

Google Dataset Search & CMR

ESIP Winter 2019

Doug Newman

NASA EED-2 Data Use Architect

douglas i newman@nasa.gov

July 9th 2013, Chapel Hill NC – around 3:30pm

What was at stake

Traditionally it has been difficult to get your collection landing pages near the top of a Google search.



First exposure

'Facilitating Dataset Discovery using new developments within schema.org' @ Summer ESIP 2013, Peter Fox

Adding schema.org 'Dataset' markup to your collection landing pages *might* solve the problem



Nothing (for 5 years)...

But...

- RDFa markup carried over from ECHO to CMR
- Refined markup based on schema.org changes and ESIP semantic web recommendations
- Waited patiently.



2013-2018 CMR and schema.org

Landing Page Markup (1 of 3)

```
<div itemscope itemtype="http://schema.org/Dataset">
  <meta itemprop="name" content ="MODIS/Terra Calibrated Radiances 5-Min L1B Swath 250m V006"/>
  <meta itemprop="alternateName" content ="MOD02QKM_6"/>
  <meta itemprop="version" content="6"/>
</div>
<span itemprop='description'>The MODIS/Terra Calibrated Radiances 5-Min L1B Swath 250m...
<meta itemprop="spatialCoverage">
  <div vocab="http://schema.org/" typeof="Place">
    <div property="geo" typeof="GeoShape">
      <meta property="box" content="90.0 -180.0 90.0 180.0" />
   </div>
  </div>
</meta>
<time itemprop="temporalCoverage" datetime="2000-02-24T00:00:00.000Z/..">2000-02-24 to present</time>
<meta itemprop="keywords" content="EARTH SCIENCE, SPECTRAL/ENGINEERING, INFRARED WAVELENGTHS"/>
```



Landing Page Markup (2 of 3)

```
itemprop="identifier" itemscope="" itemtype="http://schema.org/PropertyValue">
 <h5>D0I</h5>
 >
   <meta itemprop="propertyID" content="DOI">
   <span itemprop="value">10.5067/MODIS/MOD02QKM.006</span>
 itemprop="citation" itemtype="http://schema.org/CreativeWork" itemscope>
 <h5 itemprop="headline">MODIS/Terra Calibrated Radiances 5-Min L1B Swath 250m V006</h5>
 <span itemprop="author">MCST Team</span>
 <span itemprop="publisher">L1 and Atmosphere Archive and Distribution System (LAADS)
 <a itemprop="url" href="http://example.com">https://dx.doi.org/10.5067/MODIS/MOD02QKM.006</a>
<time itemprop='dateCreated' datetime='2012-11-05T00:00:00.000Z'>2012-11-05T00:00:00.000Z</time>
<time itemprop='dateModified' datetime='2017-12-28T00:00:00.000Z'>2017-12-28T00:00:00.000Z</time>
<h5 itemprop='provider'>GSFC</h5>
```



Landing Page Markup (3 of 3)

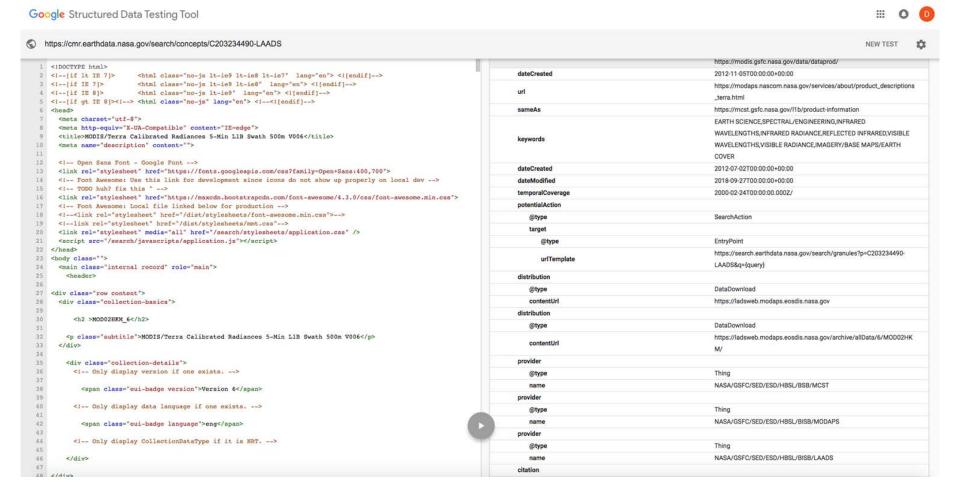
The following markup allows a search engine to visualize aspects of the CMR result to the user and facilitate acquisition of data related to that result,

