cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

Research Papers
Issue RP0195
December 2013

SCO – Scientific Computing and Operations Division

The ORIENTGATE data platform

By **Alessandra Nuzzo**
University of Salento and Scientific Computing and Operations Division, CMCC
alessandra.nuzzo@cmcc.it

Sandro Fiore
University of Salento and Scientific Computing and Operations Division, CMCC
sandro.fiore@unisalento.it

and **Giovanni Aloisio**
University of Salento and Scientific Computing and Operations Division, CMCC
giovanni.aloisio@unisalento.it

SUMMARY The ORIENTGATE project fosters concerted and coordinated climate adaptation actions across the SEE region by exploring climate risks faced by coastal, rural and urban communities; contributing to a better understanding of the impact of climate variability and change on water regimes, forests and agro-ecosystems; and analysing specific adaptation needs in the hydroelectricity, agro-alimentary and tourism sectors. The principal project results include six pilot studies of specific climate adaptation exercises, a data platform connected to the EU Clearinghouse on Climate Adaptation, capacity enhancing seminars and workshops, working partnership among the hydro-meteorological offices of the SEE countries. In particular, this document provide an overview on the activities carried out on the ORIENTGATE data platform, designed to store and manage data produced by the project partners.

Keywords: climate change adaptation, vulnerabilities and risks, impact indicators, data platform, metadata, climate datasets

INTRODUCTION

The management of large volumes of data, a secure and efficient access to environmental datasets, the definition and adoption of metadata schemas represent some of the most relevant challenges that must be faced up by scientists and researchers. A key architectural building block of an operational environment is represented by the data access layer, which has to provide an efficient, robust and secure access to the data stored on the storage devices. The data platform aims to offer a single entry point to access data produced during the project, both regarding climate simulations and impact indicators. The planned activities for the data platform are:

- *data platform design*: in the context of this project, the data platform, will provide access to a large volume of heterogeneous data. NetCDF, GRIB, CSV, ASCII Grid Data, JPG, GIF will be some of the main formats that will be managed by this platform. The design of this platform will take into consideration users needs and requirements;
- *setup of the virtual machine based environment*: a virtual machine based environment to host climate change datasets will represent the proper infrastructure to integrate several software like OPeNDAP, THREDDS, RAMADDA, Integrated Data Viewer, as well as ad hoc software developed in this task/project. The data platform will run on a virtual machine-based environment with a large amount of RAM, fast disk storage and processing power to efficiently manage the ORIENTGATE datasets;
- *data sub-setting functionality and data compression*: data sub-setting functionalities related to NetCDF data will be also provided to allow partial download of huge climate datasets. Datasets compression will be also taken into account to reduce disk-space allocation as well as the time for data delivery;
- *deployment of the main data services*: OPeNDAP/THREDDS and RAMADDA represent two middleware tools bridging the gap between data providers and data users and that will play a key role in the data platform. While the goal of the former is to simplify the discovery and use of scientific data and to allow scientific publications and educational materials to reference scientific data, the latter represents a comprehensive content repository, publishing platform and collaboration environment for the scientific domains. Adopting well-known and de-facto standard access interfaces, will make easier the integration and sharing of the data repository hosted by the ORIENTGATE data platform with other projects and institutions at an international level;
- *enhanced configurations for the data services*: to meet the efficiency requirement, different architectural solutions for the OPENDAP/THREDDS software will be evaluated and tested against the standard configuration with a single server to evaluate performance and load balancing issues;
- *design and implementation of a dashboard based monitoring and browsing tool for the data platform*: the data platform will need operational tools to check the status of the data services, as well as advanced data browsing tools providing integrated views about the available OPeNDAP/THREDDS services. A dashboard-

based system providing monitoring and multi-server browsing capabilities will be designed and developed to address both system management (through the service monitor) and user access needs (through the browsing system). From a monitoring point of view, the dashboard system will provide a data platform monitor with several charts and reports about the services availability and other relevant statistics concerning the data platform health. On the other hand, the browsing interface will allow to manage into the same context several THREDDS installations that will be deployed, configured and tuned during the project. Browsing capabilities across the THREDDS catalogues will be also provided.

THE DATA PLATFORM ARCHITECTURE

A key action in the ORIENTGATE project is the provision of a data platform which enables, at the project level, secure, efficient and transparent data sharing, access and management. The design of the data platform will take into account key requirements like:

- efficient, scalable and transparent access to large volumes of scientific data;
- metadata management.

With a data centric approach, the same data is available through different access interfaces providing different features. The data platform is a 'collection of services', running on multiple virtual machines, that are:

- 'general purpose' services, such as HTTP and FTP;
- 'domain oriented' services, like OPeNDAP/THREDDS, ESGF data node and GeoNetwork.

Figure 1 shows the overall data platform architecture.

The figure schematically shows the architecture of the data platform as consisting of a series of services, deployed on virtual machines. Schematically, HTTP, FTP, OPeNDAP and THREDDS provide support in terms of data access and download. GeoNetwork, on the other hand, allows filling out a metadata sheet for each dataset to enable data search and discovery as well as metadata browsing. Regarding the ESGF Data Node, it was not planned in the original proposal of the project, but we're evaluating the possibility to integrate it into the platform since it could be useful for managing climate data. The latest release of this software has been made available last week and we're testing the installation on a virtual machine of the project.

