



EOSDIS

NASA'S EARTH OBSERVING SYSTEM
DATA AND INFORMATION SYSTEM

Earthdata Search: Scaling, Assessing, and Improving Relevancy

IN21D - December 13, 2016

Mark Reese

mreese@element84.com

EED2 Program



The material is based upon work supported by the National Aeronautics and Space Administration under Contract Number **NNG15HZ39C**

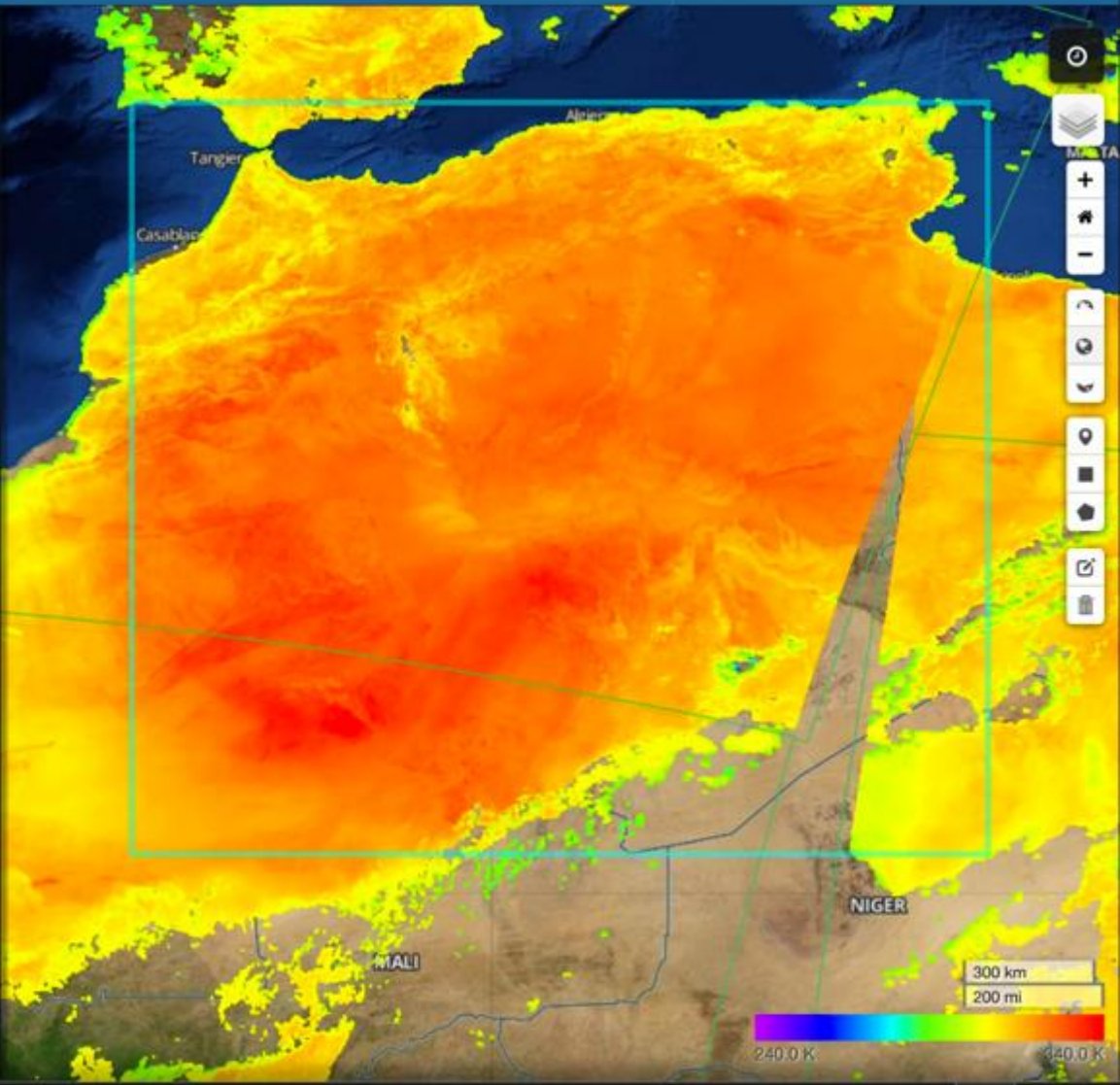
Back to Collections

MODIS/Terra Near Real Time (NRT) Land Surface Temperature/Emissivity 5-Min L2 Swath 1km

