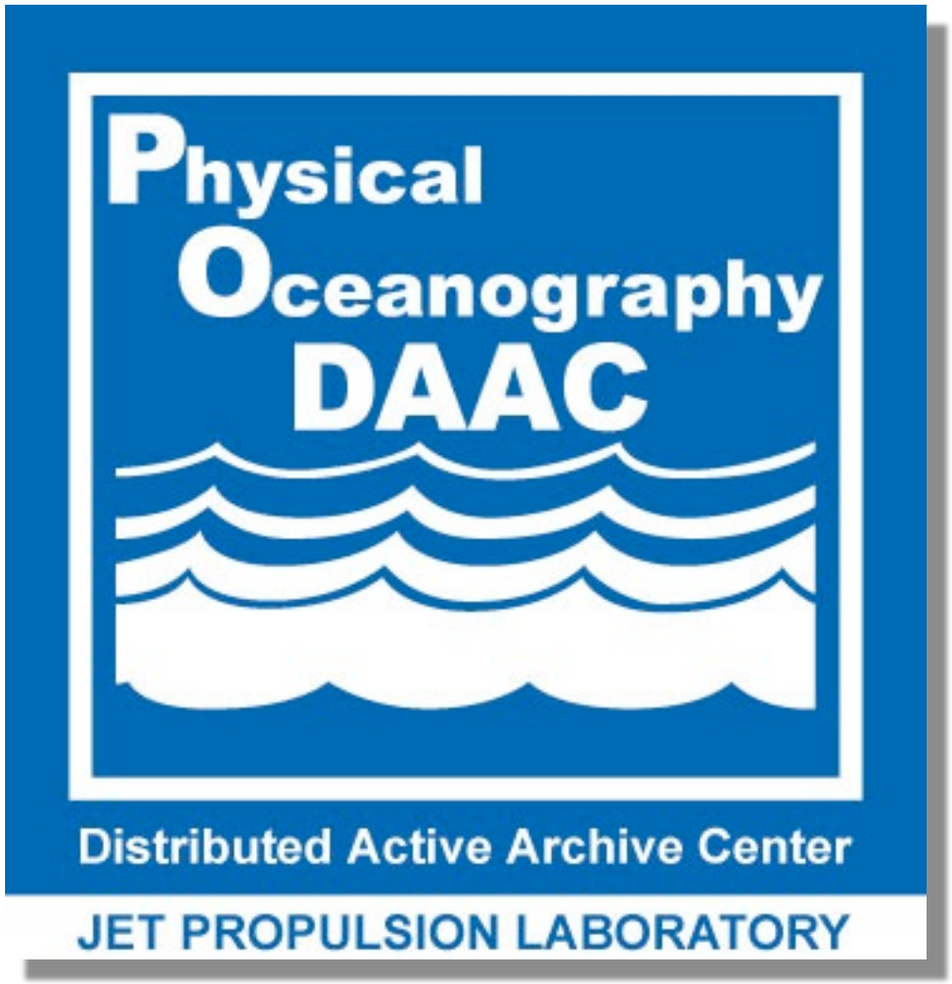
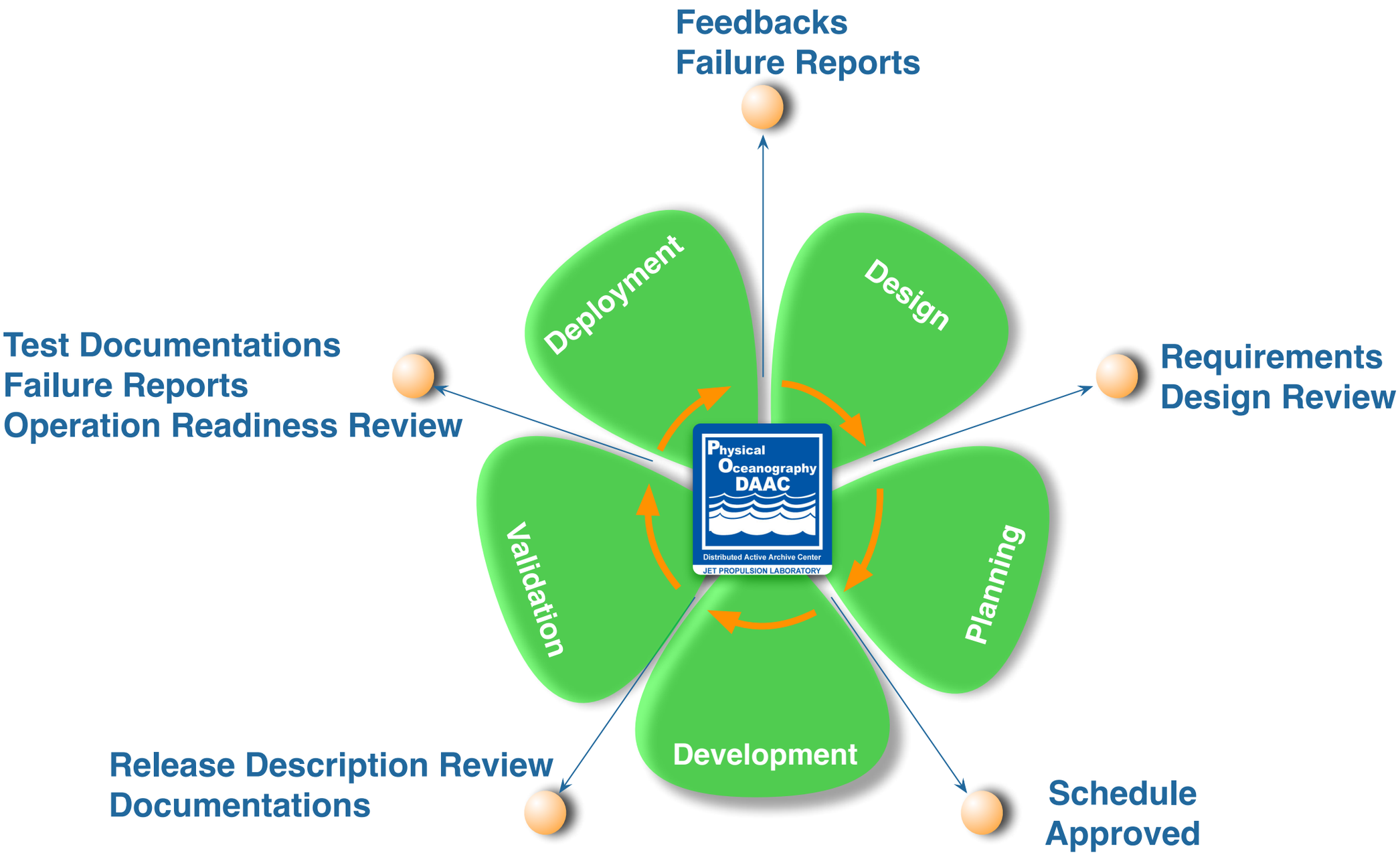

Physical
Oceanography
DAAC
Distributed Active Archive Center
JET PROPULSION LABORATORY

System v2.1.1

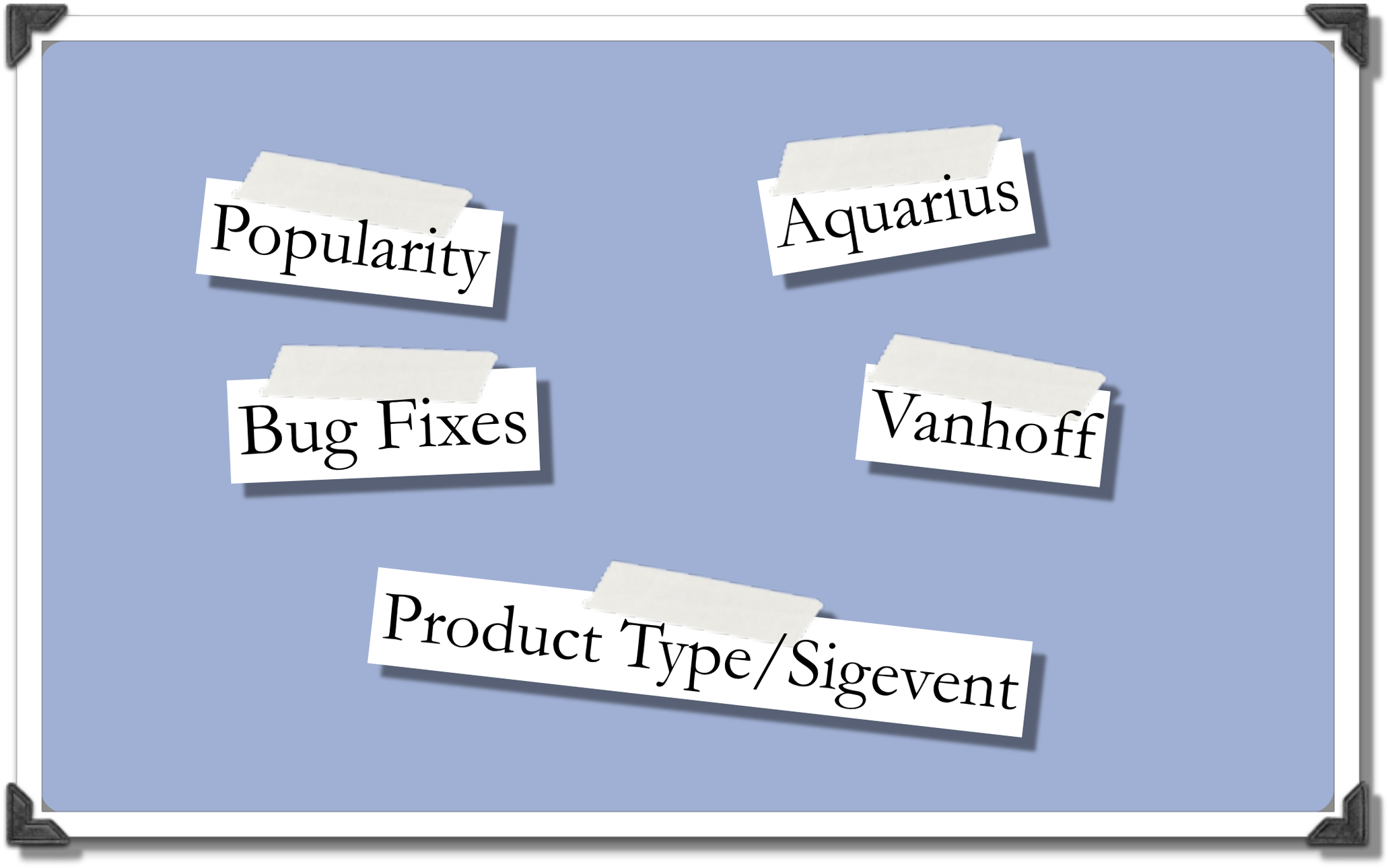
Release Description Document (RDD) Review

