
	Ice chart from operational ice service	METNO
	OSISAF sea ice concentration, ice edge and ice type	METNO
	HIRLAM air temperature forecast	METNO
2. Oil spill drift forecast and shoreline clean assessment services in France	MOTHY results (oil drift forecast)	Meteo France
	Oil map results	Cedre
	POLREP (pollution report)	Cedre
	Shoreline survey	Cedre
3. Relationships between physical and biological variables	Chlorophyll, nLw, from satellite	PML
	Surface chlorophyll (in situ)	PML
4. Ecosystem model validation	Chlorophyll, nLw, from satellite	PML
	Chlorophyll (mode hindcast)	PML
5. ICAN for coastal zone management	Many products covering areas such as: Biological environment, Physical environment, Socio-economic activity, Management, etc.	CMRC ICAN partners
6. Phytoplankton blooms in Gulf of Biscay and English Channel	Chlorophyll, nLw, from satellite - from MODIS - from MERIS - from MODIS/MERIS	Ifremer
	Chlorophyll-a, Turbidity, Temperature, Salinity, Oxygen (in situ)	Ifremer

5.2 Tools for processing services

The OGC Web Processing Service (WPS) is currently undergoing intense development, as was shown by the activity at the FOSS4G 2010 conference in Barcelona where several WPS implementations presented their newest releases.

The NETMAR project has been a major contributor to the open source Python implementation of WPS, PyWPS. Part of this contribution has been a wiki containing information for basic users and system developers (Figure 5-1). This was presented at FOSS4G 2010 and has been actively updated since.

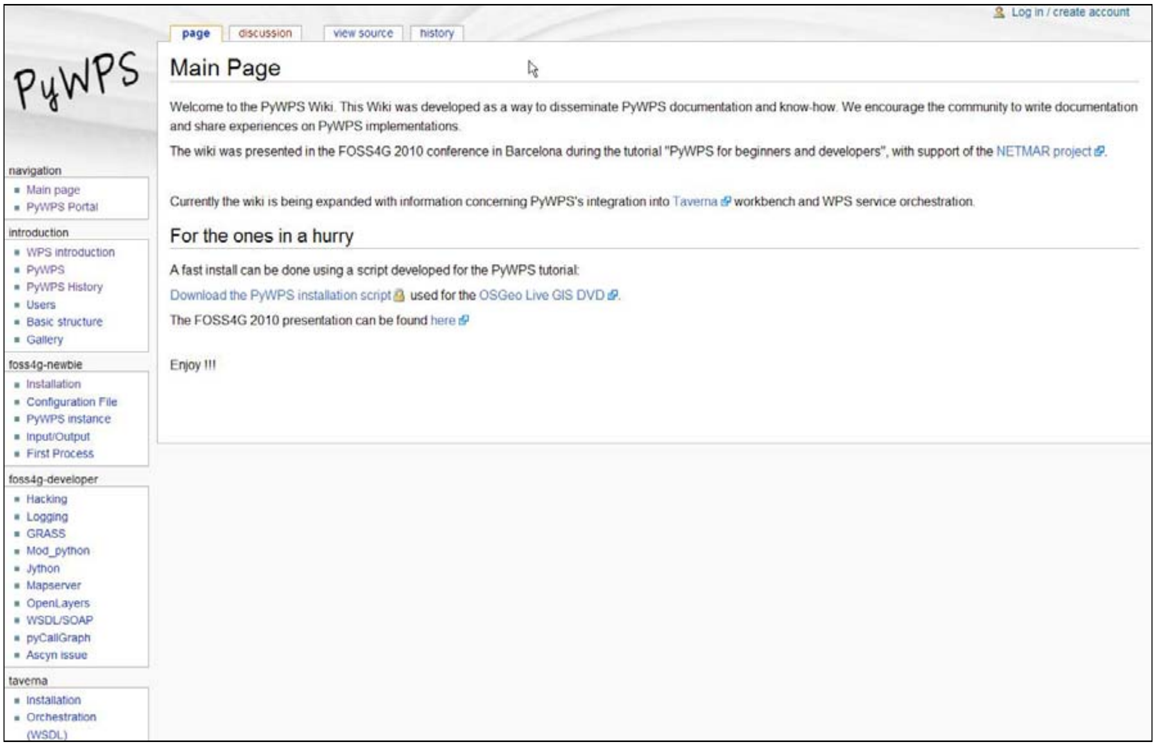


Figure 5-1 Screenshot of the pyWPS wiki maintained as part of the NETMAR project.

