

Developing an Analytical Tool to Compare Hazard-related Crowdsourced and Citizen Science Data to Official Sources

CDI FY19 Statement of Interest

Lead PI Information

PI Name: Sophia B Liu
PI Email Address: sophialiu@usgs.gov
PI ORCID: <https://orcid.org/0000-0002-8191-9582>
PI Organization: Science and Decisions Center
PI Mission Area: Energy and Minerals
PI City, State: Reston, VA

Financial Information

Total Requested Funds: 46613
In-Kind Matching Funds: 104412

Project Information

Project Description: The goal of this project is develop an analytical tool to compare hazard-related crowdsourced and citizen science data to official sources to more easily make the case as to when public-generated data is more useful, timely, and accurate than official sources.

List of Anticipated Deliverables: (1) Analytical tool, (2) Case studies, and (3) standard operating procedures for integrating the use of this analytical tool during a hazard.

SSF Element 1: Analysis

SSF Element 2: Applications

SSF Element 3: Processing

Collaborators

N/A

Title: Developing an Analytical Tool to Compare Hazard-related Crowdsourced and Citizen Science Data to Official Sources

USGS Lead PI: Sophia B Liu

Project Narrative

Many federal agencies are interested in leveraging crowdsourced and citizen science data as well as data from social media for real-time updates on hazards that effect our nation, but there are still misconceptions and distrust in the reliability of public-generated data and how to use such data for operational purposes. The goal of this project is develop an analytical tool to compare hazard-related crowdsourced and citizen science data to official sources to more easily make the case as to when public-generated data is more useful, timely, and accurate than official sources.

Crowdsourced and citizen science data pertaining to hazards are being generated from a number of avenues. Various Digital Volunteer Network (DVNs), such as Standby Task Force and CEDR Digital Corps, have assisted the FEMA Crowdsourcing Unit at the National Response Coordination Center to generate crowdsourced maps that curate information from social media, online websites, and mobile applications. USGS's Did You Feel It? (DYFI), NOAA's Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS), and NASA's Landslide Reporter are just some examples of federal citizen science projects that produce hazard-related citizen science data. Additionally, there is growing interest to leverage social media data by developing data mining tools to better understand and detect hazards, such as the USGS Tweet Earthquake Dispatch (TED) project. Some of these crowdsourcing, citizen science, and social media data mining projects (such as USGS Did You Feel It? and The National Map Corps) have directly integrated these public-generated data with official data but others have yet to determine if the quality of these data warrants this integration into official authoritative products.

In order to more easily compare these public-generated data with official data, analytical tools are needed to compare the difference in quality, timeliness, and accuracy to official data to better explain and show the value of leveraging crowdsourced, citizen science, and social media data. This project will examine and test tools like Esri's Operations Dashboard, Insights, and Time Aware app for ArcGIS, Tableau's interactive dashboards, and Microsoft's Power BI data visualization software. Public-generated spatiotemporal data will be integrated into these analytical tools along with official data to easily compare and visualize how these data differ based on timeliness, quality, and accuracy. The deliverables from this project will consist of an analytical tool or set of tools to compare certain types of crowdsourced, citizen science, and social media data with official data sources that already exist from past hazard events (possibly from Hurricane Florence and Hurricane Michael); case studies that show how public-generated data was more or less useful compared to official data based on timeliness, quality, and accuracy; and standard operating procedures for integrating the use of this analytical tool during a hazard to better inform the use of crowdsourced, citizen science, and social media data to increase situational awareness and decision support for the emergency management community.

This CDI project is intended to inform the CDI Risk Map Project and to build on authoritative national datasets for hazards. This project is designed to inform the analysis, applications, and processing elements of the Science Support Framework and support the Citizen-Centered Innovation Community of Practice. Representatives from the USGS Natural Hazards, Water, Core Science Systems, and Ecosystems mission area, as well as from representatives from FEMA, NOAA, HHS, NASA, and other federal agencies with existing crowdsourcing and citizen science projects will participate in this project.

Estimated budget table

Budget Category	Federal Funding "Requested"	Matching Funds "Proposed"
1. PERSONNEL (SALARIES including benefits):		
Federal Personnel Total:	\$20,000	\$50,000
Contract/Collaborator Personnel Total:	\$10,000	\$30,000
Total Salaries:	\$30,000	\$80,000
2. TRAVEL EXPENSES:		
Travel Total (Per Diem, Airfare, Mileage/Shuttle) x # of Trips:	\$5,000	\$2,000
Other Expenses (e.g. Registration Fees):	\$1,000	\$0
Total Travel Expenses:	\$6,000	\$2,000
3. OTHER DIRECT COSTS: (itemize)		
Equipment (including software, hardware, purchases/rentals):	\$1,000	\$2,000
Publication Costs:	\$100	\$0
Office Supplies, Training, Other Expenses (specify):	\$400	\$0
Total Other Direct Costs:	\$1,500	\$2,000
Total Direct Costs:	\$37,500	\$84,000
Indirect Costs (24.3%):	\$9,112.50	\$20,412
GRAND TOTAL:	\$46,612.50	\$104,412