DATA ARCHIVE

A DRS (Data Reference Syntax) provides a common naming system to be used in files, directories and URLs to identify datasets wherever they might be located within a distributed archive. The DRS should provide a clear and structured set of conventions to facilitate the naming of data entities within the data archive and files delivered to users. The DRS should use controlled vocabularies to facilitate documentation and discovery. Providing users with data in files with well-structured names will facilitate management of the data on the users file systems and simplify communication among users and between users and user support. The controlled vocabularies will be useful in developing category-based data discovery services. In order to better organise the data on the data platform, specific directory structures have been defined both in terms of climate simulations and impact indicators. A hierarchical directory structure allows an easy management

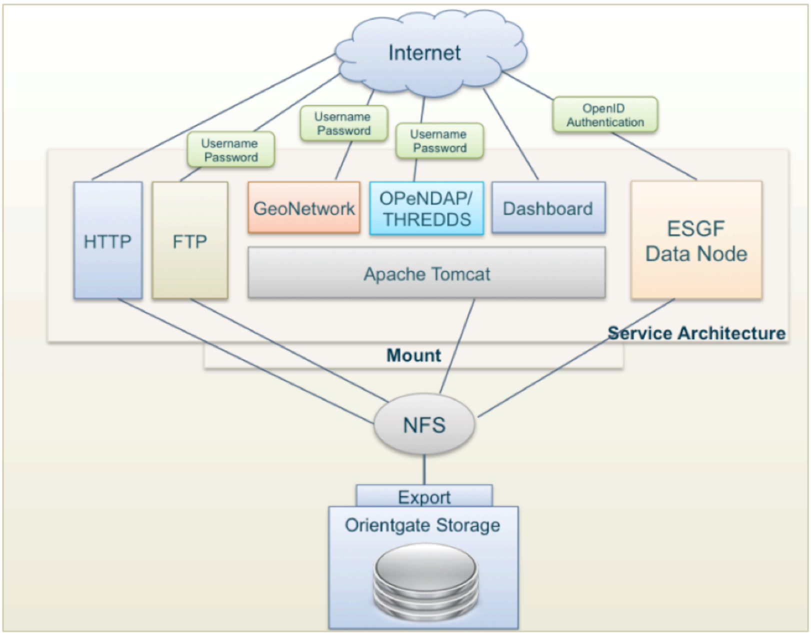


Figure 1:
ORIENTGATE Data Platform

and a simple organisation of the data, so the data regarding the output of simulations and impact indicators will be located according to the directory structure which starts from the root folder that is the project, *orientgate* and is split into two paths which will store the two different types of data.

DATA ACCESS SERVICE

The Data Access Service (DAS) provides several functionalities through the integration of different services. This service must provide the coexistence of various data services for data delivery, sub-setting and browsing. The role of the data is central, which implies that this component must necessarily be "data-centric"; the data has to be accessible from different points, although being stored in a single location on the data archive.

The following services have been installed on a virtual machine of the data platform:

- OPeNDAP;
- THREDDS;
- HTTP;
- FTP.

A brief overview of these services will be presented in the following.

OPENDAP

OPeNDAP (Figure 2) [1] is a framework that aims to simplify the sharing of scientific data on the web by making local data accessible to remote connections. It is a freely available software and allows the transformation of existing applications into OPeNDAP clients, enabling them to remote access. The advantage is that the user may not be aware of the different format in which different dataset are stored, but can equally access to the information. The uniformity with which the data are presented make the system very useful for the analysis, transfer and automatic manipulation of data. It is possible to choose the way to get the information by selecting:

- *Get ASCII*, which displays the required data in the browser window;
- *Binary (DAP) Object*, which locally stores the data in binary format;
- *Get as NetCDF*, which locally saves the file in NetCDF format.

THREDDS

THREDDS Data Server (TDS) [2] is a web server designed to provide access to scientific data and metadata using different remote access protocols. THREDDS uses Dataset Inventory Catalog in order to provide virtual directories of available data and metadata associated with them. Each catalog describes the datasets and services through which access to the data. The services offered by THREDDS provide access to the full dataset, carry out subsetting, graphically display the information, analyze the associated metadata and regenerate them according to specific standards. The TDS provides the user a series of services that represent different modes of access to data such as:

- *OGC Web Coverage Service (WCS)* that supports the retrieval of geospatial data representing phenomena with spatial and temporal variations;
- *OpenGIS Web Map Service Interface Standard (WMS)* that provides a simple HTTP interface for the request of maps from one or more distributed geospatial databases;
- *OPeNDAP* that provides access to portions of the dataset by extending the functionality of the HTTP protocol;
- *HTTP Server* that allows user to access the entire contents of the file.

