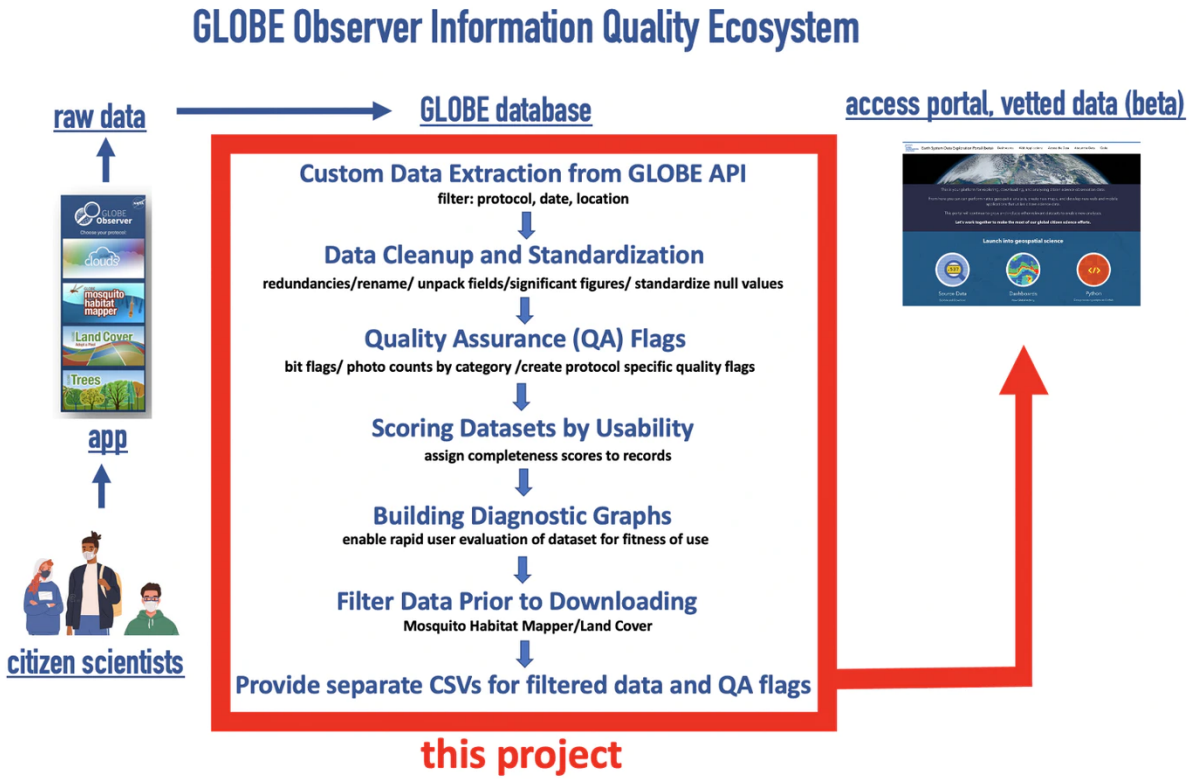
GLOBE Observer's Citizen Science Information Quality Ecosystem: Examples from the Mosquito Habitat Mapper and Land Cover Tools

Authors:

- Prachi Ingle, Westwood High School, Austin TX
- Matteo Kimura, Independence High School, Frisco, TX
- Pratham Babaria, Harmony School of Endeavor, Austin TX
- Andrew Clark, Institute for Global Environmental Strategies, Arlington VA
- Russanne Low PhD, Institute for Global Environmental Strategies, Arlington VA;
- Peder Nelson, Oregon State University, Corvallis OR,
- Rebecca Boger, Brooklyn College, NY

Abstract: *GLOBE Observer's (GO) citizen science information system provides a robust framework to both support and improve data quality. We outline the GO information quality ecosystem for the Mosquito Habitat Mapper and Land Cover tools, and describe how the citizen science participants, devices, protocols, curation, and data access and data analysis procedures contribute to GO information quality. We present a matrix that enables end users to ascertain fitness for use of GO data for their envisioned application, and we discuss how this draft matrix can be queried to iteratively improve the quality of citizen science data obtained using the GO app and the Mosquito Habitat Mapper and Land Cover tools. We invite the Citizen Science Association community to compare our citizen science information ecosystem with that of their own projects and collaboratively identify other features that we can explore to ensure and improve data quality in our respective citizen science data initiatives.*

