Joint European Research Infrastructure network for Coastal Observatory – Novel European eXpertise for coastal observaTories - **JERICO-NEXT**	
Deliverable title	Requirements specifications for the observatory operator console
Work Package Title	WP 5
Deliverable number	D5.6
Description	The document provides specifications tools that will support sea observatory operator for their data management (share, publish, and preserve dataset, document deployment in standard format…).
Lead beneficiary	IFREMER
Lead Authors	Thomas Loubrieu, IFREMER
Contributors	Loïc Petit De La Villéon, IFREMER, LeonidasPerivoliotis HCMR
Submitted by	
Revision number	V1.0
Revision Date	25/04/2017
Security	Public

The JERICO-NEXT project is funded by the European Commission's
H2020 Framework Programme under grant agreement No. 654410
Project coordinator: Ifremer, France.

History			
Revision	**Date**	**Modification**	**Author**
V0.1	03/10/2016	creation	Thomas Loubrieu
V0.2	05/10/2016	comments	Loic Petit De La Viléon
V0.3	03/01/2017	Comments	Léonidas Perivoliotis
V0.4	13/02/2017	Final submitted version (in official template)	Thomas Loubrieu
V0.6	04/04/2017	Revised version by Ifremer	Thomas LoubRIeu
V1.0	25/04/2017	Final version	Patrick Gorringe

Approvals				
	Name	**Organisation**	**Date**	**Visa**
Coordinator	Patrick FARCY	Ifremer	27/04/2017	PF
WP Leaders	Leonidas Perivoliotis	HCMR	17.04.2017	LP
WP Leaders	Patrick Gorringe	EuroGOOS	25.04.2017	PG

PROPRIETARY RIGHTS STATEMENT

Table of contents

1. Executive Summary

There is a strong necessity to improve the management process of the metadata of the deployed observatories.

As reminded in Description of Work for Task 5.3, this weakness has been highlighted during the SeaDataNet and Copernicus MEMS projects. We propose to give a major role to the observatory operators in managing the metadata. Dedicated interfaces will be implemented in order to allow operators to describe their instruments, to monitor them in real-time, to publish and to advertise the work of data acquisition. Providing these metadata publishing tools will also greatly facilitate discovery, visualization and downloading fully INSPIRE compatible services. We will take into account and will be consistent with other experiments for industry sensors such as http://www.sensorcloud.com. The standards (Sensor Web Enablement –SWE) already implemented in other projects (SeaDataNet, ODIP & JERICO) will be used.

This document defines the requirements and design of an online console for observation platforms or network operators who will use the application for both:

- Maintain and document their deployments
- Push the resulting datasets for long time preservation and availability for data services, via for example Copernicus MEMS and EMODNET.

Focus will be on a few observatory networks on which we propose to implement observatory operator advanced services: fixed moorings and non-permanent operational coastal observation networks, HF radars, flow cytometry and ferry boxes

2. Introduction

The current document reports on the requirements, high-level design and software architecture for the observatory operator console developed for JERICO-NEXT. The system is developed under the Ifremer **sensor nanny** project umbrella.

The document is organized with the following sections:

- Requirements:
  - Users and interfaces with pre-existing infrastructures
  - Core functions (shared by the different observation networks)
  - Targeted observation networks and dedicated functions
- Design:
  - Design: software architecture, technical interfaces
  - Implementation plan

3. Main report

3.1. Requirements

The current section describes the functions required for the observatory operator console.

3.1.1. *User, interfaces with pre-existing infrastructures*

The users targeted by the observatory operator console are:

- **Observatory operators**: people in-charge of organizing the deployment and maintenance of a marine in-situ observatory or a network of in-situ observatories.
- **Observation users**: Principal Investigators (PIs) or other collaborating users.
- **Long-term archive and publication infrastructure**: system (or person on account of the system) in charge of the curation of in-situ observation datasets in the long run and their publications (for example with Digital Object Identifiers). This category of infrastructure is represented by the SeaDataNet infrastructure.
- **Data Dissemination service**: system (or person on the account of the system) using in-situ observation datasets as input for a value added service (for example collection and qualification for assimilation or validation of forecasting models, climatological analysis …). This category of infrastructure is represented by the Copernicus Marine and EMODNET infrastructures.
- **Observation coordinator or decision maker**: person in charge of planning deployments or organizing future networks depending on the current situation.

Besides these core targeted users, ancillary systems are connected to the console system:

- Reference or **vocabulary services** (for example BODC vocabulary services).
- **Equipment model databases** (Sensor, Instrument, hardware, for example Fixo3 yellow pages).
- **Equipment Maintenance Management system**, for example for metrology management.
- **Dedicated observation network information system**: dedicated observatory or network oriented operator console or information system (for example ARGO GDAC, glider console).
- The **observatory** embedded system (so called smart sensors) can also be an interface of the system.

3.1.2. *Core functions*

As use cases, the common functions provided to users are:

- UC1: Describe a deployment (instrument, observatory or a network of observatories) and connected processes
- UC2: Submit a dataset in real-time or delayed mode
- UC3: Discover, Visualize and Analyze equipment and datasets
- UC4: Share information and data with colleagues or data with long term preservation and publication infrastructure or data service

The users and pre-existing infrastructures interact with use cases as follow:

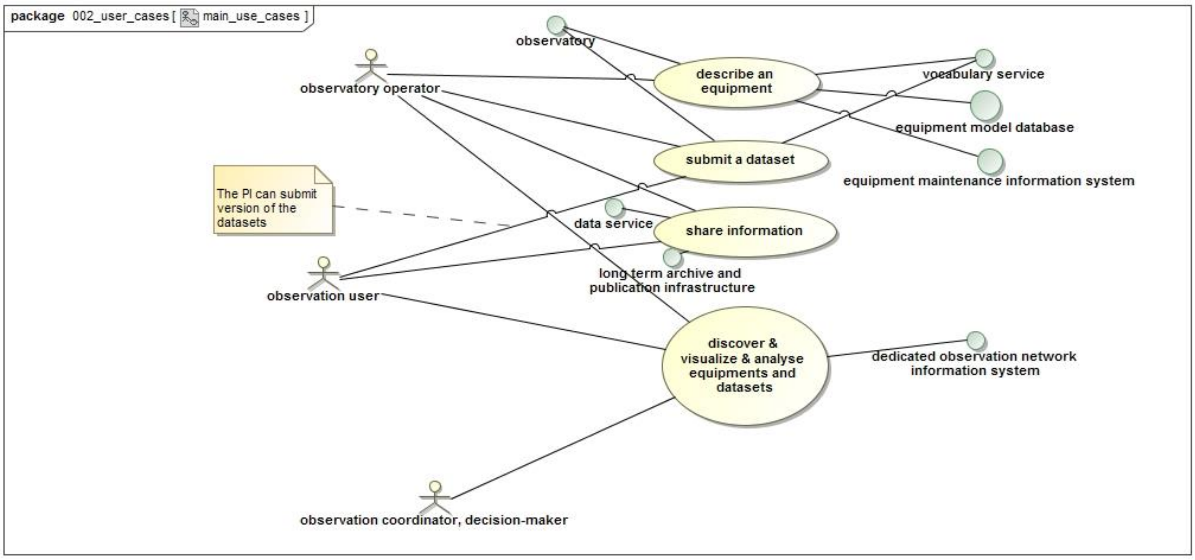

package 002_user_cases [main_use_cases]
observatory
observatory operator
describe an equipment
vocabulary service
equipment model database
submit a dataset
equipment maintenance information system
The PI can submit version of the datasets
data service
share information
long term archive and publication infrastructure
observation user
discover & visualize & analyse equipments and datasets
dedicated observation network information system
observation coordinator, decision-maker