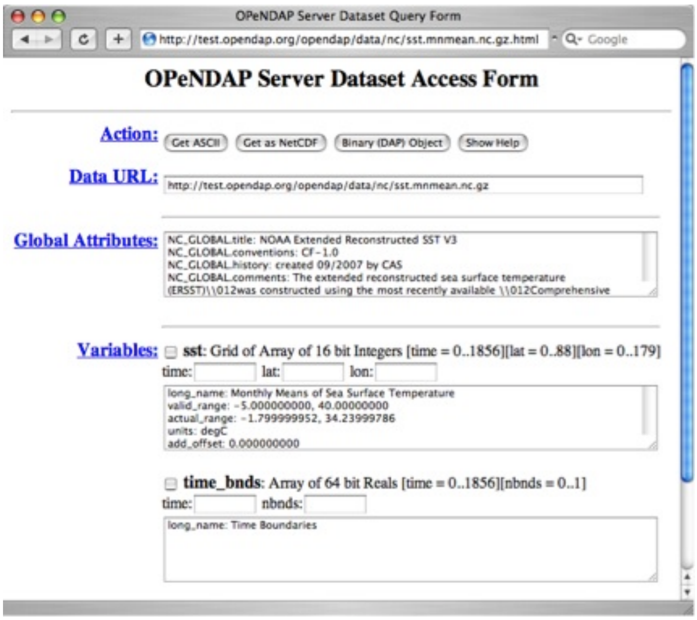
FTP

File Transfer Protocol (FTP) is an application layer protocol based on the TCP protocol for data transmission between hosts. The main objectives of FTP are:

- promote sharing of files (programs or data);
- encourage the use of indirect or implicit remote computers;
- solve in a transparent manner any incompatibilities between different file storage systems;
- transfer data in a reliably and efficiently way.

A FTP server provides several features that allow the client to interact with the remote file system:

- download / upload files;
- recovery of interrupted file transfers;
- removal of files;
- change of the file name;
- creation of directories;
- navigation between directories.



OPeNDAP Server Dataset Query Form

http://test.opendap.org/opendap/data/nc/sst.mnmean.nc.gz.html

OPeNDAP Server Dataset Access Form

Action:

Data URL: http://test.opendap.org/opendap/data/nc/sst.mnmean.nc.gz

Global Attributes:

```
NC_GLOBAL.title: NOAA Extended Reconstructed SST V3
NC_GLOBAL.conventions: CF-1.0
NC_GLOBAL.history: created 09/2007 by CAS
NC_GLOBAL.comments: The extended reconstructed sea surface temperature
(ERSST)\\012was constructed using the most recently available \\012Comprehensive
```

Variables: □ **sst**: Grid of Array of 16 bit Integers [time = 0..1856][lat = 0..88][lon = 0..179]

time: lat: lon:

```
long_name: Monthly Means of Sea Surface Temperature
valid_range: -5.000000000, 40.00000000
actual_range: -1.799999952, 34.23999786
units: degC
add_offset: 0.000000000
```

□ **time_bnds**: Array of 64 bit Reals [time = 0..1856][nbnds = 0..1]

time: nbnds:

```
long_name: Time Boundaries
```

Figure 2:
OPeNDAP Dataset Access Form

HTTP

HyperText Transfer Protocol (HTTP) is the primary system for the transmission of information in Internet. The HTTP protocol is based on TCP protocol and adopts the client-server architecture. The client sends a request message to the server for a certain resource (such as HTML files or web pages objects) or for the execution of an operation and the server returns a response message which containing information specifying whether the request was fulfilled or possible error codes. It may also include the requested contents. HTTP provides a simple authentication service based on sending the login credentials (username and password). A more advanced security service is offered by HTTPS protocol, which integrates the HTTP protocol with an encryption system like Transport Layer Security (SSL/TLS). The service includes encryption of messages exchanged between client and server and server authentication based on digital certificates.

METADATA SERVICE

Metadata are information about the data and describe the content, the quality and features of the data like origin, scope, point of contact, credits, etc. Metadata helps to search and discovery data, to better organise and use data and finally enables interoperability. The metadata service is the heart of the information system for the data stored on the data archive of the project. It manages the metadata that are used to carry out search and discovery operations on the data archive. Of course, the adoption of standards will allow the interoperability and the possibility to interact with other projects and share and compare results. To manage metadata in this project we were inspired by the international standard ISO 19115.

ISO 19115

This standard defines the schema required for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data. ISO 19115 standard [3] is part of the standards produced by ISO TC 211 Geographic Information/Geomatics and defines a common set of metadata describing information relating to the identification, the extent, the quality, the temporal and spatial pattern and the distribution of geographic data. This international standard was then developed in order to classify and describe geographic datasets, facilitate the creation and management of metadata related to geospatial data, facilitate discovery, access, evaluation and use of geographic data. An important characteristic is the possibility to apply this standard to a single dataset, to aggregations of single datasets at to single geographic features. ISO 19115 is an abstract standard because it does not define "how to" fill in metadata, but "what" a metadata should contain. In particular, the standard establishes a common core metadata that is a minimum set of elements required to describe the "geographic dataset"; this minimum set includes mandatory elements (Mandatory, M), conditional elements (mandatory under certain conditions - Conditional; C) and optional elements (optional, O), as shown in the table in Figure 3.

