In the article, Do Community Rec Improve completeness of a record(cite), LTER samples were compared against other communities in DataONE(cite). These communities used EML or CSDGM to document their resources. While it did appear that the sample from LTER was more complete than other communities using EML with respect to the conceptual version(cite) of the LTER Recommendation for Completeness(cite) in each of the five levels, surprisingly CSDGM collections were more complete in some measures than the LTER sample. This appears to be an effect of required concepts in the FGDC Recommendation that are needed to be schema valid CSDGM(cite visual FGDC)

In other work(agu cite), samples were taken from the LTER memberNode of DataONE using the publishing date (upload date?) to identify date ranges for each year, 2005-2017 starting on the first of the year and ending on the last. The expectation was that records would have an increasing rate of improvement, at least for some levels of the recommendation.

Surprisingly, this was not the case. Samples from later years had no discernible improvement in completeness, or even change in concept usage, not even when PASTA(cite) was available for record creation. There are several potential reasons for this: the sample size is closer to the record total for earlier years, or less complete records may be disproportionally represented in the sample.

Through further investigation of the LTER organization it became apparent that it would not be possible to measure the recommendations effect on records through time for the entire organization. LTER is a single member node of DataONE, but it is not a single organization. It is comprised of many sites, each with their own focus and metadata needs for their resources.

What if this isn’t a technical proficiency or tool availability problem for LTER as a whole, what if it is a social problem of information dissemination? Additionally, do the needs of the organization match the documentation needs of the metadata creators and data managers?

Are the changes in awareness of the recommendation and the availability of tools to create LTER levels complete metadata measurable for the different sites that contribute to the LTER membernode of DataONE? Do the concepts that make up the recommendation become more or less complete over time? Is there a greater effect for the elements that make up these concepts over other EML elements used at these sites?

To answer these questions, collections of records from each site were created for each year records were submitted to the LTER membernode from the set of unobsoleted records in DataONE’s holdings. The process is described in the notebook (cite) contained in the package LTERsitesThroughTime. The metadata data set is zipped up in the package for reproducibility and the solr requests to retrieve records are also in the package, with xsl to create the requests from solr queries. These queries could be changed and the results could be used to create different requests to create a new metadata dataset, making the package reusable.

This package also contains the python, xsl, and Bash required to analyze the dataset and create reports. The reports are written in Excel and uploaded to Google sheets to make sharing easy, but csv output from each type of analysis also exists if different reports or analysis are of interest for novel reuse of the results and code.