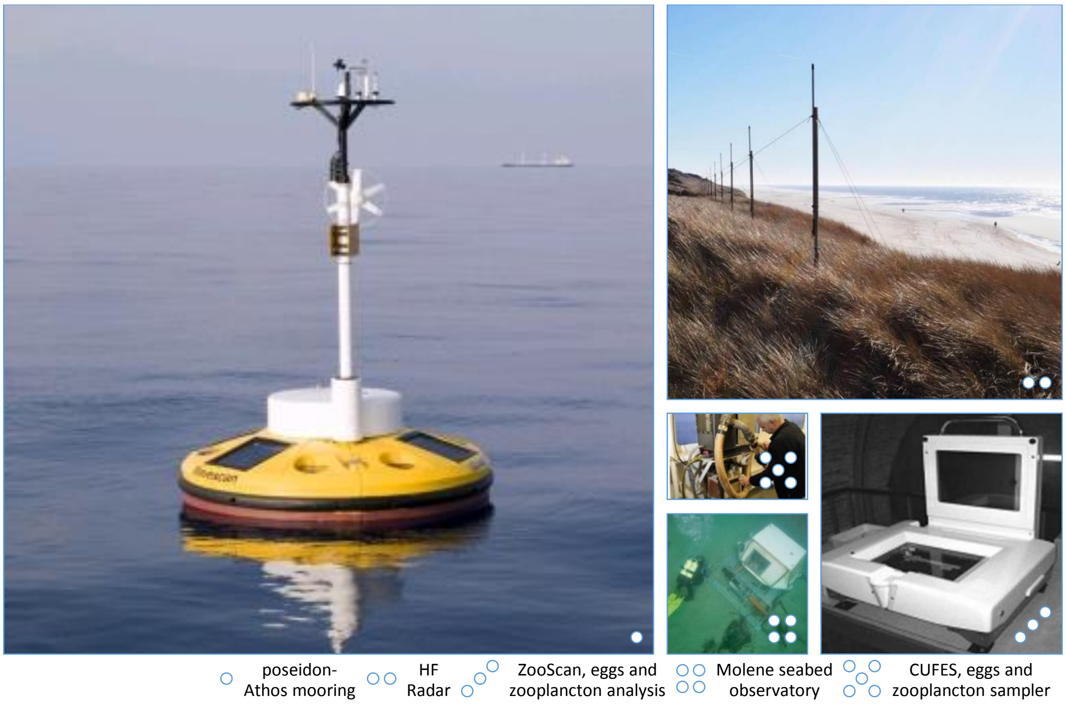
Figure 1: actors, systems and user cases

3.2. Targeted observation networks and dedicated functions

The targeted observation networks for the observatory operator console in JERICO-NEXT are:

1. Fixed observatories (seabed or moored), namely EMSO-molene and Poseidon/Athos observatory
2. HF radars deployments
3. Flow cytometry observations

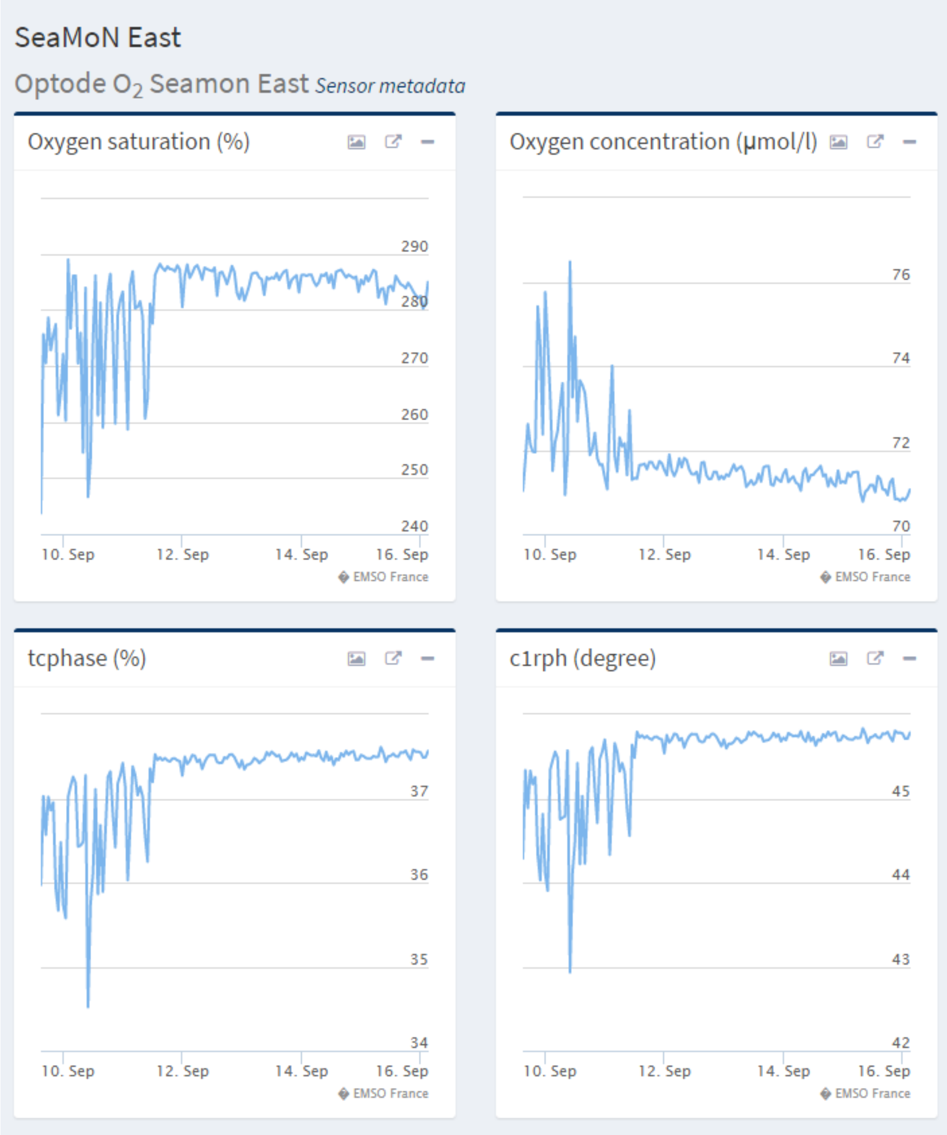
Ferry-boxes or other type of platforms might also be included in the demonstration depending on availability of resources or co-development with other projects.


