

# Citizen Science:

Even Big Data Starts Small

Dr. Lea Shanley, co-Executive Director

South Big Data Innovation Hub @ UNC-Chapel Hill

*Environmental Measurement Symposium*

*August 10, 2016*

If you had 100,000 people to  
help you with your work, what  
would you do?

# What is Citizen Science?

“The contributions of the public to the advancement of scientific and engineering *research and monitoring* in ways that may include (Federal CCS Community of Practice):

- Identifying research questions
- Designing/conducting investigations
- Collecting and analyzing data
- Developing data applications
- Developing technologies for science
- Solving complex problems.”



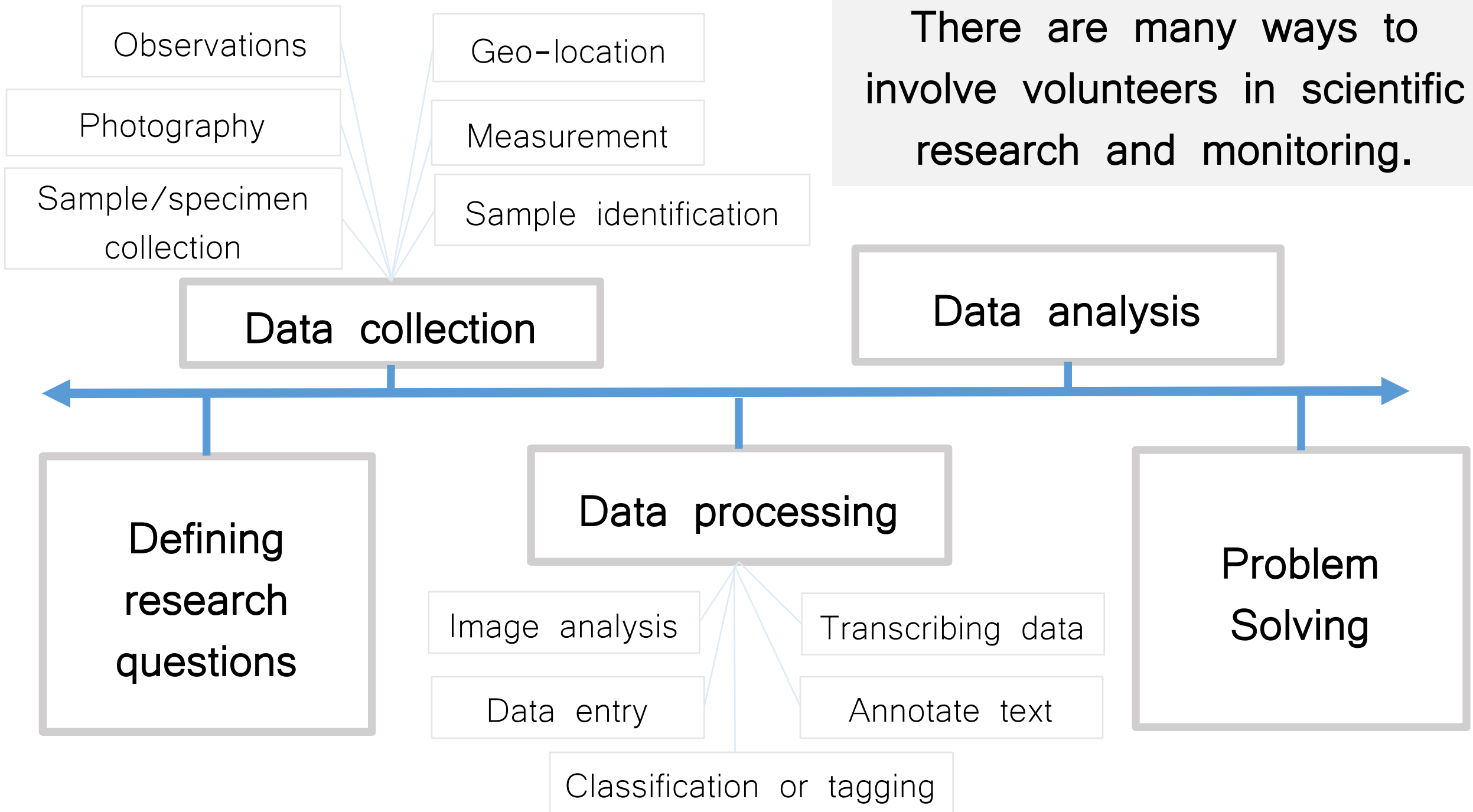
# What is Crowdsourcing?

Crowdsourcing is a process where individuals or organizations submit an open call for voluntary contributions from a large group of individuals (“the crowd”), often through an online platform.

Citizen science isn't right for solving EVERY scientific problem, but it can be valuable when applied to the right scientific or engineering problem and when designed properly.



There are many ways to involve volunteers in scientific research and monitoring.



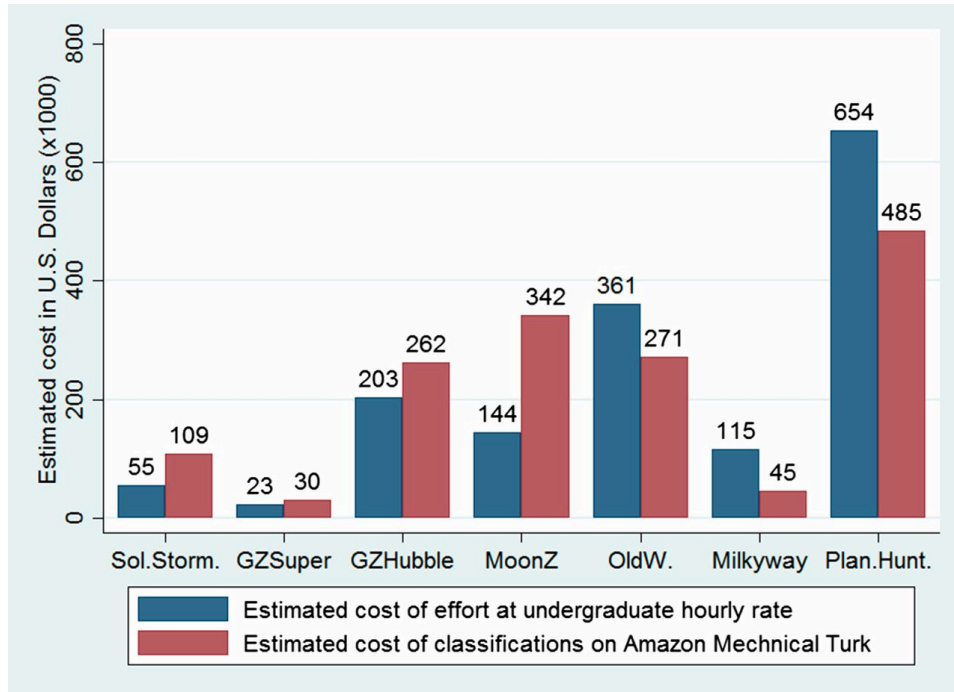
# Why consider citizen science?

