


Faceted Search and Density Information

Y. Coene, Spacebel s.a.

HMA Architecture Working Group
Frascati, 20 May 2016

- Background
 - Faceted Search proposed as FedEO future work 16/10/2014 at HMA-SE Final Presentation and at successive WGISS #39 event (JAXA):
 - Implemented by NASA (CwicSmart) in February 2015, included in Atom response.
- Geographical and temporal density or just two "facets".



Facetted Search appraisal

CWIC 2015 Annual Meeting, February 18-19, 2015

Doug Newman (NASA Earthdata – Raytheon)
Calin Duma (NASA Earthdata - XP Technologies, Inc.)

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG10HP02C.

- Applicable specifications
 - OASIS SearchRetrieve
- Schema available at <http://docs.oasis-open.org/search-ws/searchRetrieve/v1.0/os/schemas/facetedResults.xsd>
- http://www.opensearch.org/Community/Proposal/Specifications/OpenSearch/Extensions/SRU/1.0/Draft_1

7 Facet Parameters

7.1 The "facetLimit" parameter

7.2 The "facetSort" parameter

7.3 The "facetRangeField" parameter

7.4 The "facetLowValue" parameter

7.5 The "facetHighValue" parameter

7.6 The "facetCount" parameter

- <http://fedeo.esa.int/opensearch/request?parentIdentifier=XYZ&facetRangeField=geo:box>

```

<facetedResults>
  <dataSource>
    <!-- first data source -->
    <description>XYZ</description>
    <baseURL>http://fedeo.esa.int/opensearch/request?
      parentIdentifier=XYZ</baseURL>
    <facets>
      <facet>
        <!-- first facet for this datasource -->
        <description>Geographical density</description>
        <index>geo:box</index>
        <relation>intersects</relation>
        <terms>
          <term>
            <!-- first term for this facet -->
            <actualTerm>-180,80,-170,90</actualTerm>
            <count>12 </count>
          </term>
          <term>
            <!-- next term for this facet -->
            <actualTerm>-180,70,-170,80</actualTerm>
            <count>8</count>
          </term>
        </terms>
      </facet>
    </facets>
  </dataSource>
  <datasource>
    ...
  </datasource>
</facetedResults>

```

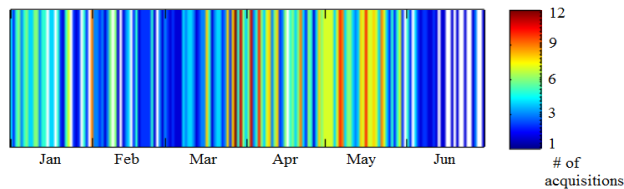
- <http://fedeo.esa.int/opensearch/request?parentIdentifier=XYZ&facetRangeField=time:end>

```

<facetedResults>
  <dataSource>
    <!-- first data source -->
    <description>XYZ</description>
    <baseURL>http://fedeo.esa.int/opensearch/request?
      parentIdentifier=XYZ</baseURL>
    <facets>
      <facet>
        <!-- first facet for this datasource -->
        <description>Temporal density (Monthly)</description>
        <index>time:end</index>
        <relation>=</relation>
        <terms>
          <term>
            <!-- first term for this facet -->
            <actualTerm>2004-05-01T00:00:00Z</actualTerm>
            <count>12</count>
          </term>
          <term>
            <!-- next term for this facet -->
            <actualTerm>2004-06-01T00:00:00Z</actualTerm>
            <count>20</count>
          </term>
        </terms>
      </facet>
    </facets>
  </dataSource>
  <datasource>
    ...
  </datasource>
</facetedResults>

```

Temporal coverage map for E1_SAR_IM__0P (1996)



Result sample

```
<sru:facetedResults
  <sru:datasource>
    <sru:datasourceDisplayLabel>ECHO REST API</sru:datasourceDisplayLabel>
    <sru:datasourceDescription>ECHO catalog holdings exposed via the ECHO REST
API
  </sru:datasourceDescription>
  <sru:baseURL>https://earthdata.nasa.gov/echo</sru:baseURL>
  <sru:facets>
    <sru:facet>
      <sru:facetDisplayLabel>Campaign</sru:facetDisplayLabel>
      <sru:index>campaign-sn</sru:index>
      <sru:relation>=</sru:relation>
      <sru:terms>
        <sru:term>
          <sru:actualTerm>ESIP</sru:actualTerm>
          <sru:count>76</sru:count>
        </sru:term>
        ...
      </sru:terms>
    </sru:facet>
  </sru:facets>
</sru:facetedResults>
```

Conclusion



- Alternative for "statistical interface" proposed for TEPs
- Encoding in (LD-)JSON to be proposed.
- Include also as Atom Link in Series search rel="index".

The image shows a document cover page with the following content:

- Logo:** ESA logo in the top right corner.
- Text:** esrin, Via Galileo Galilei, Casella Postale 84, 00044 Frascati, Italy, T +39 06 9418 01, F +39 06 9418 0280, www.esa.int
- Document Title:** DOCUMENT (in a dashed box)
- Subtitle:** Exploitation Platform Resource Catalogue Interface (in a dashed box)
- Metadata:** Prepared by: Salvatore Pinto, Issue: DRAFT, Revision: 1, Date of Issue: 22/04/2016
- Footer:** European Space Agency / Agence spatiale européenne (in a dashed box)