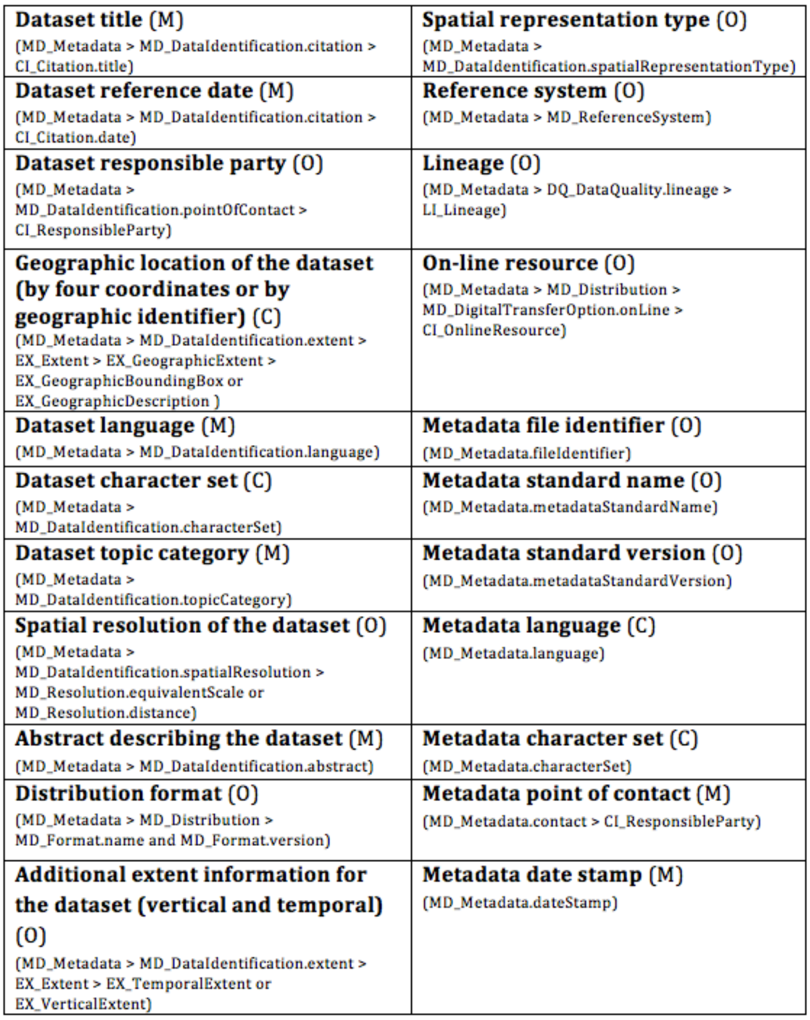
The minimum set of metadata allows the search, access, transfer and exchange of geographic data; in addition to this, the international standard also provides some extensions, in order to allow a more extensive description of geographic data and to satisfy more specialist needs. The main features that differentiate ISO 19115 from other metadata standard used for geographic data are as follows:

- adoption of UML as a standard language for modelling;
- adoption of the XML standard format for creation and management of metadata;
- ability to create profiles;
- definition of a minimum common set (core metadata);
- multi-level metadata;
- ability to integrate with other ISO standards.

The standard ISO 19115 organises the list of available metadata into several "UML packages." Each package contains one or more entities (or UML classes) that can be divided into subclasses or aggregated in superclasses. Entities contain elements identifying the unit of metadata; each entity may have different logical relationships with other entities. The following Figure 3 shows the organisation of the packages.

ISO 19139

ISO 19139 is a technical specification that defines the implementation schema of ISO 19115 according the "eXtensible Markup Language" (XML) format and can be used to describe, validate and exchange geospatial metadata. The standard ISO 19115, in fact, did not provide the specific information to generate metadata "XML compliant", that were standard not only in content but also in its structure. The implementation standard of ISO 19139 has, therefore, developed schemas in XML format in order to improve interoperability, by providing a common specification for describing, validating and exchanging metadata about geographic datasets, individual dataset, individual geographic features, attributes, types of characteristic function, etc..



Dataset title (M) (MD_Metadata > MD_DataIdentification.citation > CI_Citation.title)	**Spatial representation type** (O) (MD_Metadata > MD_DataIdentification.spatialRepresentationType)
Dataset reference date (M) (MD_Metadata > MD_DataIdentification.citation > CI_Citation.date)	**Reference system** (O) (MD_Metadata > MD_ReferenceSystem)
Dataset responsible party (O) (MD_Metadata > MD_DataIdentification.pointOfContact > CI_ResponsibleParty)	**Lineage** (O) (MD_Metadata > DQ_DataQuality.lineage > LI_Lineage)
Geographic location of the dataset (by four coordinates or by geographic identifier) (C) (MD_Metadata > MD_DataIdentification.extent > EX_Extent > EX_GeographicExtent > EX_GeographicBoundingBox or EX_GeographicDescription)	**On-line resource** (O) (MD_Metadata > MD_Distribution > MD_DigitalTransferOption.onLine > CI_OnlineResource)
Dataset language (M) (MD_Metadata > MD_DataIdentification.language)	**Metadata file identifier** (O) (MD_Metadata.fileIdentifier)
Dataset character set (C) (MD_Metadata > MD_DataIdentification.characterSet)	**Metadata standard name** (O) (MD_Metadata.metadataStandardName)
Dataset topic category (M) (MD_Metadata > MD_DataIdentification.topicCategory)	**Metadata standard version** (O) (MD_Metadata.metadataStandardVersion)
Spatial resolution of the dataset (O) (MD_Metadata > MD_DataIdentification.spatialResolution > MD_Resolution.equivalentScale or MD_Resolution.distance)	**Metadata language** (C) (MD_Metadata.language)
Abstract describing the dataset (M) (MD_Metadata > MD_DataIdentification.abstract)	**Metadata character set** (C) (MD_Metadata.characterSet)
Distribution format (O) (MD_Metadata > MD_Distribution > MD_Format.name and MD_Format.version)	**Metadata point of contact** (M) (MD_Metadata.contact > CI_ResponsibleParty)
Additional extent information for the dataset (vertical and temporal) (O) (MD_Metadata > MD_DataIdentification.extent > EX_Extent > EX_TemporalExtent or EX_VerticalExtent)	**Metadata date stamp** (M) (MD_Metadata.dateStamp)

