



Project of Strategic Interest NEXTDATA

Deliverable D2.2.3

Opening of the archives on essential climatic variables Trasmission of data to General Portal

WP Coordinator: Claudia Fratianni
INGV, Bologna

Authors
C. Fratianni, N. Pinardi
INGV, Bologna

Contents

1. Introduction.....	3
2. Definition of RR products and formats	3
3. Set up and opening of the Thredds server	5
4. Definition of user requirements for Nextdata-INGV Geoportal	6
Access to the Portal.....	6
Discover the Catalogue	7
Discover the data.....	7
Visualize data.....	7
5. Software selection for the Geoportal presentation, application and data Layers.....	8
Data Layer	8
Application Layer	9
Catalog and Search service.....	9
View service.....	10
Presentation Layer	11

1. Introduction

The activities performed during the second year were mostly dedicated to the partial implementation and design of the NextData-INGV Geoportal.

The NextData-INGV Geoportal will be a web application to discover and visualize the Reconstruction-Reanalyses (hereafter RR) data for the Mediterranean Sea for the past sixty years from 1953 up to 2012.

Activities during the second year of the project concerned:

- Definition of RR products and formats;
- Set up and opening of the thredds server;
- Definition of user requirements for NextData-INGV Geoportal;
- Software selection for the Geoportal presentation, application and data Layers.

The following sections will describe in details each activity.

2. Definition of RR products and formats

During the second year of the project, INGV completed the Reanalyses of the Mediterranean Sea starting from 1987 up to 2012, partially supported by NextData Project (see D1.3.3), and started to produce the RR since 1953.

At the moment there are two datasets:

- MedReanV4bis1987-2012;
- RR 1955-1972.

Products available for both datasets are listed in Tab.1 and Tab. 2.

Variables	Geographical coverage	Spatial Resolution	Number of levels	Temporal resolution	Temporal coverage
Temperature	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	Monthly average field	1987-2012
Salinity	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	Monthly average field	1987-2012
Sea Surface Height	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	Monthly average field	1987-2012
Horizontal velocity (meridional and zonal component)	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	Monthly average field	1987-2012
Surface heat, water and momentum fluxes	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	Monthly average field	1987-2012
Surface S, T and currents	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	Monthly average field	1987-2012

Tab.1. 1985-2012 Reanalysis products released to the project.

Variables	Geographical coverage	Spatial Resolution	Number of levels	Temporal resolution	Temporal coverage
Temperature	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	24hr average field	1955-1972
Salinity	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	24hr average field	1955-1972
Sea Surface Height	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	24hr average field	1955-1972
Horizontal velocity (meridional and zonal component)	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	33 (IODE)	24hr average field	1955-1972
Surface heat, water and momentum fluxes	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	6hr	1955-1972
Surface S, T and currents	6° W - 36.25° E 30.19° N - 45.94° N	0.0625°	1	6hr	1955-1972

Tab. 2. RR products released during the second year of the project.

Both datasets were delivered in netCDF format adopting international standards concerning in particular standard_name and CF convention, as shown in Tab. 3.

The netCDF data format is extremely flexible, self describing and has been adopted as a de-facto standard for many operational oceanography systems. The purpose of the CF conventions instead is to require conforming datasets to contain sufficient metadata that are self-describing, that is, each variable in the file has an associated description of what it represents, including physical units if appropriate, and that each value can be located in space (relative to earth-based coordinates) and time (NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.3, 2007).

Variables	Standard_name	Format
Temperature	sea_water_potential_temperature	netCDF CF1.0
Salinity	sea_water_salinity	netCDF CF1.0
Sea Surface Height	sea_surface_height	netCDF CF1.0
Horizontal velocity (meridional and zonal component)	northward_sea_water_velocity eastward_sea_water_velocity	netCDF CF1.0
Surface heat, water and momentum fluxes	<ul style="list-style-type: none"> surface_downward_eastward_stress surface_downward_northward_stress precipitation_flux water_flux_into_sea_water_from_rivers water_evaporation_flux 	netCDF CF1.0

	<ul style="list-style-type: none"> • surface_downward_latent_heat_flux • surface_downward_sensible_heat_flux • surface_net_downward_shortwave_flux • surface_net_downward_longwave_flux 	
Surface S, T and currents	<ul style="list-style-type: none"> • sea_water_potential_temperature • sea_water_salinity • northward_sea_water_velocity • eastward_sea_water_velocity 	netCDF CF1.0

Tab. 3. Product format specifications.