Retrieve Collection Data

Showing 4 of 4 matching granules for the selected day. (Show All)
Sort by: Search Time: 0.2s Report a metadata problem

- MOD11_L2.A2016162.1100.006.2016162120416.NRT.hdf**
2016-06-10T11:00:00Z to 2016-06-10T11:05:00Z
- MOD11_L2.A2016162.1055.006.2016162120412.NRT.hdf**
2016-06-10T10:55:00Z to 2016-06-10T11:00:00Z
- MOD11_L2.A2016162.0920.006.2016162103246.NRT.hdf**
2016-06-10T09:20:00Z to 2016-06-10T09:25:00Z
- MOD11_L2.A2016162.0915.006.2016162102959.NRT.hdf**
2016-06-10T09:15:00Z to 2016-06-10T09:20:00Z



Late 2015: Our catalog grew

7053 Matching Collections

Add collections to your project to compare

32960 Matching Collections

Add collections to your project to compare

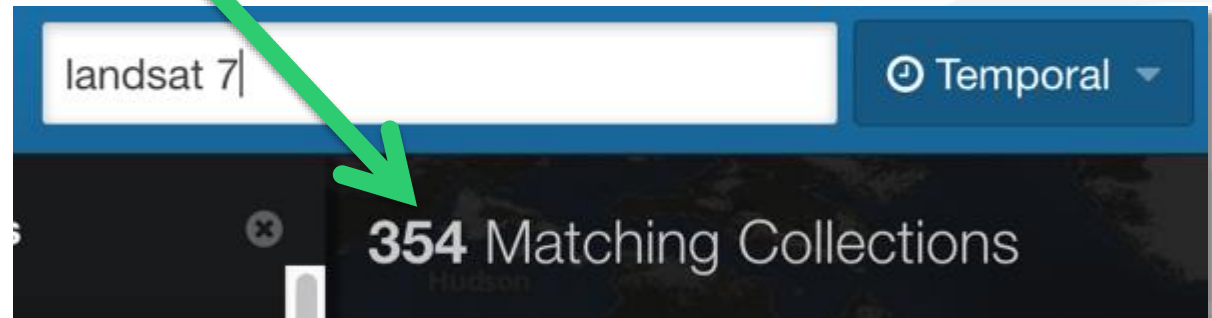
... as did our relevancy needs



landsat 7 Temporal

59 Matching Collections

This screenshot shows a search interface with a text input field containing 'landsat 7' and a 'Temporal' filter button. Below the input, a dark bar displays '59 Matching Collections' with a small 'x' icon to the left.



landsat 7| Temporal

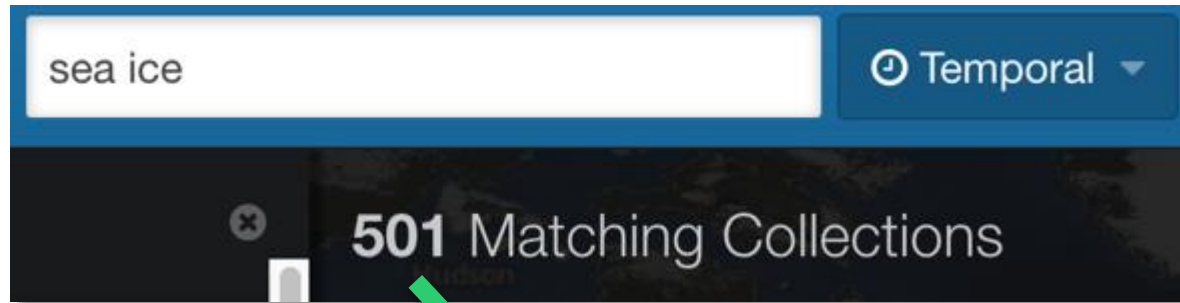
354 Matching Collections

This screenshot shows the same search interface as above, but with the text input field containing 'landsat 7|'. The results bar now displays '354 Matching Collections'. A green arrow points from the '59 Matching Collections' in the first screenshot to the '354 Matching Collections' in this one.