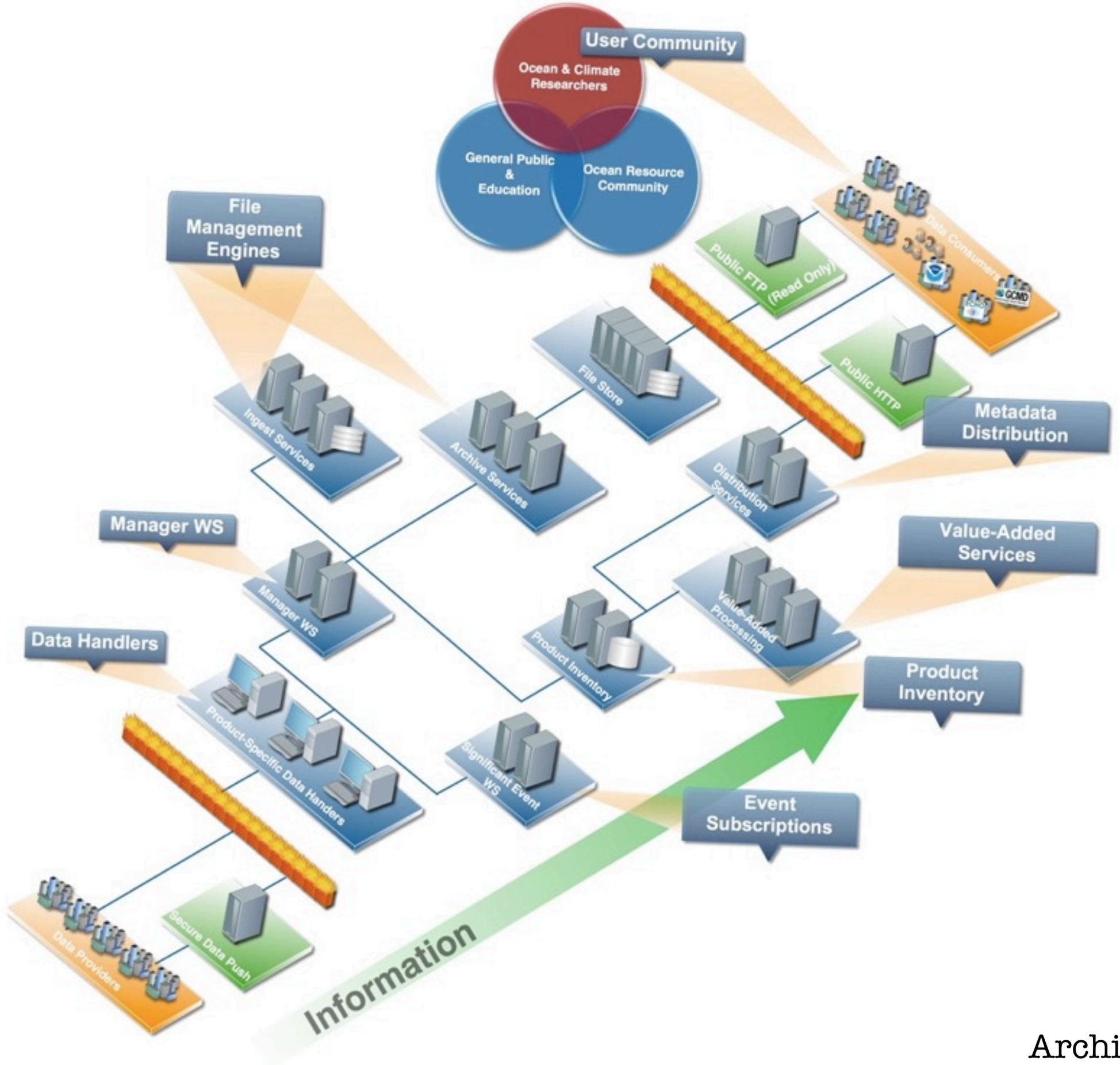
Thomas Huang
July 14, 2010


Feedbacks
Failure Reports
Test Documentations
Failure Reports
Operation Readiness Review
Requirements
Design Review
Deployment
Design
Validation
Planning
Development
Physical Oceanography DAAC
Distributed Active Archive Center
JET PROPULSION LABORATORY
Release Description Review
Documentations
Schedule
Approved

