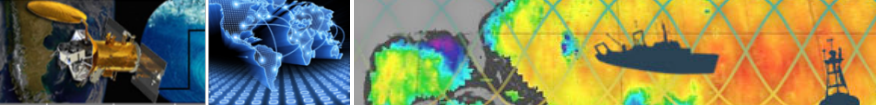


Tools & Services for Integrated Oceanographic Data Access: The Oceanographic *In situ* data Interoperability project (OIIP)

IN43A-06

Fall AGU 2018
13 December 2018, Washington DC

V. Tsontos[1], C. Lam[2], S. Arms[3], N. Quach[1], C. Thompson[1], F. Platt[1], J. Roberts[1], L. McGibbney[1]

[1] Jet Propulsion Laboratory, California Institute of Technology
[2] Large Pelagics Research Center/University of Massachusetts-Boston
[3] Unidata/UCAR

© 2018 California Institute of Technology. Government sponsorship acknowledged

CL#

Motivation, Goals & Accomplishments

- **OIIP** 2-year NASA/ACCESS15-017 funded technology development project (Oct. 2016- Sept. 2018)
- **Develop improved capacity to support NASA field campaign data** via the enhancement & integration of higher TRL technologies: ***NCEI .nc templates, CMC, ROSETTA, THREDDS, Tagbase***, *(eTUFF, MPS/SIS)*
- **R&D project addressing** some **fundamental earth science data informatics challenges**:

  Need for improved integration of **multivariate datasets** from **diverse observational platforms** in support of NASA science, satellite mission Cal/Val, decision support applications

  Challenge of *in-situ* data:
  - inherently diverse, complex, heterogeneous (esp. biological)
  - Acute Interoperability issues:
    - *Ad hoc* data file formats/insufficient adoption of self-describing formats (nc, hdf)
    - Paucity of metadata and/or non-standards compliant metadata
- **Focus on both conventional field campaign & marine animal electronic tagging data** as a representative (but also more challenging) use case
- **Open Source all software components:** source code & Docker containers on Github
  Further Info , documentation & tutorial/demo videos on **https://oiip.jpl.nasa.gov**

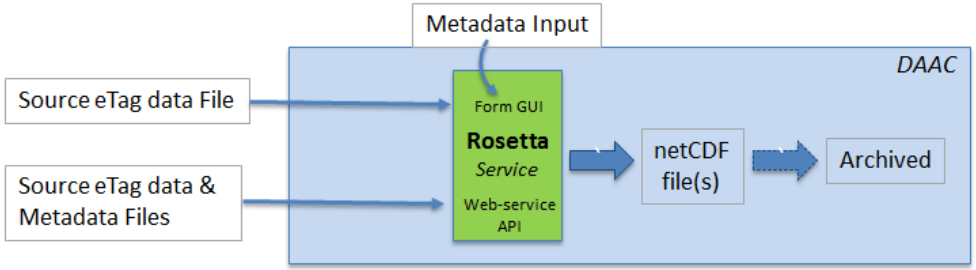
ROSETTA

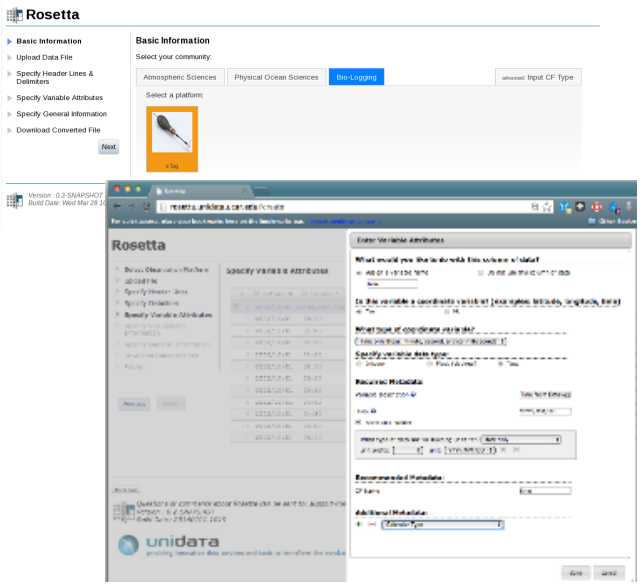
- Generalized web-based conversion tool and associated web-service/API for the production of standards compliant netCDF data files with CF/ACDD metadata from generic, columnar ASCII data file inputs.
- Facilitate conversions either interactively for single files via a Web-based GUI wizard or via API for bulk, automated processing

