Empowering Citizens to Ground Truth Science

Sophia B. Liu, Ph.D.
U.S. Geological Survey
Mendenhall Postdoc Fellow

National Geospatial Advisory Committee Meeting
Crowd-Sourced Data Spotlight Session
March 17, 2015
The Rise of Public Engagement

- **Citizen Science** is a form of open collaboration where members of the public participate in the scientific process in ways that may include identifying research questions, making new discoveries, collecting and analyzing data, interpreting results, developing technologies and applications, or problem solving.

- **Crowdsourcing** is a process where individuals or organizations submit an open call for voluntary contributions from a large group of unknown individuals (“the crowd”) or, in some cases, a bounded group of trusted individuals or experts.

- **Crowdmapping** is a process where individuals or organizations submit an open call for volunteered geographic information (VGI) or information with an associated geographic location from volunteers to produce collaborative maps.

*(Definitions from the Federal Community of Practice for Crowdsourcing and Citizen Science)*
USGS Citizen Science Projects

http://txpub.usgs.gov/myscience/
The National Map Corps

http://nationalmap.gov/TheNationalMapCorps

- Uses the Open Street Map platform but not the data due to licensing issues
- Volunteers map public building structures (e.g., police stations, post offices, schools, hospitals)
- Volunteered Geographic Information (VGI) is peer reviewed then officially integrated into The National Map
- Over 2,400 volunteers have contributed more than 116,000 points
Tweet Earthquake Dispatch

http://earthquake.usgs.gov/earthquakes/ted/

- Broadcast public Tweet alerts @USGSted indicating frequency of earthquake tweets and official USGS scientific seismic data
- Internal alert system to detect felt earthquakes based on an event detector algorithm: $C(t) = \frac{STA}{(mm \times LTA + cc)}$
- Twitter can be faster and can provide pre-seismological detections
- Detects lower magnitude felt earthquakes in populated but sparsely instrumented regions that can be missed teleseismically
- Inexpensive to develop, access, and use information from the crowd
Tweet Earthquake Dispatch

M 4.0 Hollis Center, Maine
2012-10-16 23:12:23 UTC  (Origin Time)

Tweet Frequency:
2754 tweets / min

TED Alert Time:
2012-10-16 23:13:02
39 seconds after origin time

TED Location Estimate:
Portland, Maine, USA
32 kilometers from epicenter
Tweet Earthquake Dispatch

Adam Dziewonski Observatory, Oak Ridge, MA, USA (GSN) IRIS/USGS, Harvard University
142 km from Epicenter

Origin Time
23:12:23

TED Alert Time
23:13:02
USGS lacks the personnel and capacity to analyze all the USGS aerial photos taken after every storm.

USGS coastal change prediction models need to be validated using geospatial ground truth data.

Compare and classify aerial photos of the coast before and after extreme storms like Hurricane Sandy.

Ground truth and generate coastal change predictions using a Bayesian Network model.

Educate the public about coastal vulnerability from extreme storms.

http://coastal.er.usgs.gov/icoast
iCoast – Hurricane Sandy Project

967  Total iCoast users
574  Active users (60%)

7,941  All Post-storm photos classified
462,700  Tags selected in iCoast

45,650  Total classifications (not photos)
2 min 7 secs  Average time to classify one photo
iCoast Users

- 967 Users
  - 574 (60%) users have classified at least one photo
  - 393 (40%) users have not classified any photos

- Users by Crowd Type:
  - 136 Coastal & Marine Scientist
  - 39 Coastal Manager or Planner
  - 77 Coastal Resident
  - 23 Watersport Enthusiast
  - 32 Marine Science Student
  - 30 Emergency Manager
  - 61 Digital Crisis Volunteer
  - 282 Interested Public
  - 286 Other

- Top Users:
  - Over 4,700 classifications
  - Retired Air Photo Interpreter
  - CAP Airborne Photographer
  - GISCorps Volunteers
Observer inputs and analysis have been used to train a Bayesian Network model capable of predicting the coastal process from evidence visible in the oblique aerial imagery.

The Bayesian Network model can also predict likelihood of coastal infrastructure damage.

The crowd-based analysis of the coastal change process was very similar to analysis conducted by USGS experts.
Strategic Geo-Sourcing

- Strategically crowd-sourcing the development of geospatial visualizations for Energy, Minerals, and Environmental Health USGS data and maps

- Engaging citizens at all levels of the research process particularly through geospatial storytelling

- Targeting geospatial visualization experts and enthusiasts through local hackathon meetups and the National Day of Civic Hacking June 6, 2015
Ground Truth via Geospatial Imagery
Volunteer, We Can All Be Citizen Scientists!

SOPHIALIU@USGS.GOV