Process

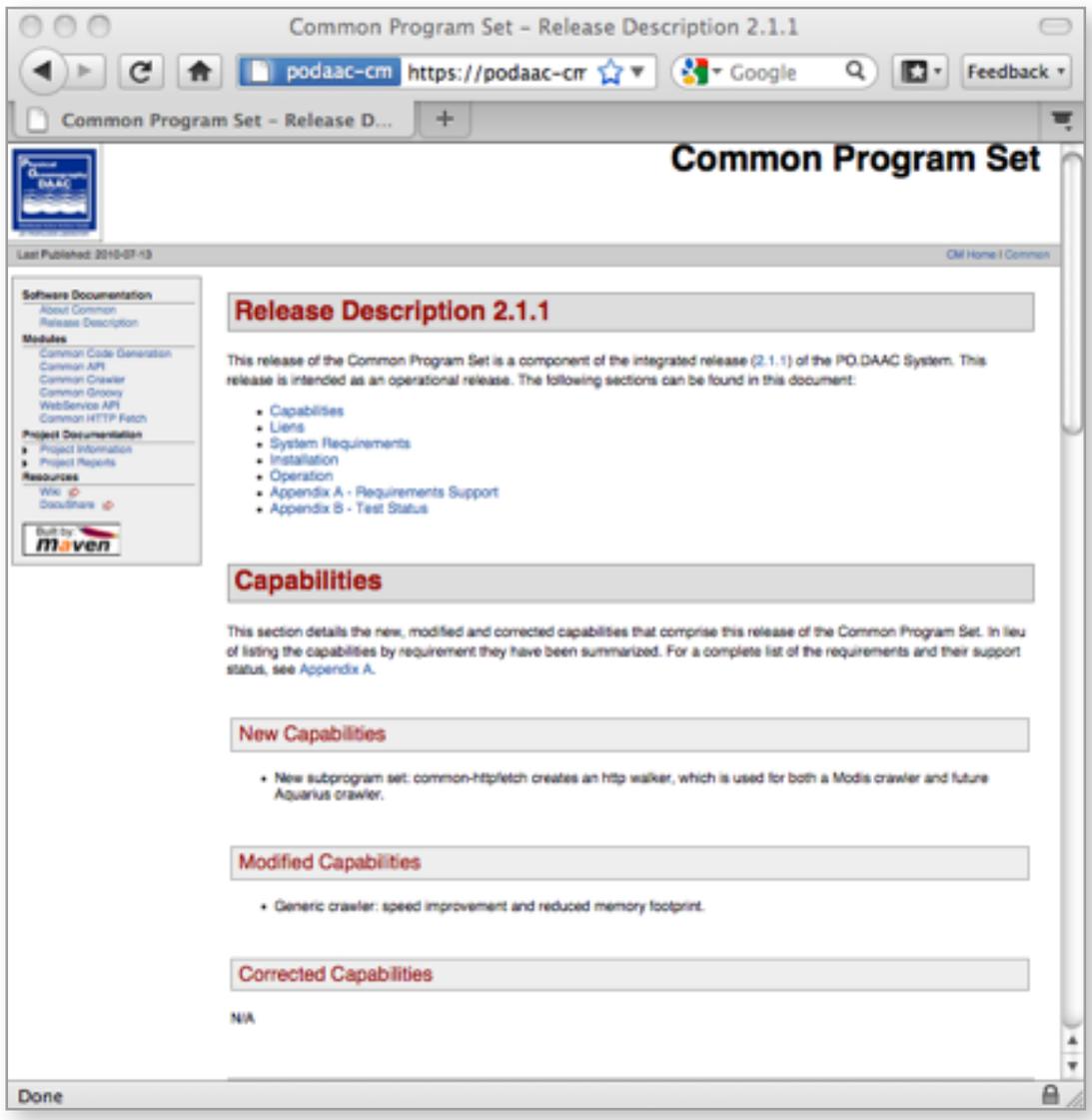
What


Popularity
Aquarius
Bug Fixes
Vanhoff
Product Type/Sigevent



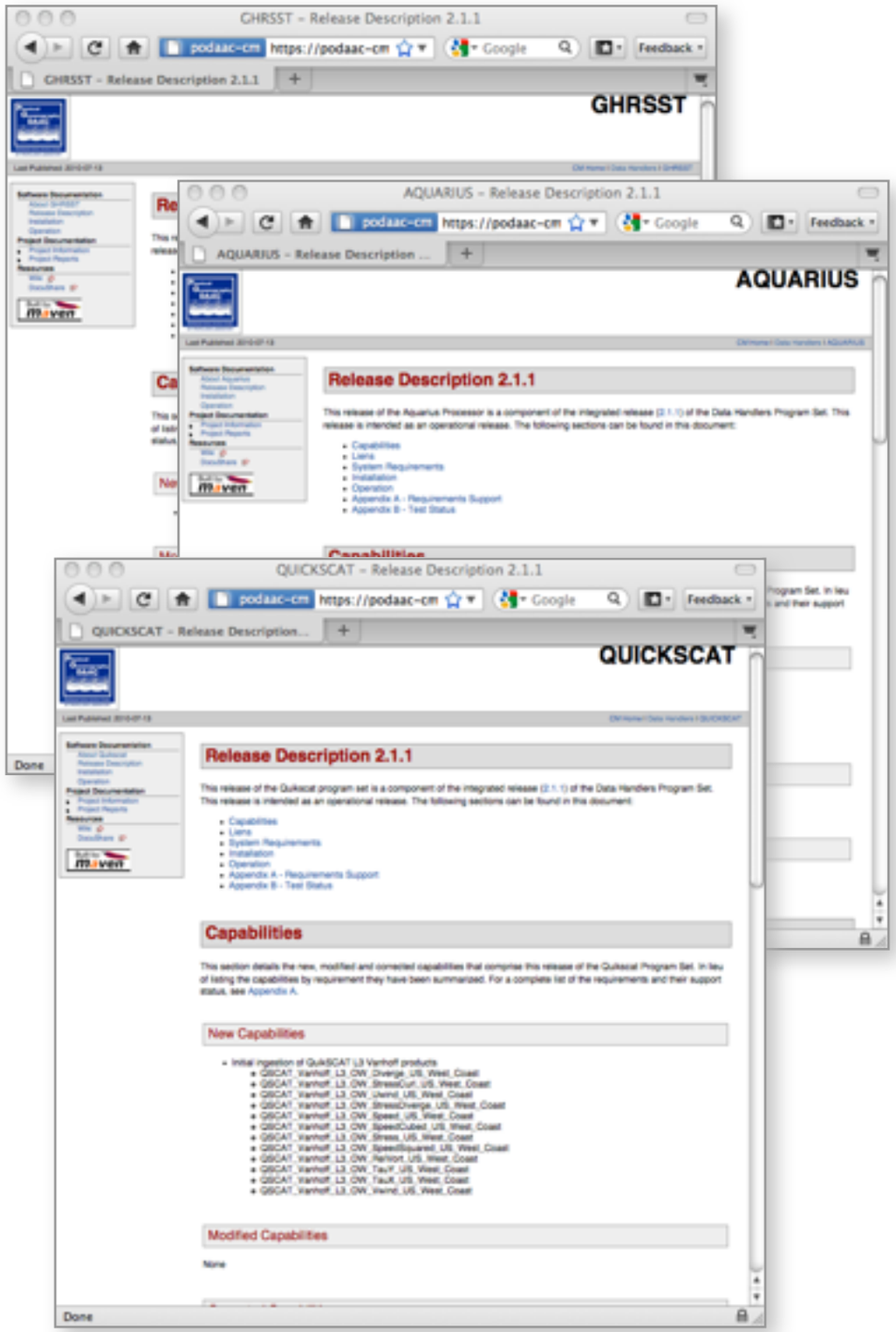
Architecture

Program Set: Common



- **New Capabilities**
  - New HTTP Data fetching library
- **Modified Capabilities**
  - Generic crawler updated to reduce memory footprint and improve data ingestion speed.

Program Set: Data Handlers



- **GHRSST**
  - **Corrected Capabilities**
    - Trac 499 - Update to GHRSST search page
    - Trac 504 - New case for relax_time_check option of Ghrsst Data handler
    - Removed any leading directory name in granule name.
    - Fixed bug with the wrong type of image file in SIP and bug with compression type.

- **Aquarius**
  - **Initial ingestion of the following datasets**
    - AQUARIUS_L2_SSS
    - AQUARIUS_L3_SSS_SMI_DAILY
    - AQUARIUS_L3_SSS_SMI_8DAY
    - AQUARIUS_L3_SSS_SMI_MONTHLY

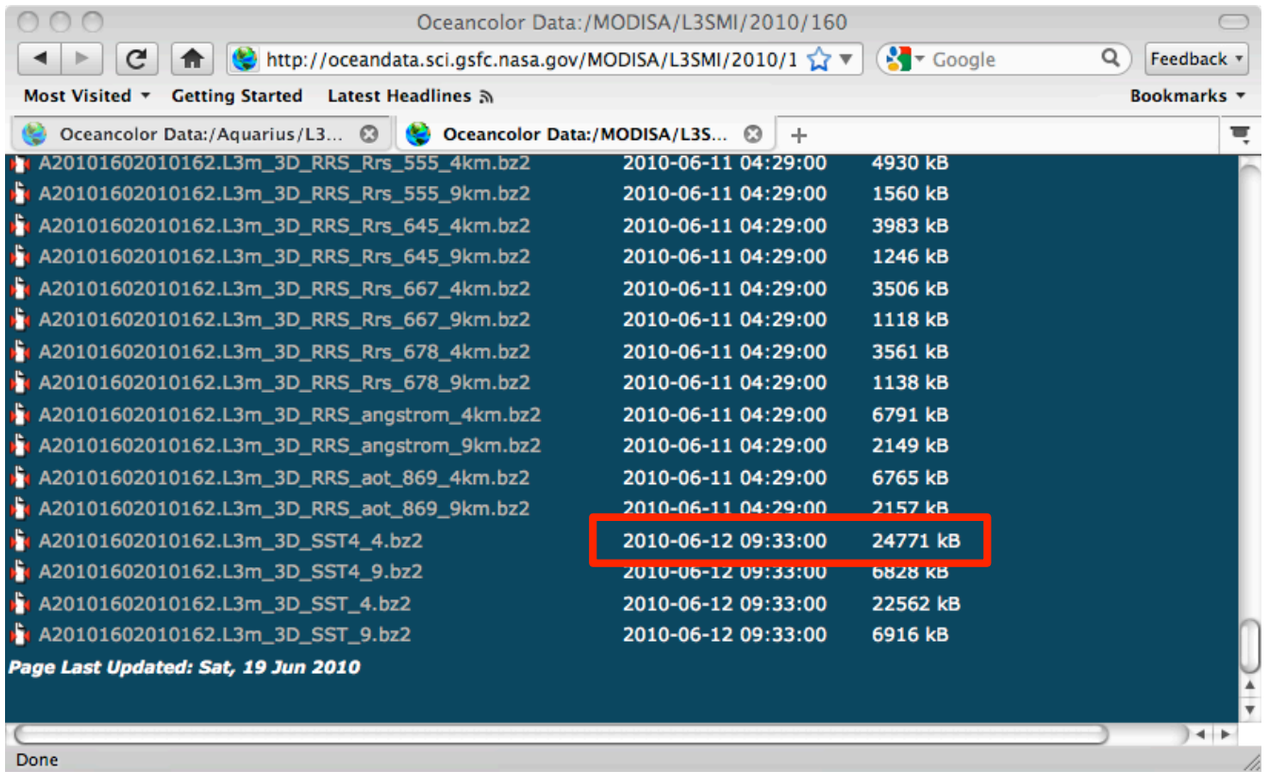
- **QuikSCAT L3 Vanhoff**
  - **Initial ingestion of the follow datasets**
    - QSCAT_Vanhoff_L3_OW_Diverge_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_StressCurl_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_Uwind_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_StressDiverge_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_Speed_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_SpeedCubed_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_Stress_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_SpeedSquared_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_RelVort_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_TauY_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_TauX_US_West_Coast
    - QSCAT_Vanhoff_L3_OW_Vwind_US_West_Coast

Issue 1: General observation on Time and size issues

- It appears the 'Last Modified' time shown on the HTML page is of local time (i.e. Eastern time zone)
- The size of the file is rounded to the nearest kilobyte

