Approaches to Development and Delivery of High-Quality Data to Facilitate Science Outcome

Jake Weltzin, USA National Phenology Network and US Geological Survey

Nature’s Notebook (www.naturesnotebook.org), a participatory monitoring program operated by the USA National Phenology Network, engages citizen scientists and professionals in longitudinal tracking of plant and animal phenology at a national scale. Data collected through Nature’s Notebook have already contributed to over 25 peer-reviewed publications related to science, conservation and natural resource management applications. This presentation will focus on tools and approaches we use to generate and deliver high-quality data that engenders confidence in the use of the data for research. For example, raw and summarized data are freely available and are accompanied by metadata, data-use and data-attribution policies, published protocols, version/change control, documentation of QA/QC, and links to publications that use the data or data products. We also provide our users a variety of tools that can be used to explore, visualize, analyze and download data for a variety of applications.

CoCoRaHS (Community Collaborative Rain, Hail and Snow Network): Abundant Science from Citizen Science

Nolan Doesken, CoCoRaHS; Colorado State University

In its origins, CoCoRaHS posed answerable questions well suited for citizen science: How spatially variable is rainfall? Are there detectable local impacts from this variability? Starting with the type of rain gauge selected, the data collection protocol, and the cyberinfrastructure that was developed to collect, analyze and share volunteer-collected data, CoCoRaHS has been “science friendly” since the beginning. By highlighting precipitation variability every day on easy-access maps and using a strategy to recruit and engage scientists and resource managers who, themselves, are both collectors of and users of these data, CoCoRaHS has been sustainable for nearly two decades. To maximize data quality and participant longevity, participant outreach includes continual reminders of the many and varied science questions and applications CoCoRaHS data contribute to.

Filling in the Blanks: Pollinator Flight Paths Using The Great Sunflower Project Data

Gretchen Lebuhn, San Francisco State University

Many biological phenomena have a periodicity that makes analyzing data collected across large geographic and temporal scales difficult to interpret, especially in cases where data coverage is patchy in space and time as is often the case in citizen science. Thousands of volunteers across the United States participate in The Great Sunflower Project (www.GreatSunflower.org) by contributing data on time-stamped counts of pollinator visits to a focal sunflower variety planted in a local garden. While these data provide a more direct measure of pollination service to a standardized plant and include a measure of effort; each location is sampled at different dates, times and frequencies as well as different points across the local flight season of the pollinators. To overcome this complication, we have used a generalized additive model to generate regional flight curves to calibrate each individual data point and to attain better estimates of pollination service at each site.
Framing the Conversation: A Citizen Science Research Lifecycle and Data Quality Tipping Points
Andrea Wiggins, University of Maryland

Typical research lifecycle models are discipline-specific and omit critical components for citizen science related to volunteer and community management. Starting from a generic high-level lifecycle model as a framework, this talk identifies many of the common tasks and decisions that are important to citizen science projects for planning and managing both research and operations. The detailed life cycle stages provide a lens on the benefits and drawbacks of process transparency for the project organizers, project contributors, and citizen science community at large. The lifecycle model immediately draws attention to topics rarely discussed in citizen science, such as how to wrap-up or hand off management of a project, funding and ethics compliance, and dissemination of assets beyond publications, such as data and software, to maximize project impact and capacity development in the citizen science community.

Guiding Principles for Success in Citizen Science from the Zooniverse
Brooke Simmons*, UC San Diego; Lucy Fortson, University of Minnesota

How did the Zooniverse go from a single unfunded website (Galaxy Zoo) asking the public to help sort data to a multi-disciplinary people-powered research platform with over 1.5 million volunteers participating in over 60 projects which have collectively produced over 100 refereed publications? A few (deceptively) simple guiding principles: (1) the research is the most important goal, (2) do the simplest thing that works, and (3) don't waste the volunteers' time. Effective research with volunteer collaborators requires fostering a healthy community of participants, which can be achieved multiple ways, all of which require active and consistent communication between project researchers and the community. Communicating research results and methods to academic peers is also key, and often requires a different strategy to maximize success compared to the avenues used to communicate the same information to volunteers.

Monarchs + Milkweed + Volunteers = Data: The Great, the non-so Great, and the Clearly Ugly
Carl Stenoien*, University of Minnesota; Karen Oberhauser, University of Minnesota

The Monarch Larva Monitoring Project (MLMP) involves volunteers from across North America who collect long-term data on immature (egg and larva) monarch butterflies and milkweed habitat. The overarching scientific goal of the project is to better understand monarch distribution and abundance during the breeding season. Volunteers choose from among several data collection activities: 1) larval host plant (milkweed) density, 2) monarch egg and larva density, 3) female choice of host plants, and 4) larval survival. Key scientific findings, and the ways in which MLMP data have been used will be described. Because the data are used frequently in scientific publications, as well as to inform public policy, it is important that data accurately represent monarch population dynamics. Because we attempt to anticipate and prevent the inclusion of inaccurate data in analysis, erroneous data are identified and culled via a range of data cleaning methods.
Synthesis and Discussion

Julia Parrish, COASST/University of Washington

Parrish will co-lead the discussion phase of the symposium, together with symposium co-organizers Weltzin and Wiggins. There will not be a formal presentation.