Earthdata Search Usability Study Process

Summer ESIP 2016

Mark Reese

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.
LP DAAC Spring User Study
MODIS Corrected Reflectance (Bands 3-6-7) - Terra
Contents

User Study Overview

Findings & Recommendations

Next Steps
User Study Overview
Our goal was to answer the following questions:

1. What are the key difficulties that people encounter when visiting the site?

2. How important is discoverability and relevance, given the way people typically search?

3. At what point do the map and timeline become useful? At what point are they in the way?

4. Is the "Download All" paradigm sufficient, or is a positive selection (shopping cart) mechanism necessary?
Our Method

Survey

Tasks & Questions

Analysis
Initial Questionnaire

Before the in-person user study, each participant was asked to complete a short survey about their work at LPDAAC and familiarity with Earthdata Search.
Questionnaire Results
Week Usage Rates

How many **hours each week** do you spend with these applications?

- 0-2 hours: 50%
- 3-5 hours: 12%
- 6-10 hours: 38%
Search & Discovery Pain Points

- Searching adjacent path and rows by date, sensor or quality for multi-decade time series.
- Removing several scenes quickly and accidentally clicking on the retrieve button. The interface often does not keep up with the user.
- Flexibility and use of KML and Shapefile searches.
- Lack of instructions and tutorials for new users.
- Identifying cloud-free scenes.
- Slow redraw/pan/zoom.
- Learning curve for new users.
Our Method

Survey

Tasks & Questions

Analysis
User Tasks

1. Find a data collection hosted at LPDAAC that has map imagery and view the latest imagery over the continental US.

2. Find and download all Landsat 8 TIRS images from January of this year which cover Sioux Falls and have no clouds.

3. Find an example of ASTER and Landsat 7 data files collected within an hour of one another over Sioux Falls.

4. Use the client to find and retrieve data which may be relevant to your work.
Notes

Summary of findings:

- We received a list of actionable items, both operational and technical, and were able to address them.

Most prominent:
- Facet quality improvement
- Collection consistency
- Ability to choose "Collection"
- Discoverability of how to view its data
- Timeline issues in figuring out the data

Note: The first 4 items were made for sure but the last 2, which are not sure.

Top features:
- Eye icon
- Granule results
- Earthdata login
- Map

Findings:

<table>
<thead>
<tr>
<th>Category</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword Search</td>
<td>Expected placename search of &quot;conterminous us&quot; to work twice</td>
</tr>
<tr>
<td>Temporal</td>
<td>Left temporal box open</td>
</tr>
<tr>
<td>Collections</td>
<td>Eye icon not behaving as expected</td>
</tr>
<tr>
<td>Granule Results</td>
<td>User could not find granule results</td>
</tr>
<tr>
<td>Earthdata Login</td>
<td>User needed to reset password for access</td>
</tr>
<tr>
<td>Map</td>
<td>User unclear as to why granules which do have map imagery in reality (Landsat) did not have &quot;Man Image&quot;</td>
</tr>
</tbody>
</table>

Estimated Task Completion Times:

<table>
<thead>
<tr>
<th>Participant</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Josh</td>
<td>10:00</td>
<td>10:00</td>
<td>15:00</td>
<td>06:00</td>
</tr>
<tr>
<td>Brian</td>
<td>NA*</td>
<td>12:00</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Jennifer</td>
<td>9:00</td>
<td>DNF</td>
<td>NA</td>
<td>10:00**</td>
</tr>
<tr>
<td>Carolyn</td>
<td>04:00</td>
<td>11:00</td>
<td>27:00**</td>
<td>NA</td>
</tr>
<tr>
<td>Danielle</td>
<td>01:00</td>
<td>03:00</td>
<td>02:00</td>
<td>NA</td>
</tr>
<tr>
<td>Rynn</td>
<td>07:00</td>
<td>05:00</td>
<td>13:00</td>
<td>06:00</td>
</tr>
<tr>
<td>Ben</td>
<td>01:00</td>
<td>06:00</td>
<td>04:00</td>
<td>05:00**</td>
</tr>
<tr>
<td>Bruce</td>
<td>15:00</td>
<td>08:00</td>
<td>NA</td>
<td>10:00</td>
</tr>
<tr>
<td>Kelly</td>
<td>05:00</td>
<td>06:00</td>
<td>05:00</td>
<td>NA</td>
</tr>
<tr>
<td>Mark***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Earthdata went down
** No ending time noted, needs video review/confirmation
*** Mark is on the Earthdata development team, not a member of LP DAAC
Our Method

Survey

Tasks & Questions

Analysis
Findings & Recommendations
Metadata Quality
Facet Quality

Users were overwhelmed by the number of facets, while simultaneously not being able to find key facets corresponding to high-value collections. This issue comes up because we only show the top 50 collections by count.
Facet Quality

Where is Landsat 8?

<table>
<thead>
<tr>
<th>Laboratory</th>
<th>691</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat</td>
<td>277</td>
</tr>
<tr>
<td>Landsat-5</td>
<td>206</td>
</tr>
<tr>
<td>Landsat-7</td>
<td>186</td>
</tr>
<tr>
<td>Maps</td>
<td>353</td>
</tr>
<tr>
<td>Meteorological Stations</td>
<td>395</td>
</tr>
</tbody>
</table>

Bioosphere?

<table>
<thead>
<tr>
<th>Biomass</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioosphere</td>
<td>2</td>
</tr>
<tr>
<td>Biosphere</td>
<td>6922</td>
</tr>
<tr>
<td>Climate Indicators</td>
<td>553</td>
</tr>
<tr>
<td>Cryosphere</td>
<td>2947</td>
</tr>
<tr>
<td>Human Dimensions</td>
<td>3872</td>
</tr>
</tbody>
</table>
Collections

Search Relevance

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

Search Relevance

Searching for *landsat*: the 28th result is the first major Landsat collection (Landsat 7 ETM+)

Searching for *sea ice*: mostly outdated collections at the top, collection-only with limited geographic scope

Searching for *modis*: the main LPDAAC and NSIDC collections are nowhere to be found in the first several pages of results, except for the two we artificially boost
Collections

Collection Visibility
Searches often produce many results which appear very similar to one another with the fields we are able to surface. Choosing the correct one is difficult, and poor relevance makes it even trickier.
Collection Recommendation

• Improve Quality
  • Facet helper
• Improve Relevancy
  • Based on version id (and other small wins)
  • Based on “has granules”
  • Relevancy ordering for facet searches
• Surfacing information that helps quickly distinguish collections from one another (map in preview image, file formats, abstract, etc...)
Next Steps
Summary: Short Term Plans

- Low hanging fruit
- Redesign Collection & Granule screens
- Reimagine Timeline
Summary: Long Term Plans

- Address metadata quality issues
- Revisit the core user and personas profiles
- My Earthdata Search™?
- Earthdata Search Lite™?
This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.