Comments

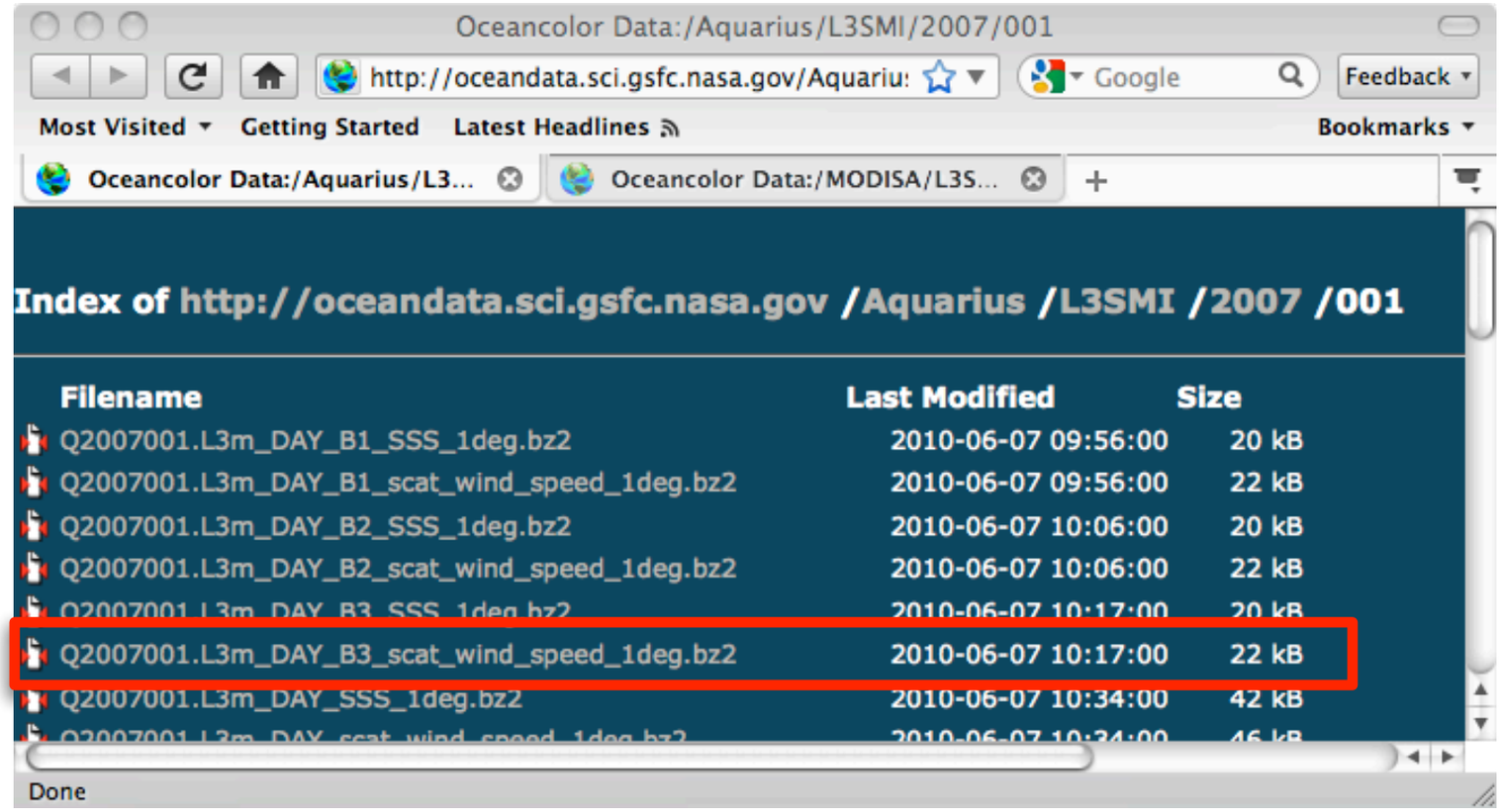
- We could perform the time conversion from Eastern Time Zone to GMT during ingestion. However, the Goddard site should consider using standard time value for scientific artifacts. At the minimum, the site should display Time Zone information.
- This size value is an issue, since the software can determine the exact size of the file unless either by physically download the file or by parsing the HTTP response header for the 'Content-Length'.




```
% ./getter.sh A20101602010162.L3m_3D_SST4_4.bz2
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100 24.1M  100 24.1M    0     0   393k      0  0:01:02  0:01:02 --:--:--  624k

HTTP/1.1 200 OK
Date: Tue, 13 Jul 2010 19:24:40 GMT
Server: OBPG/1.0
Content-Disposition: attachment; filename=A20101602010162.L3m_3D_SST4_4.bz2
Content-Length: 25365658
Last-Modified: Sat, 12 Jun 2010 13:33:02 GMT
Content-Type: application/octet-stream

-rw-r--r--  1 thuang  staff  25365658 Jul 13 12:25
A20101602010162.L3m_3D_SST4_4.bz2
```

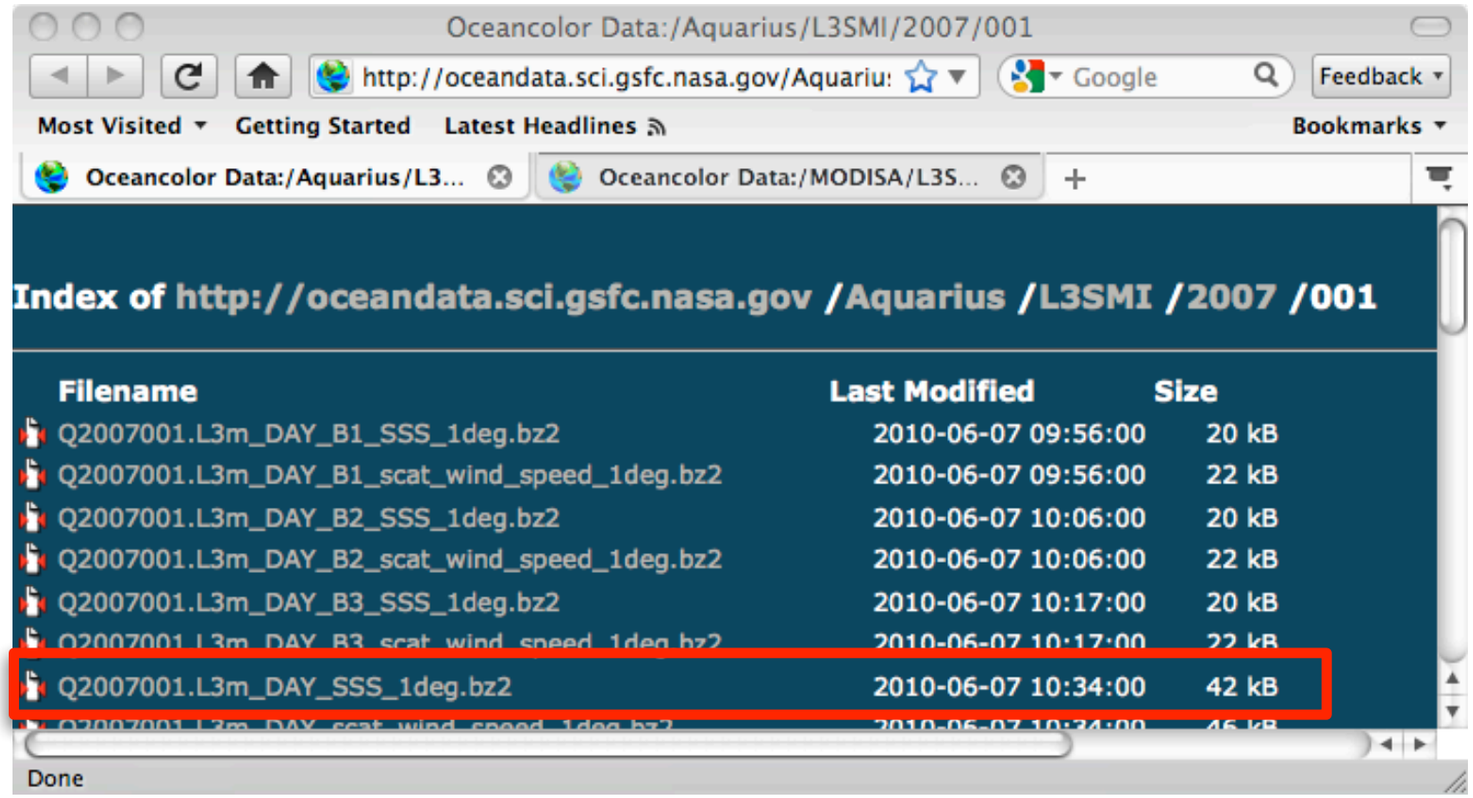
MODIS/Aquarius Issues

Issue 2: Invalid file modification time and size




```
% ./getter.sh Q2007001.L3m_DAY_B3_scat_wind_speed_1deg.bz2
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100  5511  100  5511    0     0  11133      0 --:--:-- --:--:-- --:--:-- 15745

HTTP/1.1 200 OK
Date: Tue, 13 Jul 2010 19:06:16 GMT
Server: OBPG/1.0
Content-Disposition: attachment;
filename=Q2007001.L3m_DAY_B3_scat_wind_speed_1deg.bz2
Content-Length: 5511
Last-Modified: Fri, 09 Jul 2010 16:26:14 GMT
Content-Type: application/octet-stream

-rw-r--r--  1 thuang  staff  5511 Jul 13 12:06
Q2007001.L3m_DAY_B3_scat_wind_speed_1deg.bz2
```

