# Semantically-Enabled Scientific Data Integration

Peter Fox, UCAR

Objective

- The objective of this work is to integrate information technology in support of advancing measurement-based processing systems for NASA, by integrating existing diverse science discipline and mission-specific data sources.

- This vision will be achieved using a set of technologies that feature rich semantics, that is, the precise meaning of a quantity or entity, e.g. a variable in a dataset, its units, a physical feature or phenomenon in the Earth system, how it may relate to other entities, its quality and lineage, etc.

Science data · Science data · Integration · Semantics · SWEET · Volcanoes · Climate · SESDI
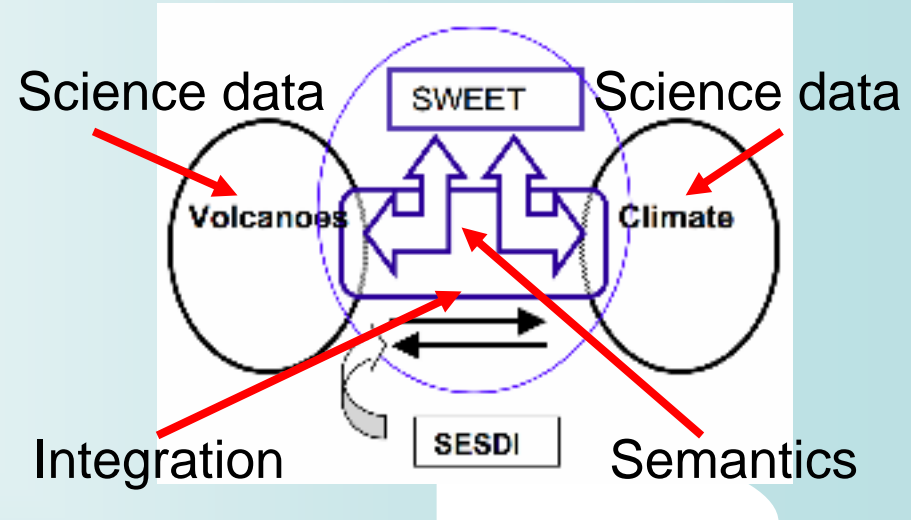
Approach

1. Evolve the existing SWEET core/domain ontology
2. Apply registry supported by an extended GEON ontology
3. Enable a range of semantically-based data services (such as data mining, validation, data integration, etc.)
4. Submit the ontology for review through the Standards and Processes Data Systems Working Group and via community feedback/collaboration
5. Demonstrate SWEET, GEON, OPeNDAP and CEDARWEB technologies within a semantic framework

Co-Is/Partners

- Rob Raskin (JPL), Deborah McGuinness (Stanford), Krishna Sinha (Virginia Polytechnic)

Key Milestones

• Plate tectonics ontology workshop, merge	12/06
• Atmospheric ontology mapping, map databases	6/07
• Publish ontologies, develop SESDI connector	12/07
• Demonstrate data integration via statistical application, generalize to solar irradiance domain/data	6/08
• Demonstrate generalized data integration	12/08

$TRL_{in} = 6$ $TRL_{current} = 7$

ACCESS
Advancing Collaborative Connections for Earth-Sun System Science

Progress since Q2

- Ontology collaboration website, community engagement
- Technology infusion leveraging VSTO and GEON ontologies, data frameworks
- Papers and presentations at conferences
- Evaluation of additions to atmospheric ontology for volcano-climate application
- Identification of existing statistical application which seeks relationships between volcanic activity and climate (C. Ammann, NCAR)
- Initial mapping of WOVODAT (World Volcano Database) schema to developed ontology, working on Nevada test site data
- Leadership and participation role in DSWG semantic web sub-group
- Development of initial Semantic Web Tech. Infusion roadmap
- Participation in ESIP, new semantic web cluster activity