The wiki covers topics like:

- Basic installation
- Basic process development
- Integration with other software such as: GRASS-GIS, Mapserver, OpenLayers
- Orchestration of services

The NETMAR project has implemented WPS 1.0.0 Annex D (Use of WPS with SOAP) and Annex E (WSDL best practice) in PyWPS, enabling WPS services to be fully integrated in BPEL engines or other workflow orchestration systems such as Taverna (http://www.taverna.org.uk/). The contributions to PyWPS also include the integration of asynchronous WPS calls into WSDL/SOAP.

The aim of the development was to provide a transparent implementation where, when the user creates a WPS service, it automatically generates a WSDL description and provides a SOAP service binding. This means that the service is immediately available for orchestration.

Sample WPS services have been coded and workflows built within Taverna. One example service generates histograms of the pixel values for supplied images. This is illustrated in Figures 5-2, 5-3 and 5-4.

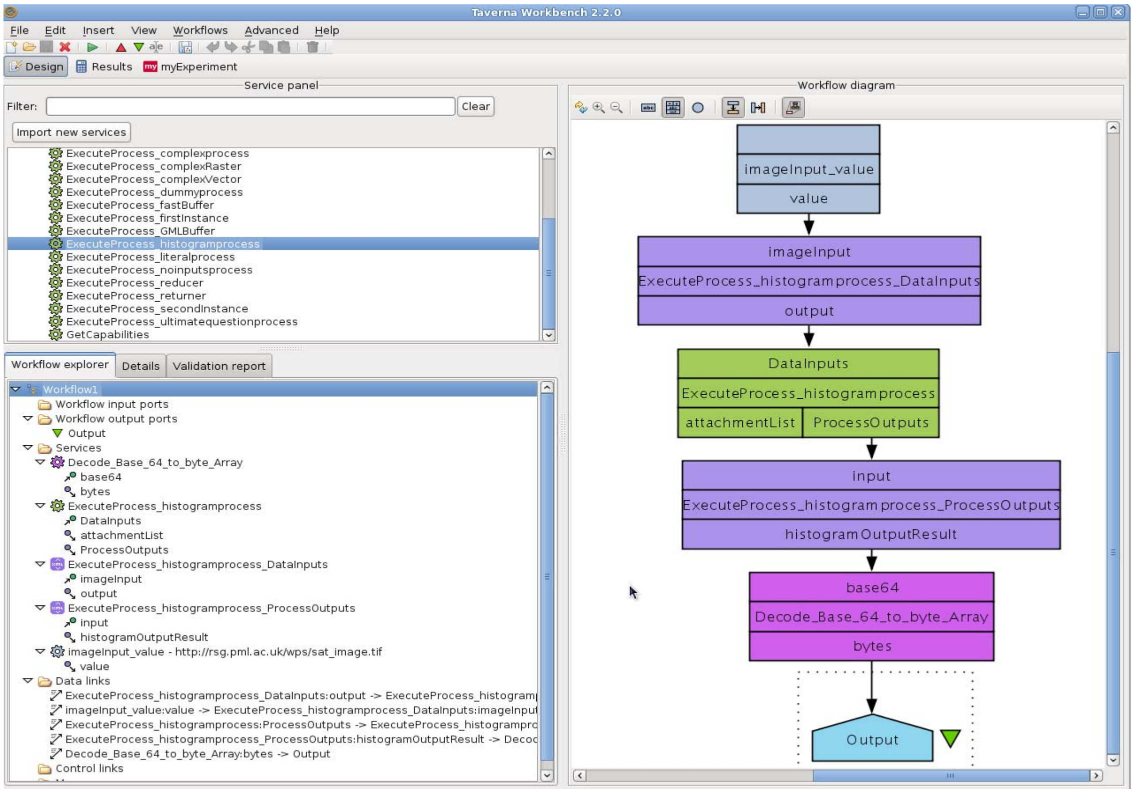


Figure 5-2 Screenshot of Taverna running a WPS service (green-box ExecuteProcess_histogramprocess) associated with extra functionalities like data input loading (grey box), data input/output (purple box) and data transformation (magenta).

Figure 5-3 Workflow image input (satellite image with Zooplankton biomass in mg-C m-3).

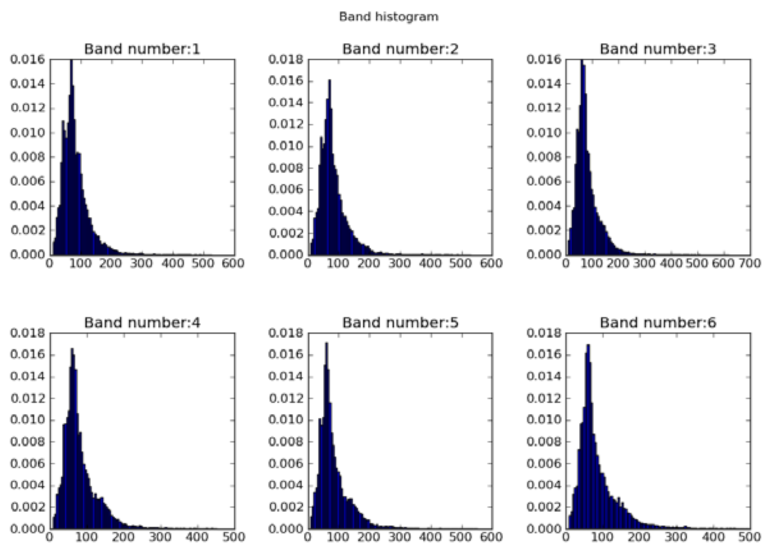


Figure 5-4 Histogram result using workflow in Figure 5-2. Each band represents a different time dimension.

The use of standard workflow descriptions and services supporting WSDL/SOAP allows workflows to be shared on collaborative sites such as myExperiment (http://www.myexperiment.org) where users can search, edit and comment on workflows (Figure 5-5).

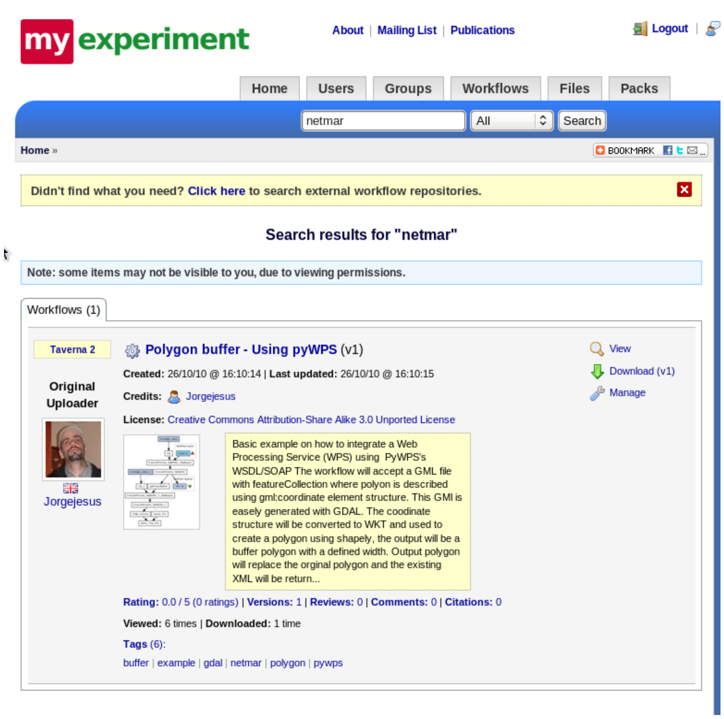


Figure 5-5 myExperiment seach using NETMAR as keyword. Workflow found is a basic polygon buffer experiment, used for testing in the project.