3. Set up and opening of the Thredds server

At the moment the products of Table 1 and 2 have been made available through an INGV dedicated server using THREDDS protocol (<http://medsearr.bo.ingv.it/thredds>).



Fig. 1. Mediterranean Sea Reanalysis/Reconstruction Thredds Server.

The Thredds Data Server catalog describes the inventory of the datasets, organized as a simple hierarchical structure with a commonly understandable name for each dataset (see Fig. 2):

- **tem:** for the variable of temperature;
- **sal:** for the variable of salinity;
- **ssh:** for the variable of sea surface height;
- **cur:** for the variable of velocities;
- **flx:** for the variable of surface fluxes.

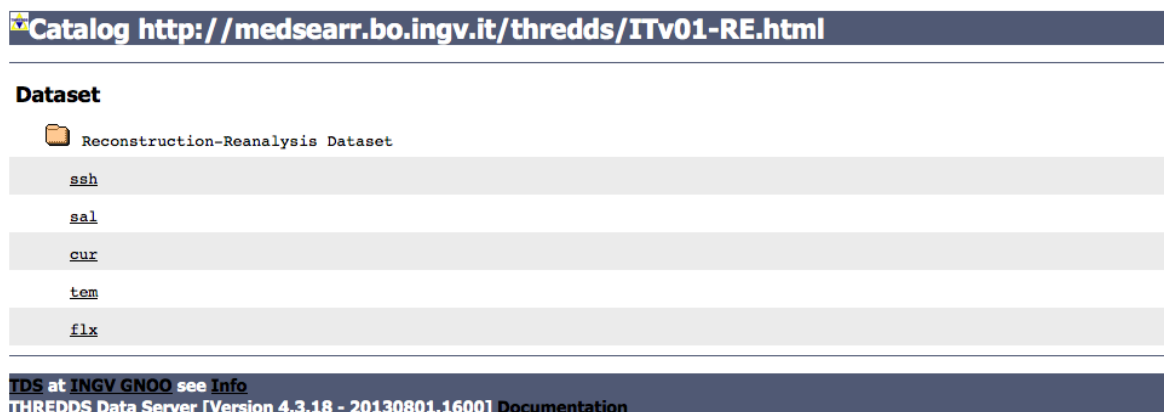


Fig. 2. Mediterranean Sea Reanalysis/Reconstruction Thredds Server structure.

4. Definition of user requirements for Nextdata-INGV Geoportal

The main actors of the NextData-INGV Geoportal are the scientific community at large, interested in ocean gridded data and applications.

On the basis of this user scenario, the Geoportal functionalities will be:

- a portal access functionality: the user access the NextData-INGV portal and selects a language (English or Italian). The portal home page gives a brief description of the INGV activities within NextData and the available products.
- The catalogue functionality: a catalogue will document the available data, method of production and quality indices.
- The discovery functionality: upon the choice of a variable/period/depth/area/ the Geoportal replies with a list of results (metadata) matching the defined criteria. Each record in the list is selectable for the data visualization procedures.
- The visualization functionality: the user will select options for the visualization of the selected RR data.

The requirement investigation was focused on the catalogue, discovery and visualization functionalities, and it is described in the following Tables.

Access to the Portal

Requirement Description	The NextData-INGV web portal will provide a web home page allowing the user to visualize information about the NextData Project and the activities of the INGV within the project. The web portal shall provide links to search, catalogue and visualization services. The web portal includes a language selector (Ita-Eng).
Justification	It is necessary for the user to understand the framework of the web portal and to access all the available services in a friendly way.
Priority	High

Discover the Catalogue

Requirement Description	The NextData-INGV web portal will provide a Catalogue web page tool to visualize available products. The catalogue provides a description of the available products. Each product will list a short description, an example image, a link to visualize metadata, a link to visualization tools.
Justification	It is important for the user to be able to access the whole list of products data, to visualize the associated metadata and to be able to select the data that users considers the most suitable for his/her own analysis.
Priority	High

Discover the data

Requirement Description	The NextData-INGV web portal will provide a search tool to discover available data. The search tool includes the following parameters to the users to easy the search: free search, geographical parameters (lat-lon boundary), time parameter (from-to), a map to identify the bounding of the domain. The result of each query is a list of data/product available with links to metadata and to access the visualization tool.
Justification	It is necessary for the user to get search tool to discover all the available data in a friendly way.
Priority	Medium

Visualize data

Requirement Description	The NextData-INGV web portal will provide a visualization tool to plot available data. The tool includes the following selectors to the users: variables, geographical selection (lat-lon), time selector (from-to), depth (if available for the selected variable), a map to identify the bounding of the domain. The result of each query is a 2d plot of the selected data with geographical references.
Justification	It is necessary for the user to get visualization tool to plot all the available data in a friendly way.
Priority	Medium

The infrastructure chosen to host the service is an INGV Linux-server infrastructure, providing the data archive and the web tool. The virtual environment is based on Ganeti, a cluster virtual server management software tool built on top of existing virtualization technologies such as KVM. It has been designed to facilitate cluster management of virtual servers and to provide fast and simple recovery after physical failures. Tab. 4 shows the server technical specifications.

