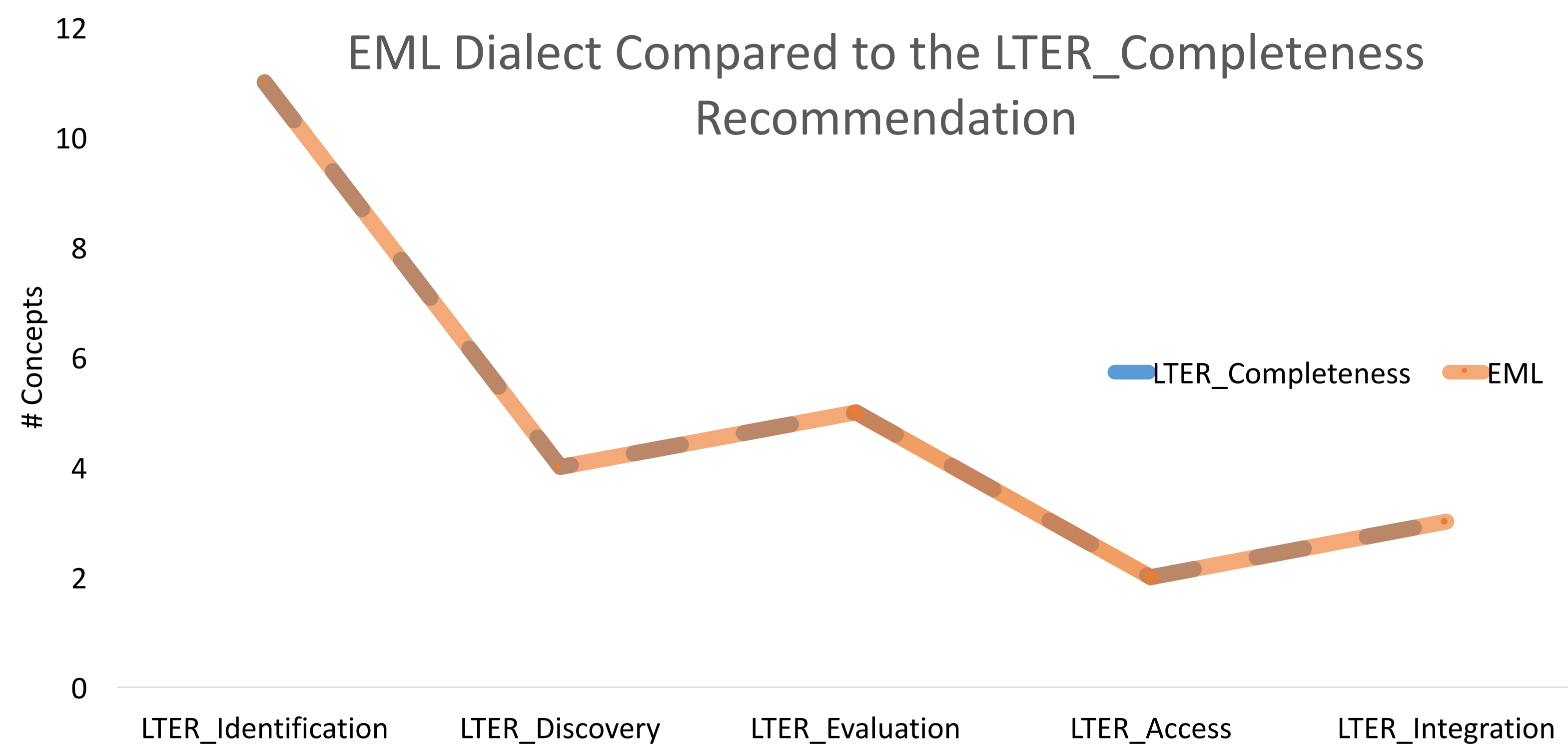


Sean Gordon (scgordon@hdfgroup.org)¹, Ted Habermann¹, Matthew B. Jones², Ben Leinfelder², Bryce Mecum², Lindsay A. Powers³, and Peter Slaughter²
 1. The HDF Group, 2. National Center for Ecological Analysis and Synthesis 3. United States Geological Society