Developing new WPS processes can be time consuming and requires basic programming skills. In order to lower the technical requirements for WPS use existing tools such as GRASS-GIS can be made available though PyWPS, using automatic scripts. NETMAR has contributed to this effort within WPS-GRASS-BRIDGE (http://code.google.com/p/wps-grass-bridge/) by improving interoperability between PyWPS and GRASS. In preliminary tests 137 modules were ported to PyWPS, becoming available through the WSDL/SOAP interface (Figure 5-6).

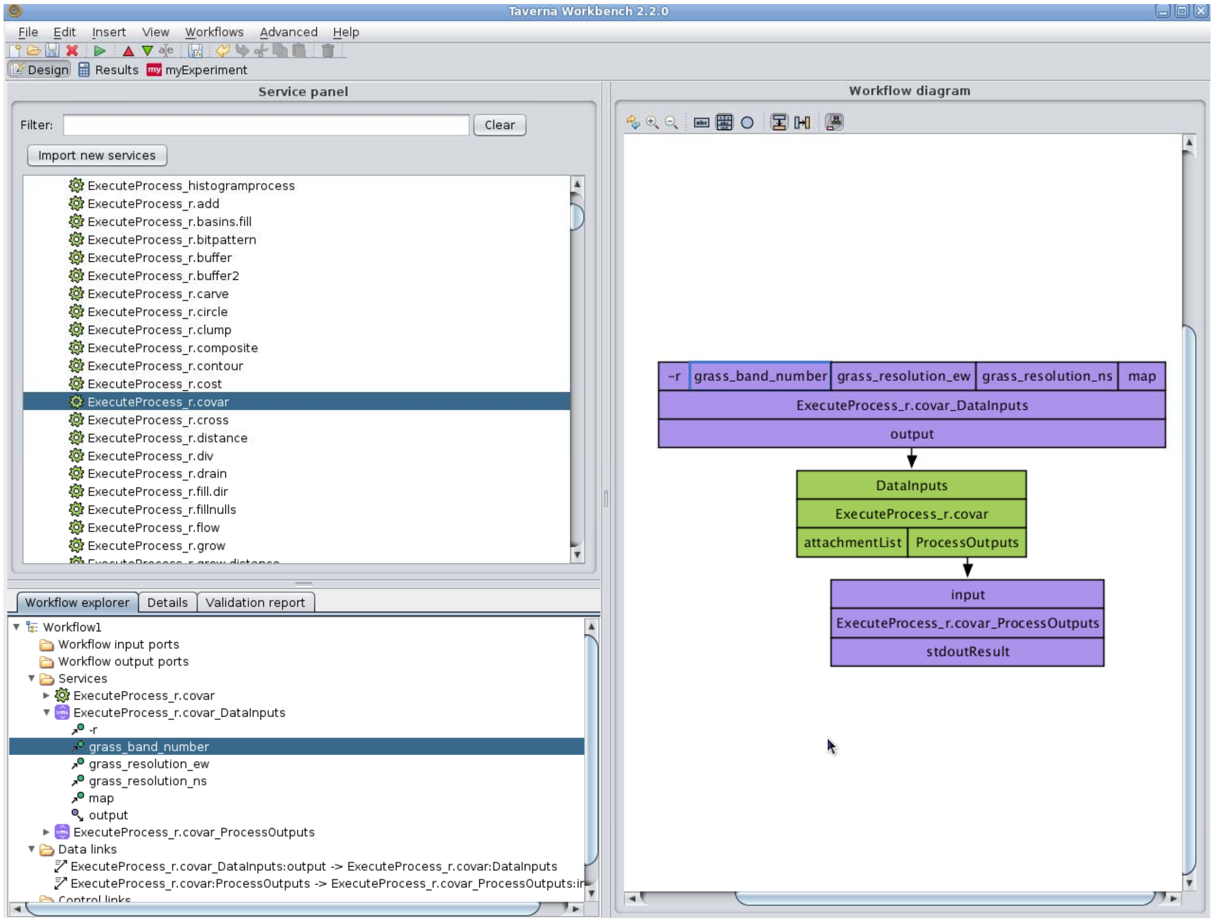


Figure 5-6 Service integration between Taverna workbench – PyWPS – GRASS-GIS bridge.

Future development will focus on workflow portlets to be integrated in NETMAR portal and extending the meta-data passed between web services to allow some semantic processing to be carried out.

5.3 Baseline NETMAR portal

A first version of the NETMAR portal was set up, using Liferay Community Edition software. This contains a wiki for each of the six pilots and additional information about technologies and tools used. The portal also contains a forum where registered users can post questions about products and services. Figure 5-7 shows one of the pages in the NETMAR wiki, for the Arctic Sea Ice Monitoring and Forecasting wiki.


Wiki

Home | Recent Changes | All Pages | Orphan Pages | Draft Pages | Search

Arctic Sea Ice and Metocean Observing System

Edit Details Print

Table of Contents [-]

Users

The users of the Arctic Sea Ice and Metocean Observing System pilot comprises of representatives from offshore and shipping companies, ship and icebreaker captains/ice pilots, national authorities, regional environmental agencies, national ice services and scientists.

The second user category, ship and icebreaker captains and ice pilots, makes both selection of sailing routes between ports of destinations as well as tactical decisions on how to penetrate through rough ice conditions when the ship (un)expectedly experience such situations. Decisions need to be made in order to operate safely and efficiently.