... as did our relevancy needs

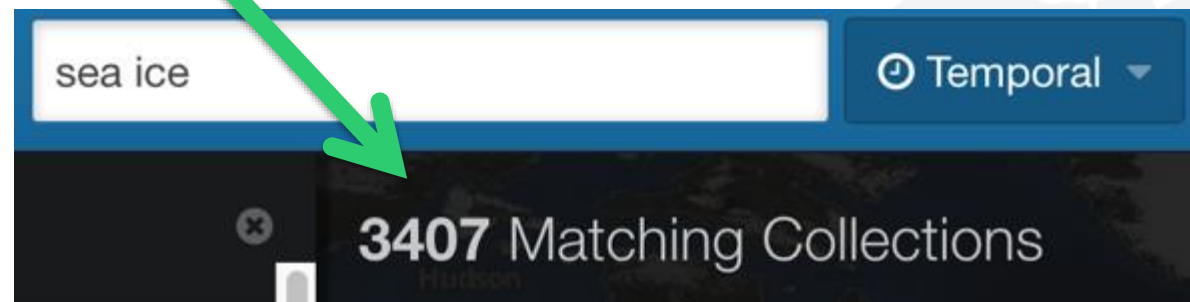
sea ice 🕒 Temporal ▾

✕ 501 Matching Collections



sea ice 🕒 Temporal ▾

✕ 3407 Matching Collections



Facets felt the strain

GROUND STATIONS	1120
GROUND-BASED OBSERVATIONS	1269
LABORATORY	708
LANDSAT	277
LANDSAT-5	206
LANDSAT-7	184
MAPS	352
METEOROLOGICAL STATION	106



Users echoed the need for better relevance

Our Method



Survey

Initial Questionnaire

Before the in-
complete a s
familiarity wit

Findings

Collections

Search Relevance

Clear-cut, highly-specific searches do not turn up what users need as the top result. Less specific searches fare worse.



If users can't find the collections
they need,
none of our other work matters



Earthdata Search

DISCOVERY CHALLENGES BEYOND RELEVANCE

Metadata is imperfect

Processing level	
1	204
1A	24
1B	219
1T	10
L1T	6
Level 1	7
NA	2
Organization	


BIOMASS	1
BIOOSPHERE	2
BIOSPHERE	839

Metadata

Image: 2 Snow Cover 5-Min L2 500m...

Center: S

Center: APS



AfSIS MODIS C
CIESIN_AfSIS_MODIS
2000-02-01 to 2012-0

No image available

AfSIS MODIS C
CIESIN_AfSIS_MODIS
2001-01-01 to 2009-1

No image available

AfSIS MODIS C
CIESIN_AfSIS_MODIS
2000-02-01 to 2012-0

No image available



AfSIS MODIS C
Release
MODIS
2014-0

S C
MODIS
2010-1

AfSIS MODIS C
CIESIN_AfSIS_MODIS
2000-02-01 to 2014-0

No image available

Interface Matters

**WIST**
Powered by 


[+Comments, Questions, or Problems?*](#)
[+User Support](#)
[+Tutorial*](#)
[+Help for this page*](#)

guest

[View all*](#) **Alert from LAADS: MODIS Level 1 and Atmospheres data products are available for ordering from WIST**

Search Results Granule List Folder Shopping Cart

Primary Data Search

 **Users may experience time out problems using Internet Explorer to login to WIST. If an IE timeout occurs while logging in, press the back button and re-login to WIST.**

[Save/Restore search](#) | [Clear search](#)

Login

User Name: Password:

[Create Account](#) | [\(Forgotten your password?\)](#) | [\(Forgotten your User Name?\)](#)

Choose Keywords for One or More Categories Text Search: [Help](#)

*Pick a discipline/topic (for example: Atmosphere:TRMM), then choose from the list of data sets.
For multiple topics: choose one topic & data sets, then the next topic & data sets.
To select/deselect* more than one data set, use Ctrl-click for PCs; Apple-click for Macintosh.*

Atmosphere:	Cryosphere:	Land:	Oceans:	Solar/Other:
<input type="radio"/> Aqua AIRS/AMSU-A/HSB <input type="radio"/> MISR	<input type="radio"/> AMSR/AMSR-E	<input type="radio"/> AMSR/AMSR-E	<input type="radio"/> AMSR/AMSR-E	<input type="radio"/> ACRIM
<input type="radio"/> AMSR/AMSR-E <input type="radio"/> MLS/Aura	<input type="radio"/> GLAS/CESat	<input type="radio"/> ASTER	<input type="radio"/> GLAS/CESat	<input type="radio"/> Field/In Situ

Earthdata Search

WHAT HAVE WE DONE?