Virtual CPU	1
OS	7.0
RAM	4Gb
Disk space	100 Gb + 1Tb (expandable)

Tab. 4: Virtual machine for Geoportal technical specifications.

5. Software selection for the Geoportal presentation, application and data layers

The selection of the software for the development of different Geoportal layers is at the basis of the development and aims to create a robust architecture (see **Error! Reference source not found.**).

The selection has been done following the user requirements and taking into account the need of using open source software or software delivered under GPL Licence. Some changes in the selection can be envisaged depending on the test results that will be run during the portal development.

Data Layer

All data is available in NetCDF format and are provided from THREDDS

(<https://www.unidata.ucar.edu/software/thredds/current/tds/>).

The THREDDS (Thematic Realtime Environmental Distributed Data Services) Data Server (TDS) is a package designed to make it as easy as possible for those who create scientific datasets to make them available to those who use them. The goal is to make datasets in many different formats and located in many different geographic locations available to users in a way that hides the data format and location information and presents only the data essential to the datasets themselves.

The THREDDS Data Server (TDS) combines catalog services with integrated data-serving capabilities, including OPeNDAP, HTTP file serving, and OGC Web Coverage Server (WCS). The OPeNDAP and WCS data-serving capabilities are built on the netCDF-Java library, which combines the netCDF-3, OPeNDAP-2, and HDF5 data models, into what is called the netCDF-4 Common Data Model. The THREDDS Data Server is 100% Java, open source, and runs as a Tomcat web server application.

The TDS has an integrated WMS server that allows the chosen data to be accessed as images via the OGC Web Map Service protocol for gridded datasets (WMS 1.3.0).

The THREDDS catalogue service is not totally INSPIRE complaint thus this THREDDS catalogue service cannot be used for Search and Catalogue Services.

Application Layer

Catalog and Search service

For this part, the choice may fall on GeoNetwork. GeoNetwork opensource is a catalog application to manage spatially referenced resources. It provides powerful metadata editing and search functions as well as an embedded interactive web map viewer.

GeoNetwork opensource has been developed to connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols (a.o. from ISO/TC211 and OGC).

Geonetwork main feature that can be used in NextData Geoportal:

- Immediate search access to local and distributed geospatial catalogues.
- Up- and downloading of data, graphics, documents, pdf files and any other content type.
- An interactive Web Map Viewer to combine Web Map Services from distributed servers around the world.
- A randomly selected Featured Map.
- Online editing of metadata with a powerful template system.
- Native support for ISO19115/ISO19119/ISO19139, FGDC and Dublin Core formatted metadata.
- Scheduled harvesting and synchronization of metadata between distributed catalogues.
- Metadata INSPIRE Compliant.

Geonetwork can be used with INSPIRE complain metadata. This is an additional feature of the application.