- Browse images
- Data download



2018 GOOGLE DATASET SEARCH

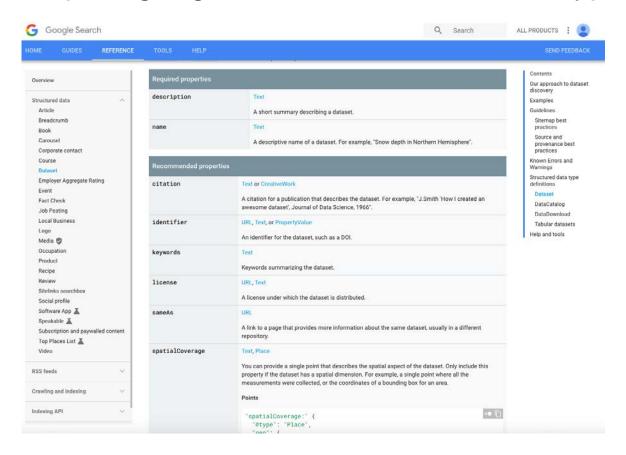
Structured data testing tool





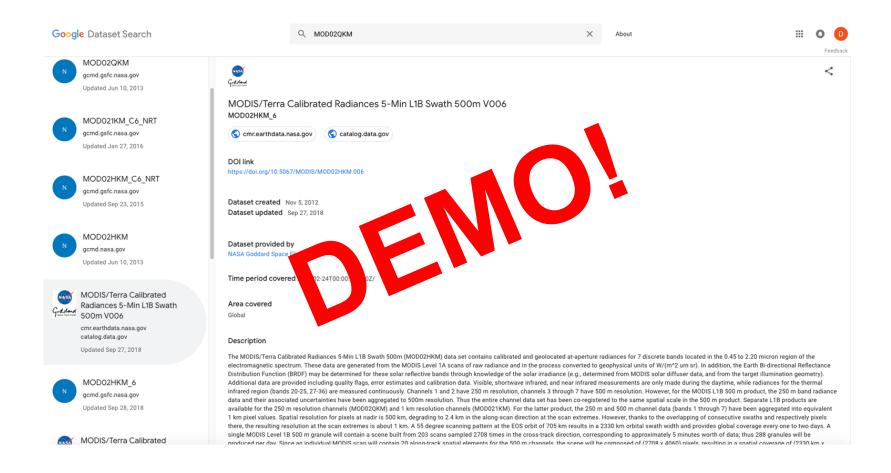
Google dataset guidelines

https://developers.google.com/search/docs/data-types/dataset





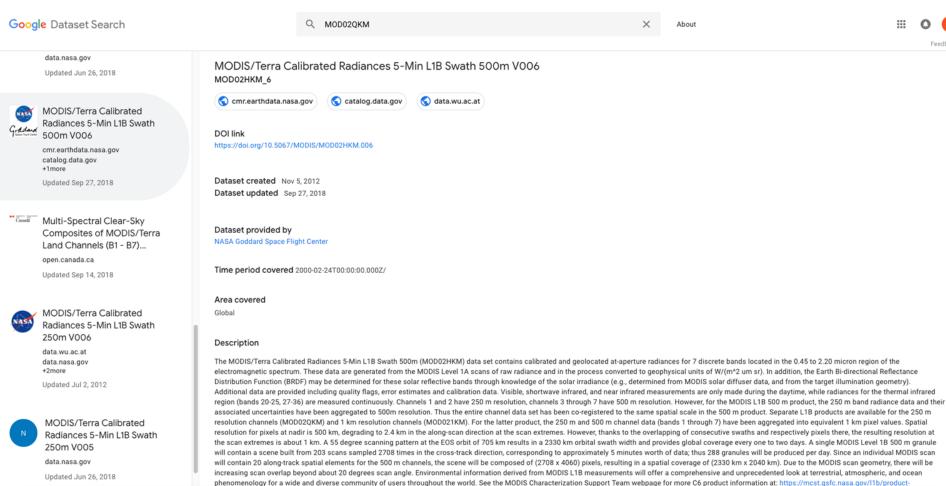
Google Dataset Search





2019 GOING FURTHER

Improving data.nasa.gov



information or visit Science Team homepage at: https://modis.gsfc.nasa.gov/data/dataprod/

