

ESIP Federation: A Model for Community Collaboration & Coordination in Big Data

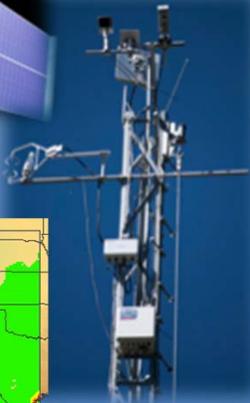
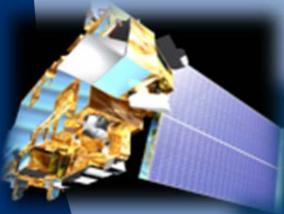
Leveraging Community Expertise in Data,
Technology and Use

Stefan Falke
Carol Meyer
ESIP

December 11, 2012

Federation of Earth Science Information Partners (ESIP) Vision

To be a leader in promoting the collection, stewardship and use of Earth science data, information and knowledge that is responsive to societal needs.



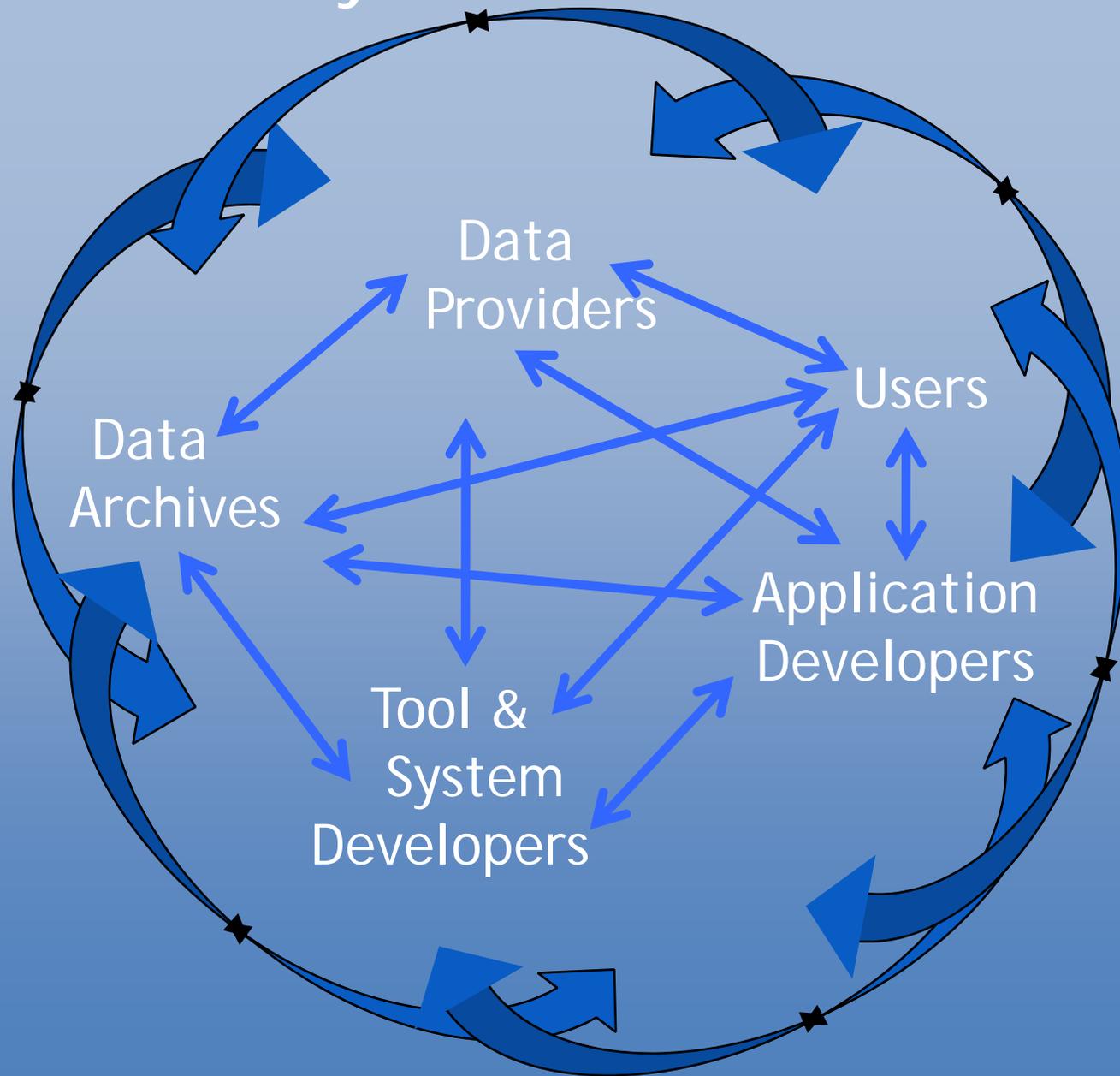
About ESIP

ESIP Federation has far-ranging expertise and provides a neutral space for Earth & environmental science data & technology practitioners to share their knowledge, experience & technology

ESIP Strategic and Vision Goals (subset)

- ESIP is the trusted community authority that supports the integration of science and data into mainstream use.
- ESIP increases use and value science data and information.
- ESIP articulates and measures the value and benefit of Earth science data.
- ESIP leads the development of the Earth science data and information field to:
 - recognize data management as a profession;
 - provide opportunities for continued professional development
 - train the research community on data management
 - engage the next generation of science data information professionals

ESIP Community



ESIP Core Values

- Agility
- Collaborative
- Collegial
- Community-driven
- Innovative
- Neutral
- Open
- Participatory
- Voluntary



Things ESIP Does

- Community-generated Best Practices (e.g. Data exchange, Citation)
- Testbed (e.g. Identifiers, Ontology)
- Community Conventions (e.g. Discovery, Data Web Services)
- Professional Development
 - Technical Workshops
 - Non-technical Workshops (e.g. Evaluation, Communication)