- Augment and enhance traditional scientific research and monitoring
  - Increase spatial or temporal frequency
  - Increase geographic extent or temporal scale (long-term)
  - Eye can be better than the algorithm
  - Reduce time and labor costs
- Provide hands-on STEM learning outside classroom
- Increase public awareness of scientific developments and methods

# Value of Citizen Science

- Theobald et al. (2014) analyzed 388 English-language biodiversity citizen science projects from around the world, estimating that **1.3 to 2.3 million people volunteered each year and made in-kind contributions worth up to US\$2.5 billion/year.**
- Mackechnie et al. (2011) reported that terrestrial biodiversity in the United Kingdom involved more than 30 different organizations to which **volunteer contributions had an estimated value of £20million (2007–08) for a government investment of £7 million.**

# Value of Citizen Science

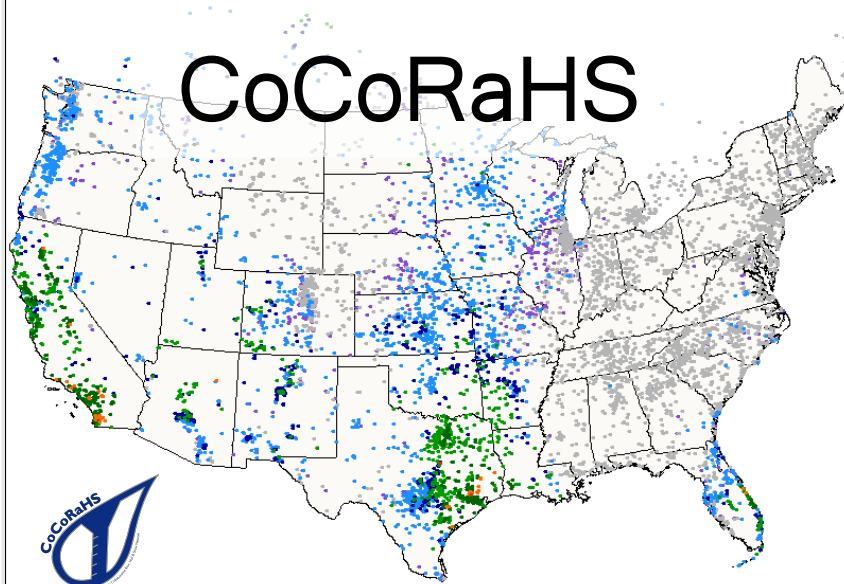


- Saurermann et al (2014) estimated that the total value of volunteer in-kind contributions over the first 6 months for a group of seven Zooniverse projects amounted to **\$1,554,474.**

Daily Precipitation (inches x.xx), for the 24 hour period ending ~7:00 am  
USA 1/7/2016

0.0 Trace 0.01-0.16 0.17-0.32 0.33-0.82 0.83-1.98 1.99-2.97 2.98-3.30

# CoCoRaHS



# Old Weather



# Nature's Notebook

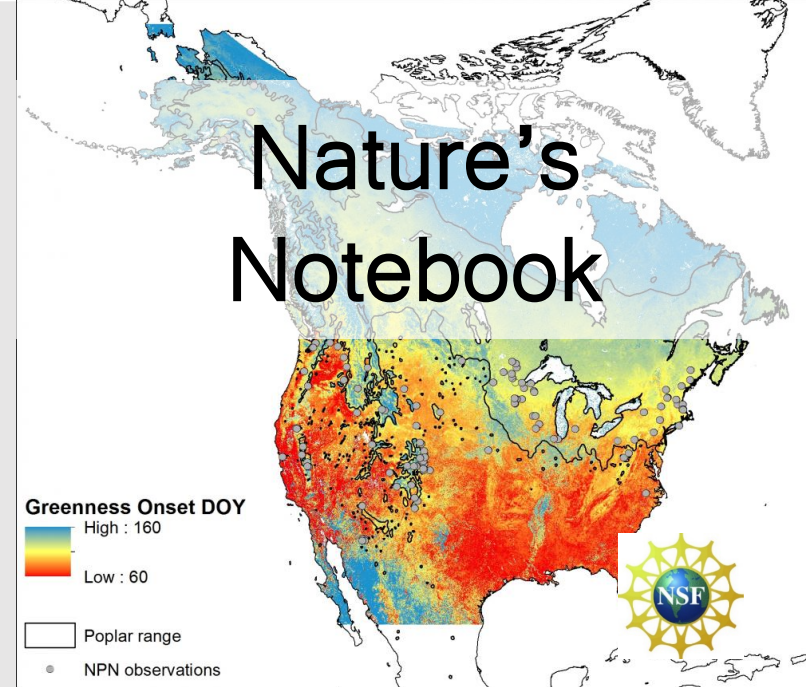
Greenness Onset DOY

High : 160

Low : 60

Poplar range

NPN observations



# CrowdHydrology



# Soil Moisture



Youth  
Learning  
As  
Citizen  
Environmental  
Scientists

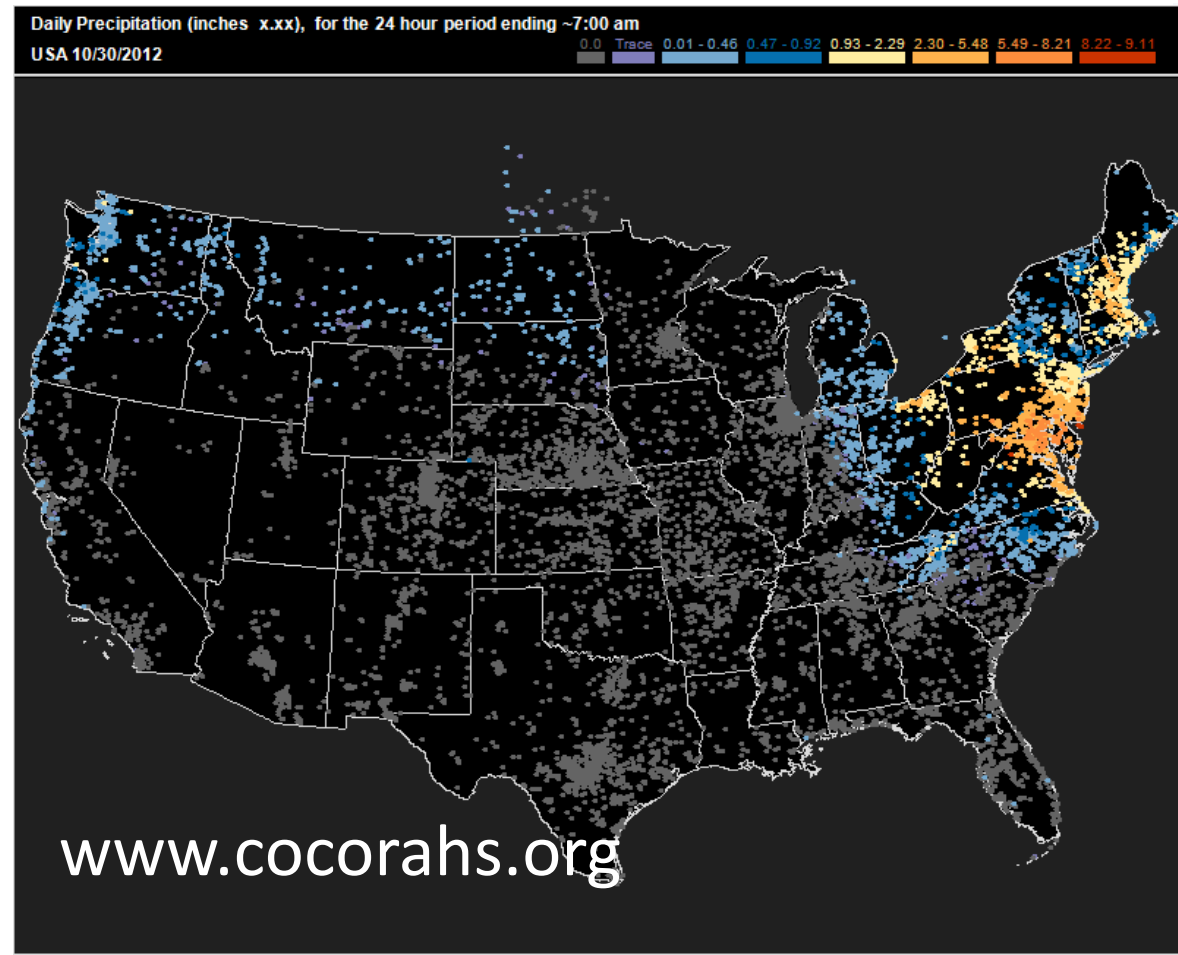
# Air Quality





# Community Collaborative Rain, Hail, and Snow Network

engages more than 35,000 volunteer observers to measure precipitation using standardized rain gauges.