Figure 3:
Minimum set of metadata for geographic datasets

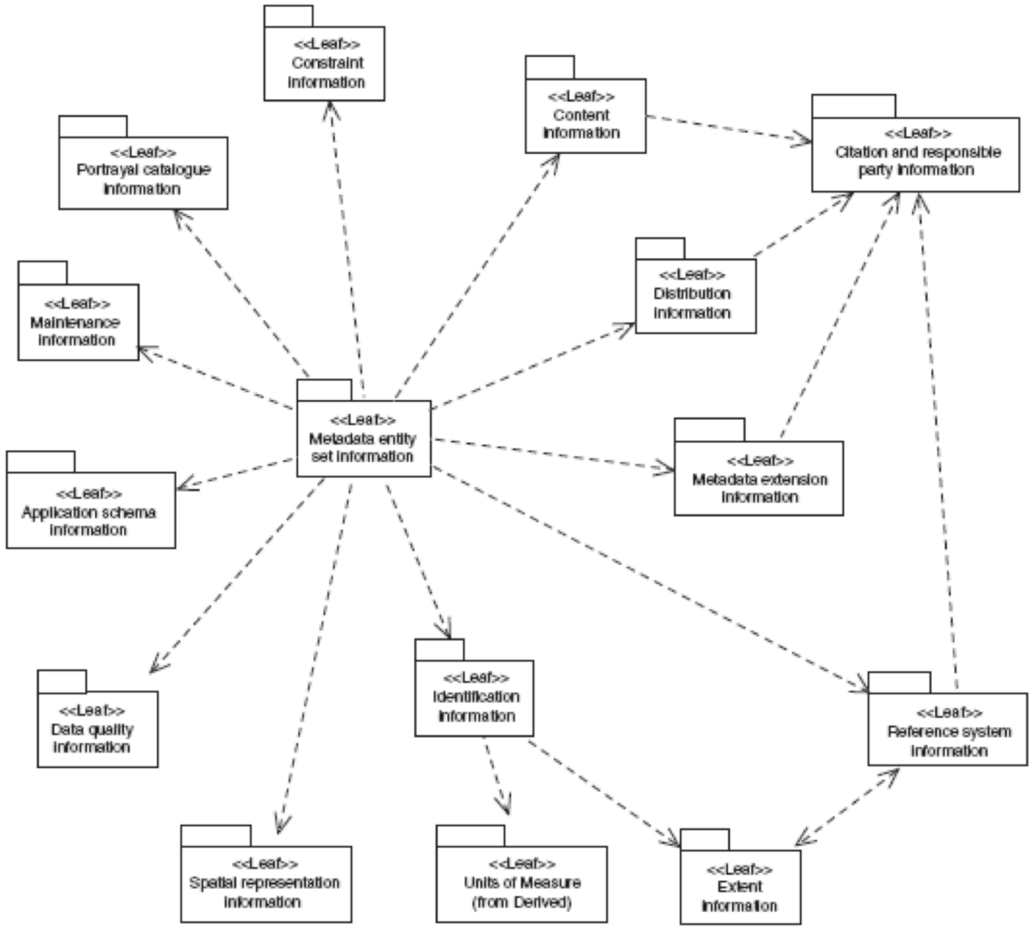

<<Leaf>>
Constraint information
<<Leaf>>
Content information
<<Leaf>>
Citation and responsible party information
<<Leaf>>
Portrayal catalogue information
<<Leaf>>
Distribution information
<<Leaf>>
Maintenance information
<<Leaf>>
Metadata entity set information
<<Leaf>>
Application schema information
<<Leaf>>
Metadata extension information
<<Leaf>>
Identification information
<<Leaf>>
Reference system information
<<Leaf>>
Data quality information
<<Leaf>>
Spatial representation information
<<Leaf>>
Units of Measure (from Derived)
<<Leaf>>
Extent information

