Data access services in MyOcean
From data production to downstream applications

Authors:
Thomas Loubrieu (Thomas.Loubrieu@ifremer.fr),
Vincent Claverie (vclaverie@cls.fr), Tony Jolibois (tjolibois@cls.fr), Rémi de Dianous (rde-dianous@cls.fr),
Jon Blower (j.d.blower@reading.ac.uk)), Guy Griffiths (guy.griffiths@reading.ac.uk)
Laia Romero (laia.romero@altamira-information.com)

MyOcean CIS
Needs, global architecture and technical challenges

On-going Challenges

- How to deal with **INSPIRE standard services** (to reach interoperability) and the **monitoring of users**:
⇒ To monitor users, an authentication system is needed, but this kind of need is not covered by OGC standards

- How to make possible a **unique user login and password** to access **different project data access services** (SeaDataNet and MyOcean for example), when user management strategy is different

- How to **distribute big data volume** through **web** services:
⇒ Oceanographic data represent very large volume of data (several Gb produced daily)

- How to provide **harmonized and standard access services** to distribute several **heterogeneous kind of data** : along-track (and swath), in-situ and gridded data

1 Web portal

- Centralized system
- Offer a single entry to the users
- Users have one login and password to access all MyOcean services

Login | Discover
View | Download

1 Information System

- Centralized system
- Provide an Inspire catalogue (OGC CSW and ISO metadata)
- Provide administration functionalities
- Host a centralized authentication system

User management | Catalogue | Monitoring
Metadata (ISO)
Authentication/Authorization

