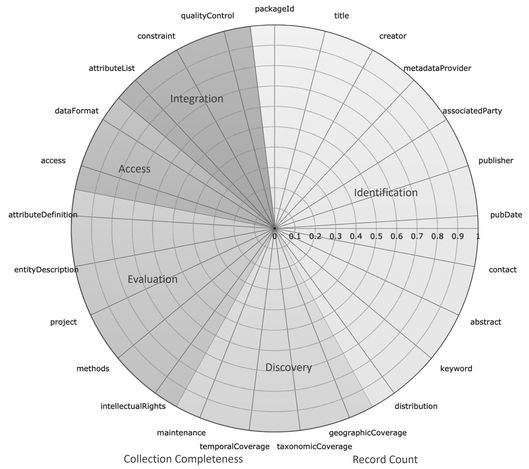
Visualizing the evolution of metadata

Sean Gordon (scgordon@hdfgroup.org), Ted Habermann (ted.habermann@gmail.com)

LTER is a community with many members. In 2004 the EML Best Practices for LTER Sites was released to help site data managers serve community information needs. Pasta provided a tool for consistent creation, evaluation and improvement starting in 2013.

We collected each sites' metadata from the past 14 years and examined what percentage of records in the yearly collections from each site used the elements in the Best Practices to identify effects of the recommendation and toolkit over time. DOI 10.5281/zenodo.1660727

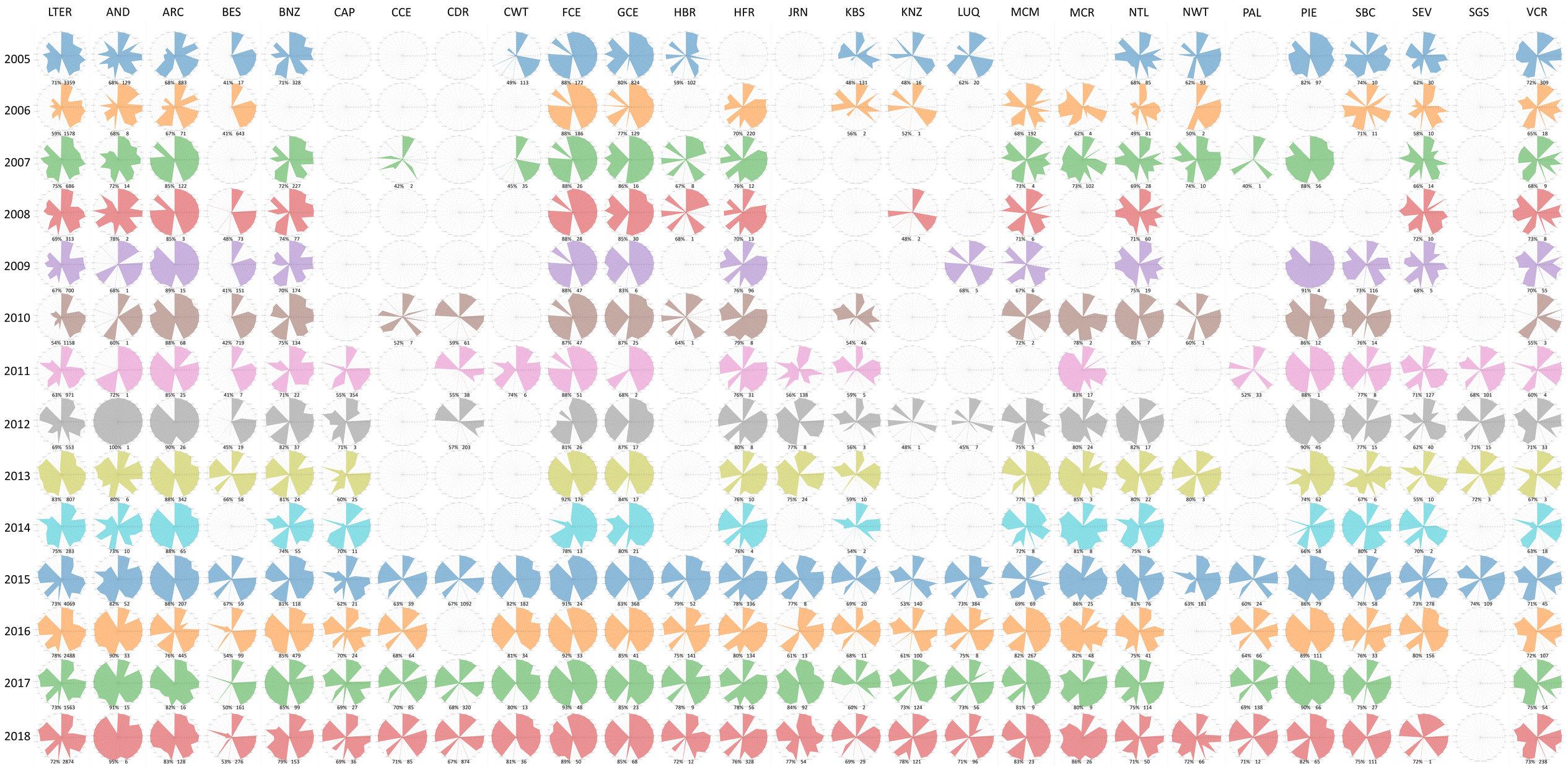
To see if collections are improving, we plot element completeness on a radar chart from zero to one. The elements and levels are plotted clockwise from twelve o'clock. Below each plot is the completeness for the collection and number of records analyzed.



During the early years, collections are generally complete for packageId, title, and creator and less complete for other concepts. Records are less structurally homogeneous and the elements are more likely from the Identification and Discovery levels of the recommendation.

In later years, concept usage is much more homogeneous for any single collection and across all sites, particularly after Pasta came into general use. A greater focus on the elements in levels beyond Discovery and Identification such as Evaluation is apparent.

2013 was the golden year for LTER Recommendation Completeness. Average completeness was 83% across the 807 records uploaded that year.



By 2015 the community decreased utilization of some elements (e.g. metadataProvider, associatedParty, and qualityControl), suggesting that community experience influences recommendation implementation.

Pasta had a profound homogenizing effect on LTER metadata. Sites that previously submitted far less complete records quickly catch up to other sites.

Some sites have consistently submitted fairly complete metadata, and the only changes have been removing the elements the community has deemed less important.