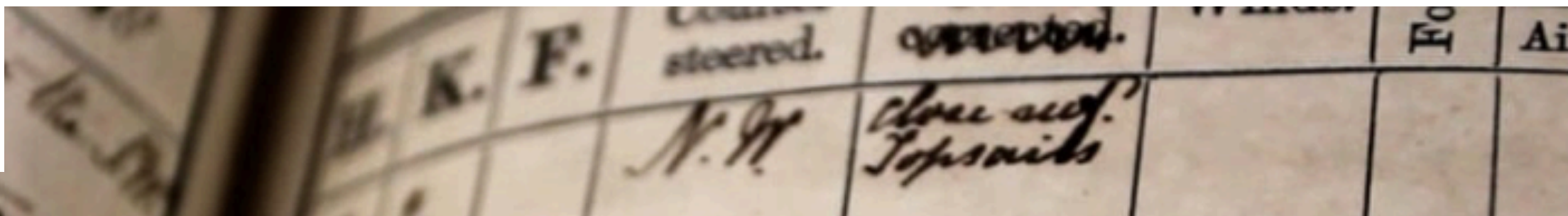




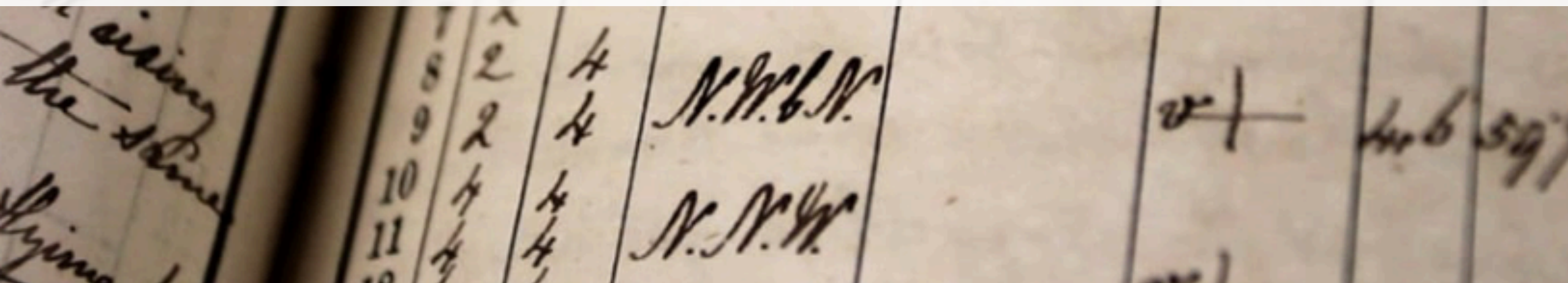


Nolan Doesken, founder of CoCoRaHS, with the rain gauge in the White House Garden

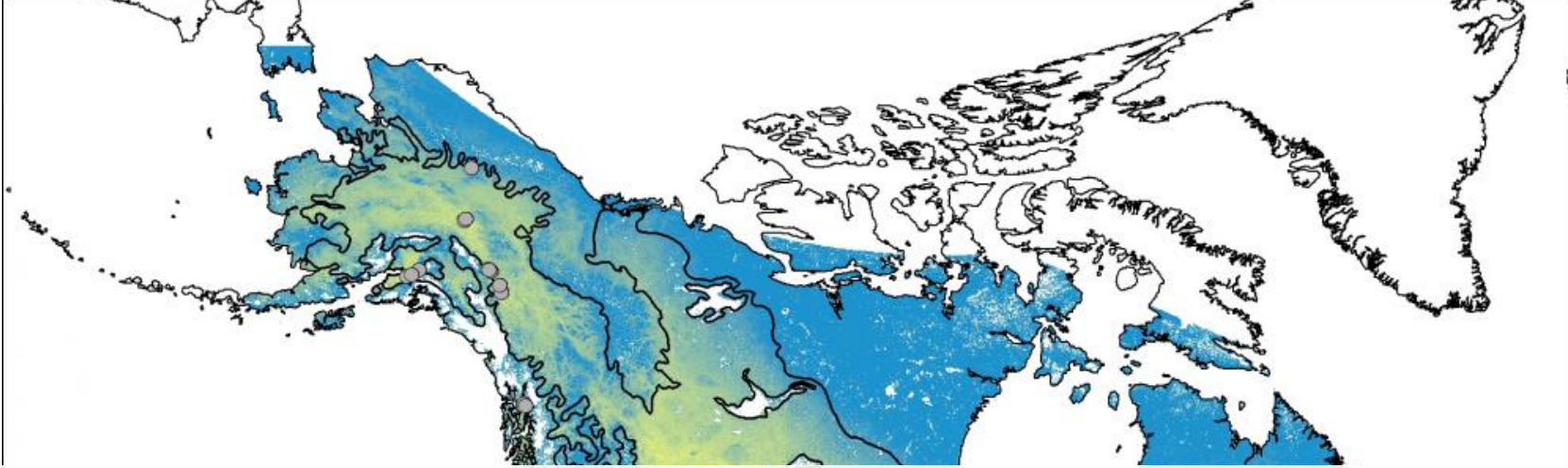




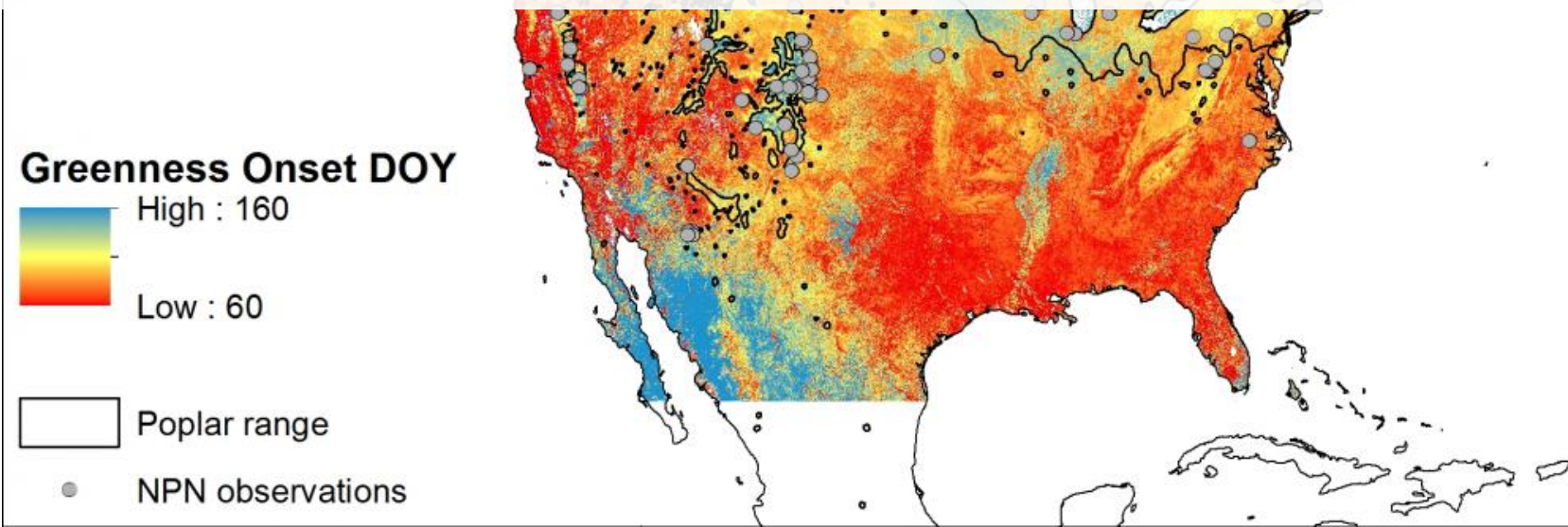
**Old Weather** coordinates crowdsourcing for annotating and transcribing ship log books for temperature, pressure, and weather events. As of Dec 2015, nearly 21,000 volunteers have transcribed over 7.5 million weather observations.