24 Dissemination Units

- Distributed all over Europe
- Provide machine to machine interfaces for viewing and download
- All DU are connected to the centralized authentication system, and are monitored (system and data access user requests)
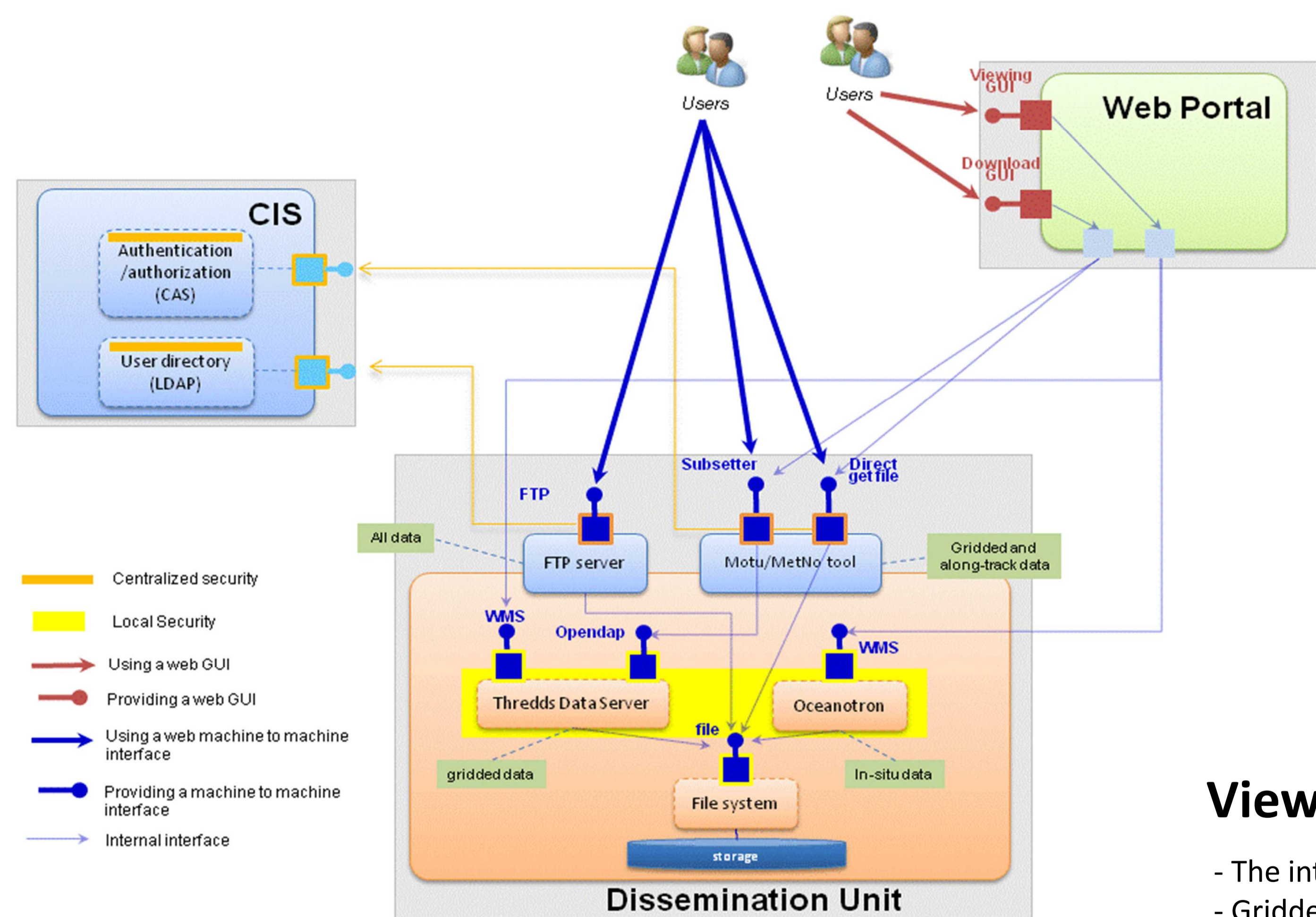
Download | View | FTP
Storage | netCDF CF
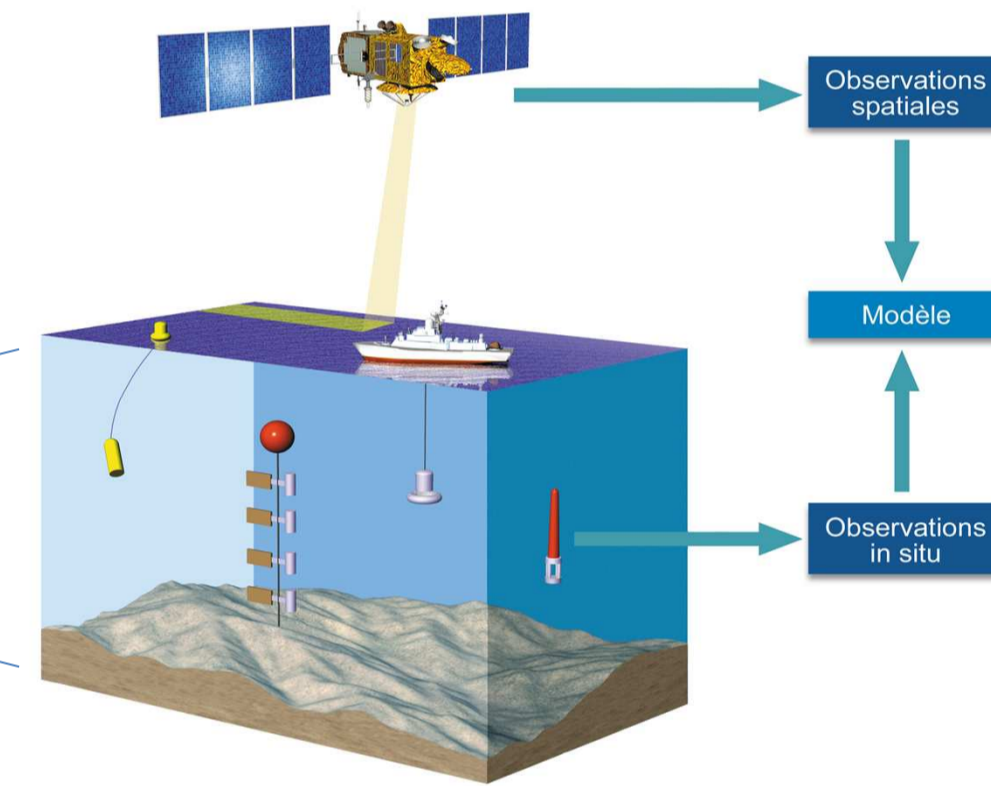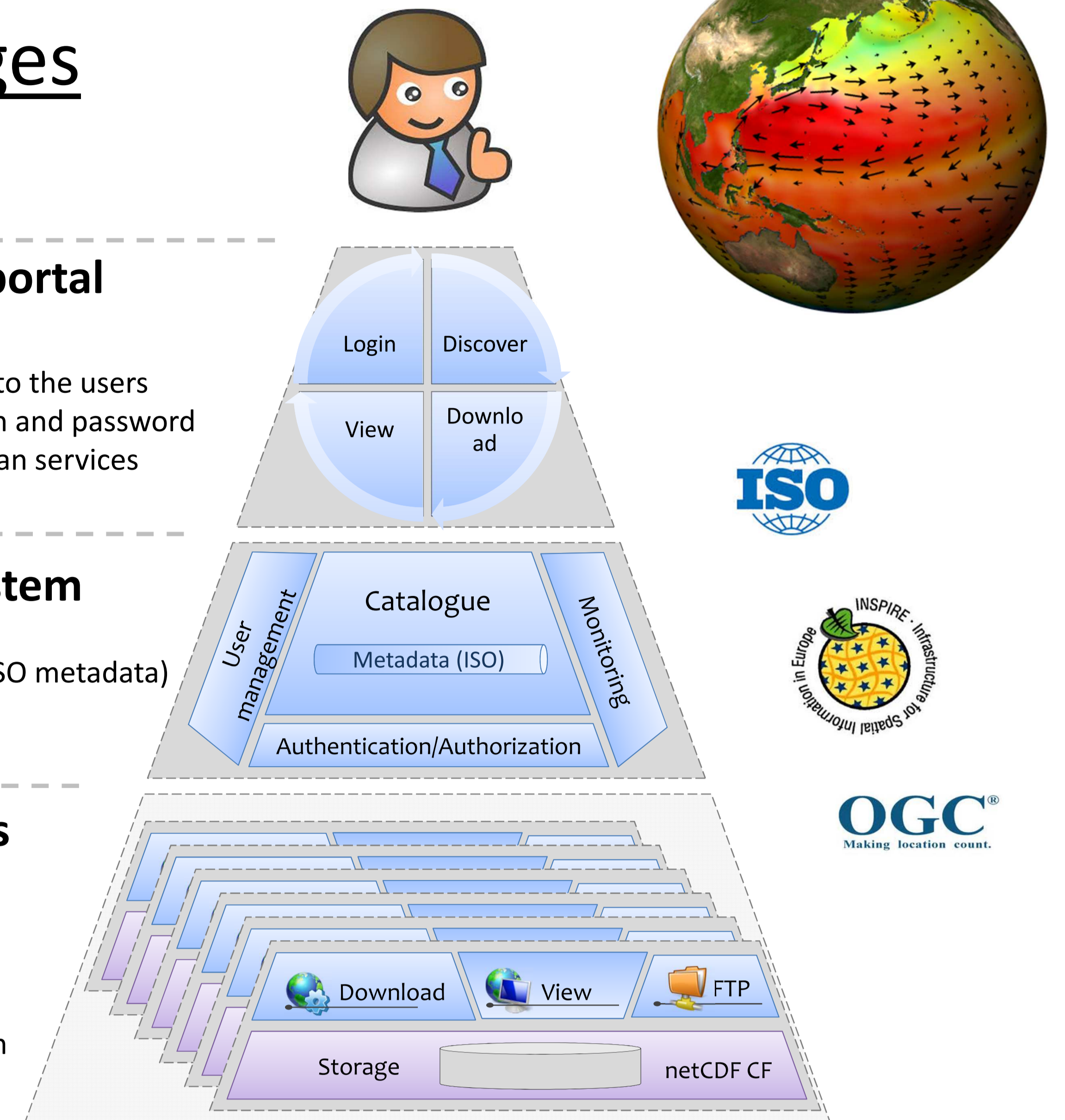
- Spatial observations
- In-situ observations
- Modelling and Forecast