○ poseidon-Athos mooring
○○ HF Radar
○○○ ZooScan, eggs and zooplancton analysis
○○○○ Molene seabed observatory
○○○○○ CUFES, eggs and zooplancton sampler

3.2.1. *Fixed observatories*

For fixed observatories, the foreseen extensions are:

- UC1-1: describe a fixed platform
- UC2-1: submit CSV time-series data files
- UC3-1: time-series visualization
- UC2-2: submit video data files
- UC3-2: visualize video



SeaMoN East

Optode O_2 Seamon East *Sensor metadata*

Figure 2: example of time-series visualization from EMSO-France web portal

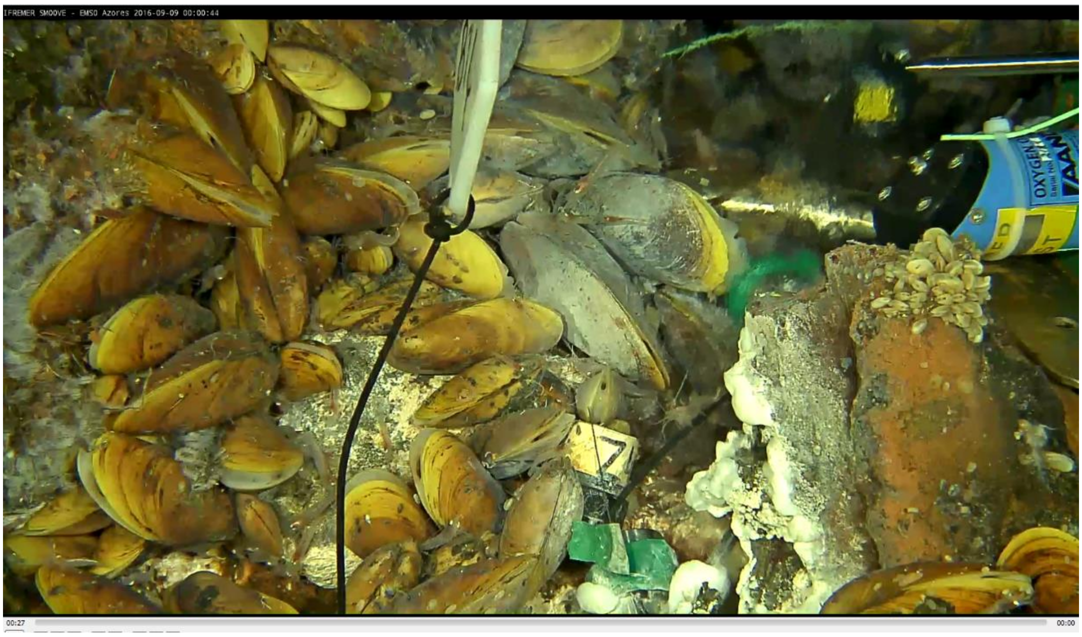
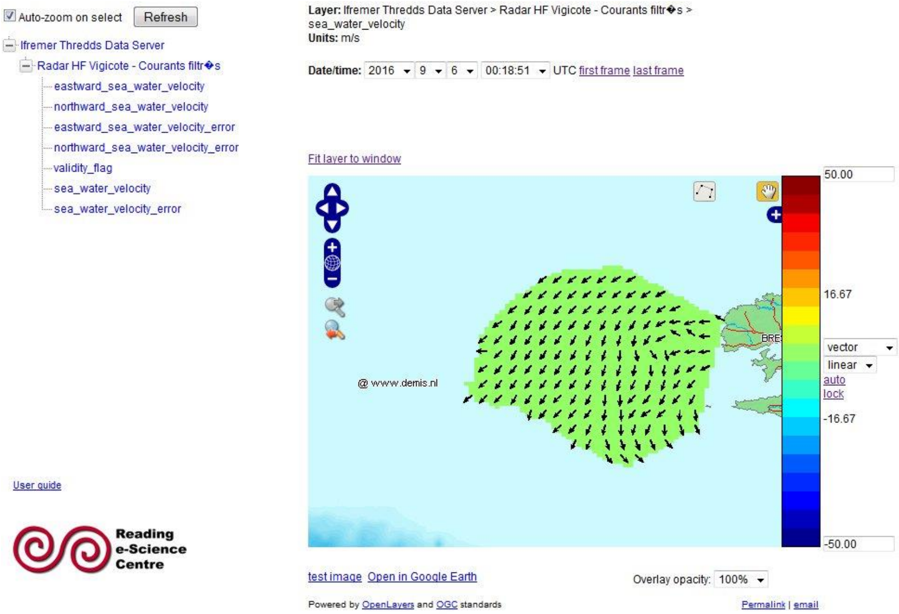


Figure 3: EMSO-Azores deep sea observatory video

3.2.2. HF radars

For HF-radars, the foreseen extensions are:

- UC1-2: Describe HF radar site and data post-processing: Description of geometry of the deployment site and description of processing chain downstream the acquisition should be supported
- UC2-3: Submit 3-4D-netCDF data files (time plus spatial grid) and pre-processing.
- UC3-3: Coverage visualization


☑ Auto-zoom on select

Layer: Ifremer Thredds Data Server > Radar HF Vigicote - Courants filtr�s > sea_water_velocity
Units: m/s

Date/time: 2016 9 6 00:18:51 UTC first frame last frame

Figure 4: example of visualization of HF radar processed observation

3.2.3. *Flow cytometry*

The foreseen extensions for flow cytometry are:

- UC2-5: Submit images (individual pictures, for example zoo-plankton)
- UC3-5: visualize images (album)
- UC2-6: submit CSV biology sampling
- UC3-6: visualize biology sampling visualization (for example pie-chart).

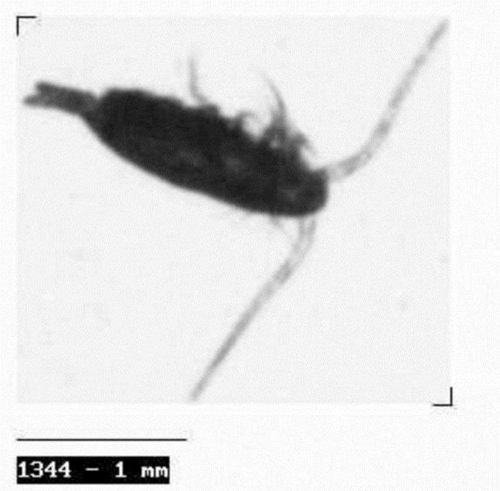


Figure 5: picture of zooplancton collected with zoocam flow citometer

3.3. Design

3.3.1. *Software architecture, technical interfaces*

The sensor nanny system has been prototyped and tested for two years, with multiple observation networks or platforms (Argo profiling floats, research vessels, autonomous sailing ship, and deep sea observatories) delivering significant volume of data (up to 1.5 million observations).

After different development iterations including the tests of different technical options, the following architecture has been stabilized. It will be used as a basis to fulfil the JERICO-NEXT requirements.

