

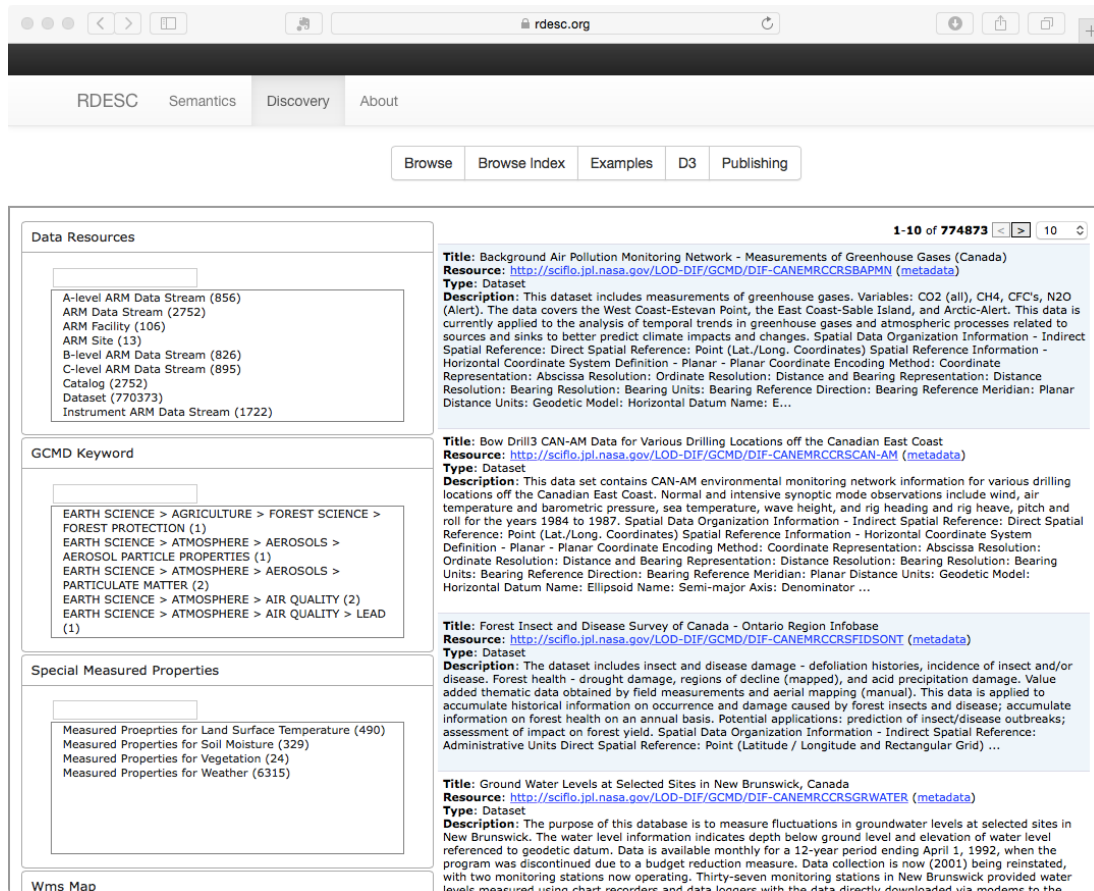
### RDESC Final report – RPI/TWC – year 3.

1. DOE award # and name of the recipient (Institution), DE-FG02-12ER26106/DE-SC0008616 to Rensselaer Polytechnic Institute
2. Project Title and name of the PI, Resource Discovery for Extreme Scale Collaboration (RDESC). Prof. Peter Fox; PI.
3. Date of the report and period covered by the report, May 29, 2015 covering the period June 1, 2014-May 31, 2015 (year 2 award period August 31, 2014)
4. A brief description of the accomplishments during the final funding period. Overall project website is <http://tw.rpi.edu/web/project/rdesc> and contains many levels of documentation of project activities and progress. At RPI PI, Prof. Peter Fox, Benno Lee (graduate student), and Patrick West (principal software engineer) participated in the project. Stephan Zednik (senior software engineer), lead for the S2S technology also worked on RDESC. RDESC activities are arranged in working groups: <http://tw.rpi.edu/web/project/rdesc/workinggroups> (see brief description and participants in Appendix A).

#### Primary activities:

- Finalized RDESC portal with ontology-based editing and discovery capability for selected resource types.

RPI maintained the rdesc.org and rdesc.net domain names for the project which hosts the discovery portal (key milestone for this year): <https://rdesc.org/browser/> (see Fig. 1) populated with over 774,000 resources. The discovery portal uses the developed queries that match to the underlying ontology encoded triples captured from the resource publishing.