Enhancements:

- support for the NOAA/NCEI v2.0 *in-situ* templates, CF/ACDD standards, discrete spatial geometry (DSG) data types
- Support the electronic tagging file and metadata specifications (eTUFF) developed by OIIP
- Complete redesign/rewrite of ROSETTA front-end
- Implementation of RESTful Web service API

https://github.com/Unidata/rosetta





Demo Videos

Interactive Conversion GUI	https://youtu.be/2IrSDTUfeNU
API Bulk conversions	https://youtu.be/_4jIDvrqiZo
eTUFF processing workflow	https://youtu.be/dYsXRQLwutc

http://rosetta.unidata.ucar.edu

THREDDS Enhancements Under OIIP as part of the Upcoming V5.0 Release

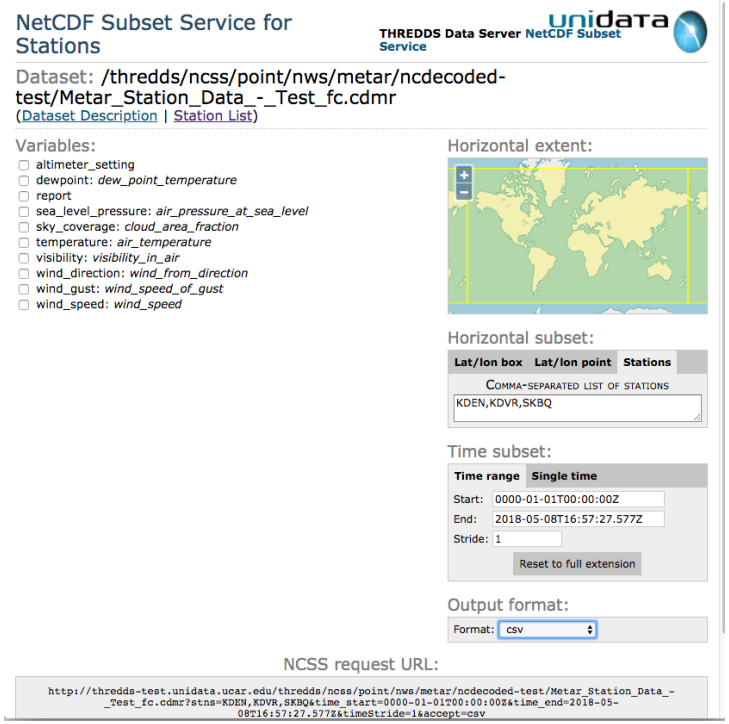
TDS: widely used web server technology providing metadata & data access for scientific datasets via a range of remote geospatial data access protocols

(eg. WCS, WMS, ncSSS, OPeNDAP)

Enhancements under OIIP

- Support for Discrete Spatial Geometry (DSG) data/feature types (point, profile, and trajectory series) at the nc-JAVA library level.
- Inclusion of the OGC Sensor Observation Service netCDF (ncSOS) plugin as part of the core TDS distribution package.

https://github.com/Unidata/thredds



Subsetting via structured THREDDS URL with parameters

HTTP service request

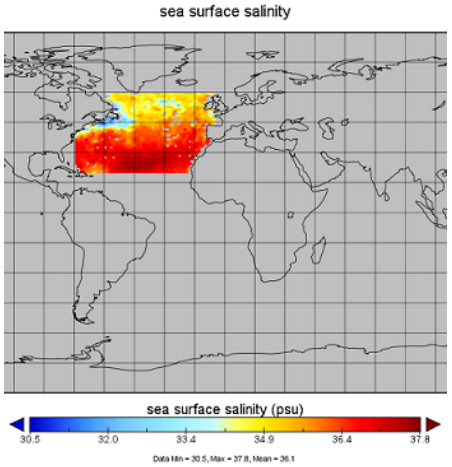
Catalog reference

http://thredds.jpl.nasa.gov/thredds/ncss/grid/ncml_aggregation/SalinityDensity/aquarius/aggregate__AQUARIUS_L3_SSS_SMI_7DAY_V4.ncml?var=l3m_data&north=60&south=20&east=-5&west=-75&time_start=2013-12-20T00:00:00Z&time_end=2014-09-24T00:00:00Z

