



DAS-TAG

Breakout session

Review of activities | GHR SST XVI | ESTEC | 22th July 2015

- More and more tools to remotely process or analyse data, reduce the data download
- Advanced functions for science and data analysis
- JPL/PODAAC
 - MUDROD : intelligent knowledge engine for data discovery => return more relevant results to users
 - data quality screening service (VQSS) : pre-filter data before download (flags, thresholds,...)
 - New visualization web GUI for data selection (developped for SMAP)
 - OceanXTreme : computational platform for data analysis (build climatology, detect anomalies, detect features,...)
 - DOMS : open public in situ to satellite collocation service sustainable and extensible => spurs, samos, icoads, argo,...
- Ifremer
 - Fast analytics tools based on Hadoop can process 100s of Gbs in a few seconds to minutes : climatology, anomalies, time series, statistics



Review on new services and infrastructures

- Big data technologies and cloud computing lead now to a game change
- NOAA partnership with major cloud service providers
 - Cloud provider form alliance with third party entities, establish list of « wanted » products
 - NOAA Provide freely their data to these clouds that make them available for free, no privileged access
 - User don't pay for data access but for using virtual machines on the cloud (remote processing of the data) – data download remains a service of provider (NOAA) and is not replaced here.
- Pathfinder processed on Amazon cloud
- Numerous similar initiatives and programs in Europe will probably change data usage in a significant manner in coming years
 - Thematic exploitation platforms (ESA)
 - H2020 Research & Innovation Program framework : calls on virtual research environment, science cloud platforms,...
 - Helix-Nebula : open market rules for science clouds



- Used for dataset identification and citation
- Several DOIs may exist for datasets published by different centers (RDAC, GDAC, LTSRF)
 - PODAAC already minted DOIs for all GHR SST datasets
 - NOAA is also required to do it
 - Issues about authorship, primary DOI, rules for associating DOIs with dynamic datasets
 - General policy
 - A new DOI should be mint for a major dataset revision (not for minor version change)
 - Dynamically extended datasets keep the same DOI all along
 - Third party publishers can mint their own DOIs, reflecting their own version of a dataset (which may be enriched or different in some way from original producer's version)
 - Authorship must be properly acknowledged in the metadata and determined by the producer
 - Action on DAS-TAG to publish GHR SST policy for DOI minting



New R/GTS organization (see Ken's)

Conference announcements :

- **2016 Ocean Science Session on Big Data, 21-26 February 2016 – New Orleans** (submit abstract before 23 September)
- **EO Science 2.0, 12-14 October, ESA/ESRIN -**
<http://eoscience20.org>

DAS-TAG future

- Merge with AUS-TAG possible way forward