 - Data Management Short Course/Workshops
- Outreach
 - Education (e.g. annual teacher workshop on climate change)
 - Professional Societies (e.g. AGU, GSA)
 - International Efforts (e.g. GEO, ISRSE)
- Provide Venue for Collaboration and Connections
 - Both virtual and in-person
 - Support with suite of collaboration tools

Case Study: Data Persistence and Publication

Problem: *Data producers get little professional recognition.*

- Traditional academic rewards system recognizes journal publications
- Good data are not always permanently available or traceable
 - & when found, attribution is not always given*
- No single permanent identifier scheme supports all data types

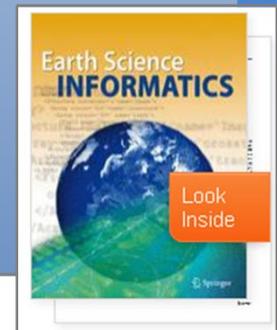
Case Study: Activities & Outputs

Activities:

- Identifiers analysis and testbed
- Community-generated consensus for data citation

Outputs include:

- Recommendations for identifiers
 - <http://springerlink.com/content/52760gq3h200gw38/?MUD=MP>
- Data citation guidelines
 - <http://commons.esipfed.org/node/308>
- Data stewardship principles
 - <http://commons.esipfed.org/node/419>



Earth Science Informatics
September 2011, Volume 4, Issue 3, pp 139-160,

[Open Access](#)

On the utility of identification schemes for digital earth science data: an assessment and recommendations

Ruth E. Duerr, Robert R. Downs, Curt Tilmes, Bruce Barkstrom, W. Christopher Lenhardt, Joseph Glassy, Luis E. Bermudez, Peter Slaughter

Q2: What has changed in your information landscape in the last 2-3 years?

- No substantial changes from a big data perspective: continued activities and trends that started at ESIP's inception
- Increased and more diverse ESIP membership
- New types of data and ways to conduct analysis, e.g. apps, Air Twitter