Figure 4:
Structure of metadata packages

GEONETWORK

GeoNetwork opensource [4] is a standard based and decentralised spatial information management system, designed to enable access to geo-referenced databases and cartographic products from a variety of data providers through descriptive metadata, enhancing the spatial information exchange and sharing between organisations and their audience, using the capacities and the power of the Internet. The system provides a broad community of users with easy and timely access to available spatial data and thematic maps from multidisciplinary sources, that may in the end support informed decision making. The main goal of the software is to increase collaboration within and between organisations for reducing duplication and enhancing information consistency and quality and to improve the accessibility of a wide variety of geographic information along with the associated information, organised and documented in a standard and consistent way. GeoNetwork's main features are:

- instant search on local and distributed geospatial catalogues;
- uploading and downloading of data, documents, PDF's and any other content;
- an interactive Web map viewer that combines Web Map Services from distributed servers around the world;
- online map layout generation and export in PDF format;
- online editing of metadata with a powerful template system;
- scheduled harvesting and synchronisation of metadata between distributed catalogues;
- groups and users management;
- fine grained access control.

The metadata profile is organised in sections and the most important, illustrated in Figure 4, are: Identification Section, Distribution Section, Reference System Section, Data Quality Section and Metadata Section.

MAIN FILE FORMATS

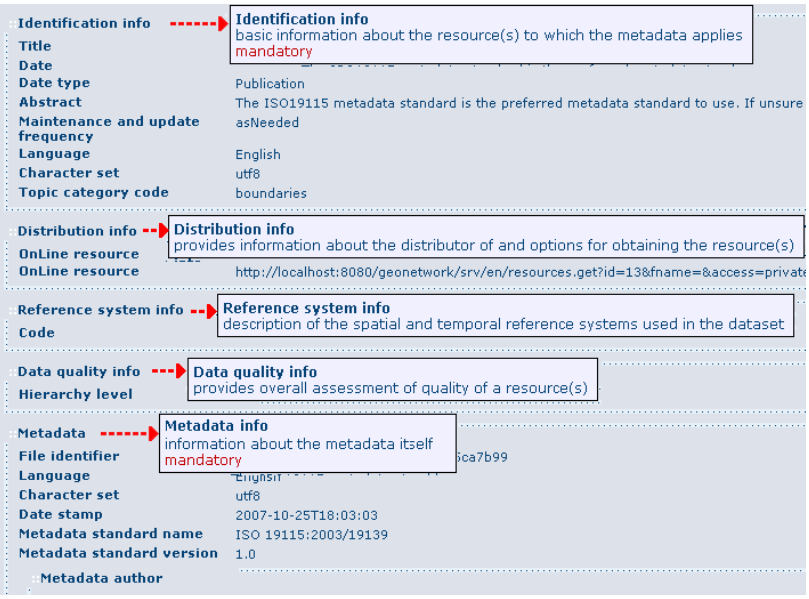
The back-end platform of the project will manage data in different formats. Below a brief overview of the main used formats.

NETCDF FORMAT

NetCDF format (Network Common Data Form) [5] is a set of software libraries and portable data format supporting creation, access, and sharing of array-oriented scientific data. It was developed and is still a project by UNIDATA, commonly used in climatological, meteorological, oceanographic and GIS applications. In NetCDF format, the dataset is stored in a binary file consists of two main parts:

- *header*, containing the information about dimensions, attributes and variables;
- *data*, which includes "fixed-size data", containing the data for variables that have limited size, and 'record-data', containing the data of the variables that have unlimited size, like the weather;

The notation used to describe an object in NetCDF format is called CDL (Common Data Language) and it is a simple tool to define the contents of a NetCDF dataset. Figure 6 shown an example of NetCDF file format.



Identification info
Title
Date
Date type — Publication
Abstract — The ISO19115 metadata standard is the preferred metadata standard to use. If unsure
Maintenance and update frequency — asNeeded
Language — English
Character set — utf8
Topic category code — boundaries

Distribution info
OnLine resource
OnLine resource — http://localhost:8080/geonetwork/srv/en/resources.get?id=13&fname=&access=private

Reference system info
Code

Data quality info
Hierarchy level

Metadata
File identifier — 5ca7b99
Language — English
Character set — utf8
Date stamp — 2007-10-25T18:03:03
Metadata standard name — ISO 19115:2003/19139
Metadata standard version — 1.0
Metadata author

Identification info
basic information about the resource(s) to which the metadata applies
mandatory

Distribution info
provides information about the distributor of and options for obtaining the resource(s)

Reference system info
description of the spatial and temporal reference systems used in the dataset