Issue 3: File listed on HTML page, but missing when trying to download

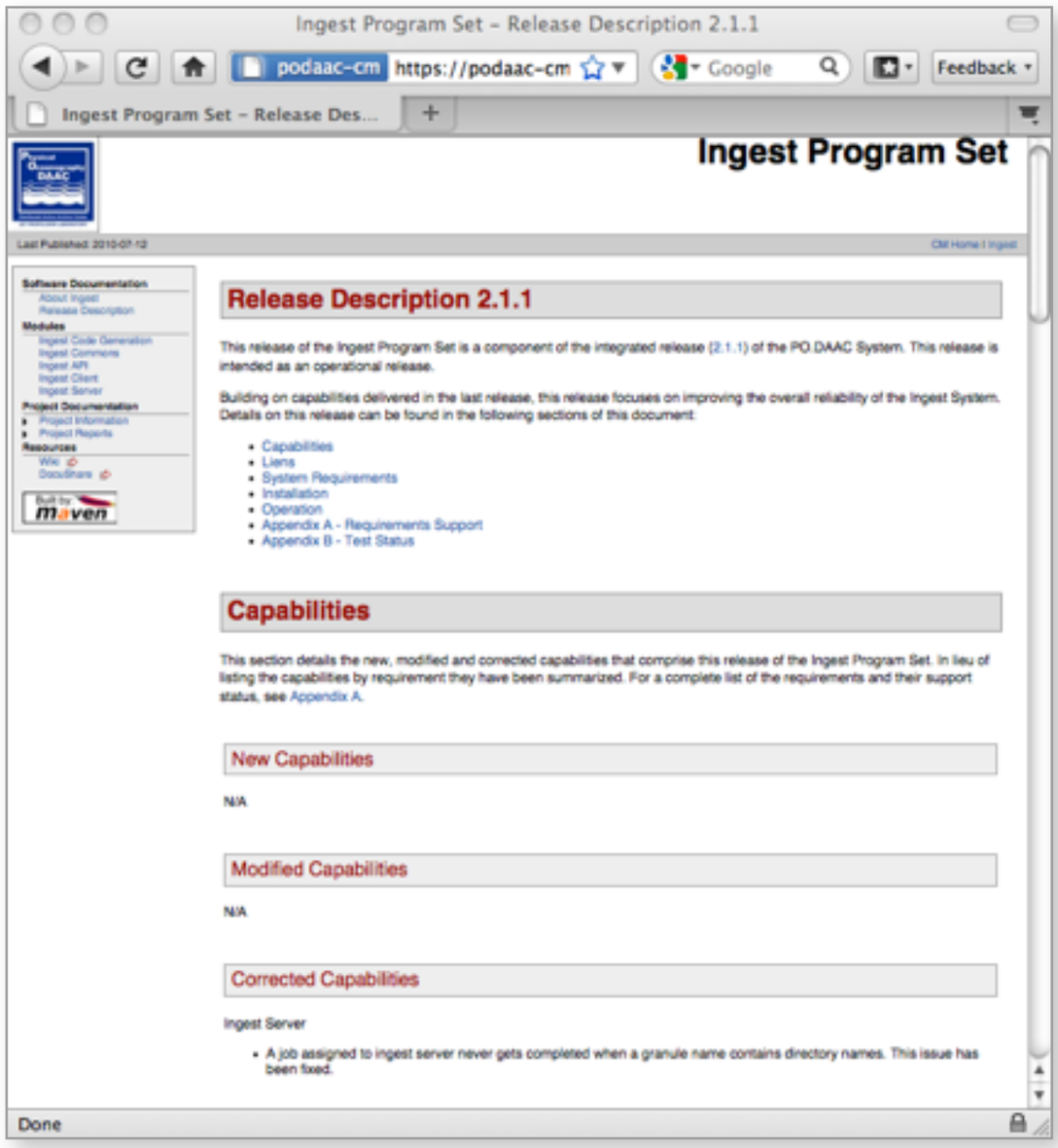



```
% ./getter.sh Q20070012007001.L3m_DAY_SSS_1deg.bz2
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0

HTTP/1.1 200 OK
Date: Tue, 13 Jul 2010 19:17:21 GMT
Server: OBPG/1.0
Content-Disposition: attachment; filename=Q20070012007001.L3m_DAY_SSS_1deg.bz2
Last-Modified: Tue, 13 Jul 2010 19:17:21 GMT
Content-Length: 0
Content-Type: application/octet-stream
X-Pad: avoid browser bug

ls: Q20070012007001.L3m_DAY_SSS_1deg.bz2: No such file or directory
rm: Q20070012007001.L3m_DAY_SSS_1deg.bz2: No such file or directory
```

The same problem with some MODIS data

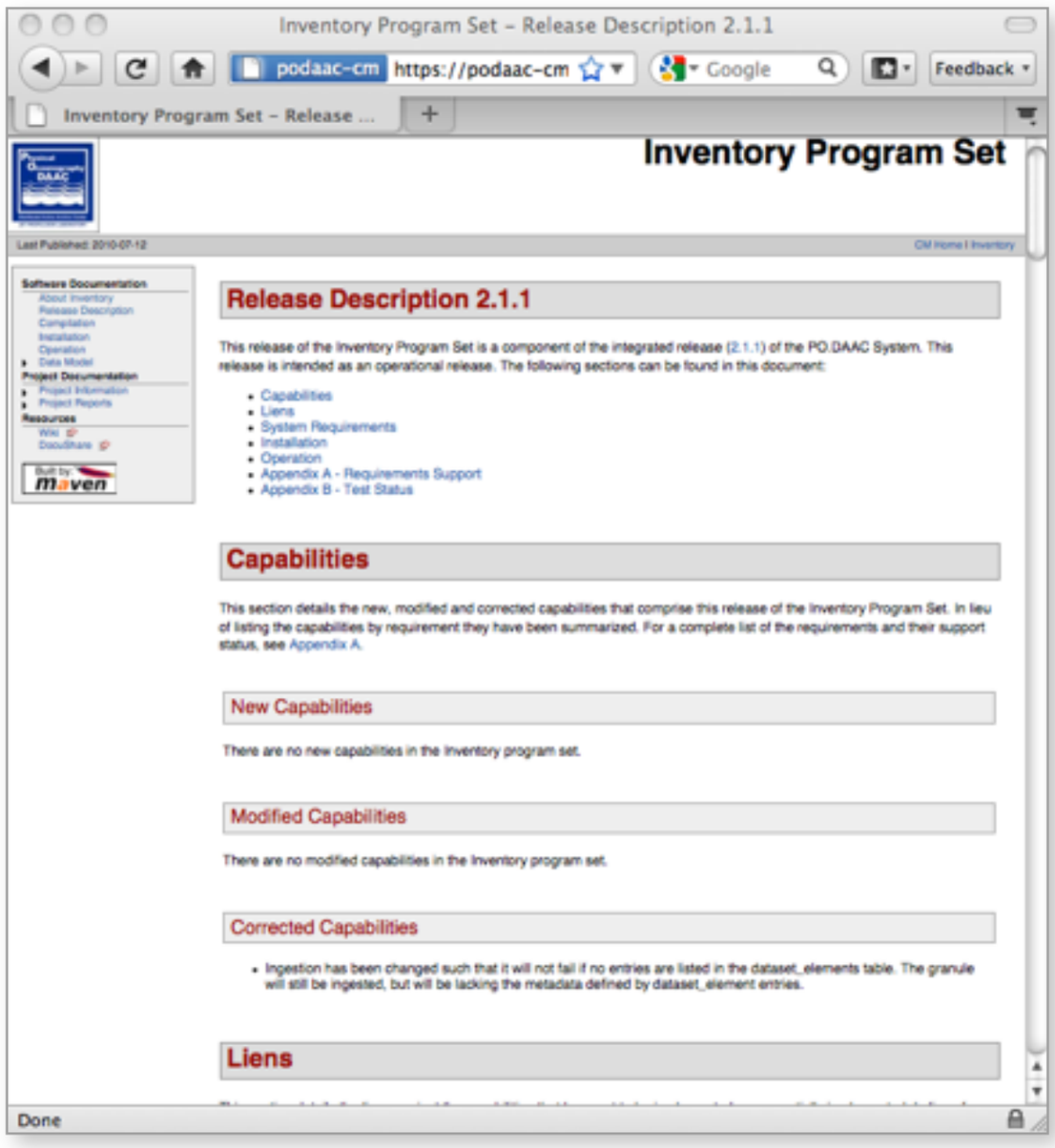
http://oceandata.sci.gsfc.nasa.gov/MODISA/L3SMI/2010/102/A2010102201010 4.L3m_3D_SST*.bz2 files.

Program Set: Ingest



- **Ingest Server**
  - **Corrected Capabilities**
    - A job assigned to ingest server never gets completed when a granule name contains directory names.

