

# Usage-based Discovery

Chris Lynnes

*with contributions by*

*ESIP Discovery cluster members: many*

*GSFC Interns: Maggie Zhu, Megan Mehta, Parth Darji, Sophia Xia*

# Current discovery interfaces present dataset characteristics

## Earthdata Search

Platforms

Instruments

Processing Levels

The screenshot shows the Earthdata Search web application. On the left is a navigation menu with categories: Features, Keywords, Platforms, Instruments, Organizations, Projects, Processing Levels, and Data Format. The 'Platforms' and 'Instruments' categories are highlighted with green boxes, and red arrows point from the external labels on the left to these boxes. The main content area includes a search bar, a 'Sort by' dropdown set to 'Relevance', and several filter checkboxes: 'Only include collections with granules' (checked), 'Include non-EOSDIS collections' (checked), and 'Advanced Search'. Below the filters, it says 'Showing 20 of 7,080 matching collections'. A table displays the following data:

Collection	Version	Start	End
GHRSSST Level 2P OSPO dataset v2.61 from VIIRS on the NOAA-20 satellite (GDS v2) (GDS version...	2.61	2018-11-07	ongoing
GHRSSST Level 2P Atlantic Regional Skin Sea Surface Temperature from the Spinning...	1.0	2013-08-01	2018-02-20
GHRSSST Level 3U OSPO dataset v2.61 from VIIRS on S-NPP Satellite (GDS v2) (GDS version...	2.61	2015-05-19	ongoing
GHRSSST Level 2P OSPO dataset v2.61 from VIIRS on S-NPP Satellite (GDS v2) (GDS version...	2.61	2014-05-19	ongoing
Global Maps of Atmospheric Nitrogen Deposition, 1860, 1993, and 2050	1	1860-01-01	2050-12-31
IRS 1C LIS3 Standard Products	1.0	1996-11-14	2007-09-20
IRS 1C Standard Geo Referenced Product	1.0	1999-10-04	2007-09-20

At the bottom of the interface, there is a footer with version information (v1.129.17), search time (0.4s), and various links including 'NASA Official: Stephen Berrick', 'FOIA', 'NASA Privacy Policy', and 'USA.gov'. A map of the Indian Ocean region is visible on the right side of the interface.

# Current discovery interfaces present dataset characteristics

floods

The screenshot displays the Earthdata Search web application. The search query is 'floods', which is highlighted in a green box. The search results page shows 15 matching collections, also highlighted in a green box. A table lists the following collections:

Collection	Version	Start	End	Granules	Provider
GPM Ground Validation Global Flood Monitoring System (GFMS) Flood Maps IFloodS V1	1	2013-03-26	2013-06-30	4,107	NASA/MSFC/GHRC
NCA-LDAS Noah-3.3 Land Surface Model L4 Daily 0.125 x 0.125 degree V2.0...	2.0	1979-01-02	2016-12-31	13,879	NASA/GSFC/SED/E...
Global Flood Hazard Frequency and Distribution	1.0	1985-01-01	2003-12-31	1	SEDAC
Global Flood Mortality Risks and Distribution	1.0	2000-01-01	2000-12-31	1	SEDAC
Global Flood Proportional Economic Loss Risk Deciles	1.0	2000-01-01	2000-12-31	1	SEDAC

The interface includes a sidebar with navigation options like 'Features', 'Keywords', 'Platforms', and 'Instruments'. The footer contains metadata such as 'Time: 0.8s', 'NASA Official: Stephen Berrick', and 'FOIA'. A blue banner at the bottom states 'Earthdata Access: A Section 508 accessible alternative'.

# What about Rainfall?

The screenshot shows the Earthdata Search interface. The search bar contains the word "rainfall". The results show 20 of 458 matching collections. The top result is "TRMM Radar Rainfall Statistics L3 1 month (5 x 5) and (0.5 x 0.5)". Below this, a table lists several other collections related to rainfall and sea surface temperature.

