

NATIONAL SCIENCE FOUNDATION
Review (PI Copy)

Proposal:1546686

PI Name:Zimmerman , Jess

Title: LTER: LTER5: Understanding Ecosystem Change in Northeastern Puerto Rico

Institution:University of Puerto Rico-Rio Piedras

NSF Program:LONG TERM ECOLOGICAL RESEARCH

Principal Investigator:Zimmerman, Jess K.

Rating:Very Good

Review:

In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

1. (intellectual merit): The data management for LUQ has created what looks to be a complete online collection of the data for the study.
2. No comment (NC)
3. The plan does indeed appear well-reasoned, well-organized, and based on a sound rationale.
4. The team has made its capabilities evident by the fact the system is indeed working today.
5. There appears to be adequate resources as long as the current level of resources is expected to remain in effect.

In the context of the five review elements, please evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

1. (broader impacts): This data collection is freely available and searchable online both at the site, as well as at the network level. It is fully compliant with the Provenance Aware System Tracking Architecture (PASTA) framework the LTER network has adopted for inter-site data aggregation. This is evident by the very fact that one can indeed obtain these datasets via the LTER PASTA Portal. PASTA is configured to reject datasets that do not meet a rigorous set of conditions. These conditions include missing or out of range values, incomplete metadata, and improper data format. Thus, these datasets are useful to the larger scientific community.
2. The use of the Drupal Ecological Information Management System (DEIMS) to catalog and share the data is potentially transformative for the LTER IM Network as a relatively unique and new tool that may see more widespread use across the network.
3. The plan, in general, matches that of the other LTER sites and thus should be effective in data collaboration and synthesis.
4. The IM for LUQ does indeed stay engaged in Network activities and contributes to various working groups and projects.
5. NC

Please evaluate the strengths and weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if applicable

NATIONAL SCIENCE FOUNDATION
Review (PI Copy)

Proposal:1546686

PI Name:Zimmerman , Jess

There were specific points related to information management. Item G on the invitation to review this proposal stated that problems existed with the website and access to data. The website is now well done, and access to the data works well. On the other hand, mentioned were specific points related to project design, such as item E, which stated that key portions of the proposed research were not motivated by long-term data, and long-term data did not inform hypotheses or predictions. Although that is clearly a data management issue, as well, it might be better evaluated by someone more familiar with theoretical ecology than data management.

Summary Statement

The website is laid out in an organized fashion and navigation was intuitively easy.

I searched for data on the website in order to see how easy that was to do and to see how well it worked. I found 188 datasets in total. Many of them, however, were multiple datasets bundled together. For example, the dataset labeled "USGS stations" contains 21 datasets from 21 separate locations. So the total number of datasets available is significantly greater than 188.

I randomly selected, and followed links to, datasets in the collection. For each one, I inspected the metadata, and I downloaded the dataset and inspected it down to the point level, and I compared it with the metadata, for traits such as column names, units, and completeness. I also looked at the form and layout of the data. In every single case the dataset was available, and matched the metadata. In total, I opened and inspected 28 datasets.

The data search results page was very easy to read.

All metadata I saw were well beyond EML level 2. Every metadata record inspected had comprehensive descriptions of the units, standardized to the LTER Unit Registry. Additionally, keywords were there, and they were standardized to the LTER controlled vocabulary.

Each metadata record had a link to the Ecological Metadata Language (EML) file used to upload the data to the PASTA system.

The plan for how the Information Manager integrates data management into the project design was clearly stated and seems feasible.

Lastly, I went to the LTER Network PASTA portal to search for LUQ data. Of the 188 datasets on the LUQ website, 145 are available via the PASTA portal. I went back to the LUQ website and got some known keywords and searched for those known keywords on the LTER PASTA portal. In each and every case, the proper corresponding dataset(s) resulted.