

# Appreciating and Archiving Present-Day Naturalists' Contributions to Science

THERESA CRIMMINS AND MICHAEL CRIMMINS

**Natural history, which is studied** through the practice of directly observing plants and animals, has been the basis for key scientific advancements, including the theories of evolution and resource partitioning. Whether the term is applied to historical figures such as Charles Darwin, Robert MacArthur, Henry David Thoreau, and John Muir or to present-day birdwatchers devotedly pursuing life lists, a *naturalist* can be characterized as an individual pursuing deeper awareness or knowledge of species and ecosystems as a passion rather than as an occupation. Many naturalists contribute valuable observations to formal citizen science programs, such as breeding-bird surveys, the Christmas Bird Count, or plant surveys; there are also many individuals who document observations of plants or animals independent of any formal program or institution. Commitments made by such independent naturalists to their passion can span years or decades and range into the tens of thousands of hours. Data sets can be valued in the hundreds of thousands of dollars. These compilations of methodical and repeated observations have the potential to be very valuable resources to science, and increasingly, data collected by naturalists independent from an agency or academic institution are being used to document environmental change and to advance science (e.g., Ledneva et al. 2004, Cook et al. 2008, Crimmins et al. 2009, Primack and Miller-Rushing 2012). Through our experience working with independent naturalists and their data, we have observed a need for a greater appreciation and more respect in the scientific community for these individuals as a resource unto themselves, as well as a need for more active efforts to engage

them in analyzing and interpreting their data.

By acknowledging the value of the data collected by naturalists working independent of institutions, scientists stand to gain much through working with these data holders to access, analyze, interpret, and archive these records, especially in cases in which the collectors are still living. Living naturalists can bring a wealth of ancillary information and ecosystem knowledge to the table, which can significantly enhance the analysis of their observations. Given the depth of commitment of some independent naturalists, these unique and potentially rich data sets may provide keys to answering questions that could not be addressed through traditional short-term research studies or even longer-term organized efforts. For example, we have worked with one such independent naturalist whose weekly observations span nearly three decades and encompass six vegetation communities along an elevation gradient. Using these data, we have been able to document clear changes in the organization of flowering communities across the elevation gradient and to link them to warming temperatures in the region (Crimmins et al. 2009). However, institutionally supported scientists and independent naturalists represent very different cultures, and bridging this gap to facilitate data sharing, long-term data use, and the archiving of those data can be difficult.

Within the scientific community, the value of data sharing is widely recognized (Parr 2007, Pullin and Salafsky 2010), and increasingly, institutions and journals require this practice. The National Science Foundation states that funded investigators are “expected to share with other researchers... the primary data, samples, physical collections, and other

supporting materials” ([www.nsf.gov/bfa/dias/policy/dmp.jsp](http://www.nsf.gov/bfa/dias/policy/dmp.jsp)). Similarly, the Ecological Society of America expects authors publishing in its journals to “make the data underlying published articles available” (<http://esapubs.org/esapubs/journals/accepted.htm>) and offers a data registry to facilitate such communication and data sharing among scientists. The recently established Dryad ([www.data-dryad.org](http://www.data-dryad.org)) is one such repository, and DataONE ([www.dataone.org](http://www.dataone.org)) is an example of a current push to develop best practices to facilitate data sharing among ecologists. However, these policies and data-sharing tools were primarily developed with professionals in mind and do not explicitly address independent naturalists' situations. Such open data sharing is not necessarily a given for private data holders.

Independent naturalists may invest huge amounts of time, energy, and money in the pursuit of their observations; paid scientists are typically supported by their institution and paid for the time they spend in the laboratory or in the field. These differences in expenditures can color attitudes regarding how openly data might be shared. Namely, individuals representing the institutional realm may—perhaps inadvertently—have unreasonable expectations for independent individuals' willingness to share the rights and access to their data. There are several reasons for which independent individuals may be wary of such openness, including fears of a loss of control over their data and its interpretation or credit for their contributions. If they do choose to share their observations with others, independent naturalists may wish to retain some level of ownership or control over their data, in the form of acknowledgement in publications, involvement in analysis

or interpretation, or authorship on publications including the data. We have also witnessed naturalists hesitate to share their observations for reasons including a lack of confidence that their data are actually useful or valuable, given their noncredentialed status; a concern over a lack of recognition of their efforts; a distrust of academic institutions; and a fear of rejection or dismissal by professional scientists.