Showing 20 of 458 matching collections

Collection

TRMM Radar Rainfall Statistics L3  
1 month (5 x 5) and (0.5 x 0.5)

Hourly Rainfall Amount	degree x 5 degree V7...							
Science Keywords: ... > Precipitation Anomalies > Rainfall Variability	GHRSSST Level 2P Global Subskin Sea Surface Temperature from TRMM Microwave Imager (TMI)...	i +	4.0	1998-01-01	2015-01-11	96,416	NASA/JPL/PODAAC	TV
Science Keywords: ... > Liquid Precipitation > Rain > 24 Hour Rainfall	GHRSSST Level 2P Regional Subskin Sea Surface Temperature from the Tropical...	i +	1.0	2005-01-31	2007-02-27	0	DOC/NOAA/NESDI...	
Science Keywords: ... > Liquid Precipitation > Rain > Average Annual Rainfall	GHRSSST Level 2P Regional Subskin Sea Surface Temperature from the Tropical...	i +	1.0	2005-01-31	2007-02-27	0	DOC/NOAA/NESDI...	gov

v1.129.17 • Search Time: 1.3s • NASA Official: Stephen Berrick • FOIA • NASA Privacy Policy • USA.gov

Earthdata Access: A Section 508 accessible alternative

# 400+ Results: too much of a good thing?

OK, what *kind* of rainfall data do you want for your flood monitoring?

- Satellite or rain gauge or merged?
- Spatial resolution?
- Temporal resolution?
- Spatial coverage?
- Temporal coverage?
- Quality?
- Latency?
- Reliability?

Let's try a new paradigm: Usage-based Discovery

**“What datasets does <*favorite flood application*> use?”**

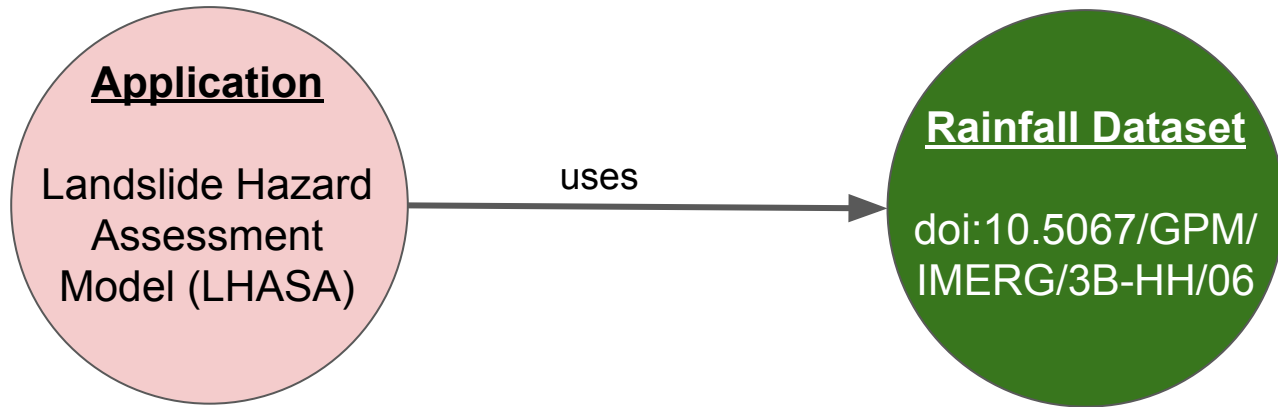
aka the

**“*I'll have what they're having*” paradigm**

aka

***Social Data Discovery***

# Graph relationship from Application to Dataset



# Usage-based Discovery User Interface

The screenshot displays a web browser window with the following elements:

