Open Data for Open Science
Advance Data Interoperability

Yan Xu, Ph.D.
Microsoft Research
Working with Environmental Data

**Grand Challenge**: vast amount of heterogeneous data
- Data from different sources
- Lack of easy-to-use tools and incentives to share data
- Lack of tools for easy adoption of (existing) standards

**Unique Challenge**: it doesn’t end with papers
- Compelling presentation of knowledge is critical
  - Influence policy makers and the general public
  - Enable citizen science to scale the effort.

Social Impact!
**ODOS: Open Data for Open Science**

**Inspiration:**
- **The Fourth Paradigm**: make scientific discoveries in big data
- **OData**: unlock data from silos of applications

**Implementation:**
- Collaborate with environmental researchers to explore new ways of using software to enable scientific discoveries.
- Take advantage of the OData compliant technologies to create solutions for data- and information-intensive problems and challenges.
  - Separate service model from data model.
  - Open up your data, not your database.
Example:

- **Project CLEO:** *Cultivating the Longtail Environmental Observatory*
  - From Sensing and Energy Research at Microsoft Research
  - A mobile sensing platform
    - ultra-portable,
    - low-power sensors,
    - phone-based data upload software
    - a location resolution and data management web service in the cloud.
    - cloud-offloaded GPS (CO-GPS) for location sensing.
    - audioport-based connector.
    - **OData compliant** (directly accessible to OData tools, e.g., WWT)
An example of Seamless Transformation from data to information to Knowledge