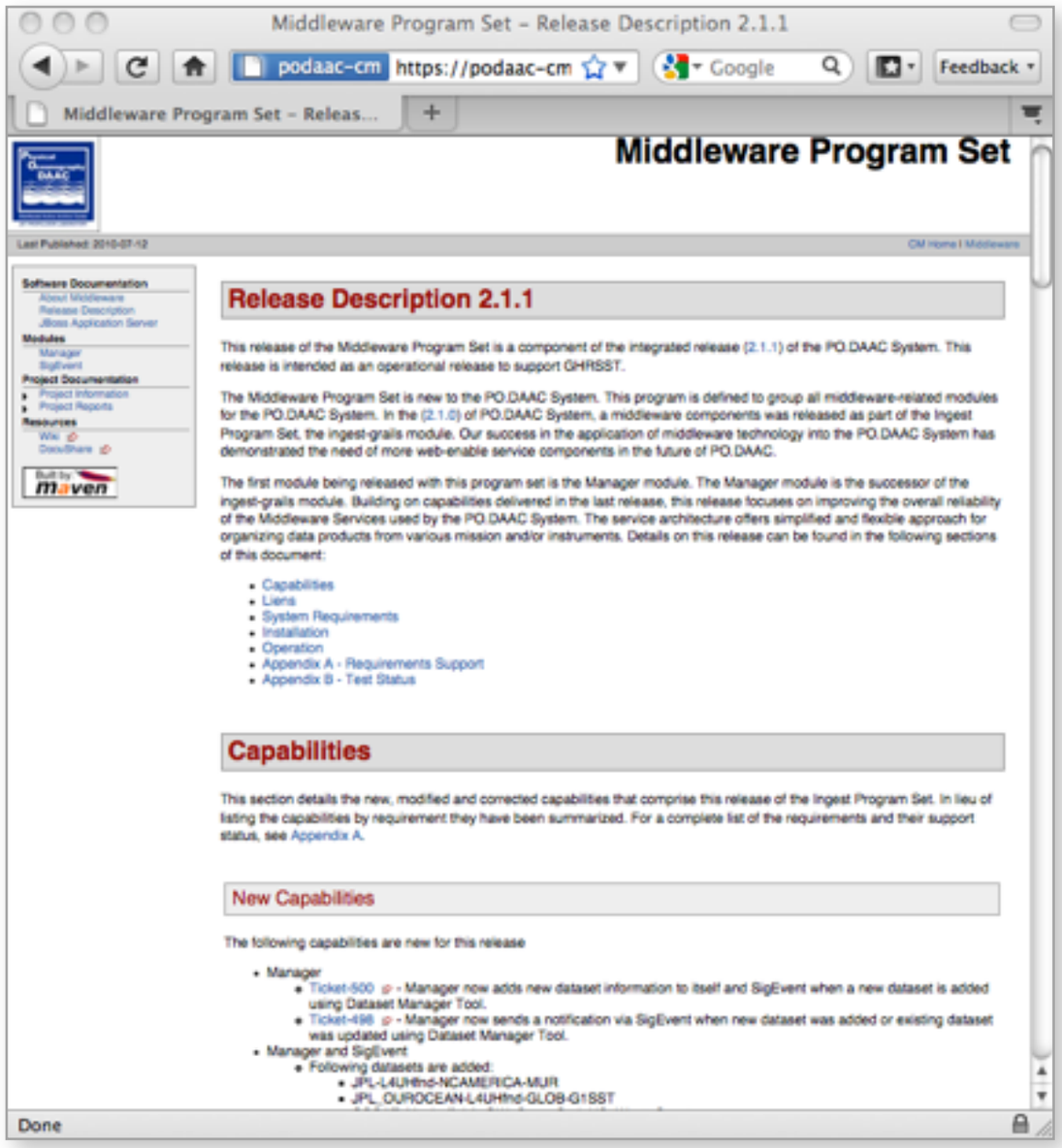
Program Set: Inventory



- **Corrected Capabilities**
  - Ingestion has been changed such that it will not fail if no entries are listed in the dataset_elements table. The granule will still be ingested, but will be lacking metadata defined by the dataset_element entries.

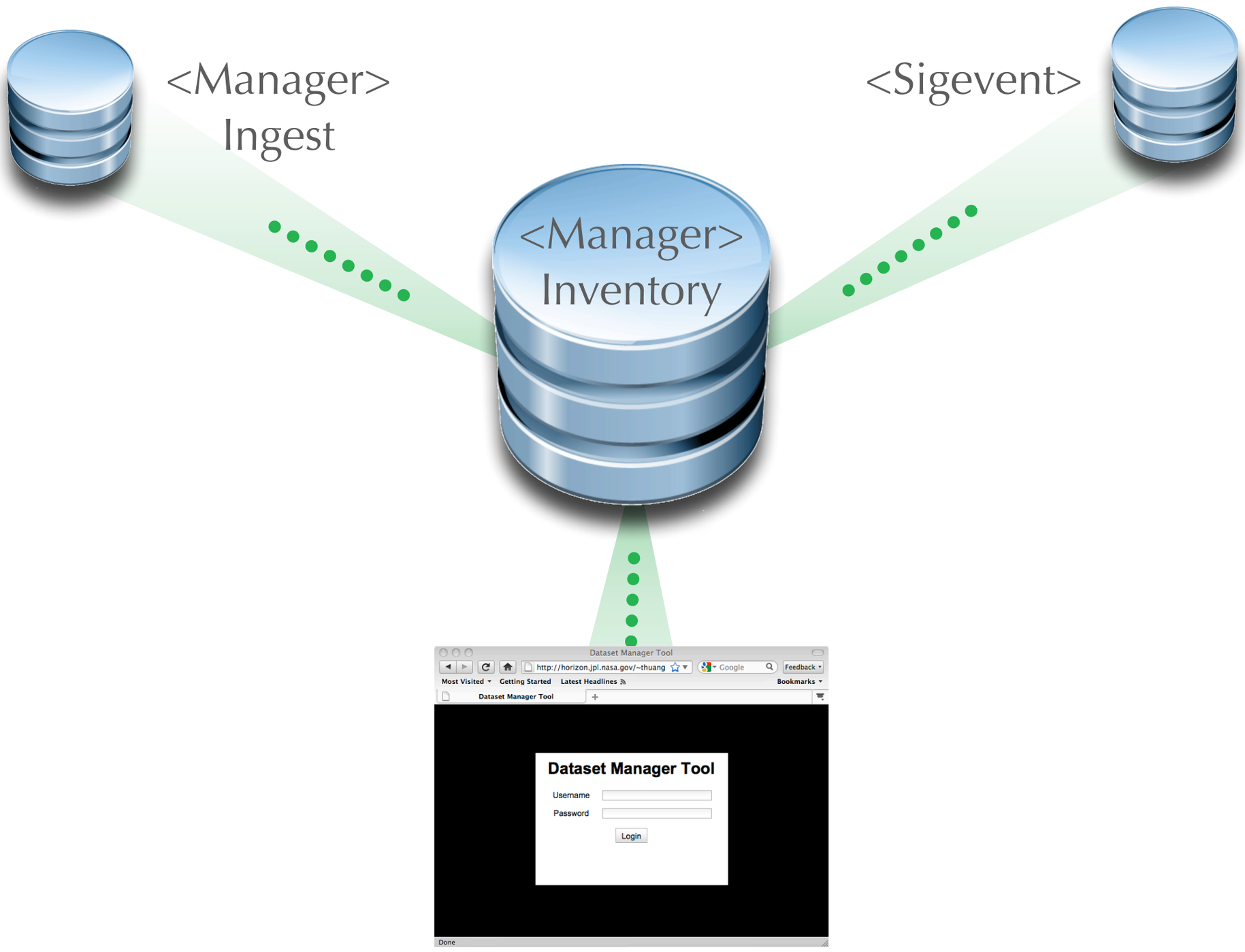
Action Item:
The Dataset Manager Tool still needs to be updated to allow DEs to specify granule-level data element. A ticket will be created for this missing feature.

Program Set: Middleware

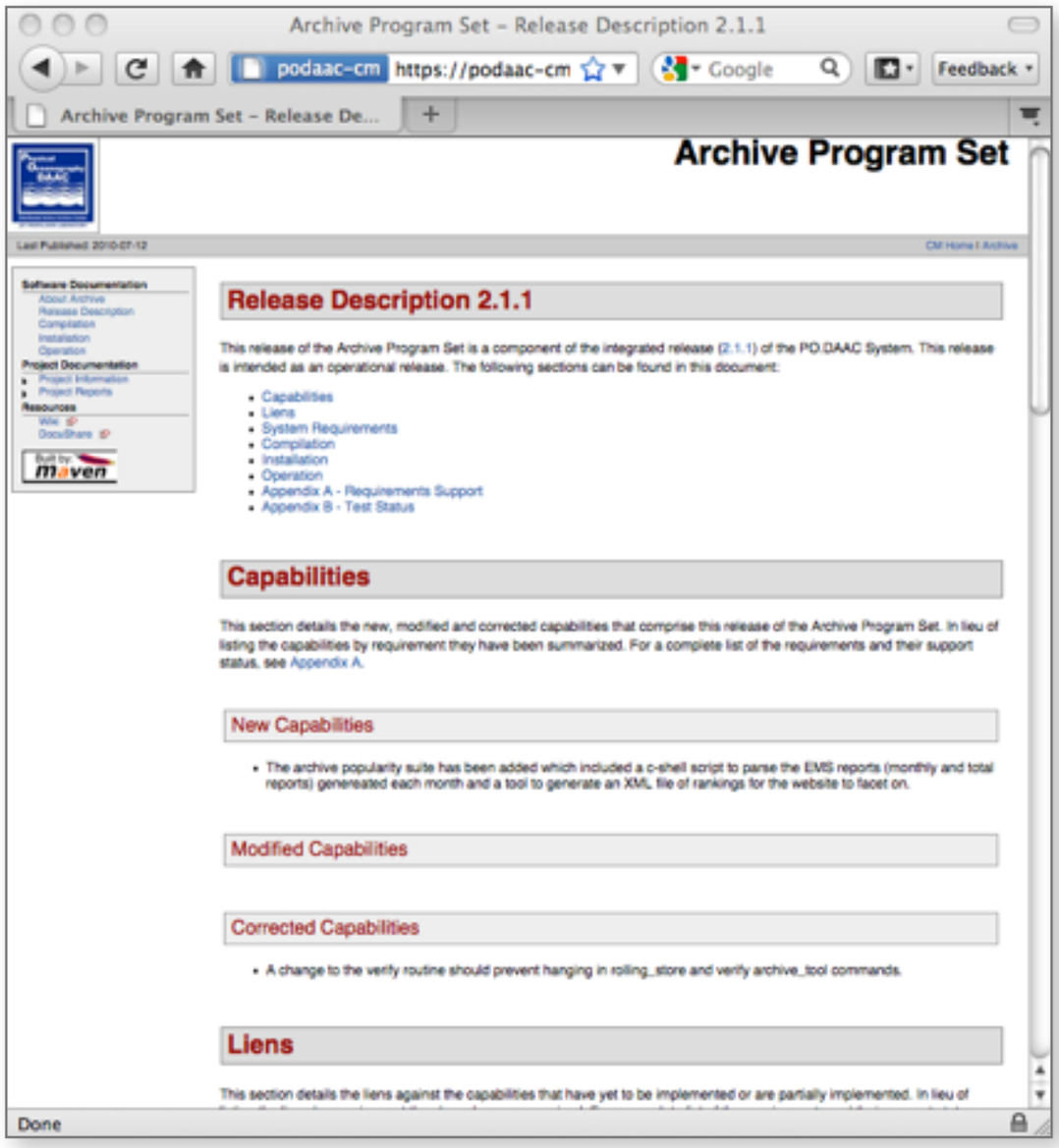


- **New Capabilities**
  - **Manager WS**
    - Trac 500 - Manager now adds new ingest product types and Sigevent category when new dataset is added using the Dataset Manager Tool
  - **Manager WS and Sigevent WS**
    - Added support for Aquarius, Vanhoff, and the new GHRSST datasets.
- **Modified Capabilities**
  - **Manager WS and Sigevent WS**
    - GHRSST ingest product types have been renamed to drop the GHRSST- prefixe. The product types are now the same as dataset short names.
- **Corrected Capabilities**
  - Trac 511 - A bug was found in the 1.4.2 to 2.0.0 SQL time conversion routine. An SQL patch file has been provided to correct this issue