In 2014, **Nature's Notebook** volunteers recorded more than 1 million observations on plants and animals that scientists use to analyze environmental change.



Chris Lowry, University at Buffalo, and Mike Fienen, USGS WI Water Science Center

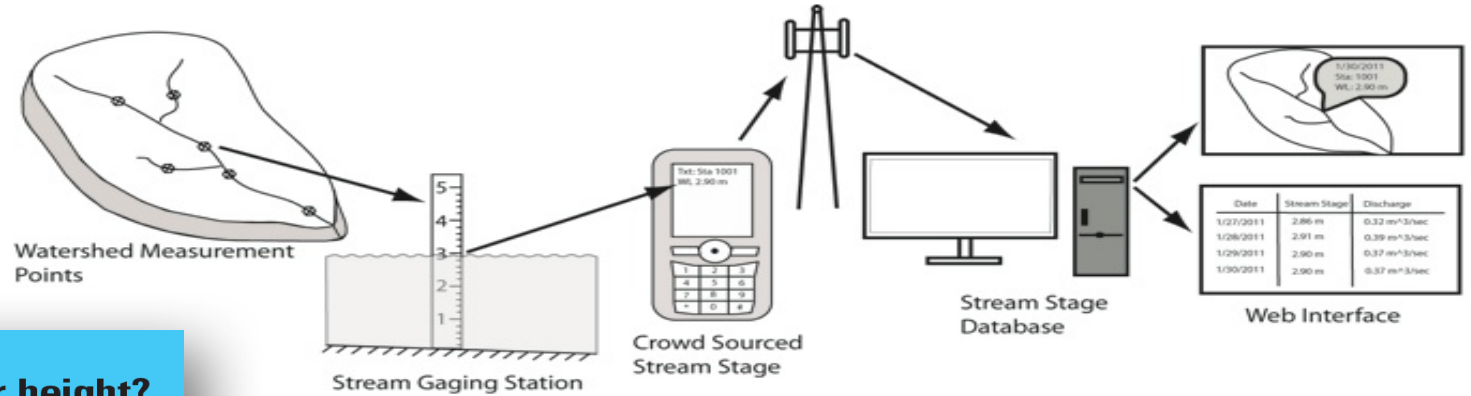


## What's the water height?

Text "WI1009" and the  
current height to:  
**608-514-1889**



CrowdHydrology



## Hydrograph at NY1000



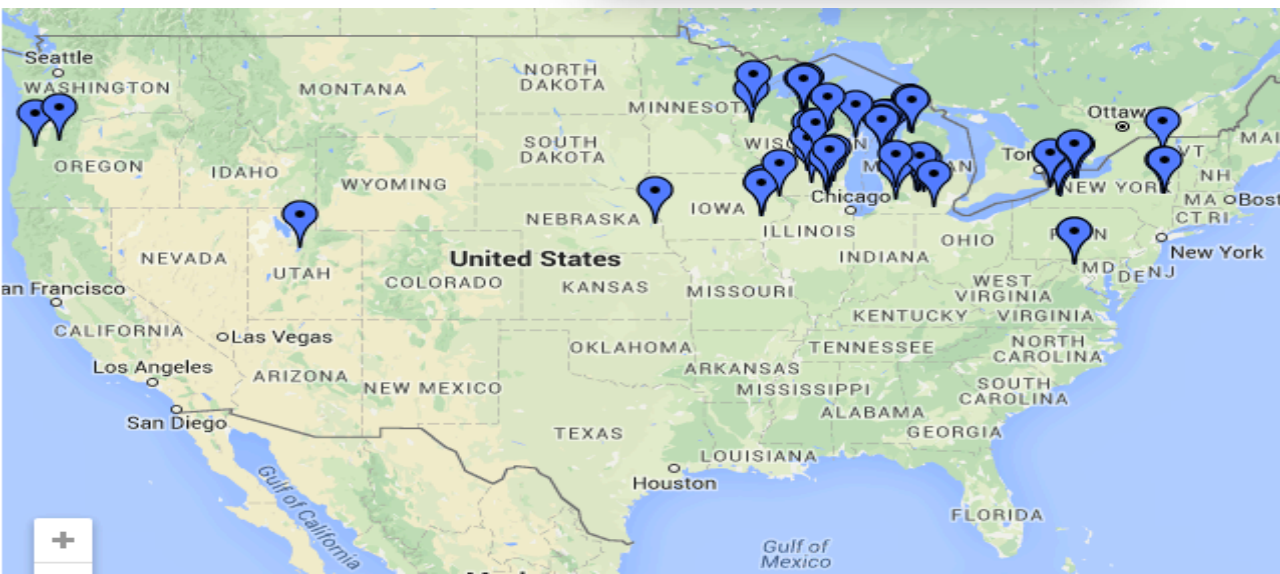
Total number of gages: >60

Usable messages: 90%

Total Observations: >10,000

Total unique users: >4500

(Fienen and Lowry, 2012; Lowry and Fienen, 2013)