The following components have been identified:

- CM1: File sync, share server and plugins to manage datasets in various formats and observatory description files. This is the cornerstone for the users submitting resources to the system (datasets and observatory descriptions).
- CM2: Deployment editor
- CM3: Indexation engine
- CM4: Discovery, visualization and analysis portal

	CM1: File sync and share server and plugins	CM2: Deployment editor	CM3: Indexation engine	CM4: Discovery, visualization and analysis portal
UC1: Describe a deployment				
Core	X	X		
UC1-1: fixed platform		X		
UC1-2: Site geometry		X		
UC2: Submit a dataset				
Core	X		X	
UC1-3: Data post-processing	X		X	
UC2-1: CSV time series	X		X	
UC2-2: vidéos	X		X	
UC2-3: 3-4D NetCDF	X		X	
UC3: Discover, Visualize and Analyze				
Core			X	X
UC2-4: images				X
UC2-5: CSV biology sampling				X
UC3-1: CSV time series				X
UC3-2: vidéos				X
UC3-3: 3-4D netcdf				X
UC2-4: images				X
UC2-5: CSV biology sampling				X
UC4: Share information				
individuals	X		X	X
functions (long term archive, data service)	X		X	X

Figure 6: implementation matrix, use cases vs components

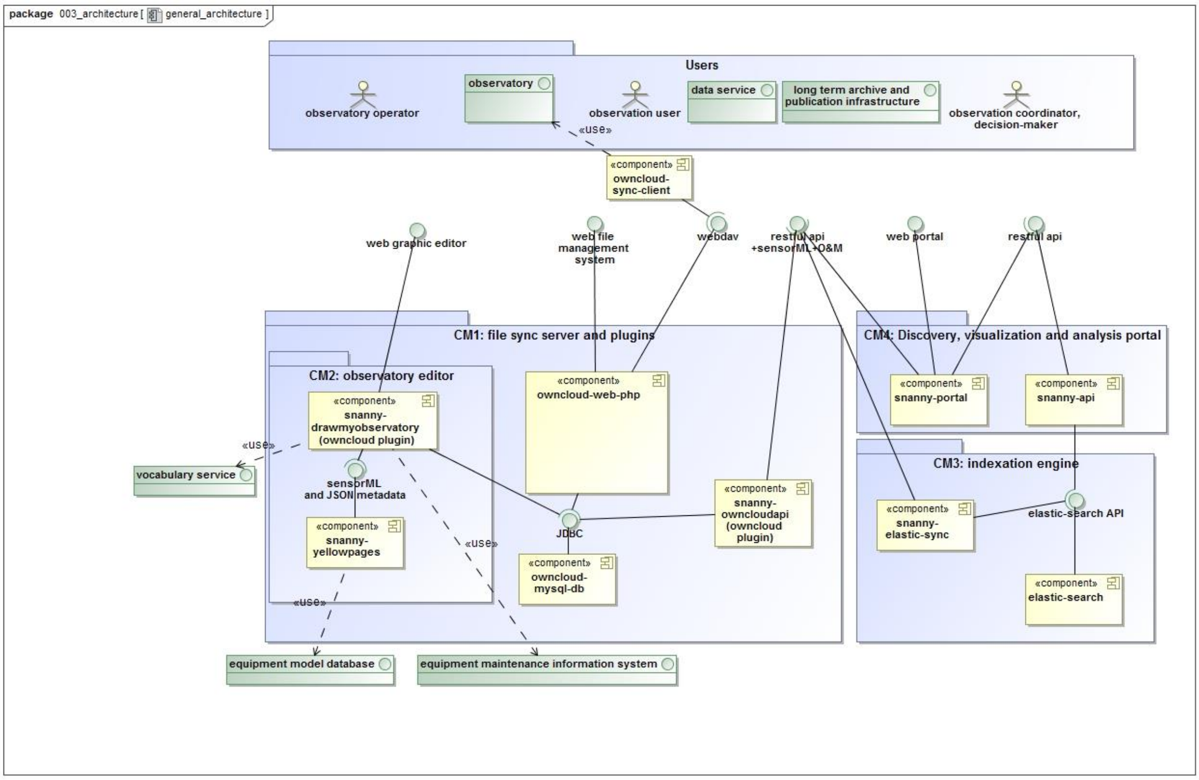

package 003_architecture [general_architecture]
Users
observatory operator
observatory
observation user
data service
long term archive and publication infrastructure
observation coordinator, decision-maker
«use»
«component» owncloud-sync-client
web graphic editor
web file management system
webdav
restful api +sensorML+O&M
web portal
restful api
CM1: file sync server and plugins
CM4: Discovery, visualization and analysis portal
CM2: observatory editor
«component» snanny-drawmyobservatory (owncloud plugin)
«component» owncloud-web-php
«component» snanny-portal
«component» snanny-api
«use»
vocabulary service
sensorML and JSON metadata
CM3: indexation engine
«component» snanny-owncloudapi (owncloud plugin)
«component» snanny-elastic-sync
elastic-search API
«component» snanny-yellowpages
JDBC
«use»
«component» owncloud-mysql-db
«component» elastic-search
«use»
equipment model database
equipment maintenance information system

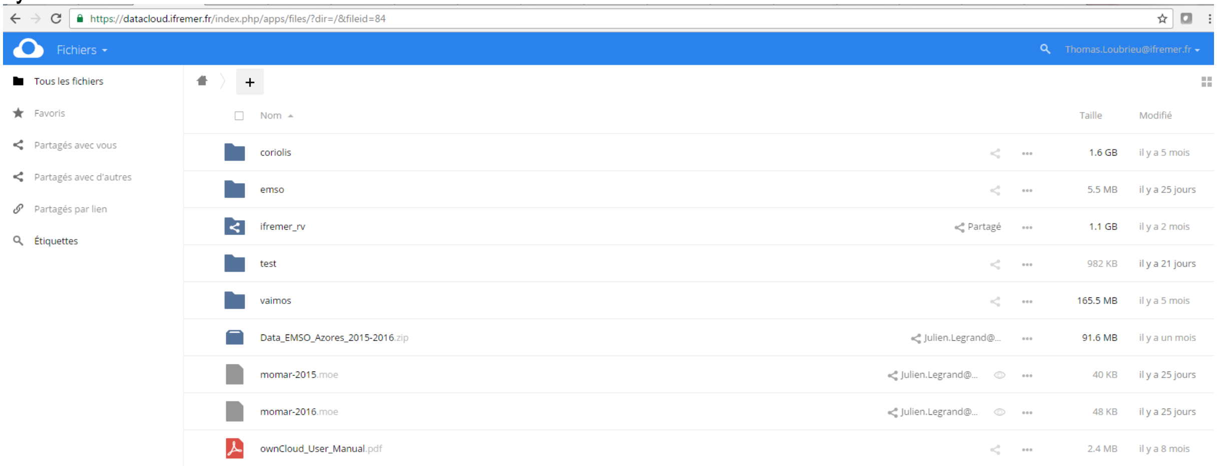
Figure 7: the components, sub-components and technical interfaces in-between

The components used can be on-the-shelf components (as much as possible open-source) or dedicated components. The dedicated component sources are managed using GitHub hosting facility (https://github.com/) For each reference are given in the detailed architecture section.