Variable list

Spatial bounding box

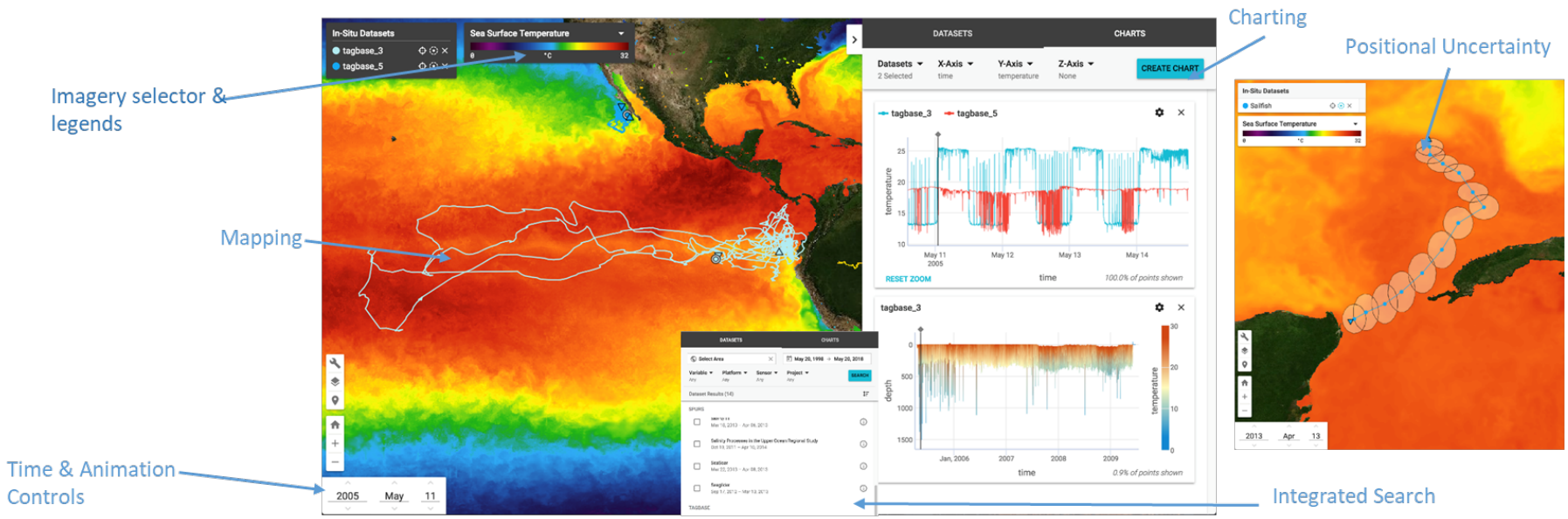
Time range filters



Web-based Oceanographic Data Visualization Tool

- Integrated visualization of raster and vector-based *in situ* & satellite datasets
- Synchronized horizontal and vertical views of data and their evolution over time
- Accurate temporal representation of data & time interval control
- Integrated Data Search capability
- Performant, high data volume throughput (data decimation algorithm)
- Open Source technology components (JPL-CMC, GeoServer, Solr, postgreSQL)

https://github.com/oiip/oiip-data-viewer



Conclusions

- **Data interoperability consideration are vital** because they **enable efficient access & consumption of data** by diverse applications/services **sustainably at reduced cost**

- **Acute data file level interoperability problem** associated with in-situ & electronic tagging datasets **is tractable given** community **convergence around data standards encapsulated by NCEI** .nc templates & **eTUFF** and a **conversion tool such as ROSETTA**

- **OIIP has produced a range of useful outputs** for the **Earth Science Data Community**:
  - **standards enhancements** & open source **tools addressing in-situ data interoperability** issues
  - **demonstrate the value added of multivariate data integration** via enhanced visualization tool
- General **approach to community metadata & vocabulary development**

- Interest in **community comment**

- Future **collaborative development opportunities** (eg. NASA/CEOS COVERAGE project)