

Environmental Data Initiative Summer Fellowship: An Experience at Archbold Biological Station

Gabriel E. Kamener^{1,2}, Vivienne L. Sclater²

¹Environmental Data Initiative, environmental datainitiative.org

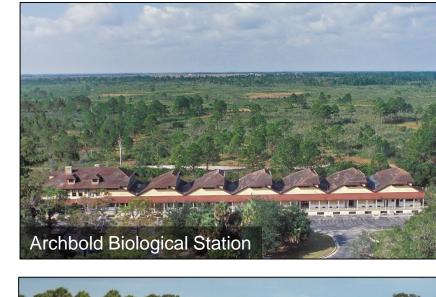
²Archbold Biological Station, Venus, FL



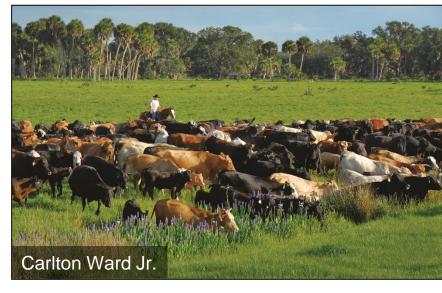
Abstract

Publishing standardized ecological data and high-quality metadata is of increasing importance across multiple NGOs, research institutions, and government agencies; however, publishing high-quality data packages can be burdensome and time consuming for researchers and information managers. The Environmental Data Initiative (EDI) aimed to assist in overcoming such challenges through supporting five Summer 2018 (June 4 – August 13, 2018) Fellowships hosted and mentored by OBFS field stations in its ecological data management training program. This poster explores the experience of summer EDI fellow Gabriel Kamener, his mentor, Data Manager Vivienne Sclater, and researchers at Archbold Biological Station (ABS) and how they benefited from EDI's Ecological Data Management Summer Fellowship Program.

Background









- For more than 75 years, Archbold has served as an internationally celebrated not-for-profit dedicated to science, education, and conservation. Founded in 1941 by Richard Archbold, it is located in one of the most ecologically diverse and threatened regions in the world, the ancient scrub ridge in the heart of Florida.
- Archbold works on the ridges, ranches and rivers of the region and manages over 20,000-acres (with over 19 federally listed threatened and endangered species found throughout).
- Research at Archbold has resulted in 2,300 published papers, but many datasets need to be curated and published for easy access to other researchers.
- Partnering with EDI offered a promising way to move forward with publishing high-quality data and metadata.



Gabriel had a background in biology and ecological fieldwork (including a herpetology internship at ABS), but he was interested in gaining more experience working with data and metadata as an EDI Fellow at ABS.



As the Data Manager at ABS, Vivienne needed someone to take on tidying, quality checking, and publishing high-quality data and metadata to EDI's data portal.

The EDI Process

Create

- Package tables
- QA/QC
- Software tools

Document

- Metadata
- Support standardized metadata (EML)

Publish

- Data repository
- DOIs, data as a valued scientific product

Discover

EDI makes discovery easier by:

- Centralized data with standardized metadata and vocabularies
- Excellent search interface (both manual and API)
- DataONE member node
- EDI works with journals and standards setting organizations

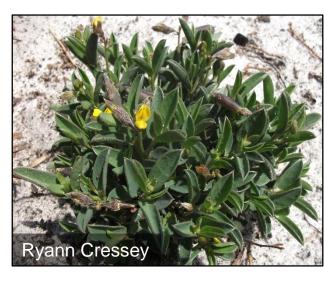
Reuse

EDI facilitates reuse:

- EDI is composed of scientists and data managers experienced in reproducible science
- DOIs, data as a valued scientific product
- Data provenance is documented
- Data package versioning supports dataset updates

Results and Further Work

- Three datasets published during fellowship (including plantdemography and restorationecology data)
- Trained employee hired by ABS at conclusion of fellowship
- Two additional datasets
 published within a month of
 fellowship completion
- ABS plans to publish at least ten more datasets by March 2019





Conclusion

"My EDI Fellowship provided the opportunity to transition from collecting ecological data in the field to managing data and publicly archiving datasets with standardized metadata. EDI's training workshop introduced me to skills needed to tidy, document, and publish datasets, and it offered the opportunity to work with data professionals and other fellows to overcome the challenges of messy data. Additionally, the hands-on experience of working with Vivienne and researchers at ABS allowed me to apply what I learned at the workshop, and it allowed me to expand my data-management skillset. The experience led to a Research Data Assistant position in the station's Data Management department." – Gabriel Kamener, Summer 2018 EDI Fellow

Archbold Biological Station Staff

"EDI's Ecological Data Management Summer Fellowship program has been a great benefit to Archbold Biological Station. Gabriel was able to publish multiple datasets as well as clean one of our larger, more difficult datasets. Gabriel was able to give insightful feedback to researchers and was able to find answers to any questions they had. EDI's commitment to the fellows benefitted Gabriel by providing mentorship and insight as well as Archbold by providing a trained employee to focus on data archiving and documenting the procedure for future use." – Vivienne Sclater, Data Manager

"The EDI fellowship allowed the Plant Ecology Program to make great strides in publishing our long-term plant demographic data. We worked with fellow Gabriel Kamener and Vivienne Sclater of Archbold to get two complex datasets organized and posted. In addition, the procedures and coding that were created will make posting of additional datasets more feasible and improve the quality of those datasets." – Eric Menges, Senior Research Program Director and Research Biologist

"It has been a real bonus for Archbold, hosting EDI Fellow Gabriel Kamener in summer 2018. Everyone at Archbold always has the best of intentions about managing and posting scientific data, but Gabriel helped turn our wishes into reality by publishing five of our field-station datasets this summer (ranging from endangered-plant demography to wetland-community dynamics). Gabriel also built a well-documented, useful library of R scripts and procedures, and he transferred his training from EDI by mentoring our scientific staff, students and research interns. Our scientists are now better prepared to think ahead about data management and to anticipate data publication as a logical final step in the process of science at Archbold." – Hilary Swain, Executive Director

Acknowledgements

Thank you to EDI and Archbold Biological Station for making the fellowship experience possible.

This material is based upon work supported by the National Science Foundation under grants #1565103 and #1629233. Any opinions, findings, conclusions, or recommendations expressed in the material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.