3.3.2. Detailed architecture

3.3.2.1. CM1: File sync and share server

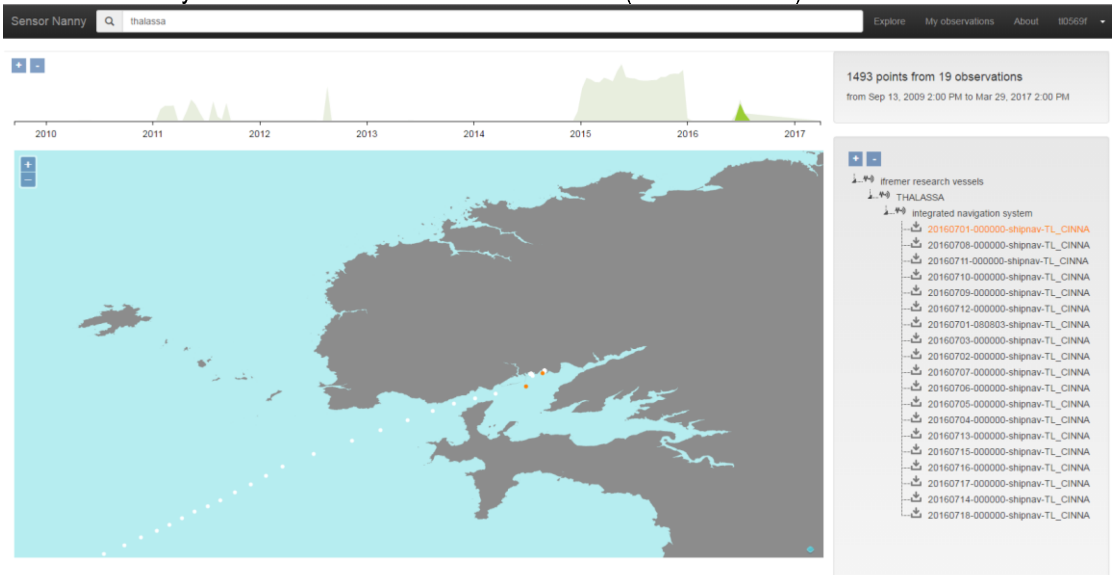
The open source, on the shelf component, **owncloud** has been selected to implement the file sync and share systems.



Nom		Taille	Modifié
coriolis		1.6 GB	il y a 5 mois
emso		5.5 MB	il y a 25 jours
ifremer_rv	Partagé	1.1 GB	il y a 2 mois
test		982 KB	il y a 21 jours
vaimos		165.5 MB	il y a 5 mois
Data_EMSO_Azores_2015-2016.zip	Julien.Legrand@...	91.6 MB	il y a un mois
momar-2015.moe	Julien.Legrand@...	40 KB	il y a 25 jours
momar-2016.moe	Julien.Legrand@...	48 KB	il y a 25 jours
ownCloud_User_Manual.pdf		2.4 MB	il y a 8 mois

Figure 8: owncloud web interface

The tool is now widely used in companies as an in-house dropbox system. It allows authenticated users to manage their files in their web browsers or to synchronize their local directories with the cloud system, through a webdav interface (http://www.webdav.org/). This latest feature is useful for automated, real-time data submission. It has been successfully tested with IFREMER research vessels (see screenshot).



1493 points from 19 observations

from Sep 13, 2009 2:00 PM to Mar 29, 2017 2:00 PM

- ifremer research vessels
  - THALASSA
    - integrated navigation system
      - 20160701-000000-shipnav-TL_CINNA
      - 20160708-000000-shipnav-TL_CINNA
      - 20160711-000000-shipnav-TL_CINNA
      - 20160710-000000-shipnav-TL_CINNA
      - 20160709-000000-shipnav-TL_CINNA
      - 20160712-000000-shipnav-TL_CINNA
      - 20160701-080803-shipnav-TL_CINNA
      - 20160703-000000-shipnav-TL_CINNA
      - 20160702-000000-shipnav-TL_CINNA
      - 20160707-000000-shipnav-TL_CINNA
      - 20160706-000000-shipnav-TL_CINNA
      - 20160705-000000-shipnav-TL_CINNA
      - 20160704-000000-shipnav-TL_CINNA
      - 20160713-000000-shipnav-TL_CINNA
      - 20160715-000000-shipnav-TL_CINNA
      - 20160716-000000-shipnav-TL_CINNA
      - 20160717-000000-shipnav-TL_CINNA
      - 20160714-000000-shipnav-TL_CINNA
      - 20160718-000000-shipnav-TL_CINNA

Figure 9: Thalassa research vessel navigation submitted in real time (every 5-minutes) through the owncloud client, as seen in the sensor nanny web portal

In the sensor nanny system, the resources (system or observation description and data) submitted by a user on owncloud must comply with rules regarding:

- file naming
- format convention (based on OGC/Sensor Web Enablement standards)

Using these standards, the resources are considered by the system as deployments, observations and datasets. The convention especially enables to structure a hierarchy of deployment, from network to sensors (see screenshot).

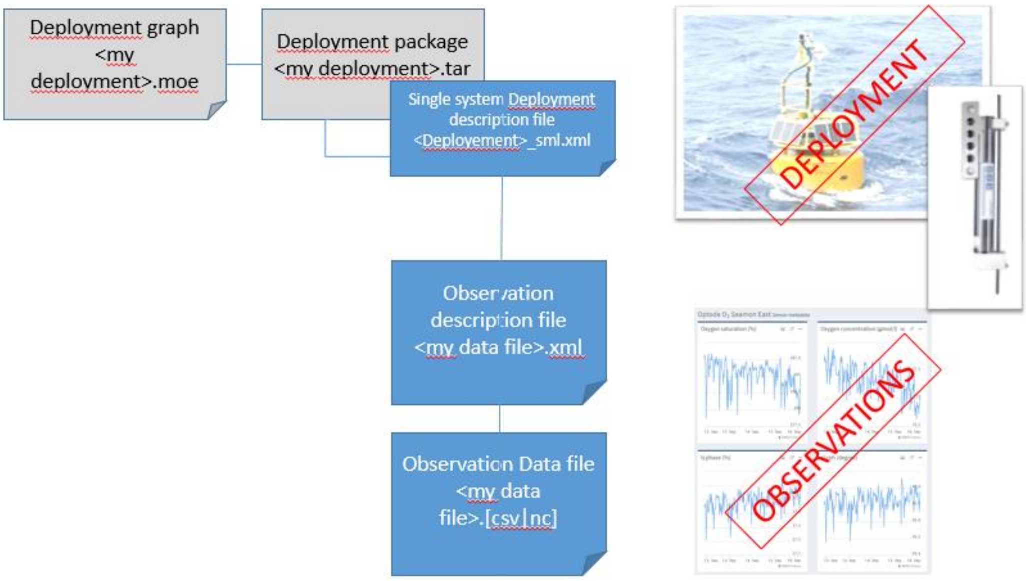
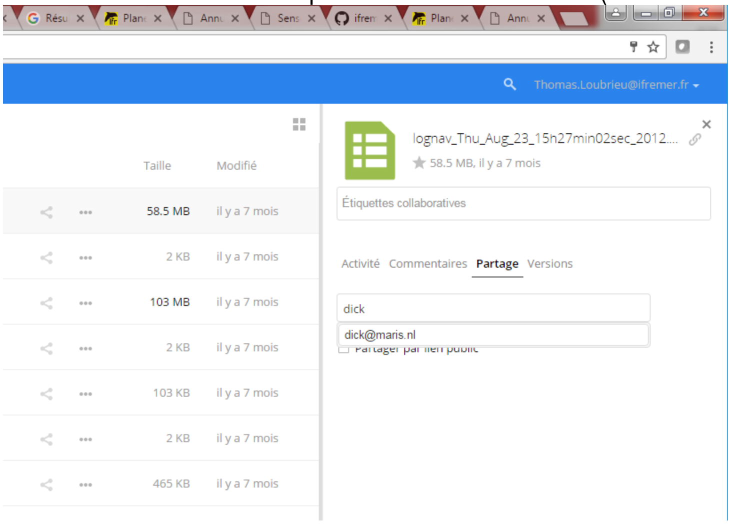

