

GHRSSST Regional/Global Task Sharing

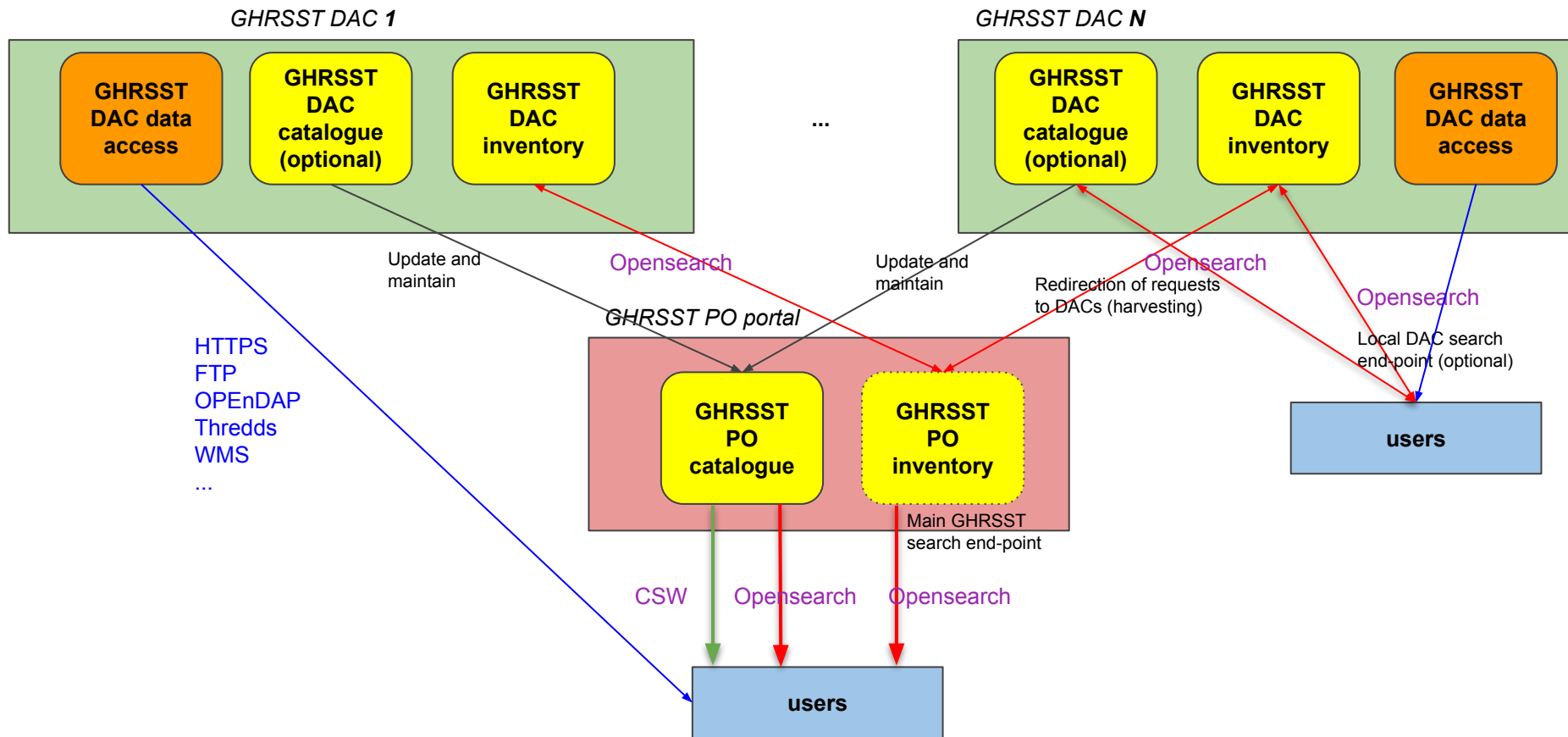
Task Team Report - GHRSSST XXII



R/G TS Task Team

- Long active task team, follow-on to DAS TAG
- **Main activity in last two years:**
 - specifying and releasing the new R/G TS framework
 - supporting its implementation
- **Participants:** Jean-François Piollé (Ifremer), Ed Armstrong (PO.DAAC), Wen-Hao Li (PO.DAAC), Yongsheng Zhang (NOAA), Huai-min Zhang (NOAA), Korak Saha (NOAA)
- Anyone wants to join, just email me: jfpiolle@ifremer.fr