To ease the clash of cultures that can occur between independent naturalists who may not be comfortable with sharing their life's work and professional scientists who may have expectations for ready access to such observations, it may be best to treat the relationship as a partnership. Involving a living data collector in every stage of a project concerning their data may seem like an obvious choice, but our experience has been that this practice is frequently planned but not actually carried out. We call for a more overt effort in this direction on the behalf of scientists, engaging naturalists, to the extent that they desire, at every stage of an effort, including idea and question formulation, analysis, and the drafting of manuscripts. Independent naturalists, who have invested a great deal in the pursuit of their observations, can offer a wealth of ecosystem knowledge that may be valuable in analysis and interpretation, as well as in formulating appropriate research questions. Such deep organismal or system familiarity has the potential to truly enhance the work of professional scientists, who typically do not have the luxury of spending hundreds or thousands of hours in the field. Professional scientists can provide context, supplementary data sets, analytical tools, and a means for making results more readily accessible to the broader scientific community. Frequent contact among members of the independent-naturalist-professional-scientist partnership may also have the added benefit of increasing the naturalist's familiarity with the scientific process.

Another improvement that can be made to facilitate independent-naturalist-professional-scientist data

sharing is the development of data-use and data-attribution policies. Data-use policies establish ground rules for the use or sharing of data among parties, whereas data-attribution policies define expectations for acknowledgement and credit for the use of data. Such policies formalize the partnership and can be written to directly address any specific concerns that members of the partnership harbor. Ideally, they should be revisited periodically over the course of the project and respectfully carried out. We have established such policies with colleagues and an independent naturalist in several instances and feel that in addition to clarifying roles and expectations, they have helped the naturalist to feel more included.

There is also a clear need to develop mechanisms for a long-term archiving of independent naturalists' data that can ensure that the back-and-forth communication characteristic of a true partnership between the data owner and the analyst can be upheld. What we envision would expand on existing archives that house data and suggest proper use and acknowledgement by also facilitating direct involvement of the data provider in analysis and interpretation. The establishment of these mechanisms will not only demonstrate recognition among the scientific community of the potential value of these observations but will also ensure the observations' long-term preservation. This could possibly be accomplished through data-sharing nodes such as DataONE, the Avian Knowledge Network, or the USA National Phenology Network. This will require creativity and enhancements to the current types of data-sharing policies that are in place at data-sharing nodes.

Independent naturalists and their observations have a demonstrated value to science, although these resources are at risk from a cultural divide and from a lack of mechanisms to discover and work with the data holders and to share and archive their data. We are calling for a serious treatment of this situation; science and society have much to lose by not valuing both the data and the collectors. It is imperative to find ways

to discover and archive independent naturalists' observations. Naturalists have much to offer to science; consider the lasting contributions of naturalists such as Alfred Wallace and Georges Cuvier. It is also important to engage these individuals as a resource, when they are willing, while we still can, because their direct experience can greatly enhance their raw data. Just imagine what sort of ancillary information we could gain if we could work with Henry David Thoreau or Aldo Leopold while analyzing their meticulous records of plant phenology!

### Acknowledgments

We are thankful to C. David Bertelsen for sharing his rich and extensive data set; our experience working together provided the impetus for this piece. Jake F. Weltzin, Elizabeth Wolkovich, Alyssa Rosemartin, and Ellen Denny provided valuable comments that significantly improved the manuscript. The USA National Phenology Network's National Coordinating Office staff also provided helpful input.

### References cited

- Cook BI, Cook ER, Huth PC, Thompson JE, Forster A, Smiley D. 2008. A cross-taxa phenological dataset from Mohonk Lake, NY and its relationship to climate. *International Journal of Climatology* 28: 1369–1384.
- Crimmins TM, Crimmins MA, Bertelsen CD. 2009. Flowering range changes across an elevation gradient in response to warming summer temperatures. *Global Change Biology* 15: 1141–1152.
- Ledneva A, Miller-Rushing AJ, Primack RB, Imbres C. 2004. Climate change as reflected in a naturalist's diary, Middleborough, Massachusetts. *Wilson Bulletin* 116: 224–231.
- Parr CS. 2007. Open sourcing ecological data. *BioScience* 57: 309–310.
- Primack RB, Miller-Rushing AJ. 2012. Uncovering, collecting, and analyzing records to investigate the ecological impacts of climate change: A template from Thoreau's Concord. *BioScience* 62: 170–181.
- Pullin AS, Salafsky N. 2010. Save the whales? Save the rainforest? Save the data! *Conservation Biology* 24: 915–917.

---

*Theresa Crimmins (theresa@usanpn.org) is a staff member of the USA National Phenology Network, in Tucson, Arizona. Michael Crimmins (crimmins@u.arizona.edu) is in the Department of Soil, Water and Environmental Science at the University of Arizona, in Tucson.*

doi:10.1525/bio.2012.62.6.3