Deployment graph <my deployment>.moe
Deployment package <my deployment>.tar
Single system Deployment description file <Deployement>_sml.xml
Observation description file <my data file>.xml
Observation Data file <my data file>.[csv|nc]
DEPLOYMENT
OBSERVATIONS

Figure 10 : file convention for a resource indexation

The system provides a “share” function which allows file owners to share their resources publicly or with identified users. This feature will be used and extended to implement the use case UC4 (share information).



Taille	Modifié
58.5 MB	il y a 7 mois
2 KB	il y a 7 mois
103 MB	il y a 7 mois
2 KB	il y a 7 mois
103 KB	il y a 7 mois
2 KB	il y a 7 mois
465 KB	il y a 7 mois

lognav_Thu_Aug_23_15h27min02sec_2012....

58.5 MB, il y a 7 mois

dick

dick@maris.nl

□ Partager par lien public

Figure 11: share resources

In addition, owncloud allows hosting plugins dedicated to specific file management directly in the web browser (e.g. display a pdf, play a video …). This feature has been especially chosen to embed the deployment editor in owncloud: as in windows desktop, a user can double-click on a resource (a file) and automatically open the appropriate software to edit the resource (for example Microsoft word for a .docx document), in owncloud, the user can open a resource and automatically open the appropriate plugin to edit the resources in the web browser (for example GIS system for a GPS navigation file, the deployment editor for an deployment description file…)
The “File sync and share server” component is divided into the following sub-components:

- **Owncloud-web-php:** a web application developed in php which provides the human interface and webdav machine to machine interface.
- **Owncloud-MySQL-database:** MySQL schema for file metadata. It can be extended by the plugins to manage extra information (e.g. deployment or observation metadata).
- **Owncloud-sync-client:** enables file synchronization from the local directories of the user or from the system of the observatory if the system is connected and necessary bandwidth is available.
- **Snanny-owncloud-api:** dedicated plugin which enables to easily access the deployment and observation metadata in sensorML and O&M OGC standard formats (for a machine to machine interface). This interface is especially provided for interoperability purpose with other data services (Copernicus, EMODNET) and long term archive infrastructure (SeaDataNet). The plugin also triggers generation of O&M metadata when the user submits a data file (for example a CSV file).

For owncloud resources see https://owncloud.org/
For snanny-owncloud-api see on GitHub https://github.com/ifremer/snanny-owncloudapi
For sensorML and O&M OGC’s Sensor Web Enablement standard see http://www.opengeospatial.org/ogc/markets-technologies/swe. The implementation of these standards in sensor nanny is also coordinated in the marine community through the marine profile see http://adsabs.harvard.edu/abs/2016EGUGA..1814690J

3.3.2.2. CM2: deployment editor

The observatory editor is implemented as an owncloud plugin. This plugin is activated whenever a user wants to create, view or update a file with extension ".moe" (for marine observatory editor format).

The .moe file is a specific json file containing metadata on the observatory and on its graphic representation.

The user can edit an observatory description by drag-and-drop of predefined hardware or instruments. The user can link them together to describe complex deployments. For each component a set of attributes (name, description, identifiers, properties, events, contacts, location …) can be described. One component can be exported to be re-used in a different description.

When the deployment description is saved in the editor it is exported in a tar archive containing standard sensorML XML files (one for each instrument or hardware component of the deployment).

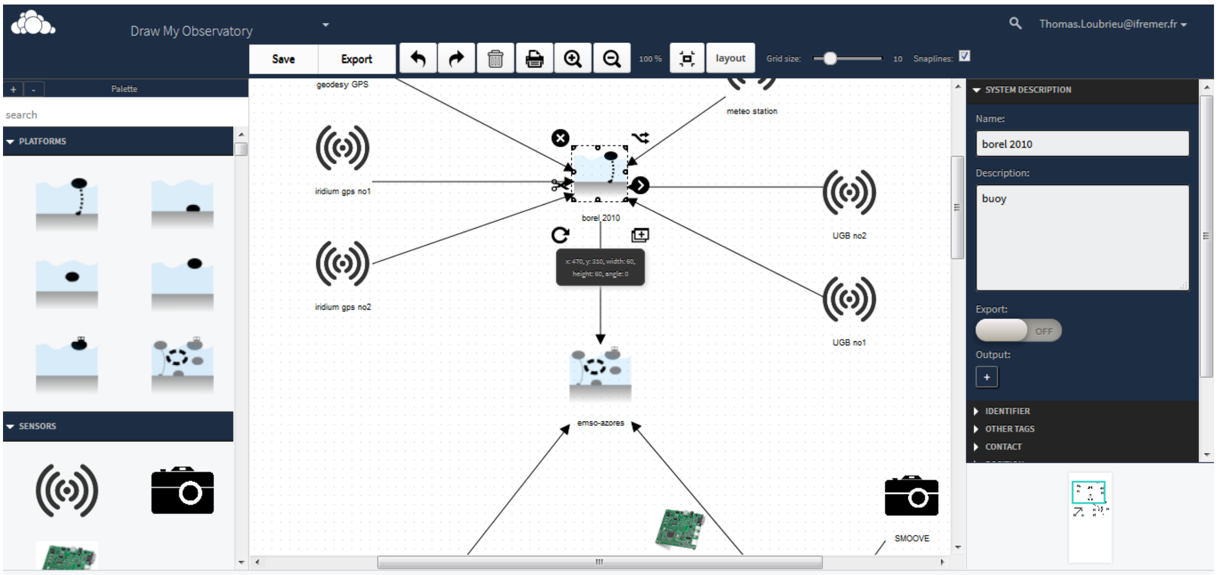
The pre-defined components available in the palette are generated from the equipment database known as Fixo3 yellow pages (http://www.esonetyellowpages.com/). The information is imported in the editor through sensorML and Json format using the snanny-yellowpages components.