R/G TS refined data discovery, search and access system



Activities

- **Semi-regular telco of R/G TS team**
- Pilot project demonstrated central catalogue and federated query system based on OpenSearch
- Released system architecture document :
https://www.ghrsst.org/wp-content/uploads/2020/02/GHRSST-Regional_Global-Task-Sharing-R_G-TS-v1.0rev1.pdf
- Survey of DACs and Producers on R/G TS readiness :
<https://forms.gle/8Ws3NJA2L462rHM8>
- **Updated list of DAC and GDP contacts**
- **Defined content of central catalogue's dataset metadata profile**
- **Copernicus / Eumetsat funding R/G TS Central Catalogue and Inventory Search service (opensearch) granted to Ifremer**
- **Started implementation**

Status of R/G TS Architecture document

Document finalised and released, available on GHRSSST web site, googlable

https://www.ghrsst.org/wp-content/uploads/2020/02/GHRSSST-Regional_Global-Task-Sharing-R_G-TS-v1.0rev1.pdf

Basis for implementation of R/G TS

Describe the roles of **GHRSSST Data Producers [GDPs]** and **Data Assembly Centers [DACs]**, services to be implemented in the next one to two years

Central services from GHRSSST PO

Implementation of R/G TS central services

Copernicus / EUMETSAT ITT released in 2020 / Kicked off end of March 2021

12 months schedule => March 2022

- Implement and operate central catalogue (Ifremer)
- Implement and operate Opensearch End-Point Access (Ifremer)
- GHRSSST Web Site (DMI - GHRSSST PO)
- 10 months for complete software implementation and testing
- Beta service release at K0+10 (Dec.2021)

Implementation of Central catalogue

- based on Pilot Project initial version
- ISO 19115/3 metadata profile
- profile defined from GHRSSST L2P/L3/L4 products based on ISO 19115/3, based on different catalogue examples (PODAAC, NOAA, Ifremer/CERSAT)
- The GHRSSST central catalogue will not be federated but will be populated from existing catalogues whenever possible (already tested in Pilot Project) - import to be implemented per source center. **Mapping GHRSSST Dataset Id (defined by producer) ⇔ DAC to be done**
- None of the inputs have the same specification & content : most relevant and most common metadata extracted to defined the GHRSSST common metadata model
- Description of common metadata model circulated to R/G TS : **being reviewed**
- Including cloud access informations

GHRSSST Dataset metadata profile (1/2)

| TITLE | Definition | Use | Content example |
|-------------------------|-------------------------------------------------------------------|-----|-------------------------------------------------------------------------------------------------|
| Description | | | |
| Title | Title of the described resource | 1 | Brazil/Tropical Atlantic High-Resolution Sea Surface Temperature Gridded Level 4 Daily Analysis |
| Abstract | Abstract of the described resource | 1 | This L4 SST product is produced at ultra-high resolution (UHR) on a 0.02 x 0.02 degree grid... |
| Project | Project name(s) linked to the resource | 0:n | GHRSSST |
| Collection ID | Resource identifier (linked to producer for exemple) | 0:1 | CER-SST-BRA-1D-002-ODY-MGD |
| DOI | Digital Object Identifier | 0:1 | 10.17882/52804 |
| Instrument | Instrument name | 1:n | AVHRR |
| Platform | Platform name | 1:n | METOP-A |
| Level | Production level | 1 | L4 |
| Acquisition pattern | Acquisition type | 1 | composite |
| Compositing | Composition method if any | 1 | Optimal interpolation |
| Latency | Delay from acquisition to distribution | 0:1 | less than 24 hours |
| Begin date | Date of first observation available | 1 | 08/29/2010 |
| End date | Date of last observation available (empty if still in production) | 0:1 | |
| Temporal Resolution | Time between two observations | 1 | 1 day |
| Spatial Resolution | Spatial resolution | 1 | 0.02 degree |
| Projection | Geographic projection | 1 | ETRS89 (EPSG:4258) - Equirectangular projection |
| Geographic area | Geographic area covered with the observation | 1 | Tropical Atlantic |
| Geographic bounding box | Bounding box of the geographic area | 1 | westBoundLongitude -75.00 EastBoundLongitude -15.00 southBoundLatitude -25.00 |