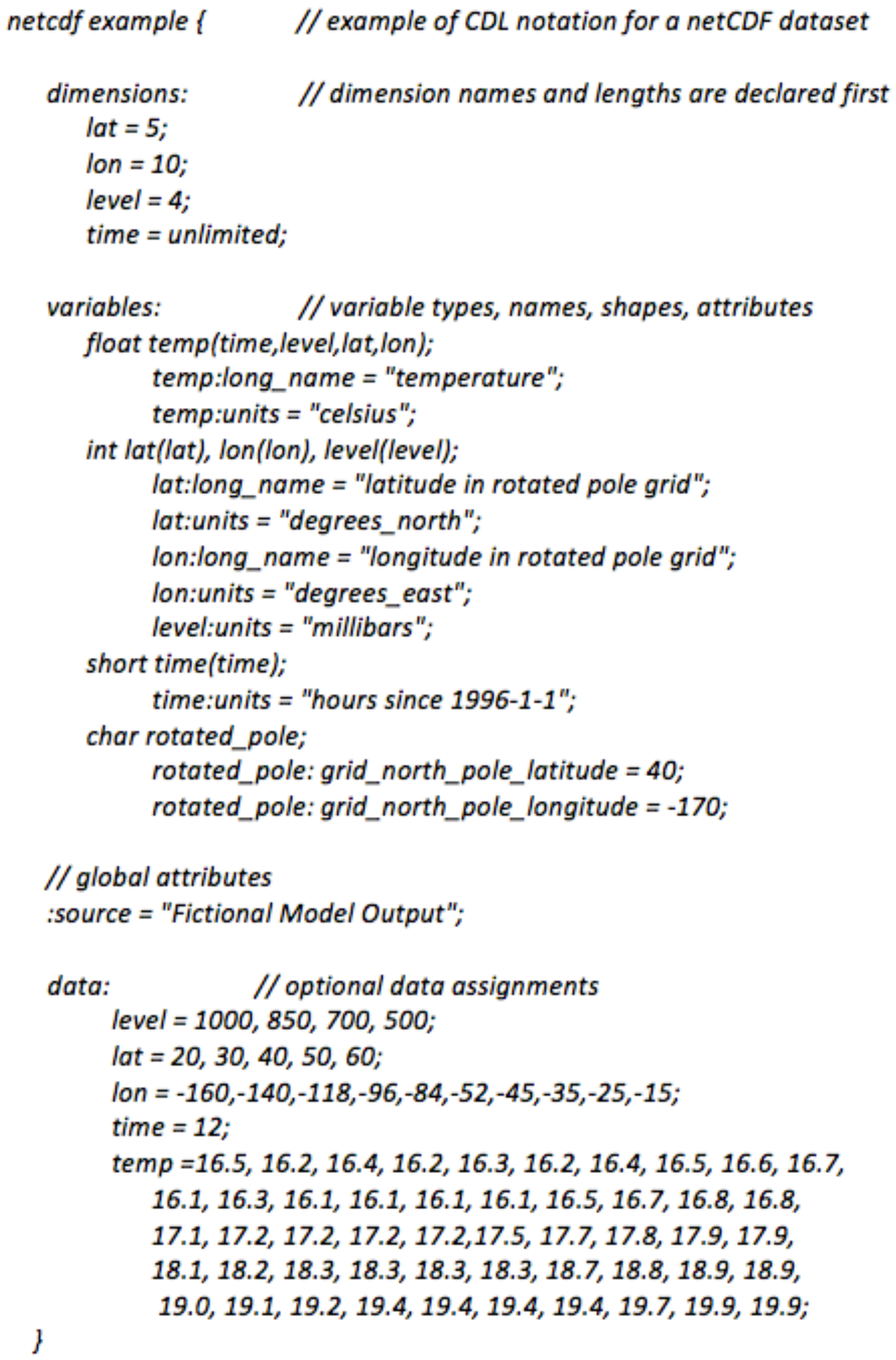
Data quality info
provides overall assessment of quality of a resource(s)

Metadata info
information about the metadata itself
mandatory

Figure 5:
Main metadata section

ESRI SHAPEFILE FORMAT

ESRI shapefile [6], or simply a shapefile, is a popular geospatial vector data format for geographic information system. It is developed and regulated by ESRI as an open specification for data interoperability among ESRI and other GIS software products. Shape files spatially describe vector features: points, lines and polygons, representing, for example, water wells, rivers and lakes. Each item usually has attributes that describe it, such as name or temperature. ESRI is an international supplier of Geographic Information System (GIS) software and geo-database management applications. A shape file stores non-topological geometry and attribute information for the spatial features in a data set. The geometry for a feature is stored as a shape comprising a set of vector coordinates. Since shape files do not have the processing overhead of a topological data structure, they have advantages over other data sources such as faster drawing speed and edit ability. Shape files handle single features that overlap or that are non-contiguous. They also typically require less disk space and are easier to read and write. Shape files are simple because they store the primitive geometric data types of points, lines, and polygons. Therefore, a table of records will store properties/attributes for each primitive shape in the shape file. Shapes (points/lines/polygons) together with data attributes can create infinitely many representations about geographic data. Representation provides the ability for powerful and accurate computations. While the term "shape file" is quite common, a "shape file" is actually a set of several files. Three individual files are *mandatory* to store the core data that comprises a shape file: *.shp*, *.shx* and *.dbf*. The shape file relates specifically to .shp files but alone is incomplete for distribution, as the other supporting files are required. There are further *optional* files which store primarily index data to



```
netcdf example {          // example of CDL notation for a netCDF dataset

   dimensions:            // dimension names and lengths are declared first
      lat = 5;
      lon = 10;
      level = 4;
      time = unlimited;

   variables:             // variable types, names, shapes, attributes
      float temp(time,level,lat,lon);
            temp:long_name = "temperature";
            temp:units = "celsius";
      int lat(lat), lon(lon), level(level);
            lat:long_name = "latitude in rotated pole grid";
            lat:units = "degrees_north";
            lon:long_name = "longitude in rotated pole grid";
            lon:units = "degrees_east";
            level:units = "millibars";
      short time(time);
            time:units = "hours since 1996-1-1";
      char rotated_pole;
            rotated_pole: grid_north_pole_latitude = 40;
            rotated_pole: grid_north_pole_longitude = -170;

   // global attributes
   :source = "Fictional Model Output";

   data:              // optional data assignments
         level = 1000, 850, 700, 500;
         lat = 20, 30, 40, 50, 60;
         lon = -160,-140,-118,-96,-84,-52,-45,-35,-25,-15;
         time = 12;
         temp =16.5, 16.2, 16.4, 16.2, 16.3, 16.2, 16.4, 16.5, 16.6, 16.7,
             16.1, 16.3, 16.1, 16.1, 16.1, 16.1, 16.5, 16.7, 16.8, 16.8,
             17.1, 17.2, 17.2, 17.2, 17.2,17.5, 17.7, 17.8, 17.9, 17.9,
             18.1, 18.2, 18.3, 18.3, 18.3, 18.3, 18.7, 18.8, 18.9, 18.9,
              19.0, 19.1, 19.2, 19.4, 19.4, 19.4, 19.4, 19.7, 19.9, 19.9;
  }
```