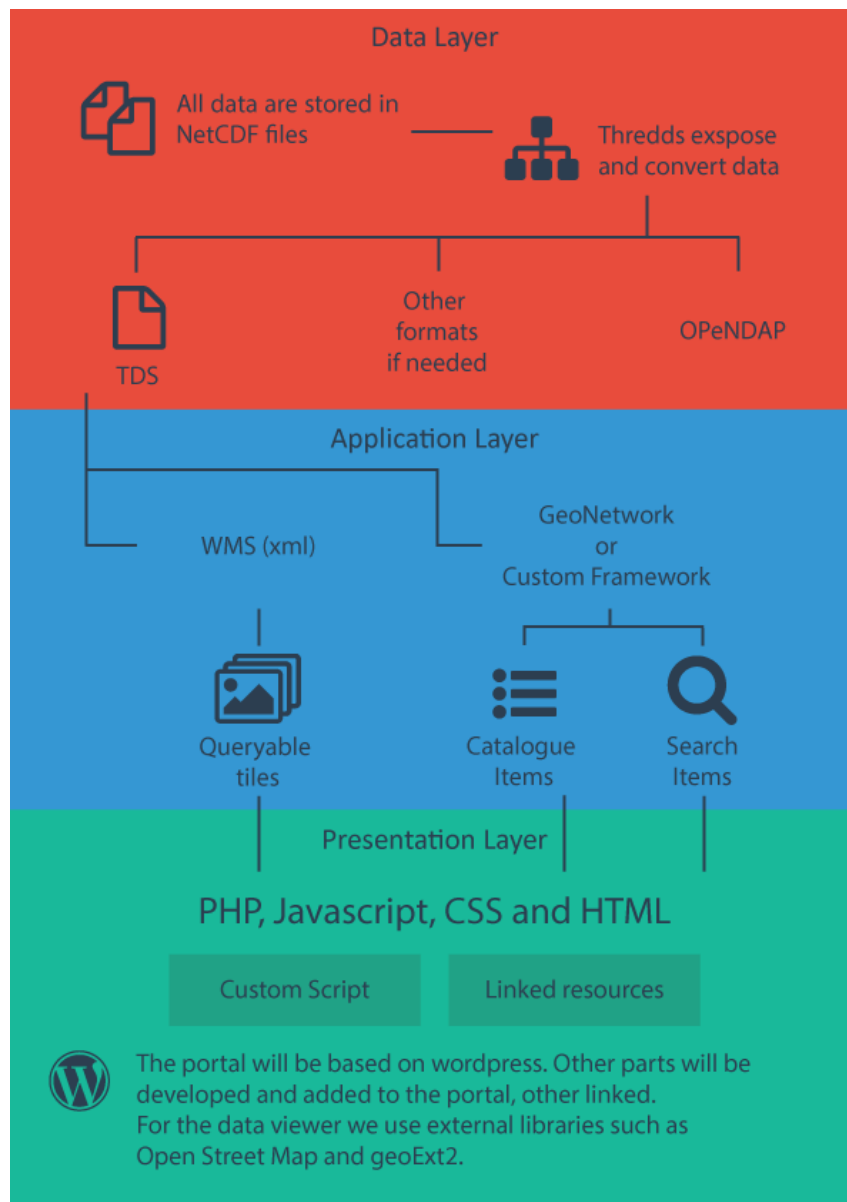


Fig. 3. Graphical representation of work flow in the portal.

View service

The View Service will be developed using OpenLayers and WMS (Web Mapping Server).

OpenLayers is a JavaScript web map library that will be used to create, manage and display web based maps. This unit will provide examples of OpenLayers codes that can be used to produce your own simple web maps using various OGC standard services.

OpenLayers was selected as the web mapping API the because of its flexibility. OpenLayers separates the API from the underlying base layer data. This allows the programmer to select base layer data from a number of sources without having to convert the application another mapping API. OpenLayers also supports a wide arrangement of Open Geospatial Consortium (OGC) formats. <http://openlayers.org/>

A WMS server (embedded in THREDDS) will be used to visualize data. The WMS server only returns the data which are in the view box of the client application instead of the entire dataset. Data are also returned as a transparent image, which easily be layered. These advantages made WMS the content delivery method of choice.

Presentation Layer

The entire presentation layer will be developed using technologies such as: PHP, Javascript, CSS. One of the selected applications is GeoExt. This is Open Source and enables building desktop-like GIS applications through the web. It is a JavaScript framework that combines the GIS functionality of OpenLayers with the user interface of the ExtJS library.

The portal will be built using Wordpress (Content Management System, <http://wordpress.org/>), and will implement the necessary modules to the map view that contains the graphical representations of data.

The major advantages of using Wordpress are:

- Ease of Use: WordPress is very easy to use and has an intuitive interface. Adding new pages, blog posts, images, etc. on a regular basis is a breeze and can be done quickly.
- Manage Website from Any Computer: WordPress is browser-based. The administrator can login from any Internet connected computer and manage the site.
- No HTML Editing or FTP Software Required: WordPress is a self-contained system and does not require HTML editing software. The site administrator can create a new page or blog post, format text, upload images (and edit them), upload documents, video files, image galleries, etc. all without the need for additional HTML or FTP software.
- The Design of the portal is 100% Customizable: WordPress acts as the engine for the website. The look and feel of the site can be 100% customized.
- Extend the Functionality of Your Site with Plugins: WordPress makes this possible with plugins.

The major advantage is that Wordpress can be easily integrated with all the other application that will be developed for the NextData Portal. All the foreseen services, as well as the product catalogue, can be integrated into Wordpress assuring interoperability.