| | | | | | |
|---------------------------------------------|------------------------------------------------|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|------------------------------|
| | | | southBoundLatitude 25.00 | | |
| Main instrumental or geophysical parameters | Keyword from an internal thesaurus | 0:n | Ocean Temperature | | |
| Keywords (GCMD) | Keyword from GCMD thesaurus | 0:n | /Ocean Temperature/Sea Surface Temperature | | |
| Contacts | | | | | |
| Point of contact (2 fields) | Information about resource point of contact(s) | 0:n | Institution /Name | e-mail | URL |
| | | | Helpdesk Cersat | cersat@ifremer.fr | https://cersat.ifremer.fr |
| Principal Investigator (2 fields) | Information about the PI | 0:n | Institution Name | e-mail | URL |
| | | | Emmanuelle Autret | eauret@ifremer.fr | https://www.ifremer.fr/lops/ |
| Producer (2 fields) | Who is in charge of producing the data | 0:n | Institution Name | e-mail | URL |
| | | | Ifremer/CERSAT | cersat@ifremer.fr | https://cersat.ifremer.fr |
| Distributor (2 fields) | Who is in charge of distributing the data | 0:n | Institution Name | e-mail | URL |
| | | | Ifremer/CERSAT | cersat@ifremer.fr | https://cersat.ifremer.fr |
| Funder (2 fields) | Who is the funder | 0:n | Institution Name | e-mail | URL |
| | | | ESA | | https://esa.int |
| Access and Usage | | | | | |
| Helpdesk (2 fields) | Information about the Helpdesk | 0:1 | Institution /Name | e-mail | |
| | | | HelpDesk CERSAT | cersat@ifremer.fr | |
| Access policy | Access policy from a list of restrictions | 1 | OtherRestrictions | | |
| Usage policy | Usage policy | 1 | | | |
| Required citation | Required citation | 1 | These data are produced for ESA/Medspiration project and were obtained from the Centre de Recherche et d'Exploitation Satellitaire (CERSAT), at IFREMER, Plouzané (France) | | |
| Distribution Format | Name of the format | 0:1 | Netcdf | | |

GHR SST Dataset metadata profile (2/2)

| | | | | | |
|------------------------------|----------------------------------------------|-----|---------------------------------------|---------------------|---------------------|
| Format Version | Version of the format | 0:1 | 3 | | |
| Format Convention | Convention name (from a list) | 0:n | CF-1.4 | | |
| Format Amendment | Amendment number | 0:n | | | |
| History | | | | | |
| Status | Status of the dataset (active/completed) | 1 | active | | |
| Updates (2 fields) | Description /date of an update | 0:n | Information | Date | |
| | | | | | |
| Issues (3 fields) | Description /temporal interval of an issue | 0:n | Information | Start Date | End Date |
| | | | No acquisition, calibration manoeuvre | 2021/01/12 06:23UTM | 2021/01/17 08:45UTM |
| Versions (3 fields) | Date description and number of a new version | 0:n | Version Information | Version Number | Date |
| | | | New sensors added. | V5.4 | 2020/01/03 |
| Discover the resource | | | | | |
| Overview | Preview | 0:n | png, jpeg file | | |
| Direct download | | | | | |
| FTP | FTP information | 0:1 | ftp.ifremer.fr/osisaf/data/amr2 | | |
| HTTPS | HTTPS information | 0:1 | https://ifremer.fr/data/osisaf/amr2 | | |
| THREDDS | THREDDS information | 0:1 | tds0.ifremer.fr/osisaf/amr2 | | |
| Cloud | Public cloud access information | 0:n | | | |
| Service | | | | | |
| WMS | WMS server link | 0:1 | wms.ifremer.fr/osisaf/amr2 | | |
| OPeNDAP | OPeNDAP link | 0:1 | opendap.ifremer.fr/osisaf/amr2 | | |
| Documents | | | | | |
| User guide | Link(s) to user guide(s) | 0:n | | | |
| Processing and | Link(s) to Processing | 0:n | | | |

| | | | |
|----------------------------|------------------------------|-----|--|
| validation and document(s) | validation document(s) | | |
| Other document(s) | Link(s) to other document(s) | 0:n | |