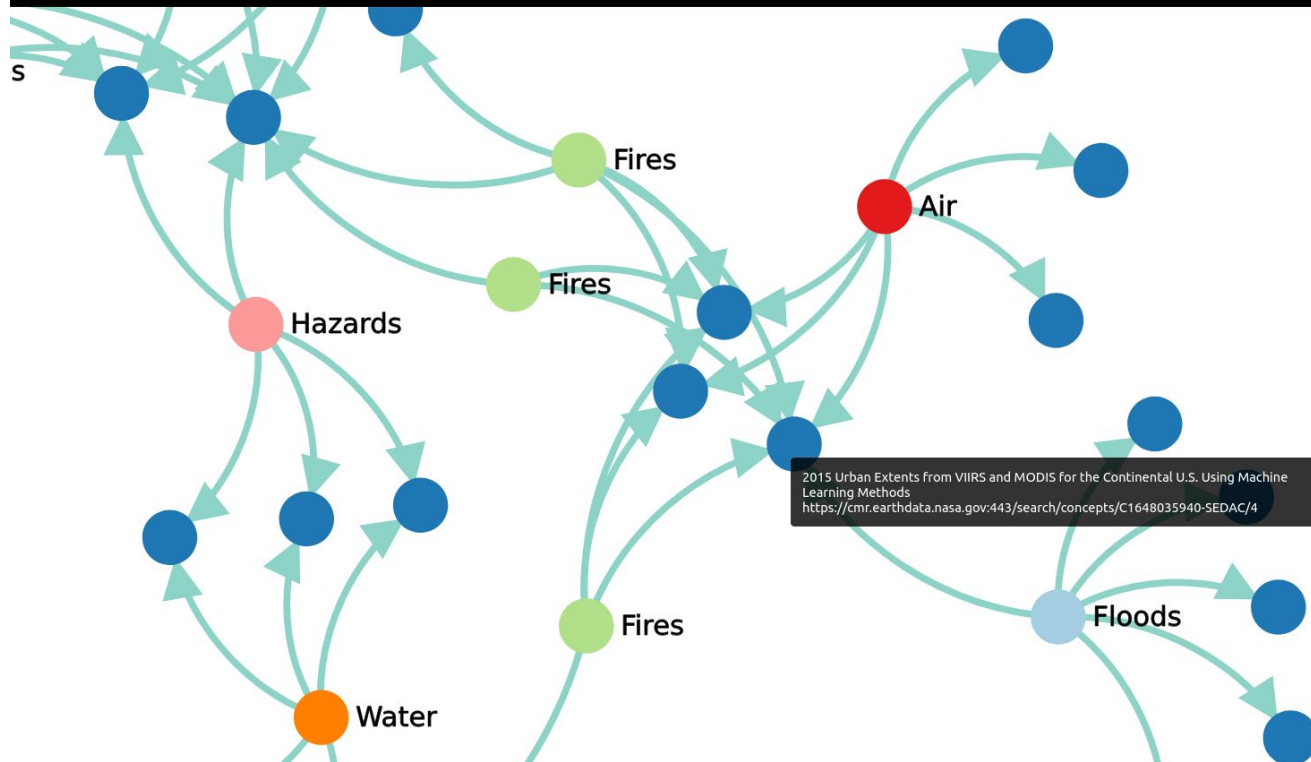
- Browser Address Bar:** `g8e443wkm7.execute-api.us-west-1.amazonaws.com/production/Landslides/SLIP-DRIP%`
- Page Header:** "EODIS Usage Based Discovery" with a "Landslides" dropdown menu.
- Main Content Area:**
  - SLIP-DRIP Product** section with "Website" and "Publication" buttons.
  - Text: "The Sudden Landslide Identification Product (SLIP) combs through Earth imagery and analyzes consecutive images of the same location to spot changes in soil moisture, muddiness, and other surface features...."
  - Preview of a NASA article: "Automating the Detection of Landslides" (Monday, July 11, 2016).
- Right Sidebar:** "Datasets" section listing "GPM IMERG Early Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06 (GPM\_3IMERGHE)" with a "See Dataset" button.



# Hmmm...

EOSDIS Usage Based Discovery

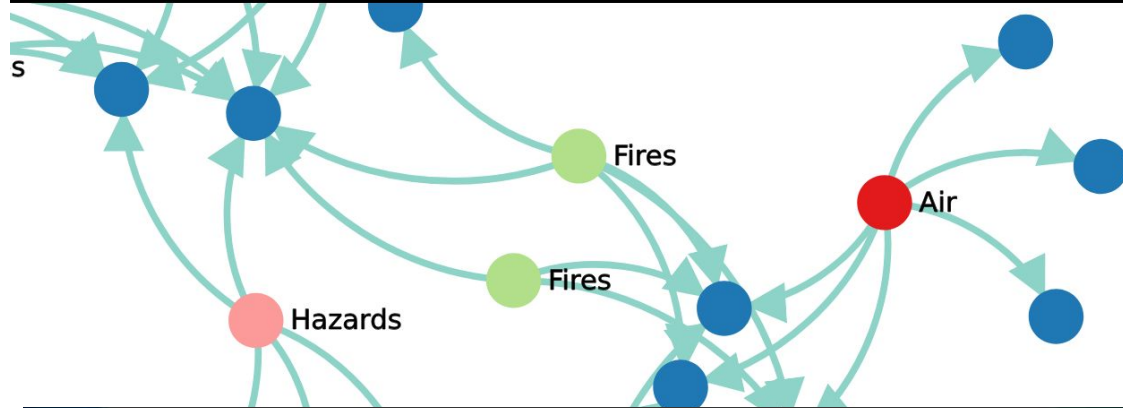
[Explore](#) [About](#)