Relevance: Temporal Overlap

“Aerosols” November 1st – 30th, 2016

315 Matching Collections

Add collections to your project to compare and retrieve their data. [? Learn More](#)

Search Time: 0.8s [Report a metadata problem](#)

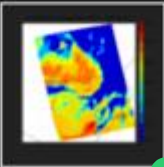


AIRS/Aqua L1B Infrared (IR) geolocated and calibrated radiances V005 (AIRIBRAD) at GES DISC

AIRIBRAD v005 - NASA/GSFC/SED/ESD/GCDC/GESDISC

2002-08-30 ongoing | 7177 Granules

[i](#) [+](#)




AIRS/Aqua L1C Infrared (IR) resampled and corrected radiances V006 (AIRICRAD) at GES DISC

AIRICRAD v006 - NASA/GSFC/SED/ESD/GCDC/GESDISC

2015-11-22 ongoing | Granules

[i](#) [+](#)



CAL_LID_L1-Standard-V4-00

CAL_LID_L1-Standard-V4-00 vV4-00 - Atmospheric Science Data Center

2006-06-12 to 2016-07-01 | 0 Granules

[i](#) [+](#)

**PLACEHOLDER! PLACEHOLDER!
PLACEHOLDER!**

The right hand side of this slide will show search results after temporal overlap has been implemented in CMR which will alter the result order to exclude the second result.

**PLACEHOLDER! PLACEHOLDER!
PLACEHOLDER!**

Relevance: Newer versions first

Version 6 Near-Realtime

Thumbnail	Title	ID	Date	Granules
No image available	MODIS/Terra Near Real Time (NRT) Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km	MOD05_L2 v6NRT - NASA/GSFC/EOS/ESDIS/LANCEMODIS	2015-12-06 ongoing	2873
	MODIS/Terra Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km V006	MOD05_L2 v6 - NASA/GSFC/SED/ESD/HBSL/BISB/LAADS	1999-12-18 ongoing	1682654
	MODIS/Terra Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km V5.1	MOD05_L2 v5.1 - NASA/GSFC/SED/ESD/HBSL/BISB/LAADS	1999-12-18 ongoing	1680564
No image available	MODIS/Terra Near Real Time (NRT) Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km (Collection 005)	MOD05_L2 v5NRT - NASA/GSFC/EOS/ESDIS/LANCEMODIS	1999-12-18 ongoing	2873


Version 6

Version 5.1

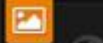
Version 5 Near-Realtime

Relevance: Collections with granules first


2,862 Granules

 **MODIS/Aqua Near Real Time (NRT) Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km (Collection 005)**
No image available
MYD05_L2 v5NRT - NASA/GSFC/EOS/ESDIS/LANCEMODIS
2002-05-04 ongoing | 2862 Granules


1,459,806 Granules

 **MODIS/Aqua Total Precipitable Water Vapor 5-Min L2 Swath 1km and 5km V5.1**
No image available
MYD05_L2 v5.1 - NASA/GSFC/SED/ESD/HBSL/BISB/LAADS
2002-05-04 ongoing | 1459806 Granules

Collection only

 **MODIS/Terra Granule Level 2 Water Vapor Infrared Jpeg image**
No image available
MOBWIR v6 - NASA/GSFC/SED/ESD/HBSL/BISB/LAADS
2000-02-25 ongoing | Collection only

Collection only

 **MODIS/Terra Granule Level 2 Water Vapor Near Infrared Jpeg image**
No image available
MOBWSW v6 - NASA/GSFC/SED/ESD/HBSL/BISB/LAADS
2000-02-25 ongoing | Collection only