Implementation of Opensearch end-point

— — —

- Federates existing end-points
 - NASA: <https://cmr.earthdata.nasa.gov/search/site/docs/search/api.html>
 - NOAA: <https://data.noaa.gov/onestop/collections?q=GHRSSST>
 - Eumetsat: <https://codas.eumetsat.int/>
 - Ifremer / OSI SAF: <https://opensearch.ifremer.fr>
- No common syntax for queries and returned result : **Mapping GHRSSST Dataset Id**
↔ DAC Id to be done
- Minimum set of search keywords
 - count, startPage
 - datasetId
 - timeStart, timeEnd
 - geoBox
- Description of common syntax and result to be circulated to R/G TS : **R/G TS**
TT member to review it so that we agree on the content
- Output: ATOM, JSON
- Fault-tolerance: service unavailable, timeout exceeded, no response
- Open source python package (can be used for other purposes)
- Early implementation of Opensearch service completed

Opensearch search request example

http://visi-cersat-web1.ifremer.fr/opensearx/granules.raw?datasetId=AVHRR_SST_METOP_B-OSISAF-L2P-v1.0&startPage=0&count=100&timeStart=2020-01-01T00:00:00Z&timeEnd=2020-01-01T23:59:59Z&geoBox=-180.0,-90.0,180.0,90.0

- No assumption on where the data are served (here both at P0.DAAC and Ifremer)
- Single access point
- Results from both DACs are returned at once
- Returned as XML or JSON

Opensearch result

```
▼ entries:
  ▼ 0:
    ▼ id: "https://opensearch.ifremer.fr/granules.atom?uid=avhrr_sst_metop_b-osisaf-l2p-v1.0:20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
    ▼ title: "20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
    ▼ summary: "Granule metadata for 20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
    updated: "2021-06-07T08:14:00+00:00"
    dc_date: "2020-01-01T21:31:03Z/2020-01-01T21:34:03Z"
    ▼ geobox: "-82.8310828355 64.400221112 55.9890096531 87.0992535141252"
    geobox_where: null
    ▼ links:
      ▼ 0:
        title: "DATARMOR GPFS"
        rel: "enclosure"
        type: "application/octet-stream"
        ▼ href: "/home/ref-osisaf-public/data/sst/l2p/global/avhrr_metop_b/2020/001/20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
      ▼ 1:
        title: "FTP"
        rel: "enclosure"
        type: "application/octet-stream"
        ▼ href: "ftp://eftp1.ifremer.fr/cersat-rt/project/osisaf/data/sst/l2p/global/avhrr_metop_b/2020/001/20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
      ▼ 2:
        title: "HTTPS"
        rel: "enclosure"
        type: "application/octet-stream"
        ▼ href: "https://osi-saf.ifremer.fr/sst/l2p/avhrr_metop_b/2020/001/20200101213103-OSISAF-L2P_GHRSSST-SSTsubskin-AVHRR_SST_METOP_B-sstmgr_metop01_20200101_213103-v02.0-fv01.0.nc"
```

First item of result list

Properties of returned granule

URL of granule, in different protocols (FTP, HTTPS, cloud,...)

Opensearch

Opensearch service is mostly intended for machine to machine exchanges:

- From a script (e.g. daily download over an area)
- From a scientific workflow (e.g. jupyter notebook)
- From an application: visual display of available data
- ...

Libraries, scripts and packages exist to ease the encoding/decoding of opensearch requests. Example will be provided.

Example: search API in python working with different opensearch services: https://gitlab.ifremer.fr/naiad/eo_harvest/

Conclusion and way forward

Implementation of central catalogue and federated opensearch service is now on its way and will be completed by next GHRSSST workshop

Engagement of GHRSSST data producers will start early next year to populate and update the GHRSSST catalogue

Engagement of GHRSSST DACs will start this year to interconnect the existing open-search end-points

Challenge for the next two years is to have a comprehensive GHRSSST catalogue and all DACs interconnected