52 Production Units

Download and view gridded, in-situ and along-track data

• A unique login & password to access the download services

- Subsetter and Diret get file (http) connected to a single sign on service implemented with CAS technology
- FTP servers are connected directly to the central LDAP

• Transaction monitoring, to know who access what and when

- Full transaction accounting is done for authenticated access services (ie the download services in MyOcean). Project and producers know who access which dataset, how and when
- A simple transaction accounting (based on IP address) is done on free access services (ie the view service in MyOcean).

• Data files in NetCDF 4

- As TDS input, the netCDF4 format should be accepted in MyOcean V4 (to be confirmed). A careful analysis with all DUs will be conducted to confirm the full compatibility
- As output of subsetter, it will be done in the next future version

View service : all the data served through WMS (client and server)

- The interface is not secured, because usual WMS clients can't manage single sign on
- Gridded data are served through WMS by ncWMS, which is embedded in Thredds Data Server
- In-situ data will be served through WMS by Oceanotron

Download service for gridded data : several services depending on the needs

- Subsetter service, implemented by Motu and Met No tool : allow user to download a geospatial and temporal subset of a dataset
- Direct Get File service, implemented by Motu and Met No tool : allow user to download a temporal subset of a dataset (files)
- FTP service

Download service for in-situ data : FTP, vsftpd connected to the central LDAP

Download service for along-track data : FTP or Direct Get File

NCWMS
View netCDF files (gridded data) through WMS. ncWMS is embedded in Thredds Data Server

OCEANOTRON
Download in-situ data through Opendap.
View in-situ data through WMS

MOTU
Download netCDF files with subsetting through a robust and secured server

- Reads Opendap interface with java NetCDF API
- Secured with CAS
- Authorization : allow a user profile (managed in LDAP) to access certain datasets and not others (configurable)
- Robustness : Motu makes use of block oriented techniques: data are read according to a pre-configured and fixed size (data blocksize) in RAM
- Robustness : a queue server to manage incoming requests
- Functionality : Deal with data either side of the longitudinal boundary
- Functionality : scale_factor and add_offset attributes are not lost in result file
- Request size threshold configurable