Submission Summary

Conference Name

16th International Conference on Biological and Biomedical Ontology

Track Name

ICBO2025 short talks

Paper ID

47

Paper Title

METPO: A Pragmatic Ontology for Microbial Ecophysiological Traits

Abstract

Microbial trait databases (BacDive, BactoTraits, etc.) are most useful when semantically normalized by an appropriate ontology. BioPortal and OLS have several microbial trait ontologies, but most are unmaintained or have ROBOT validation errors; no single ontology provides full coverage of the necessary domains. The best, MicrO, covers only 4 of our benchmark domains (like 'temperature preference') at ≥80% classes per domain, but it has 103 validation errors and has not been updated since 2018. This motivated the creation of the Microbial Ecophysiological Trait and Phenotype Ontology (METPO), which is predominantly a simple 250 class trait hierarchy. To build METPO, we employed batchwise- and targeted LLM workflows along with searches across the entire corpus of OLS class embeddings. Annotations like IAO:0000119 (definition source) and skos:closeMatch are used to acknowledge precedent while maintaining simplicity, stability and a focused scope.

METPO can ground statements in knowledge graphs such as KG-Microbe (KGM), which integrates strain-level trait data in order to predict optimal growth media and conditions for uncultured microbes. For example, METPO enables KGM to express statements like "E. coli K-12 has the mesophilic phenotype" <NCBITaxon:83333 METPO:2000102 METPO:1000615> or "...grows optimally at 37°C" (more complex). METPO also includes properties to express chemical interactions, with KGM using chemical CURIes from ChEBI, so they don't need to be imported into METPO. Reification classes are created for non-binary relations like "reduction of X to Y". METPO and KGM extensions about interactions with lanthanides are in development.

METPO also contributes to and learns from literature mining: e.g., if an abstract states "Microbe X has phenotype Y" but Y lacks representation in METPO, the manual addition of Y to METPO is triggered. Thus, literature is used like database content to drive this KG-optimized ontology.

Created

10/26/2025, 9:53:08 PM

Last Modified

10/26/2025, 9:53:08 PM

Authors

Mark Miller (Lawrence Berkeley National Laboratory) <MAM@lbl.gov>

Marcin Joachimiak (Lawrence Berkeley National Laboratory) <mjoachimiak@lbl.gov>

Nomi Harris (Lawrence Berkeley National Laboratory) <nlharris@lbl.gov>

Primary Subject Area

Scientific Domains (all biomedicine, biological, and life science related domains are welcome)