OntologySummit2014: (Track-D)
Tackling the Variety Problem in Big Data

Co-Champions
Ken Baclawski, ...

The Variety Problem

- Big Data faces many challenges, some of which could benefit from the use of ontologies.

- The focus of this track is on managing the complexity and variety of data being processed and analyzed in Big Data projects.

- Examples of sources of complexity and variety

  - Background knowledge of the domain
  - The structure of the data
  - Provenance of the data
  - Mapping data and metadata from multiple sources
  - Processing workflows
  - Privacy concerns

Mission

- Identify potential solutions for tackling the sources of complexity and variety in Big Data.

- Develop specific use cases for the integration of ontologies with Big Data tools and techniques.

- Reach out to the Big Data community to foster communication and collaboration.

Plan and Deliverables

1. Elaborate the list of sources of Big Data variety that have the greatest potential for benefiting from the use of ontologies.

2. Identify potential solutions for the list developed in the first phase.

3. Find examples of Big Data projects for which we can propose how ontology techniques can be applied to the project.

4. Express these as use cases for ontologies in Big Data.

5. Develop the write-up (a section) that synthesizes the results of the discourse in this track as a contribution to the Ontology Summit 2014 Communique.

6. Gather citations to relevant resources for the Ontology Summit 2014 Community Resources Library.