Figure 6:
Example of NetCDF file

improve performance, locate under the same folder. Mandatory files are:

- *.shp* - shape format: stores the feature geometry;
- *.shx* - shape index format: a positional index of the feature geometry;
- *.dbf* - attribute format: columnar attributes for each shape.

Optional files are:

- *.prj* - projection format: the coordinate system and projection information;
- *.sbn* and *.sbx* - projection format: coordinate system and projection information;
- *.fbn* and *.fbx* - a spatial index of the features for shapefiles that are read-only;
- *.ain* and *.aih* - an attribute index of the active fields in a table;
- *.ixs* - geocoding index for read-write shapefiles;
- *.mxs* - a geocoding index for read-write shapefiles (ODB format);
- *.shp.xml* - geospatial metadata in xml format;
- *.atx* - attribute index for the .dbf file in the form of shapefile.columnname.atx;
- *.cpg* - used to identify the code page for identifing the character encoding to be used.

GRID RASTER FORMAT

An ESRI grid s a raster GIS file format developed by ESRI, which defines geographic space as an array of equally sized square. grid points arranged in rows and columns. Each grid point stores a numeric value that represents a geographic attribute (such as elevation or surface slope) for that unit of space. Each grid cell is referenced by its x,y coordinate location. An ESRI grid has two formats:

- a proprietary binary format, also known as an ARC/INFO GRID, ARC GRID and many other variations;
- a non-proprietary ASCII format, also known as an ARC/INFO ASCII GRID.

The formats were introduced for ARC/INFO. The binary format is widely used within ESRI programs, such as ArcGIS, while the ASCII format is used as an exchange, or export format, due to the simple and portable ASCII file structure. In particular, a binary ESRI grid is stored in several files contained in at least two directories: the *name* directory and an *info* directory, where *name* has strict naming conventions. The *name* subdirectory is composed of different types of files containing information on the geographical location and attributes of the relative grid:

- *dblbnd.adf:* - contains the geographical boundaries of the raster grid;
- *hdr.adf* - contains information on the resolution of the grid, the compression, the number and size of cells;
- *sta.adf* - contains the statistical values of the raster;
- *vat.adf* - contains the attributes associated with the areas of the grid;

- *w001001.adf* - contains the values of the cells;
- *w001001x.adf* - index file containing pointers to each tile of the raster file w001001.adf.

The *info* subdirectory, instead, contains other information on the grid:

- arcNNNN.dir (where N is any integer): list of records;
- arcNNNN.dat (where N is any integer): tabular data;
- arcNNNN.nit (where N is any integer): fields definition.

CONCLUSIONS

This report described the ORIENTGATE data platform. It is composed of a set of services through which data produced during the different work packages will be stored, managed and published providing multiple, specific functionalities. Moreover, this report provides an overview of the data platform architecture, describing the main services and data formats as well as the metadata part responsible for enabling the data search and discovery.

Bibliography

[1] OPeNDAP - http://www.opendap.org

[2] THREDDS - https://www.unidata.ucar.edu/software /thredds/current/tds/

[3] http://www.iso.org/iso/catalogue_detail.htm ?csnumber=26020/

[4] http://geonetwork-opensource.org/

[5] NetCDF - http://www.unidata.ucar.edu/software/netcdf/

[6] http://www.esri.com/library/whitepapers/pdfs/ shapefile.pdf

Visit **www.cmcc.it** for information on our activities and publications.

The Euro-Mediteranean Centre on Climate Change is a Ltd Company with its registered office and administration in Lecce and local units in Bologna, Venice, Capua, Sassari, Viterbo, Benevento and Milan. The society doesn't pursue profitable ends and aims to realize and manage the Centre, its promotion, and research coordination and different scientific and applied activities in the field of climate change study.

cmcc
Centro Euro-Mediterraneo
sui Cambiamenti Climatici