Background

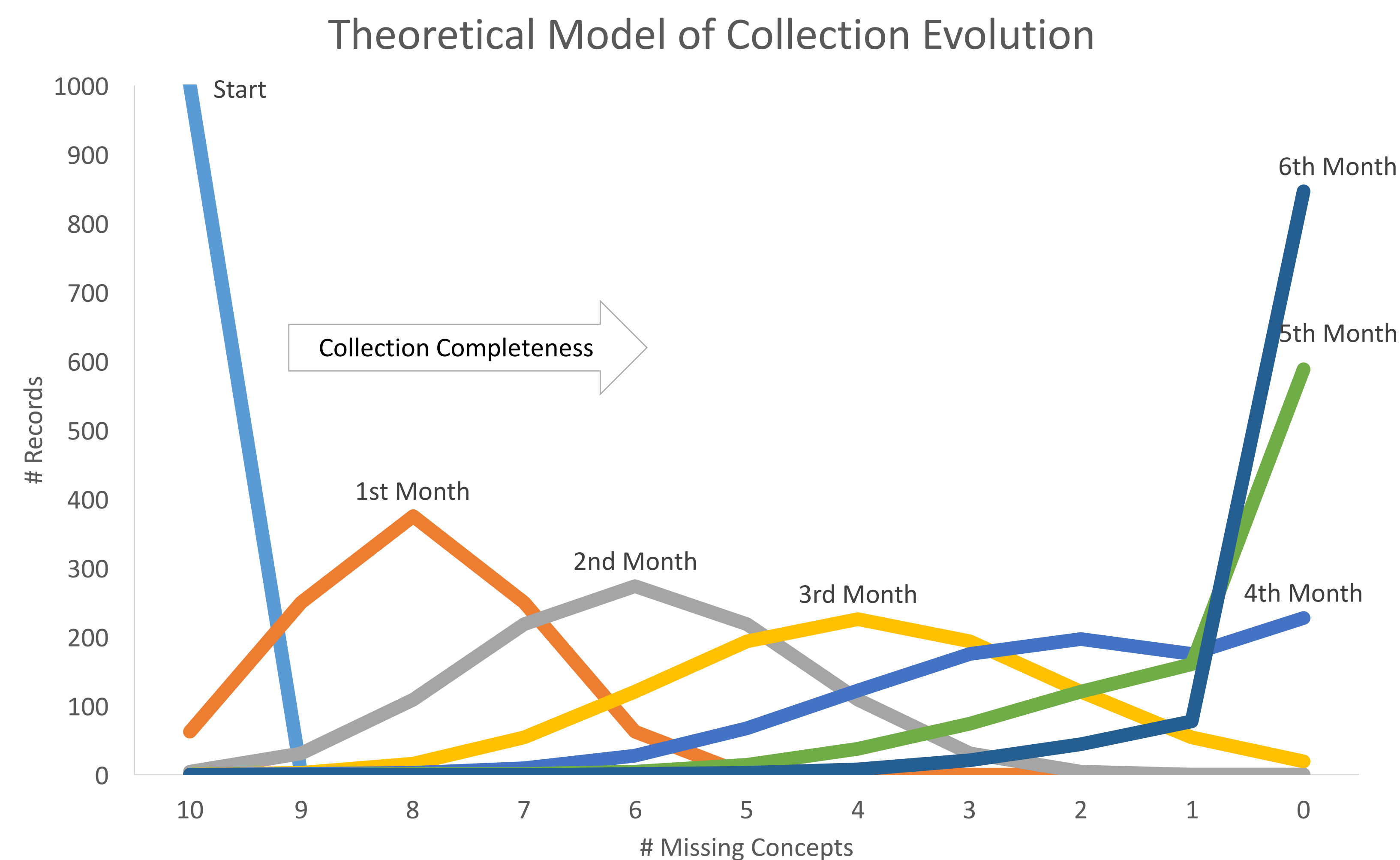
The Long Range Ecological Network uses the Ecological Markup Language metadata dialect for documentation and created their recommendation for use with EML.

There are five documentation use cases in the LTER recommendation: Identification, Discovery, Evaluation, Access, and Integration. As shown below, the dialect and recommendation have no missing concept gap.