The "deployment editor" component is thus divided into the following subcomponents:

- **Snanny-drawmyobservatory**: the graphic editor for observatory
- **Snanny-yellowpage**: the gateway to feed the palette with records from the Fixo3 yellow page.

For snanny-drawmyobservatory see https://github.com/ifremer/snanny-drawmyobservatory

For snanny-yellowpages see https://github.com/ifremer/snanny-yellowpages



Name: borel 2010

Description: buoy

Export: OFF

Output: +

Figure 12: snanny-drawmyobservatory web interface

The interoperability of the description provided with already deployed solution for data management of the targeted observations will be considered. Especially the HF Radar facility developed and operated by SOCIB[1] will be considered.

[1] http://www.socib.es/?seccion=observingFacilities&facility=radar http://www.socib.es/?seccion=observingFacilities&facility=radarInfo

3.3.2.3. CM3: indexation engine

The indexation engine parses the resources submitted to owncloud. It pushes the relevant information on deployments and datasets (hierarchy, spatio-temporal coordinates, keywords and sharing properties) into a high performance and scalable no-SQL database (elasticSearch).
The current system is implemented by a cluster of 3 nodes which has been successfully tested with up to 2.5 million 4D-points index (time, x, y , z, keywords and ancestors components).
The “indexation engine” component is divided into the following subcomponents:

- **elasticSearch**: the no-SQL database in which indexes are stored
- **snanny-elastic-sync**: the batch process which runs every 5 minutes to import or update records from owncloud system to elasticSearch.

For elasticSearch see https://www.elastic.co
For snanny-elastic-sync see https://github.com/ifremer/snanny-elastic-sync

3.3.2.4. CM4: Discovery, visualization and analysis portal

The discovery, visualization and analysis portal provides fluent and user-friendly access to the resources submitted to the system. It is designed to allow a synthetic view on numerous observations by providing:

- Spatio-temporal and keyword search criteria
- Synthetic observation density map and timeline (see screenshot)
- Equipment and observation tree (see screenshot).

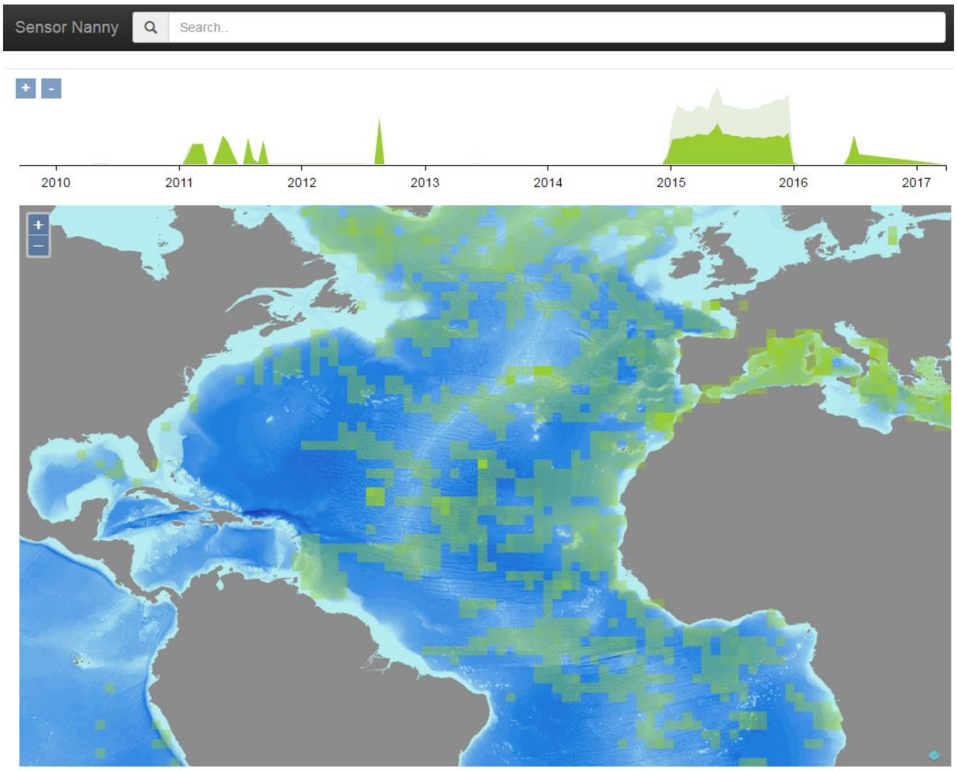
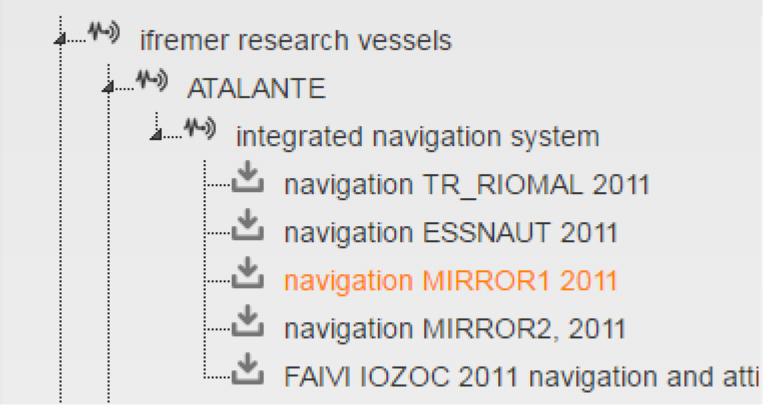


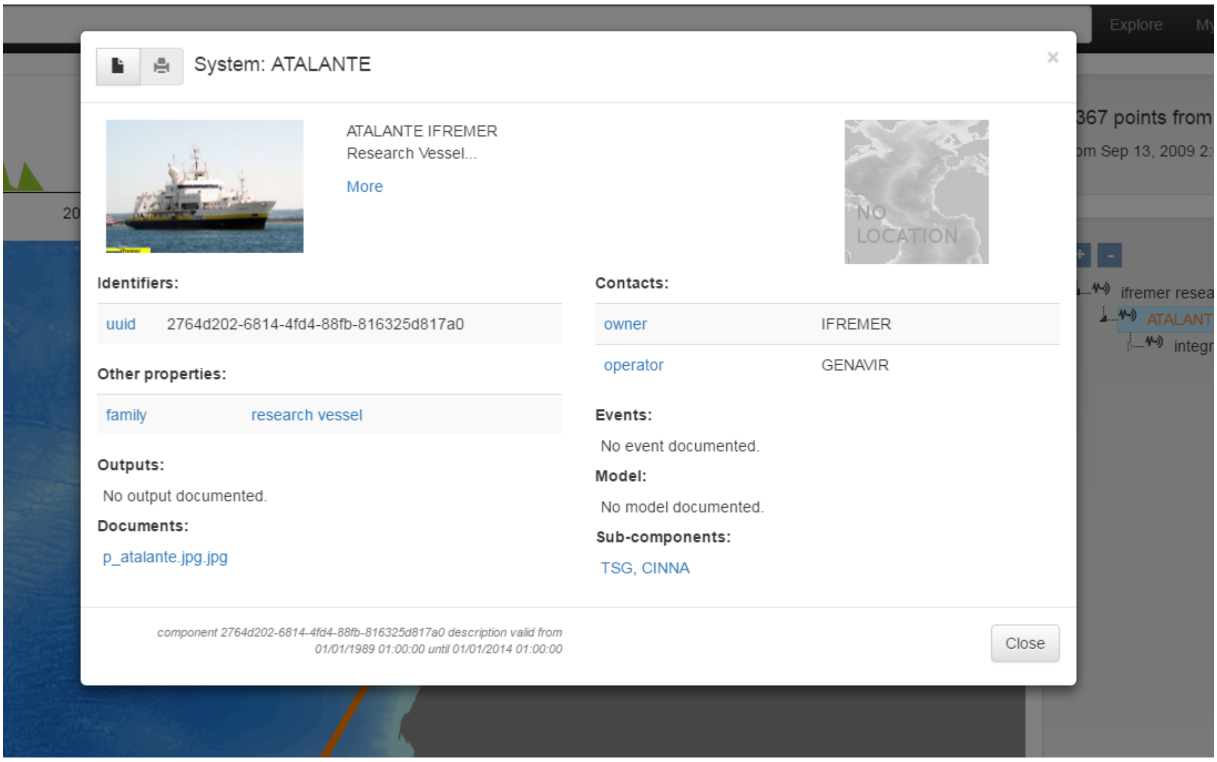
Figure 13: observation density timeline (on top) and map