# Hmmm...

EODIS Usage Based Discovery

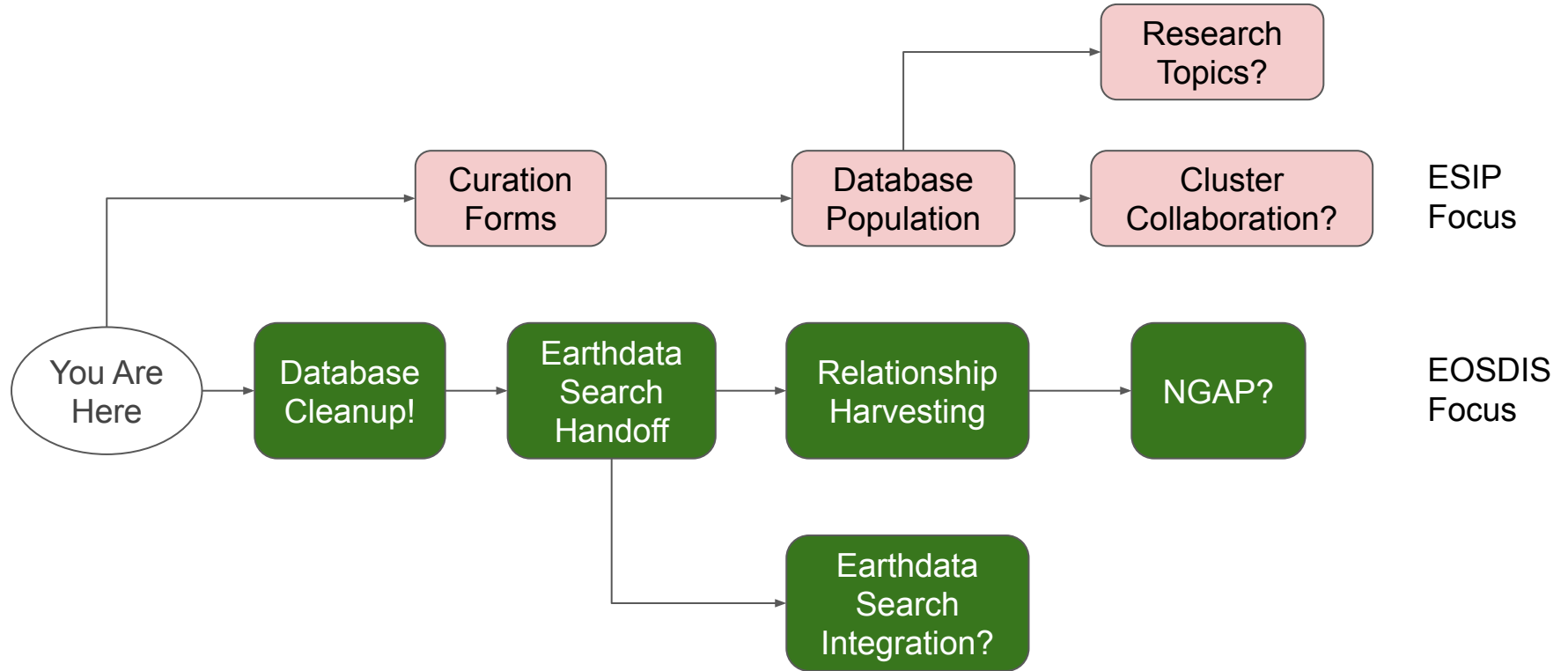
[Explore](#) [About](#)



2015 Urban Extents from VIIRS and MODIS for the Continental U.S. Using Machine Learning Methods  
<https://cmr.earthdata.nasa.gov:443/search/concepts/C1648035940-SEDAC/4>

**Water**

# Current Plan



# Speculation

- Minimal dependence on metadata consistency could help federate heterogeneous resource repositories
  - CMR, Earth Science Data Pathfinders, EONET, ESCCOR, IMPACT KG...
  - NASA: SDMWG?
- Usage  $\Leftrightarrow$  Dataset relationships have policy benefits
  - Impact metrics: what/how many applications/articles use dataset X?
  - Identification of collaboration opportunities