**Figure 1. RDESC portal – faceted discovery browsing for Resources. Facets on the left include data resources, linked category vocabularies based on the Global Change Master Directory (GCMD) and special measure properties.**

Completed contributions to small-scale and large-scale resource exploitation (<http://tw.rpi.edu/web/project/rdesc/workinggroups/aggregation>) and data services access and harvesting (<http://tw.rpi.edu/web/project/rdesc/workinggroups/dataservices>) which guide the mapping of queries on to underlying data and metadata models. The queries, which enable the discovery user application are tied to a variety of ontologies (see <http://tw.rpi.edu/web/project/rdesc/workinggroups/ontologies>) that moved beyond the focus on the previous ARM project resources (based on third generation use cases).

- Completed test-bed development

The majority of the test-bed work was on resource publishing, which is documented at: <http://tw.rpi.edu/web/project/rdesc/resources>. The architectural framework is at: <http://tw.rpi.edu/web/project/rdesc/architecture>.

- Team meetings and dissemination of results

Regular team meetings were being held. These included monthly virtual all-hands meetings for overall status, milestones and directions. Weekly a technical call addressed specific issues.

Meeting notes for regular meetings can be found at:  
<http://tw.rpi.edu/web/project/rdesc/meetingnotes>. Results from the project were disseminated at the ASCR PI monthly meeting (Weaver/ Stephan presenting).

Further information on presentations is available at:  
<http://tw.rpi.edu/web/project/rdesc/presentations>.  
RPI participated in the development of the RDESC paper appearing in Computers and Geosciences in 2015.

RPI milestones are maintained at: <http://tw.rpi.edu/web/project/rdesc/milestones>.  
Year 3 activities included (RPI involved in all):

- Developed mechanisms to incorporate alternative ontologies
  - Linked concepts in related ontologies enabling cross-domain discovery queries.
  - Implemented query interface for requested query patterns and query expansion.
  - Implemented crawlers to harvest metadata from tagged web resources.
  - Demonstrated RDESC resource discovery including the full range of resources through orchestrated set of queries.
  - Reported and disseminated major findings in national meetings and via journal publications.
5. Unexpended funds: None.

#### Appendix A – RDESC Working groups:

##### Ontology Development

**members:** [Alan Chappell](#), [Jesse Weaver](#), [Peter Fox](#)

**description:** Focuses on the aggregation of existing ontologies for use in scientific resource description; also covers ontology development to fill in the gaps left by existing ontologies.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/ontologies>

##### User Interface and Information Display Experiments

**members:** [Patrick West](#), [Benno Lee](#)

**description:** Investigating the display of large-scale information in a user interface

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/ui>

##### Data Service Publishing Investigations

**members:** [Eric Rozell](#), [Benno Lee](#), [Alan Chappell](#), [Jesse Weaver](#), [Sumit Purohit](#)

**description:** Focuses on the trade-offs between different mechanisms for data service publishing (e.g., SADI vs. SSWAP).

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/dataservices>

##### Resource Storage Investigations

**members:** [Eric Rozell](#), [Alan Chappell](#), [Jesse Weaver](#)

**description:** Focuses on the options for storing scientific resource descriptions and their corresponding provenance.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/storage>

#### Evaluation Options and Metrics

**members:** [Karen Schuchardt](#), [Alan Chappell](#), [Peter Fox](#)

**description:** Focuses on the approaches that will be used to evaluate the outcomes of this project.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/evaluation>

#### Discovery Demonstrations

**members:** [Eric Rozell](#), [Karen Schuchardt](#), [Alan Chappell](#), [Jesse Weaver](#), [Patrick West](#), [Benno Lee](#)

**description:** Focuses on the development of prototypes to allow users to discover scientific resources.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/demos>

#### Large Scale Resource Re-Publishing

**members:** [Karen Schuchardt](#), [Eric Rozell](#), [Benno Lee](#), [Jesse Weaver](#), [Sumit Purohit](#)

**description:** Focuses on the development of data pipelines for ingesting batches of scientific resources.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/aggregation>

#### Small Scale Resource Re-Publishing

**members:** [Benno Lee](#), [Karen Schuchardt](#), [Eric Rozell](#), [Jesse Weaver](#)

**description:** Focuses on the publication of resources, scraping of metadata, and form-based approaches to collection metadata at the individual item level.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/publication>

#### Metadata Usage Experiments

**members:** [Sinan al-Saffar](#), [Eric Rozell](#), [Benno Lee](#), [Jesse Weaver](#), [Sumit Purohit](#)

**description:** Investigating the usage of metadata in the wild, including CF conventions for NetCDF datasets, HDF conventions, and other forms of metadata.

**website:** <http://tw.rpi.edu/web/project/rdesc/workinggroups/metadata>