Premise

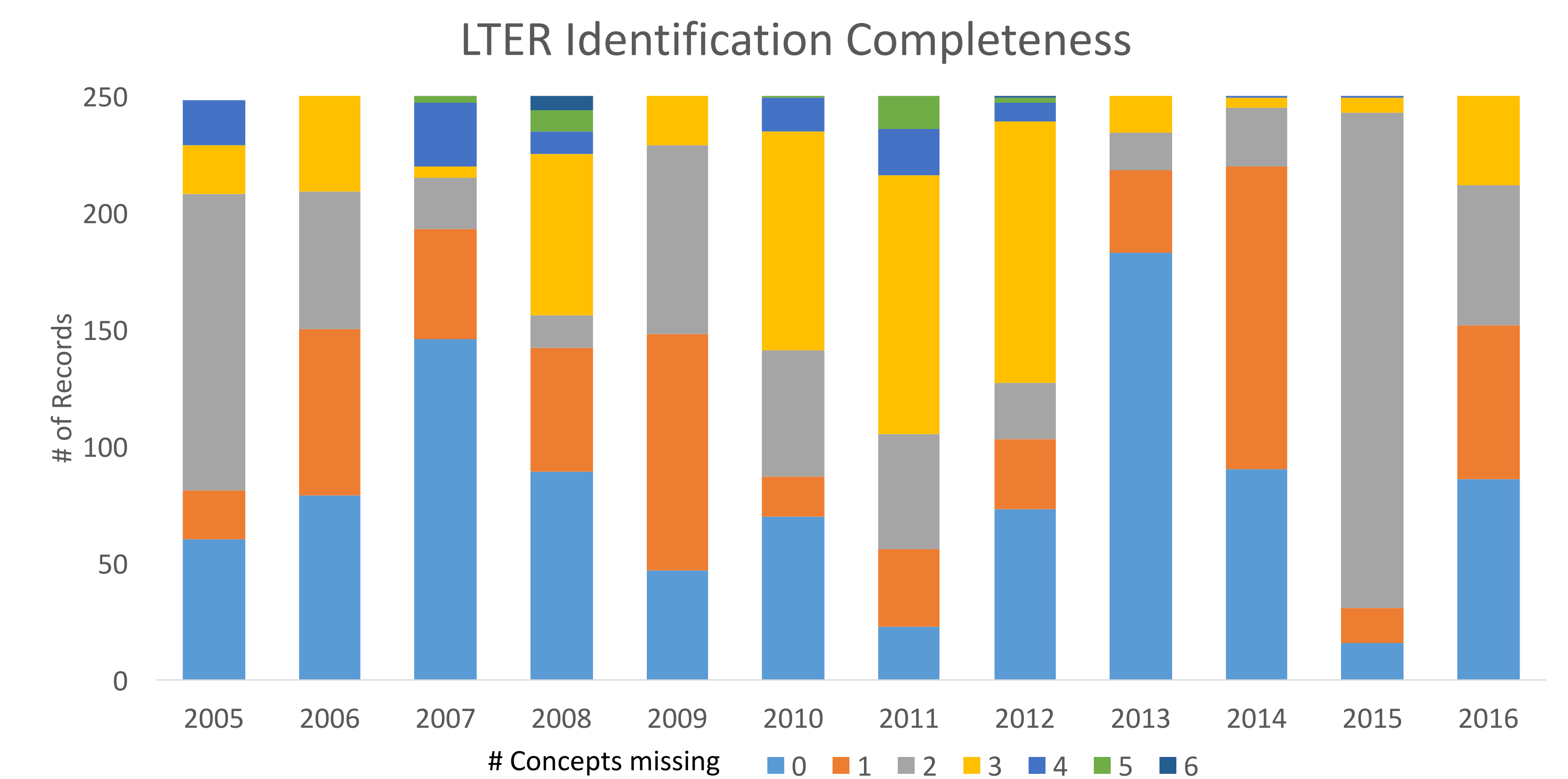
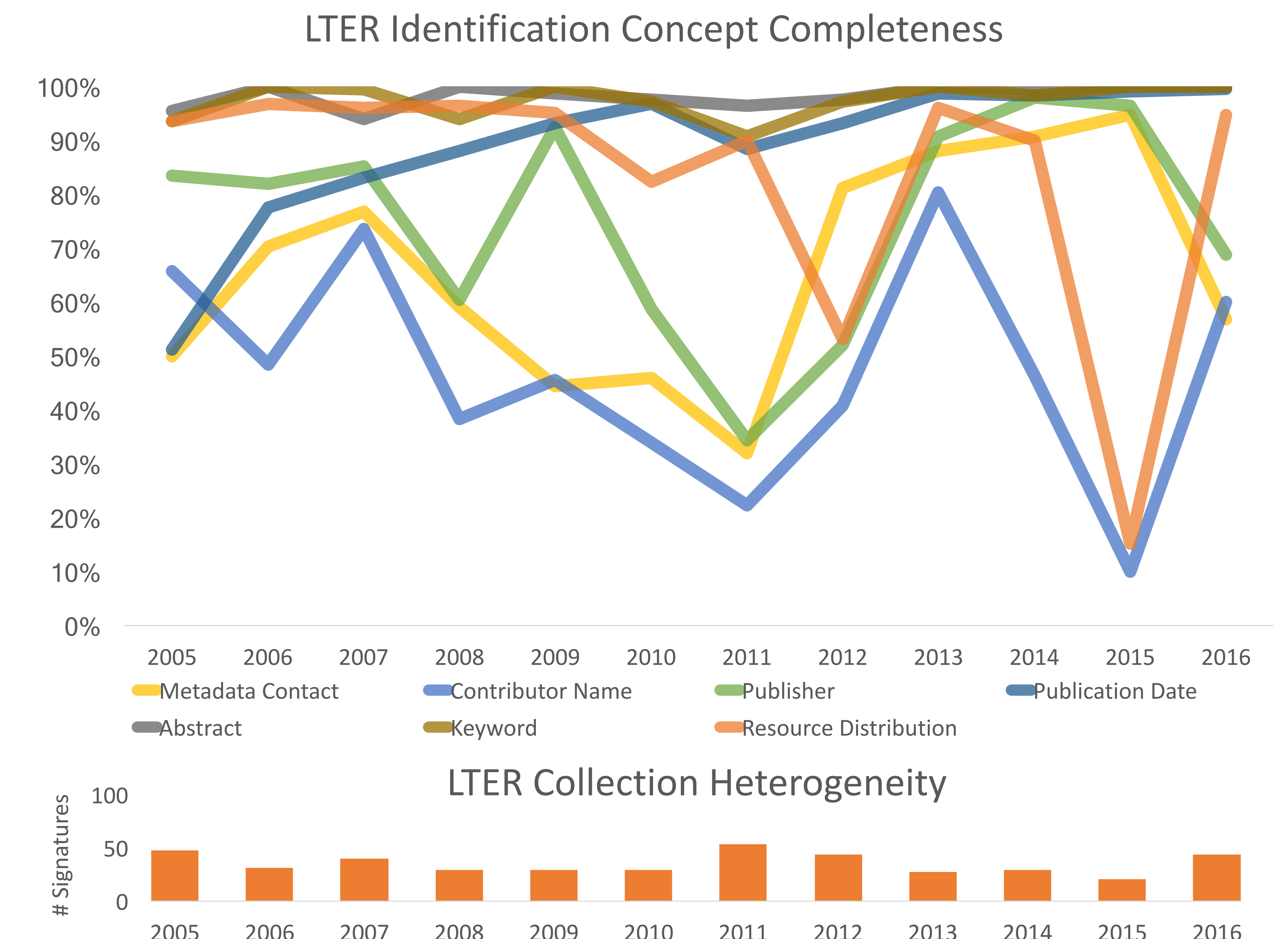
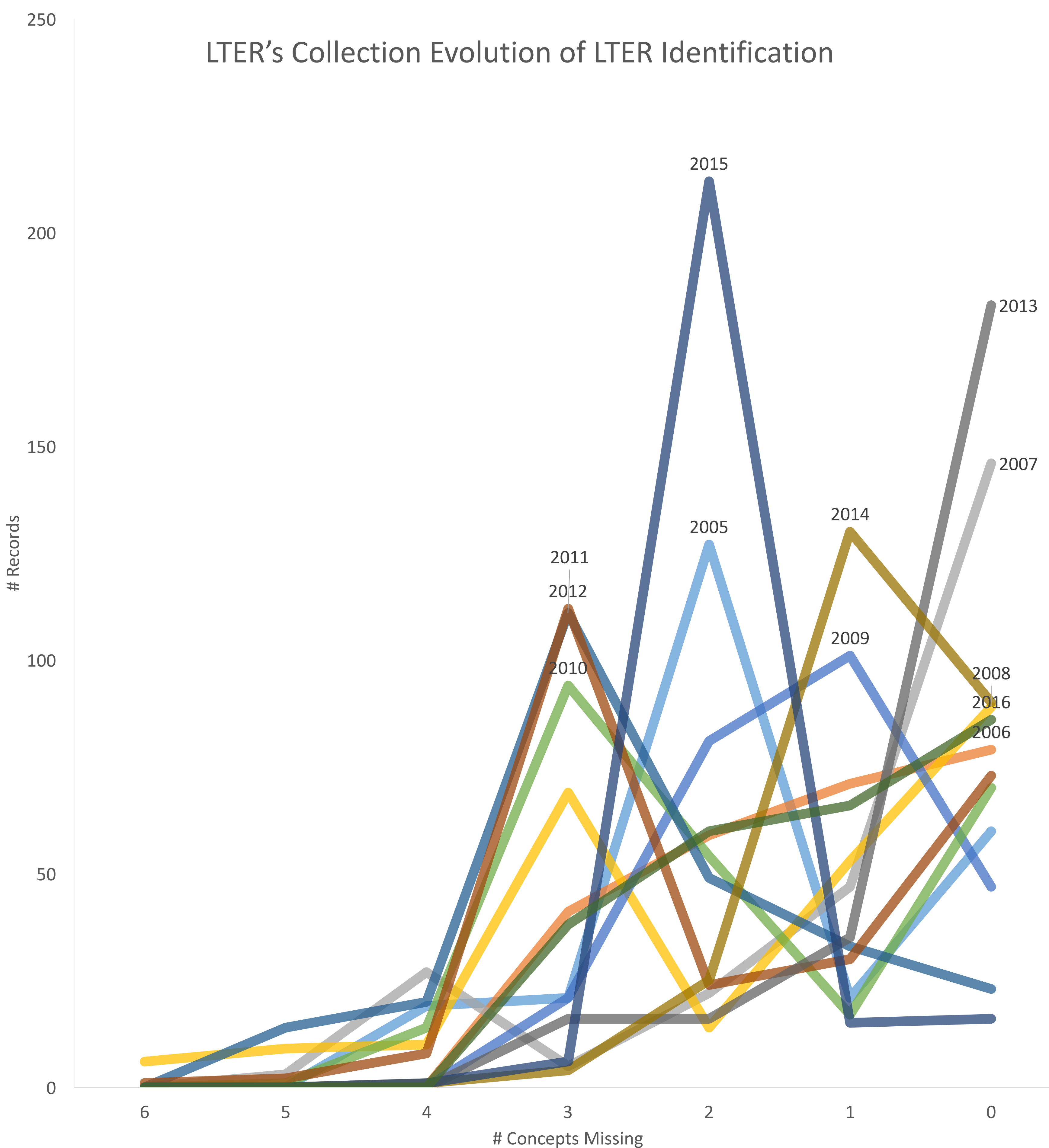
The LTER Completeness Recommendation includes concepts the LTER community considers important for creating high quality EML metadata. Ideally the completeness of LTER metadata should improve over time. The graph below uses a theoretical model to illustrate how metadata completeness is improved over time. This model output uses a metadata concept improvement rate of 50%, and displays every fourth time step to simulate a 6 month period of curation for 1000 records.



Process

- Utilized a python sampling tool that leveraged DataONE's SOLR index to identify and create XML collections of 250 LTER metadata records from each year 2005-2016.
- Used XSL rubrics to determine conceptual content in each record.
- Analyzed results for completeness of 25 concepts in the Recommendations Analysis Dashboard¹ for each years collection.
- Compared analyses across time periods using collection evolution² analysis variations.
- Compared heterogeneity of each collection to completeness using signature score groups.

Results



Observations

- No clear temporal progression towards completeness of a recommendation over entire time period.
- Clear adherence to dialect schema required concepts: Resource Title, Resource Identifier, Author / Originator, Resource Contact.
- Varying degrees of adoption of the other concepts.
- Collection heterogeneity has no clear effect on completeness.

Limitations

- Not a set of records through time.
- Sampling proportion vs sampling size.
- No ethnographic perspective.