ifremer research vessels
- ATALANTE
  - integrated navigation system
    - navigation TR_RIOMAL 2011
    - navigation ESSNAUT 2011
    - navigation MIRROR1 2011
    - navigation MIRROR2, 2011
    - FAIVI IOZOC 2011 navigation and atti

Figure 14 : deployment and dataset tree

The equipment can be displayed in a pop-up which uses the sensorML description managed in owncloud as a back-end (see screenshot).



System: ATALANTE

ATALANTE IFREMER
Research Vessel...
More

Identifiers:

uuid 2764d202-6814-4fd4-88fb-816325d817a0

Other properties:

family research vessel

Outputs:

No output documented.

Documents:

p_atalante.jpg.jpg

Contacts:

owner IFREMER

operator GENAVIR

Events:

No event documented.

Model:

No model documented.

Sub-components:

TSG, CINNA

component 2764d202-6814-4fd4-88fb-816325d817a0 description valid from 01/01/1989 01:00:00 until 01/01/2014 01:00:00

Figure 15: equipment description pop up

Datasets are displayed using plugins that are triggered depending on the dataset file type (CSV time series, video, 3 or 4D NetCDF coverage, image, CVV biology sample). The analysis tools (for example quality assessment tools) can be implemented as plugins also.

Depending on the type of datasets, the relevant expertise in Europe will be considered and as much as possible re-used. For HF radar, collaboration with the EuroGOOS HF Radar task team with for example SOCIB inputs on data processing[2] will be considered. For visualization, standard components will be integrated, for example THREDDS Data Server and ncWMS which are commonly used now to enable HF radar visualization (SOCIB, ETT, IFREMER, SHOM)..

[2] http://www.socib.es/?seccion=observingFacilities&facility=radar#reports

The authenticated user sees in the interface 2 categories of the deployments and observations:

- Those he/she submitted (which he/she owns)
- and those which have been shared with him/her by their owner

The “Discovery, visualization and analysis portal” component is divided into the following sub-components:

- **Snanny-api**: restFul service back-end requesting the elasticSearch database to retrieve use request results
- **Snanny-portal** and plugins: web application presenting the resources to answer the observation user and observation coordinator in a user-friendly layout

For snanny-api see https://github.com/ifremer/snanny-api
For snanny-portal see https://github.com/ifremer/snanny-portal

3.3.2.5. Implementation plan

While prototyping the system and proof-testing it against EMSO-Azores deep-sea deployment and others, a significant number of functions has been implemented and stabilized (in green). However a number of developments still need to be done to fulfil the JERICO-NEXT objectives (in red).

	CM1: File sync and share server and plugins	CM1-2: Observatory editor	CM3: Indexation engine	CM4: Discovery, visualization and analysis portal
UC1: Describe an equipment	□			
Core	🟩 X	🟥 X		
UC1-1: fixed platform		🟩 X		
UC1-2: Site geometry		🟥 X		
UC1-3: Data post-processing		🟥 X		
UC2: Submit a dataset				
Core	🟩 X		🟩 X	
UC2-1: CSV time series	🟩 X		🟩 X	
UC2-2: videos	🟥 X		🟥 X	
UC2-3: 3-4D netcdf	🟥 X		🟥 X	
UC2-4: images	🟥 X		🟥 X	
UC2-5: CSV biology sampling	🟥 X		🟥 X	
UC3: Discover, Visualize and Analyze				
Core				🟩 X
UC3-1: CSV time series				🟥 X
UC3-2: videos				🟥 X
UC3-3: 3-4D netcdf				🟥 X
UC2-4: images				🟥 X
UC2-5: CSV biology sampling				🟥 X
UC4: Share information				
individuals	🟩 X		🟩 X	🟩 X
functions (long term archive, data service)	🟥 X		🟥 X	🟥 X

The development will be done sequentially as follow:

- Support fixed-platform (milestone 1):
  - UC1-Core, support BODC vocabularies in the observatory editor
  - UC2-2, UC3-2 (optional): support videos
- Support for HF radar (milestone 2):
  - UC1-2: site geometry in observatory editor
  - UC1-3: post-processing description in observatory editor
  - UC2-3, UC3-3: submit and visualize 3-4D NetCDF files
- Support for flow cytometry (milestone 3):
  - UC2-4, UC3-4: submit and visualize images
  - UC2-5, UC3-5: submit and visualize biology sample

The planned schedule is:

- **Milestone 1:** May 2017, tests with EMSO-Molène and Poseidon/Athos
- **Milestone 2:** June 2017, tests to be defined (with JRA tasks)
- **Milestone 3:** December 2017, complementary test to be defined (with JRA-tasks)

4. Conclusions

The documents design a flexible tool which can fulfil platform or network operator requirements. The chosen architecture has been proof tested with deep sea observatory and research vessels platforms for deployment documentation and data synchronization and sharing.

The proposed architecture based on open-source solutions (for example owncloud), open standards (OGC/Sensor Web Enablement) and the flexibility and modularity of it allows to adapt the solution to the targeted networks, fixed platforms, HF radar and flow cytometry. Moreover it allows thinking of wider adoption thanks to collaboration with other projects and initiatives, for example for Ferry Boxes.

The plan for development, deployment and testing is defined in the document and will be applied in task 5.3, Work Package 5 of project JERCIO-NEXT.

5. Annexes and references

Sensor Web Enablement: http://www.opengeospatial.org/ogc/markets-technologies/swe
Sensor Nanny source code on github: https://github.com/ifremer/
Sensor Nanny portal: http://snanny.ifremer.fr/
elasticSearch: https://www.elastic.co
owncloud: https://owncloud.org/
fixo3 yellow pages: https://www.esonetyellowpages.com/