

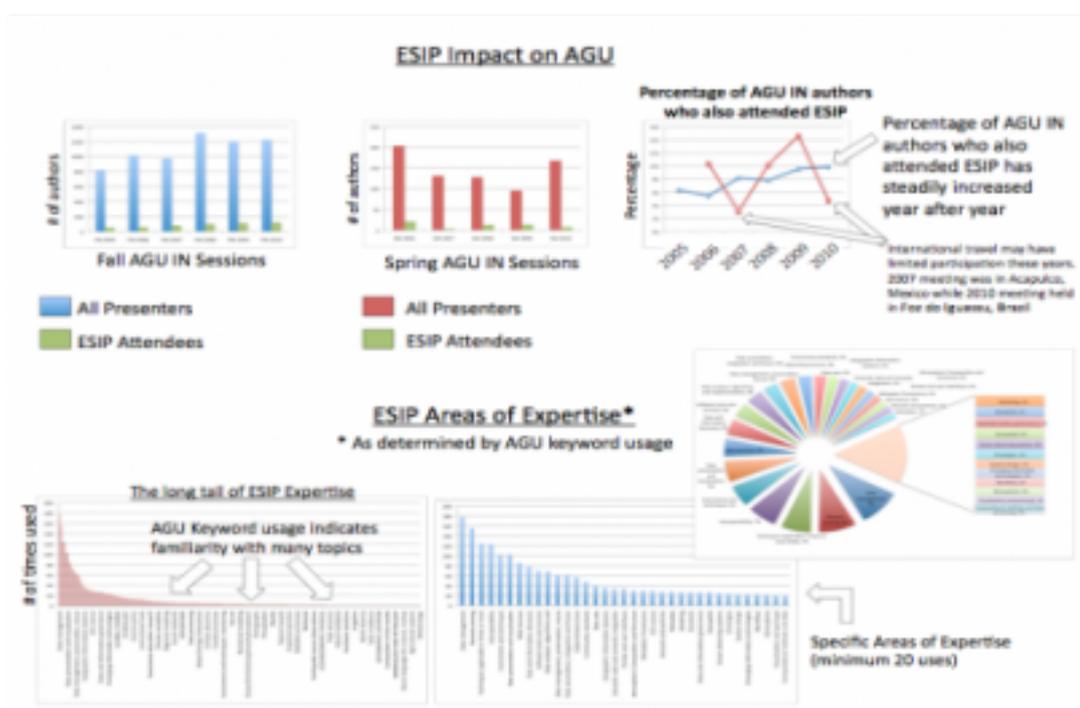
## [Linking Open Research Data for Earth and Space Science Informatics](#)

[1]

Submitted by superadmin on Fri, 2012-02-03 10:56 **Event:** [Winter Meeting 2012](#) [2]

### Abstract:

Earth and Space Science Informatics (ESSI) is inherently multi-disciplinary, requiring close collaborations between scientists and information technologists. Identifying potential collaborations can be difficult, especially with the rapidly changing landscape of technologies and informatics projects. The ability to discover the technical competencies of other researchers in the community can help in the discovery of research partnerships. In addition to collaboration discovery, this data can be used to analyze trends in the field, which will help project managers identify emerging, irrelevant, and well-established technologies and specifications. This information will help keep projects focused on the technologies and standards that are actually being used, making them more useful to the ESSI community. We present a two-part solution to this problem: a pipeline for generating structured data from ESSI abstracts and an API and Web application for accessing the generated data. We use a Natural Language Processing (NLP) technique, Named Entity Disambiguation, to extract information about researchers, their affiliations, and technologies they have applied in their research. The extracted data is encoded in the Resource Description Framework using Linked Data vocabularies, including the Semantic Web for Research Communities ontology and the Friend-of-a-Friend ontology. The data is exposed in four ways: a SPARQL query-able endpoint, linked data, Java APIs, and a Web application. We also capture the provenance of the data transformations using the Proof Markup Language, including confidence scores from the NLP algorithms used. Our implementation has used only open source solutions, including DBpedia Spotlight and OpenNLP. We plan to set up an open source project for this work so that it can continue to evolve through community contributions. Submitted by: Eric Rozell, Rensselaer Polytechnic Institute, [rozele@rpi.edu](mailto:rozele@rpi.edu) [3]



### Images:

[4]

**Attachments for download:**  [narock\\_rozell\\_esip\\_poster\\_winter\\_2012.pdf](#) [5]

**Author(s):**

**Name:** [Eric Rozell](#) [6]

**Organization(s):** [Rensselaer Polytechnic](#)

[Institute](#) [7]

**Name:** [Tom Narock](#) [8]

**Organization(s):** [University of Maryland  
Baltimore County](#) [9]

**Keywords:** [Funding Friday Winner](#) [10]

**Source URL:** <https://commons.esipfed.org/content/linking-open-research-data-earth-and-space-science-informatics>

## Links

[1] <https://commons.esipfed.org/content/linking-open-research-data-earth-and-space-science-informatics>

[2] <https://commons.esipfed.org/event/winter-meeting-2012>

[3] <mailto:rozele@rpi.edu>

[4] [https://commons.esipfed.org/sites/default/files/narock\\_rozell\\_esip\\_poster\\_winter\\_2012.png](https://commons.esipfed.org/sites/default/files/narock_rozell_esip_poster_winter_2012.png)

[5] [https://commons.esipfed.org/sites/default/files/narock\\_rozell\\_esip\\_poster\\_winter\\_2012.pdf](https://commons.esipfed.org/sites/default/files/narock_rozell_esip_poster_winter_2012.pdf)

[6] <https://commons.esipfed.org/node/292>

[7] <https://commons.esipfed.org/taxonomy/term/221>

[8] <https://commons.esipfed.org/node/293>

[9] <https://commons.esipfed.org/taxonomy/term/205>

[10] <https://commons.esipfed.org/tags/funding-friday-winner>