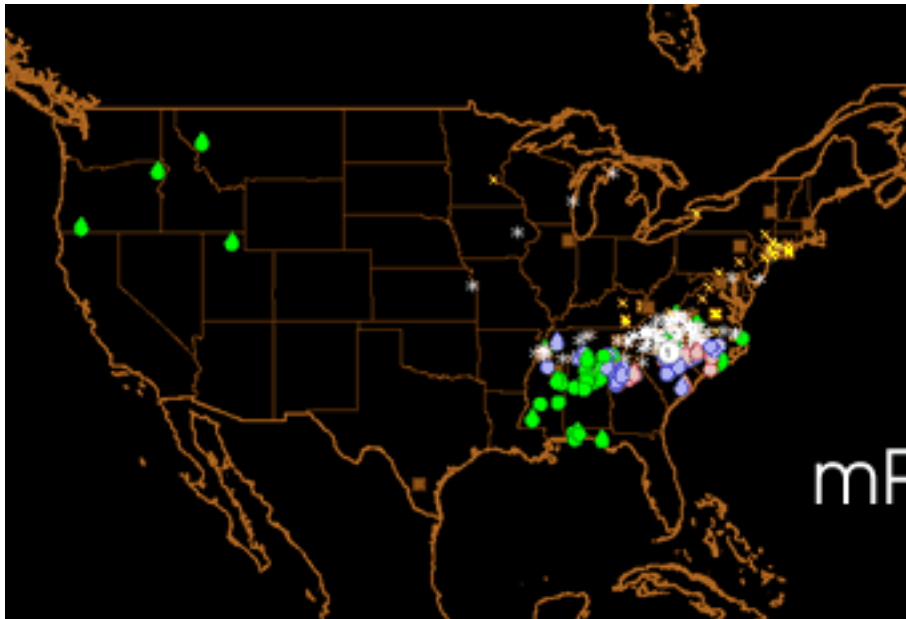




**NOAA**

NATIONAL OCEANIC AND  
ATMOSPHERIC ADMINISTRATION  
UNITED STATES DEPARTMENT OF COMMERCE

**mPING** mobile app has collected more than 860,000 ground-based observations that help verify weather models.



mPING — crowdsourcing weather observations



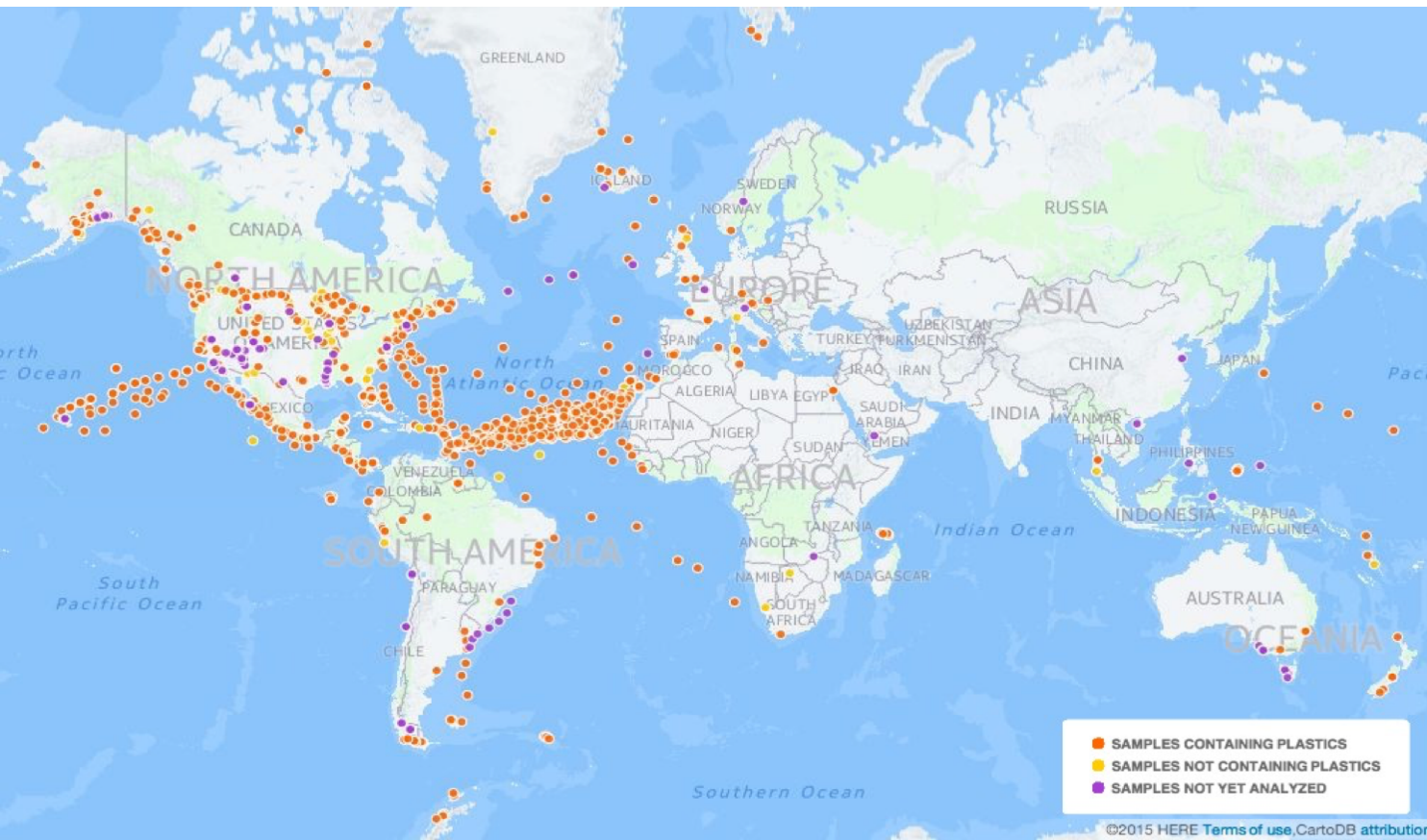
F BOURGOUIN



The Clean Air Coalition of Western NY and  
Tonawanda Bucket Brigade detected high levels of  
benzene in the air, spurring an official investigation of  
Tonawanda Coke that lead to prosecution and an  
**86% reduction in benzene.**



Adventurers and Scientists for Conservation has collected 1,346 marine and 429 freshwater 1L surface grab samples worldwide, finding microplastics (91%, 62%).



# Challenges

- **Data quality assurance** (e.g., DataONE/Wiggins 2013)
- **Legal and Ethical** (Privacy, Liability, Intellectual Property, IRBs)
- **Security of open systems**
- **Cyber Infrastructure** to support long-term management and sharing of citizen science data, metadata, and related media.
- **Project Evaluation and Impact**
- **Value of citizen science** (social, economic)
- **Social science research** (e.g., volunteer motivations, incentives)



# Overcoming Barriers

- **Serve as a Connector:** Hosted monthly live webcasts, roundtables, and workshops, connecting government staff with researchers, practitioners, and industry.
- **Identify and Engage Champions.**
- **Highlight Best Practices Case Studies:** *New Visions in Citizen Science* report & workshop.
- **Grow the Networks.** CCS, CSA.





# People

The US Federal Community of Practice on Crowdsourcing and **Citizen Science** engages 300 federal staff from 40 departments and agencies. <https://www.citizenscience.gov/community/>



# Overcoming Barriers

- **Align with Priorities:** Included citizen science and crowdsourcing projects in the Administration's Open Government National Action Plan. Highlighted in White House Science Fair.
- **Provide Top Level Support and Guidance:** Written endorsement by Agency officials. Collaborated with White House to organize the Citizen Science Forum and shape the President's Strategy for American Innovation and a White House Memo on Citizen Science.