Products

Some examples of products offered for planning of sailing routes and convoy operations in sea ice covered areas are shown below. Satellite Synthetic Aperture Radar (ASAR) images provide detailed information about sea ice conditions as illustrated in Figure 1. SAR images can be classified by automatic algorithms (Figure 2) and trained operators of an ice service (Figure 3). Other products that are useful for captains and ice pilots include ice drift predictions, met-ocean forecasts and observations from ships and airplanes.

Figure 1

Figure 2

Figure 3

Services

These and other products will become available through the Arctic Sea Ice and Metocean Observing System Pilot.

Figure 5-7 Screenshot of the NETMAR wiki for Pilot 1: Arctic Sea Ice Monitoring and Forecasting.

6 References

[Dun10] Dunne, Declan, et al. (2010). D2.1 Review of projects, initiatives & technologies addressing system architectures for distributed Environmental Information Systems (EIS). NETMAR Deliverable D2.1 (grant agreement number 249024).

[Las10] Lassoued, Yassine et al. (2010). D4.1 Review of semantic frameworks. NETMAR Deliverable D4.1 (grant agreement number 249024).

[LC10] Leadbetter, Adam and Oliver Clemens (2010). D3.2 Review of available ontology tooling. NETMAR Deliverable D3.2 (grant agreement number 249024).

[LLH+08] Lassoued Y., Hamre T., Hackett B., Groom S. and O´Grady E. (2008), Interrisk ontologies (revised). NERSC special report, No. 82, 2008.

[LL10] Lassoued, Yassine and Adam Leadbetter (2010). D4.3.1 NETMAR Semantic Framework specification – Version 1. NETMAR Deliverable D4.3.1 (grant agreement number 249024).

[LLC10] Lowry, Roy, Leadbetter, Adam and and Oliver Clemens (2010). D3.3 Review of available ontologies and their interfaces. NETMAR Deliverable D3.3 (grant agreement number 249024).

[Val01] Vallecillo, Antonio, 2001. RM-ODP: The ISO Reference Model for Open Distributed Processing