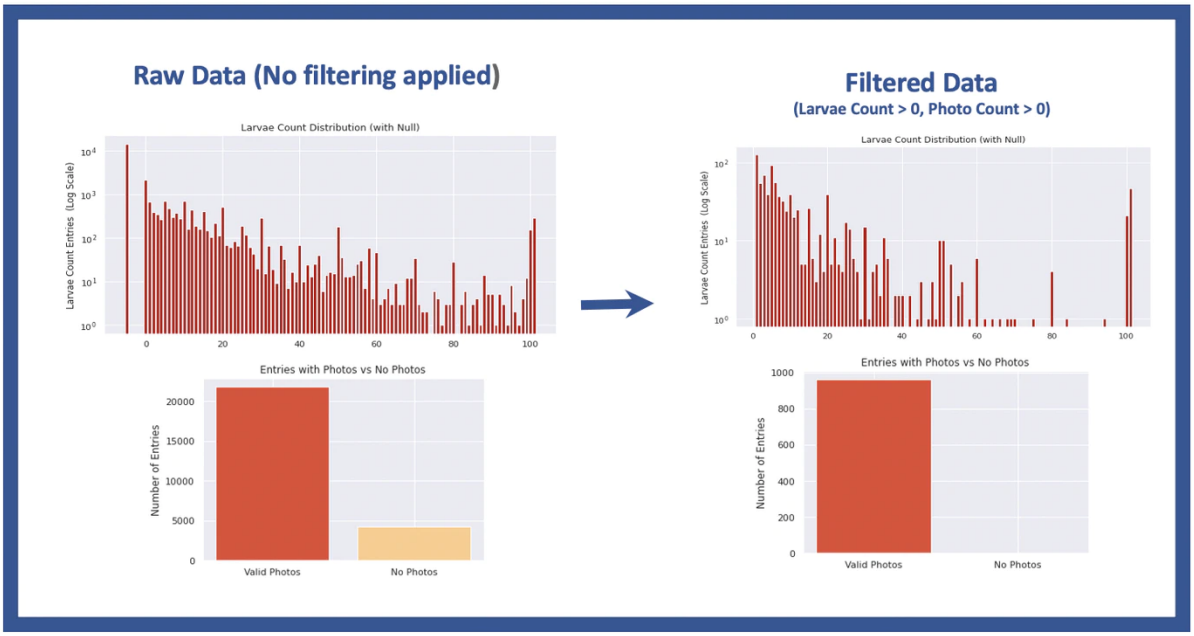


Here we describe the sequence of steps we used in completion of the project. First, we identified those data issues that constrained GLOBE

Observer data accessibility and overall quality for the Mosquito Habitat Mapper protocol and for the Land Cover protocol. We implemented changes needed, and identified future tasks that will improve the user experience:

Identify common data problems	**Establish standard naming convention** **Remove redundant homogenous columns** **Standardize null values** **Limit each field to one value** **Report realistic precision (not machine generated)**
Address critical data issues	**Cleanup and standardize data** **Build in attribute validation procedures** **Develop Quality Assurance (QA) flag system** **Generate data completeness scores** **Enable human and machine-readable metadata**
Identify future data operations	**Refine and update data clean up and QA procedures** **Strengthen applicability of GLOBE legacy datasets** **Ensure alignment to CSA Metadata best Practices** **Tag extraneous or illogical data entries** **Enrich citizen science GLOBE data with satellite data** **Build machine learning models using GLOBE data**

With big data, it is an onerous task for scientists to understand the attributes of a data set and evaluate its fitness for use for a desired application. As an example, here is the result of filtering Mosquito Habitat Mapper data to to assess the whether there are sufficient photo images of larvae. As you can see, this routine reduces the number of records available, but all records include the larvae photos:



Since data collection is voluntary, Citizen Science Data is not always complete. So, we created a Completeness Score that indicates the percentage of data fields filled out by the User. In the map below, larger, darker circles represent more complete entries in the Mosquito Habitat Mapper database. You can see the reach GLOBE Observer Mosquito Habitat Mapper: there are more than 28K observations obtained from citizen scientists from a wide variety of environments:

One of the issues we encountered in conducting research using GLOBE Observer Land Cover and Mosquito Habitat Mapper data was connecting these data sets to each other and enriching with satellite data. We plan to work on this task in the coming months.

Please view our full poster presentation: https://youtu.be/cRWKYfDWR5I

Authors: We are high school citizen scientists who have been working this past year on data accessibility and data quality issues associated with the GLOBE Observer Mosquito Habitat Mapper and Land Cover data.

Mosquito Mappers was sponsored by NESEC, through an award to the Institute for Global Environmental Strategies, Arlington VA. by NASA, award NNX6AE28A.

Boulder, CO · Updated Wed, April 21

Rusty Low