Facets: Humanizers

Keywords	
AGRICULTURE	1620
ATMOSPHERE	4265
BIOLOGICAL CLASSIFICATION	2999
BIOOSPHERE	4
BIOSPHERE	5025
CLIMATE INDICATORS	321
CRYOSPHERE	1879
HUMAN DIMENSIONS	2990
HYDROSPHERE	235
LAND SURFACE	3847
OCEANS	4511
PALEOCLIMATE	662
SOLID EARTH	2275
SPECTRAL ENGINEERING	1
SPECTRAL/ENGINEERING	1518



Keywords	
Agriculture	1620
Atmosphere	4265
Biological Classification	2999
Biosphere	5025
Climate Indicators	321
Cryosphere	1879
Human Dimensions	2990
Hydrosphere	235
Land Surface	3847
Oceans	4511
Paleoclimate	662
Solid Earth	2275
Spectral/Engineering	1518

Facets: Relevance

2D Coordinate System ^	
CALIPSO	54
MISR	40
MODIS T	8
MODIS SIN	95
WELD Alaska	5
WELD CONUS Tile	6
WRS-1	2
WRS-2	9



Interface Updates

The screenshot displays the EarthData Search interface. At the top, there are navigation links for "Data Discovery", "DAACs", "Community", and "Science Disciplines". The main header includes the "EARTHDATA Search" logo and a search bar with the placeholder text "Type any topic or collection name". To the right of the search bar are "Feedback" and "Login" buttons. Below the search bar is a map of Africa and the Middle East. A blue box indicates "19446 Matching Collections". Below the map, there are sorting options: "Sort by Relevancy" and a checkbox for "Hide collections without granules". A "Report a metadata problem" button is also present. The search results are listed below, showing two datasets:

- A minimal dif dataset**: 1 Granule • 1970-01-01 ongoing • summary of the dataset. Includes a "SUBSETTING" tag and a "minimal_dif_dataset - 52N" identifier. Buttons: "View Granules" and "+ Add to project".
- ACRIM III Level 2 Daily Mean Data V001**: 4 Granules • 1999-12-20 ongoing • This file contains Level 2 total solar irradiance in the form of daily means gathered by the ACRIM instrument on the ACRIMSAT satellite. Buttons: "View Granules" and "+ Add to project".

At the bottom of the interface, there is a footer with "searchlab-deploy-4", "Search Time: 3.3s", "NASA Official: Andrew Mitchell", "FOIA", "NASA Privacy Policy", "USA.gov", and "Earthdata Access: A Section 508 accessible alternative".

Earthdata Search

MEASURING SUCCESS

Improved Metrics

```
{"event": "search",  
  "data":{"page_size":"20","free_text":"sea ice","page_num":"1"}}
```

```
{"event": "filter",  
  "data":{"echo_collection_id":"C1000001538-PROV","page_size":"20","page_num":"1"}}
```

```
{"event": "search",  
  "data":{"page_size":"20","free_text":"modis sea ice","page_num":"1"}}
```

```
{"event": "search",  
  "data":{"page_size":"20","free_text":"mod29","page_num":"1"}}
```

```
{"event": "filter",  
  "data":{"echo_collection_id":"C1000001160-NSIDC_ECS",  
  "page_size":"20","page_num":"1"}}
```

```
{"event": "access",  
  "data":{"collections":["C1000001160-NSIDC_ECS"]}}
```



Metric Results

5 minutes plus session (Goal 1 Completions)

38,715



Visits data acquisition page (Goal 2 Completions)

20,178



Data Access Initiation (Goal 3 Completions)

18,128



Data Access Completion (Goal 4 Completions)

12,483



	May 2016	December 2016
Average Search Rank of Accessed Data	N%	M%

PLACEHOLDER! PLACEHOLDER! PLACEHOLDER!

This slide will contain additional information about average search ranking of collections that users access

PLACEHOLDER! PLACEHOLDER! PLACEHOLDER!

Metric Challenges

- Statistical significance with a small population
- Definitions of “success” in a free site with multiple uses
- Tail wagging the dog



Earthdata Search

QUESTIONS?

Thank You!

Mark Reese

mreese@element84.com



This material is based upon work supported by the National Aeronautics and Space Administration under Contract Number **NNG15HZ39C**.

Raytheon