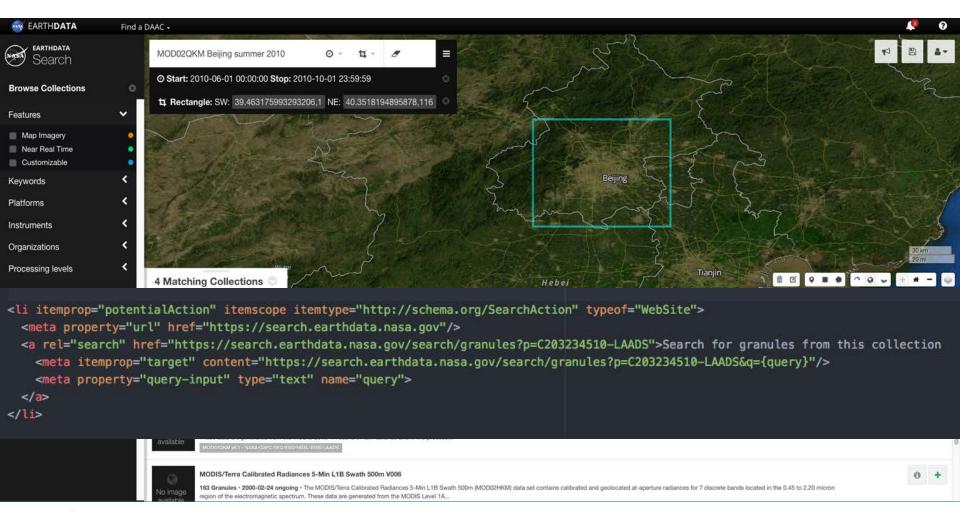


Smart handoffs from Google

- If we have a collection with granules the landing page will provide a link to Earthdata Search constrained to that collection.
- That link is tagged as a schema.org search action

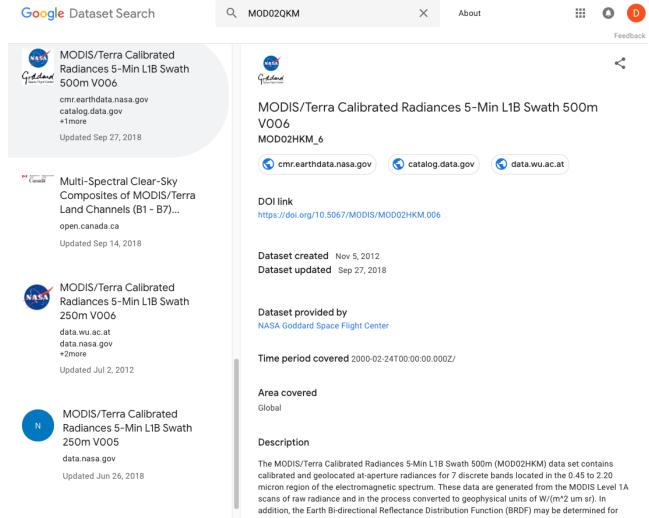


For example





Search to Dataset Search





QUESTIONS

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.

Raytheon

in partnership with

