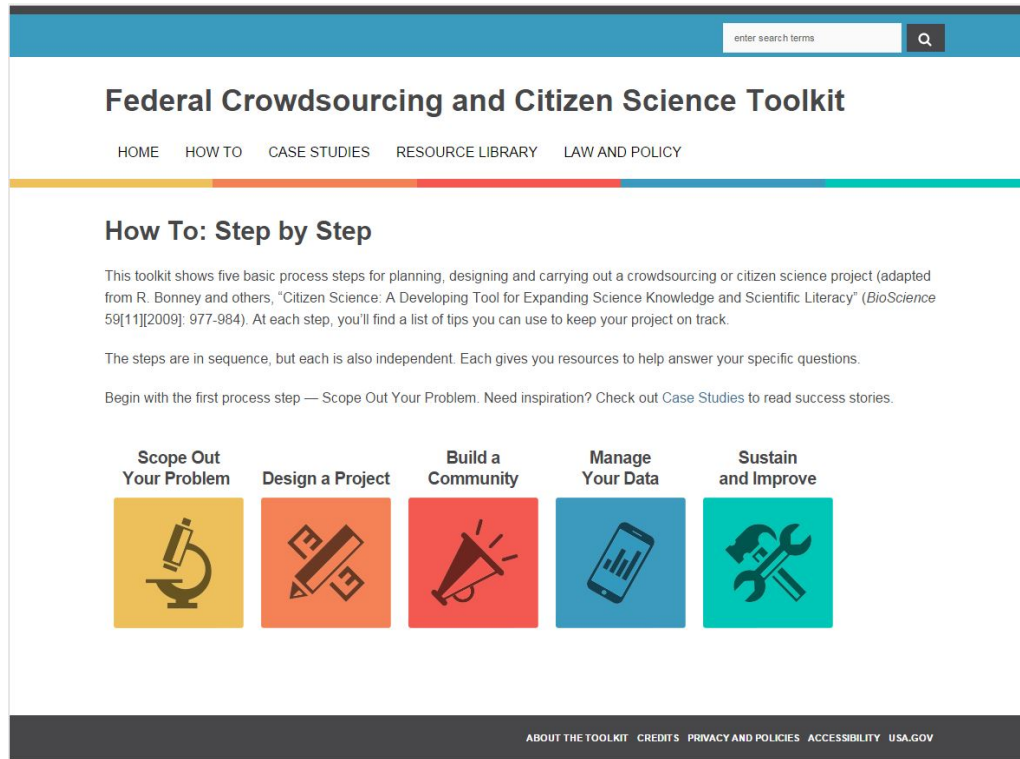


# Develop and Share Best Practices

- Citizen Science Central Toolkit
- Federal Citizen Science Toolkit
- Zooniverse Project Builder
- EPA Air Sensor Toolkit for Citizen Science
- DataONE Data Management Guide for Citizen Scientists



# How can we engage 2.8 million Federal staff?



- Diverse set of case studies
- “How to” information on program design, data issues, working with participants, etc.
- Information on resources, support and partner organizations
- Law and policy guidance

*Developed by 125 members of the Federal Community of Practice on Crowdsourcing and Citizen Science, in collaboration with the White House:*

<https://crowdsourcing-toolkit.sites.usa.gov>

# How do we ensure data quality?

- Quality Assurance Project Plan
- Repeated sample/tasks
- Participant tasks involving control items
- Uniform or calibrated equipment
- Personal knowledge of participant skills/expertise
- Participant training
- Participant testing
- Rating participant performance
- Filtering of unusual reports
- Contacting participants about unusual reports
- Automatic recognition techniques
- Expert Review
- Paper data sheets submitted in addition to online entry
- Digital vouchers
- Data triangulation
- Data normalization
- Data mining
- Data quality documentation

# Overcoming Barriers

- **Understand and Address Legal Issues:** Commissioned legal analyses of **privacy, IP, and liability** implications of citizen science. Wilson Center's legal interactive tool. Connected the legal and innovation communities.
- **Increase Funding Opportunities:** NSF Core Priority Area for 2016; NASA, IARPA, NOAA, and NIH call for proposals; and the FY'17 White House Budget Memo for R&D.

# Overcoming Barriers

- **Incorporate into Policy:** Open Government National Action Plans. White House Memo on Crowdsourcing and Citizen Science, issued Sept 30, 2015. President's Strategy for American Innovation.
- **Inform Legislation:** Sen. Coons introduced the Crowdsourcing and Citizen Science Act 2015.





# US Open Gov National Action Plan 2013



“...Recognizing the value of the American public as a strategic partner in addressing some of the country’s most pressing challenges, the United States will work to more effectively harness the expertise, ingenuity, and creativity of the American public by enabling, accelerating, and scaling the use of open innovation methods across the Federal Government...”



# Holdren Memorandum, Sept 30, 2015

- Core Principals: ( 1 ) Data Quality/Fitness for Use, ( 2 ) Openness, ( 3 ) Meaningful Public Participation
- Requirements
  - Improve coordination of and support for citizen science and crowdsourcing within and between federal agencies, designate agency coordinators.
  - **Projects:** Contribute to a public database of Federal citizen science and crowdsourcing projects.

[https://www.whitehouse.gov/sites/default/files/microsites/ostp/holdren\\_citizen\\_science\\_memo\\_092915\\_0.pdf](https://www.whitehouse.gov/sites/default/files/microsites/ostp/holdren_citizen_science_memo_092915_0.pdf)

# President's Strategy on American Innovation, Oct 2015

“A core component of the [Strategy for American Innovation](#) is increasing the ability of agencies to deliver better results at lower costs for the American people—through an Innovation Toolkit. These approaches can increase the effectiveness and agility of the government through improvements in its core processes and ability to solve problems.” This includes: Prizes and Challenges; **Citizen Science and Crowdsourcing (pps 70–73)**; Open Data; Innovative Procurement; etc.

[https://www.whitehouse.gov/sites/default/files/strategy\\_for\\_american\\_innovation\\_october\\_2015.pdf](https://www.whitehouse.gov/sites/default/files/strategy_for_american_innovation_october_2015.pdf)

# Incorporate into Policy

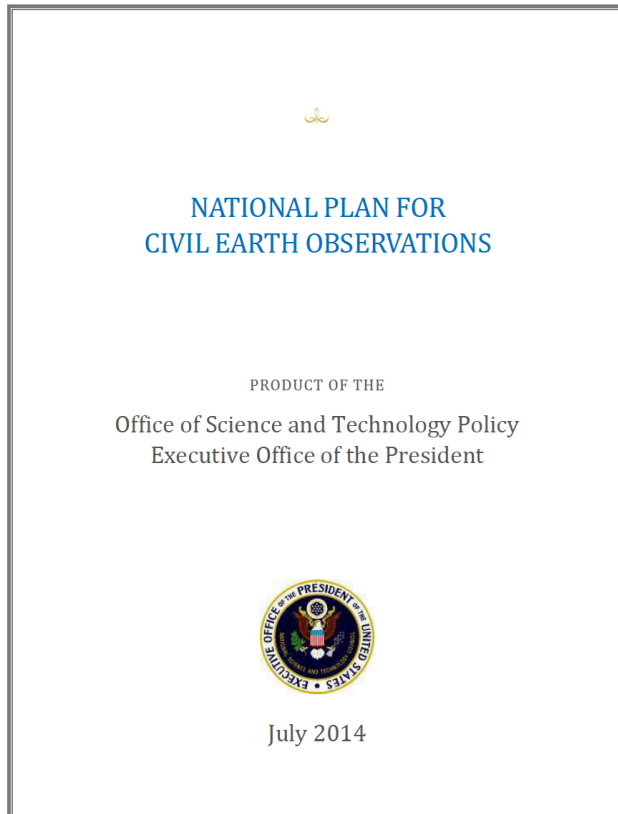
- **National Climate Assessment:** Collaborated with US Global Climate Change Program to identify ways in which citizen science is/could contribute to climate indicators.
- **National Civil Earth Observation Strategy and Plan:** Worked with US Group on Earth Observations and White House Office of Science and Technology Policy to identify where citizen science could augment and enhance satellite-based earth observations.

# US Global Climate Change Program Strategic Plan

- “...observation of ecological and social systems can be dramatically improved by collecting new kinds of data or using new data collection methods, including **emerging opportunities to vastly scale-up the use of non-traditional data sources and “citizen science” research programs...however** it will be challenging to integrate these measurement networks into broader observational systems.”
- “Distributed computing, applications for mobile technology, and social networking have the potential to **dramatically scale up citizen science where interested members of the public serve as observers, modelers, and analyzers of the Earth system**, contributing to the scientific enterprise and broadening the meaning of global change in their own lives.”