Automation -
Creation of Product Type and Sigevent Category



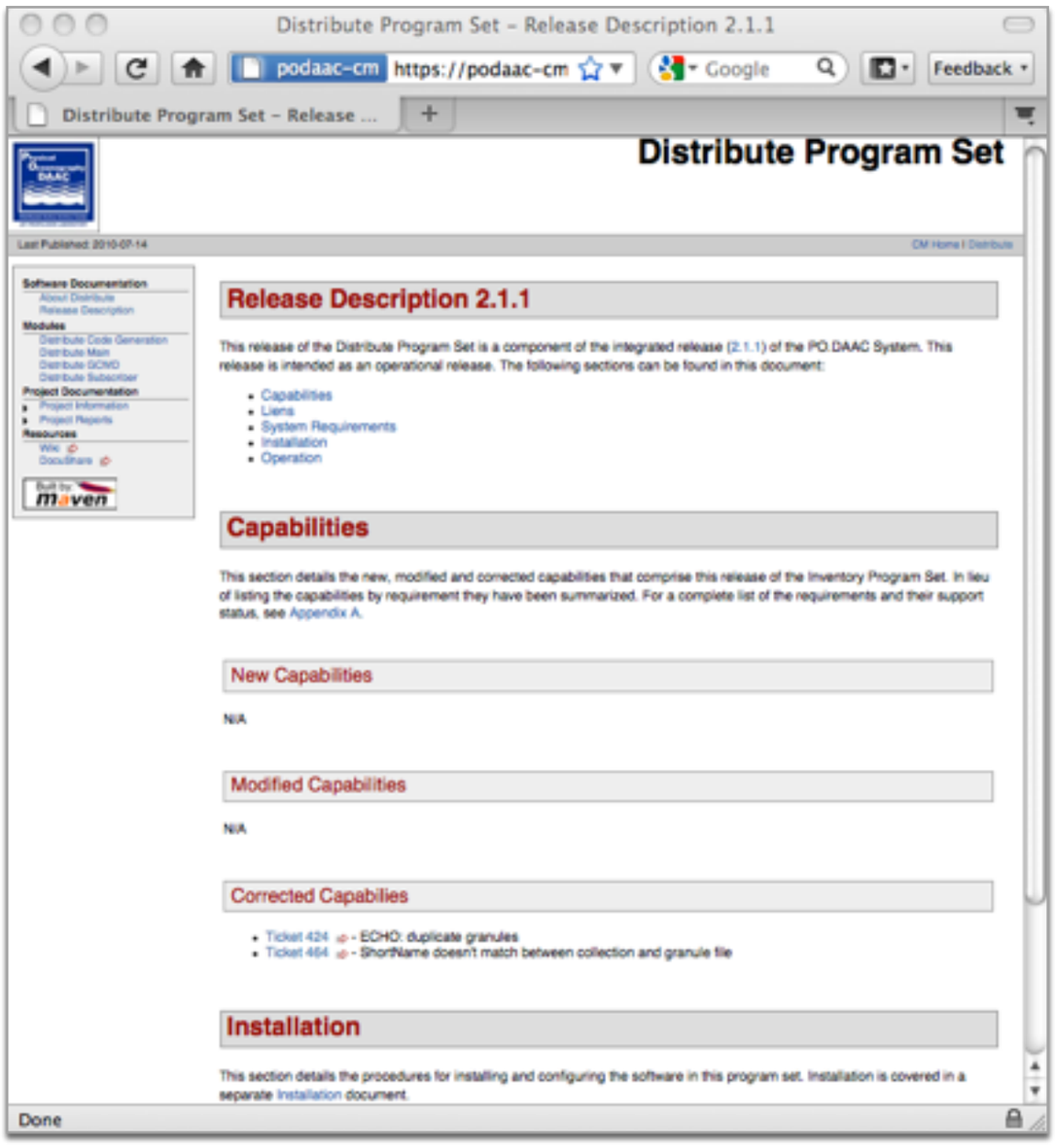
Program Set: Archive



- **New Capabilities**
  - The archive popularity suite has been added which included a C-shell script to parse the EMS reports (monthly and total reports) generated each month and a tool to generate an XML file of rankings for the Solr index generation.

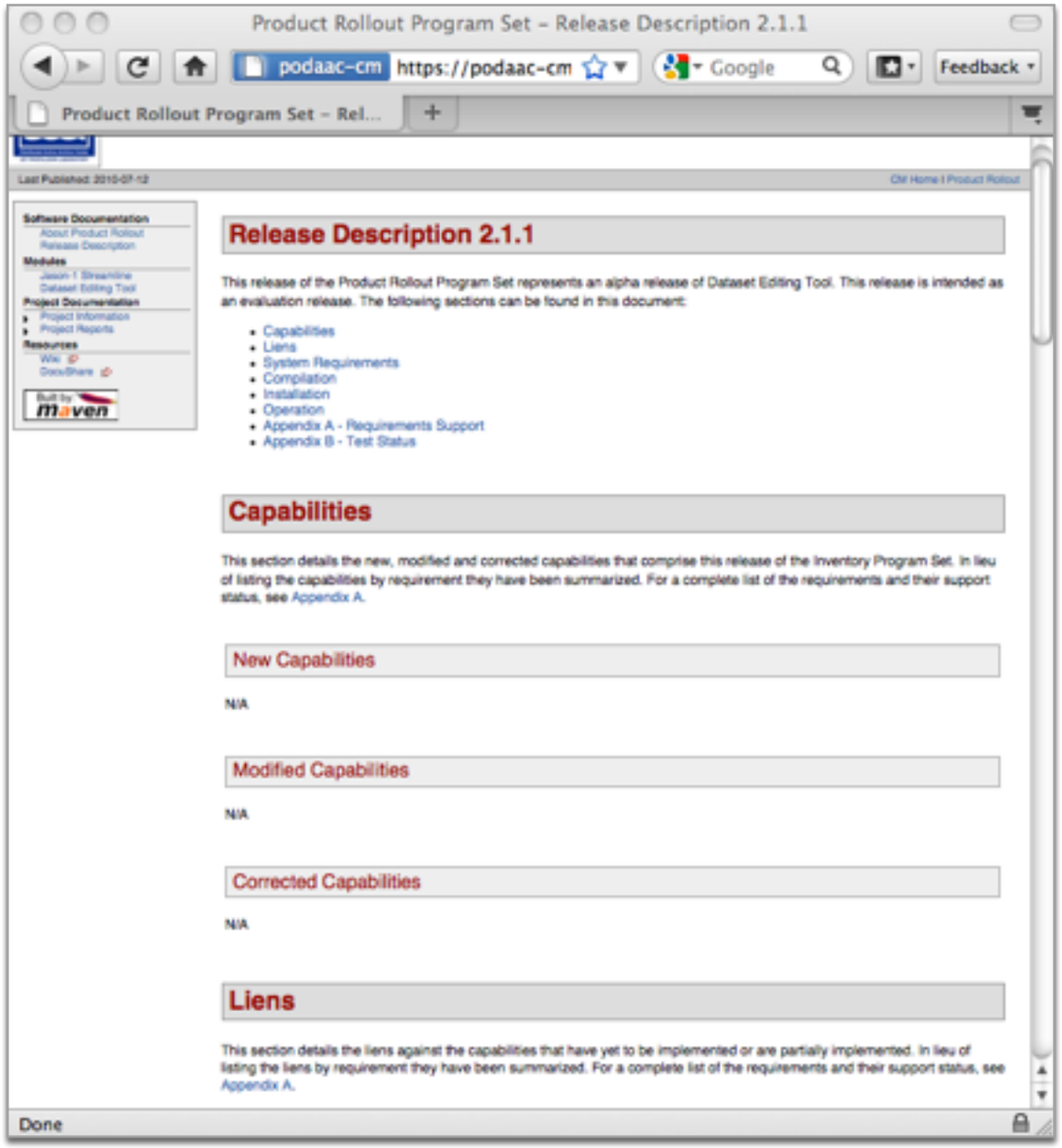
- **Corrected Capabilities**
  - Updated the verify routine to prevent hanging in rolling_store and verify archive_tool commands.