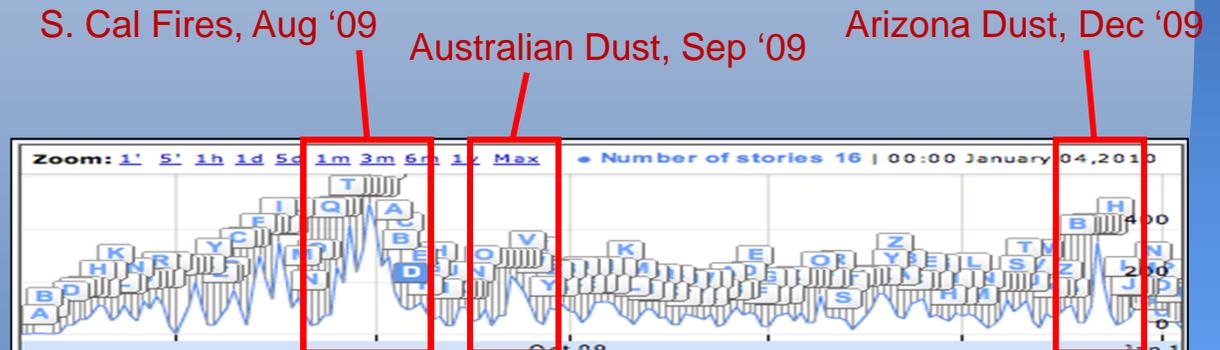
Forbes.com Home Page for the World's Business
U.S. EUROPE ASIA
Home Business Investing Technology Entrepreneur

Social Media
Air Twitter: Social Media Meets Science
Marsha Walton, 11.08.10, 06:00 AM EST
Smog making you choke? Do you have the flu? Your tweets could be valuable to researchers.

mnn mother nature network
7 times social networking saved lives
UK police publish all incidents on Twitter
Greenpeace vs. Facebook

Science information channels used to be pretty limited. Scientists would publish their findings in obscure journals, and only other researchers would ever see them.

Now a new generation of scientists, armed with new technology tools, has led to some big changes.



Tweets for 2009-8-28

- I'm sleepy, but I want news. I heard there is a crucial piece of news I need to watch before I go to bed. Something about air quality.
- @alopink lmao. air quality's already shit. but i guess the fire adds greatly.
- Contemplating going out for a bike ride but wondering how stupid that would be with the current air quality being what it is.
- Only four helicopters. Air quality didn't seem too bad. Three more concerts though...
- Two new fires in La Canada and Ranchos Palos Verdes. Air quality is going to deteriorate again.
- @typeffend Oh, well in that case, damn those workmen! No A/C in this weather (and with this air quality) must absolutely suck.
- Man, the air quality is going from bad to worse.
- We are surrounded by fires here in Hollywood. The air quality will be simply marvelous tomorrow.
- gosh its so flippin hot here and the air quality is horrible! please pray for all of us that are near the fires in souther cali...
- Cornell Chronicle: Beijing air quality study :-P
- @kch917 I am in torrance watching on ch4 right now. I haven't been outside yet but I am sure the air quality is horrible.

Robinson et al., 2010

Q3: What are your biggest challenges?

- Getting right information to right person at right time in right context
- Advancing community-oriented approaches (cross-organization, cross-discipline)
- Making interoperability easier

Q4: What changes do you see in the next five years?

- Simpler ways to share data and conduct analysis
- Increased role of citizen scientists
 - Volunteered geographic information
- Big data is not just about the Data
 - Data processing and analysis
 - Data lineage and provenance
- Big data challenges don't end after the 'first phase' of big data analysis

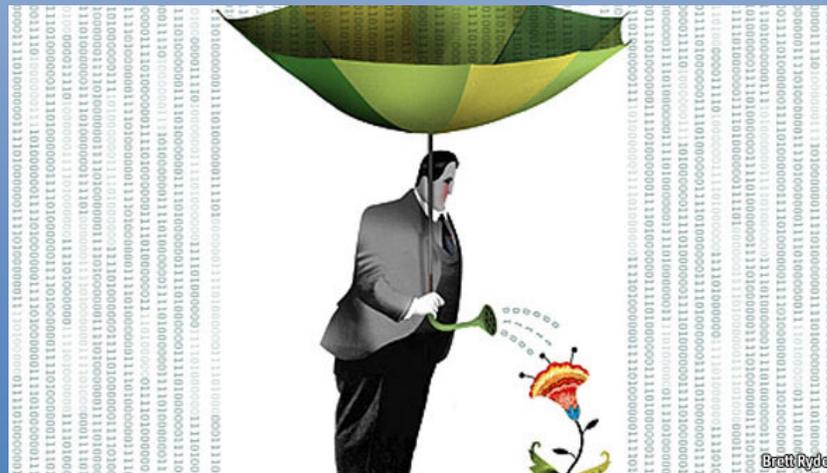


Image source:
"The Data Deluge", *The Economist*, Feb. 25, 2010