Source: Emily Therese Cloyd, USGCRP

# US Nat'l Plan for Civil Earth Observations



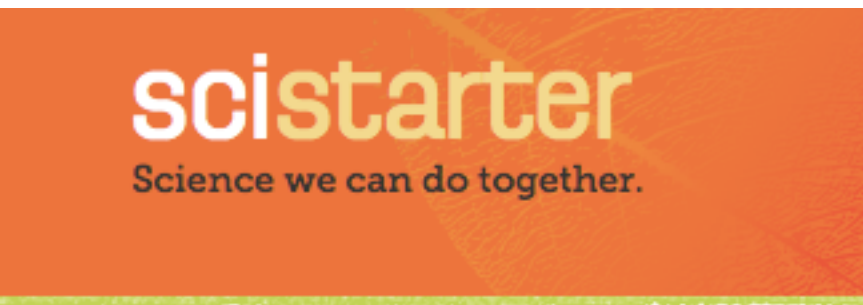
- Focus on user needs and measurements
- Crowdsourcing and citizen science included for: improving data management; increasing efficiency and cost savings; improving observational density and sampling
- Expanding availability and use of open data

# Incorporate into Environmental Policy

- Investing in Citizen Science Can Improve Natural Resource Management and Environmental Protection (McKinley et al)  
[https://www.researchgate.net/publication/282186399\\_Investing\\_in\\_Citizen\\_Science\\_Can\\_Improve\\_Natural\\_Resource\\_Management\\_and\\_Environmental\\_Protection](https://www.researchgate.net/publication/282186399_Investing_in_Citizen_Science_Can_Improve_Natural_Resource_Management_and_Environmental_Protection)
- Citizen Science and Policy: a European Perspective (Muki Hakley) [https://www.wilsoncenter.org/sites/default/files/Citizen\\_Science\\_Policy\\_European\\_Perspective\\_Haklay.pdf](https://www.wilsoncenter.org/sites/default/files/Citizen_Science_Policy_European_Perspective_Haklay.pdf)
- Next Generation Compliance is Here (David Elam)

# Overcoming Barriers

- Improve Discovery of Projects: SciStarter.com



## project finder

pick an activity



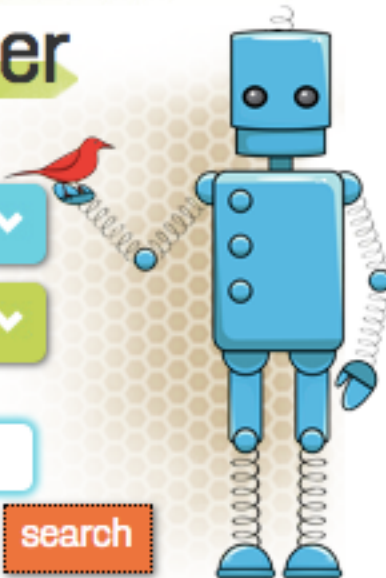
pick a topic



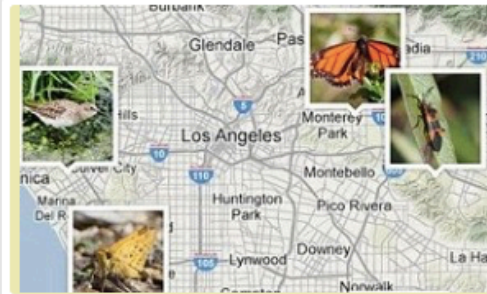
search keywords

advanced search

search



## project search results



### LA. Nature Map

The L.A. Nature Map hosted by the Natural History Museum of Los Angeles is an interactive map that displays local plant and wildlife observations.



### Los Angeles Butterfly Survey

The Natural History Museum of Los Angeles County is partnering with Butterflies and Moths of North America (BAMONA) to share data and learn more about L.A. butterflies and moths. Help us find and photograph them in Los Angeles.



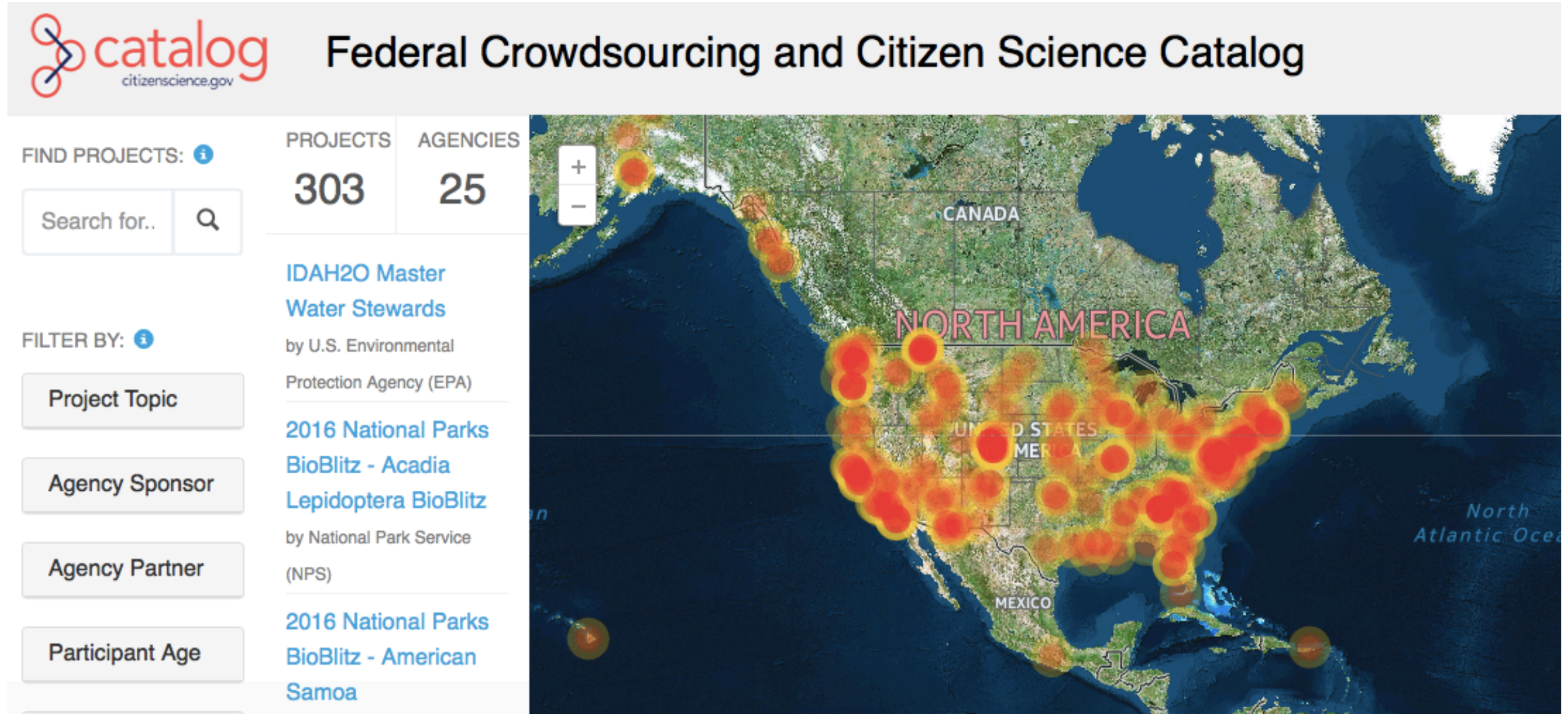
### LA Spider Survey

In order to conduct a large-scale survey of urban spiders, we need the help of the public. We are asking people to collect spiders in their homes and gardens, fill out a simple data sheet about their collection, and send or bring the spiders and forms to the Natural History Museum.