Program Set: Distribute

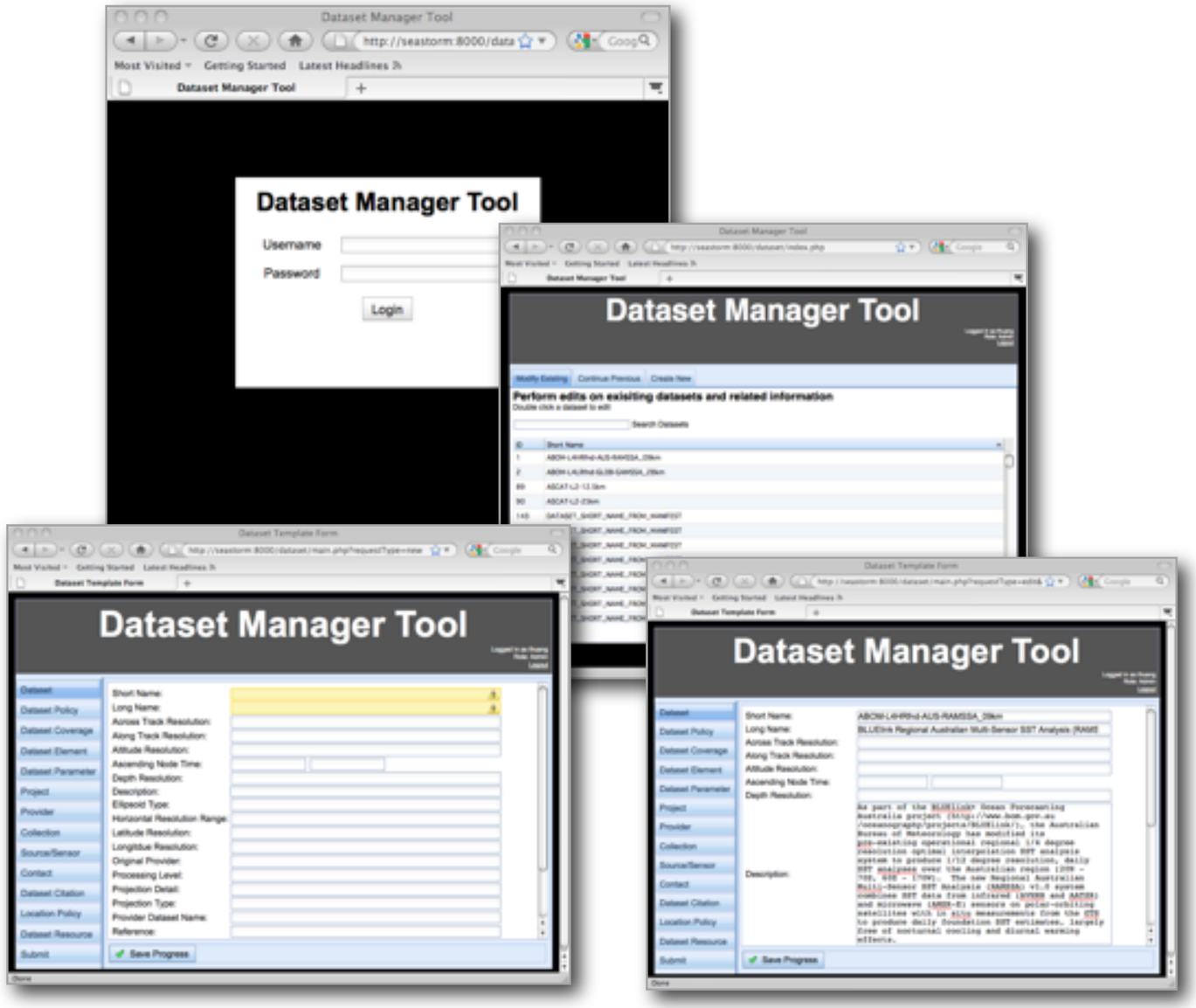


- **Corrected Capabilities**
  - Trac 426 - ECHO: duplicate granules
  - Trac 464 - ShortName doesn't match between collection and granule file

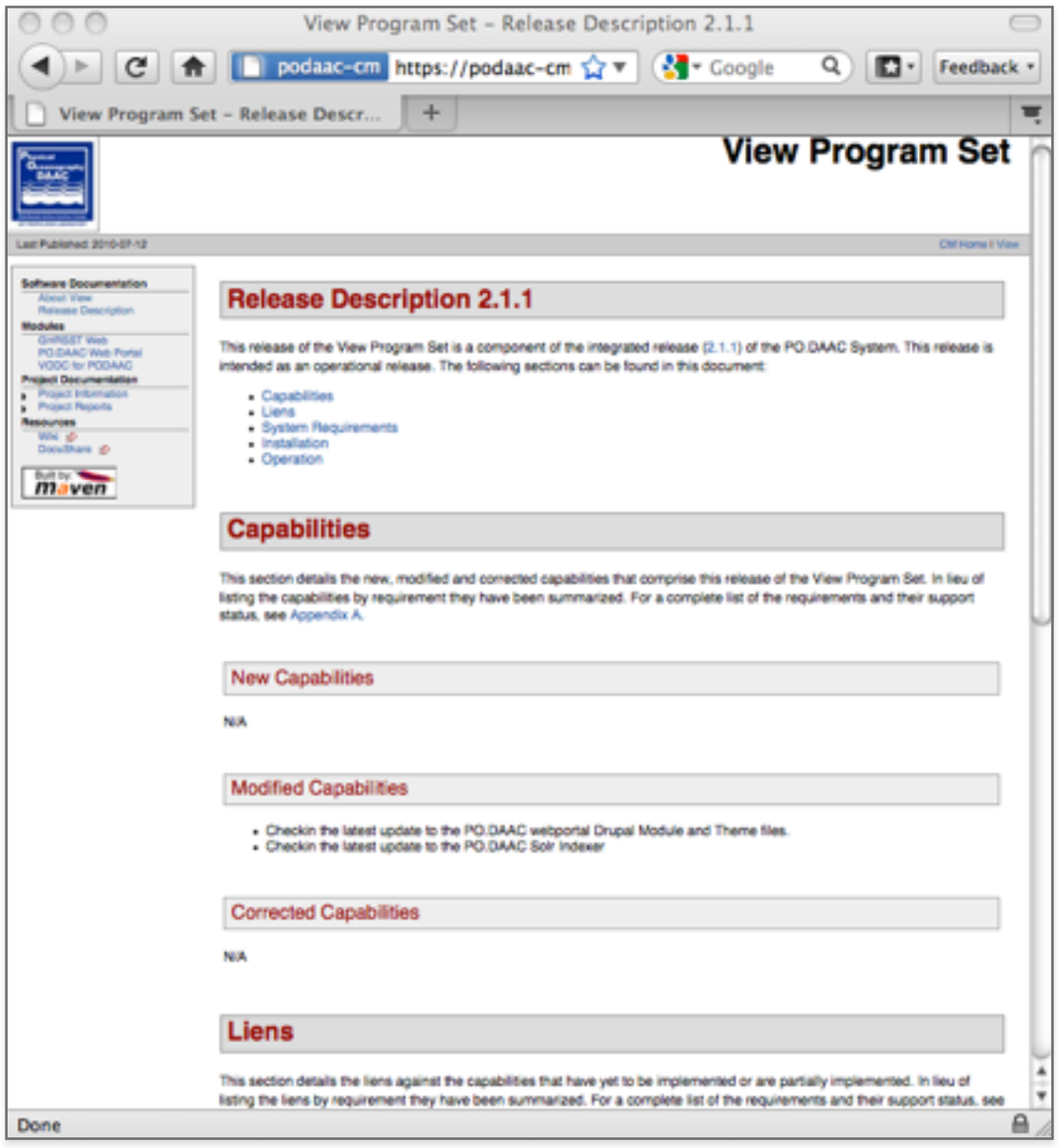
Program Set: Product Rollout



- **No significant changes**



Program Set: View

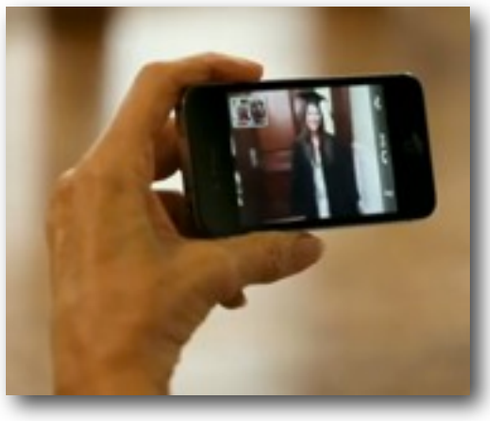


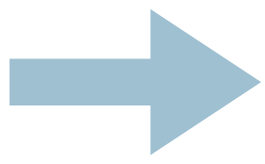
- **Modified Capabilities**
  - Checkin the latest updates to the PO.DAAC web portal Drupal Module and Theme files.
  - Checkin the latest update to the PO.DAAC Solr indexer
- **Work-in-progress**
  - http://seastorm.jpl.nasa.gov:8000/drupal

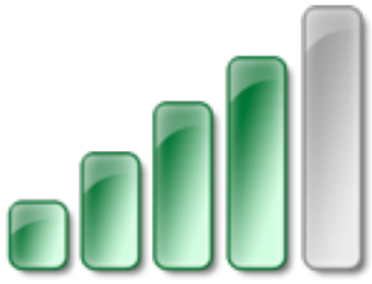
Internal Review

List of Closed Trac Tickets

Num	Title
378	GHRSST specific remove noaa symlink after granule rolled off
424	ECHO: duplicate granules
464	ShortName doesn't match between collection and granule file
498	Dataset manager Tool - submit notification
499	Update to GHRSST search page
500	add new datasets to non-inventory tables
504	New case for relax_time_check option of Ghrsst Data handler
511	incorrect time values for non-inventory tables
512	Add missing Product IDs in collections to support dataset popularity







?

Physical Oceanography DAAC

Distributed Active Archive Center

JET PROPULSION LABORATORY