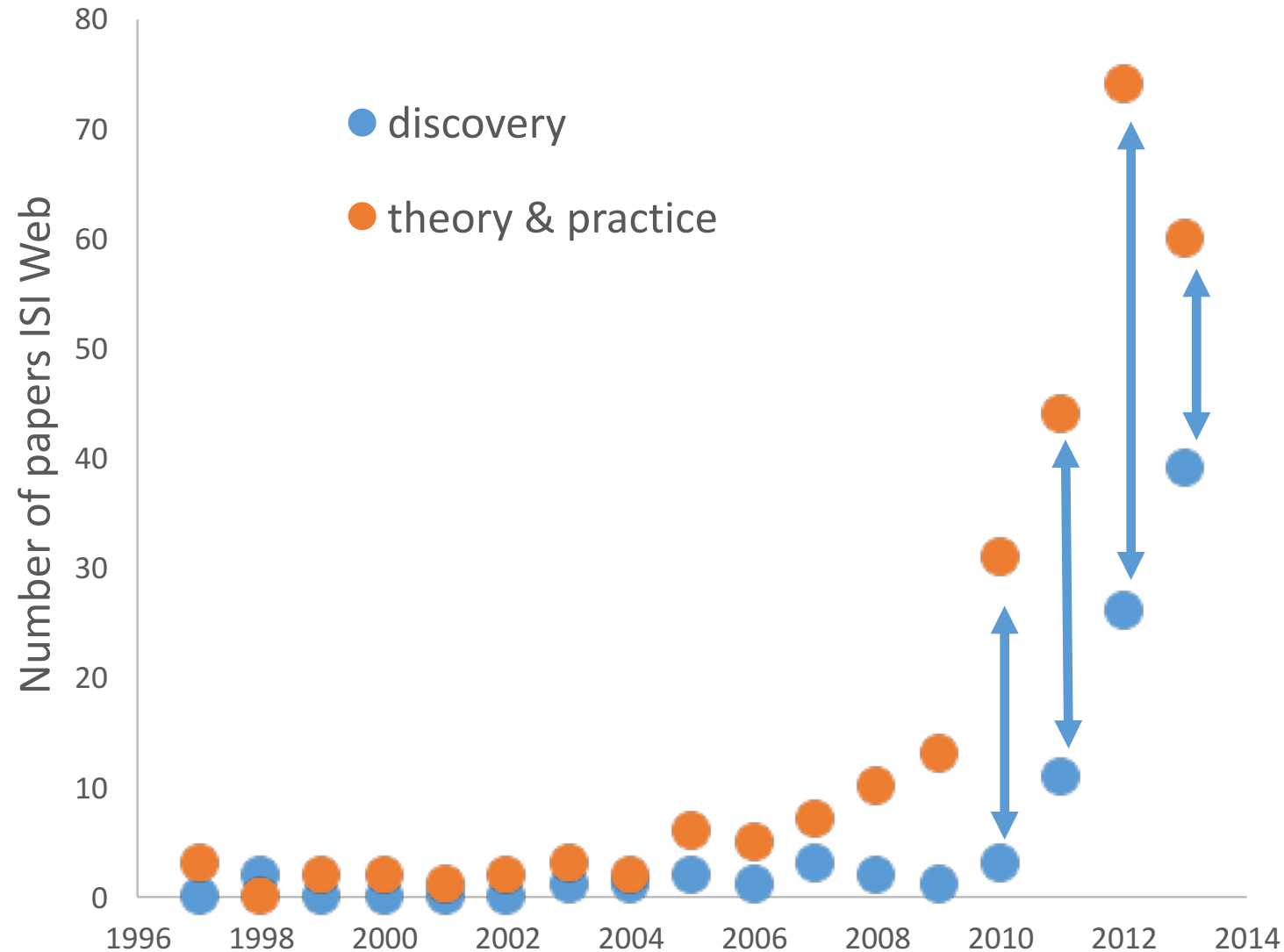
# Overcoming Barriers

- Improve Discovery of Projects: [citizenscience.gov](https://citizenscience.gov)

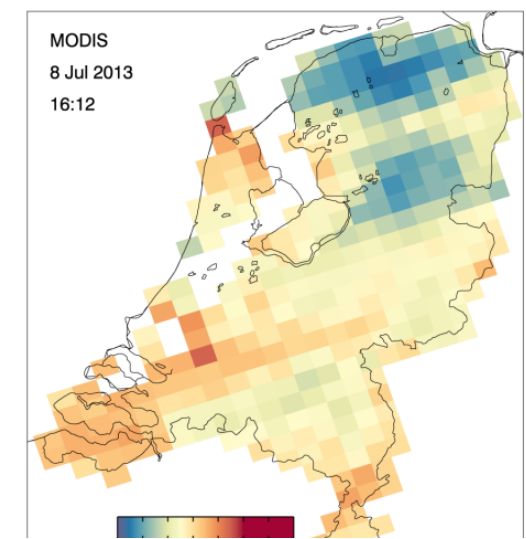
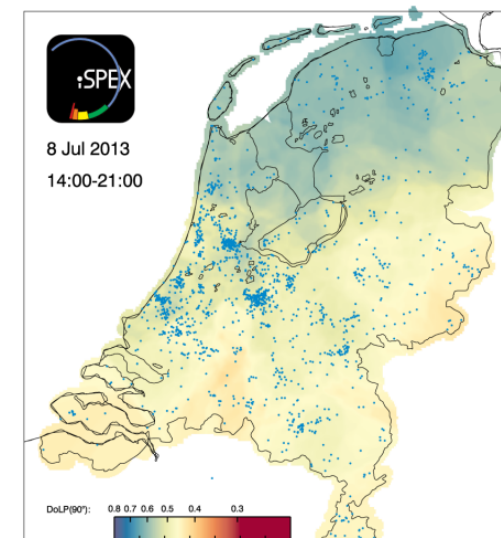




- **Improve Access to Articles:** Citizen Science Journal



60–80%  
of total  
(visible)  
growth  
are  
papers  
about  
theory &  
practice



Advances in technology are enabling and enhancing citizen science projects







**Citizen science at EPA:** 1) Work with **communities** to understand local problems; 2) **Monitor** the environment for environmental protection; 3) Engage volunteers in **research** relevant to EPA's mission; 4) **Educate** the public about environmental issues.





Air Sensor Toolbox for Citizen Scientists provides guidance on affordable, next-generation air quality sensors.





## ***RESEARCH AND DEVELOPMENT HIGHLIGHTS:*** **MOBILE SENSORS AND APPLICATIONS** **FOR AIR POLLUTANTS**



# Big Data Innovation Hubs



## Goals

- To stimulate an agile and sustainable national Big Data innovation ecosystem; and
- To accelerate partnerships among people in business, academia, and government who apply data science and analytics to advance science, address societal challenges, and spur economic development in the region.
- <http://www.southbigdatahub.org>

# Big Data Innovation Hubs



## Cyberinfrastructure Working Group

- Mapping existing collaborations, services, and resources, and identifying Hub member needs.
- Facilitating distributed data sharing, computing, and analysis across institutions and partners.
- **\$3M in Microsoft Azure cloud computing credits!**



**Lea Shanley, PhD**

co-Executive Director, South Big Data Innovation Hub

@ Renaissance Computing Institute

University of North Carolina at Chapel Hill

Lshanley@renci.org

<http://southbigdatahub.org>