

Nancy J. Copley, M. Dickson Allison, Cynthia L. Chandler, Robert C. Groman, Adam I. Shepherd, Peter H. Wiebe, and David M. Glover Woods Hole Oceanographic Institution

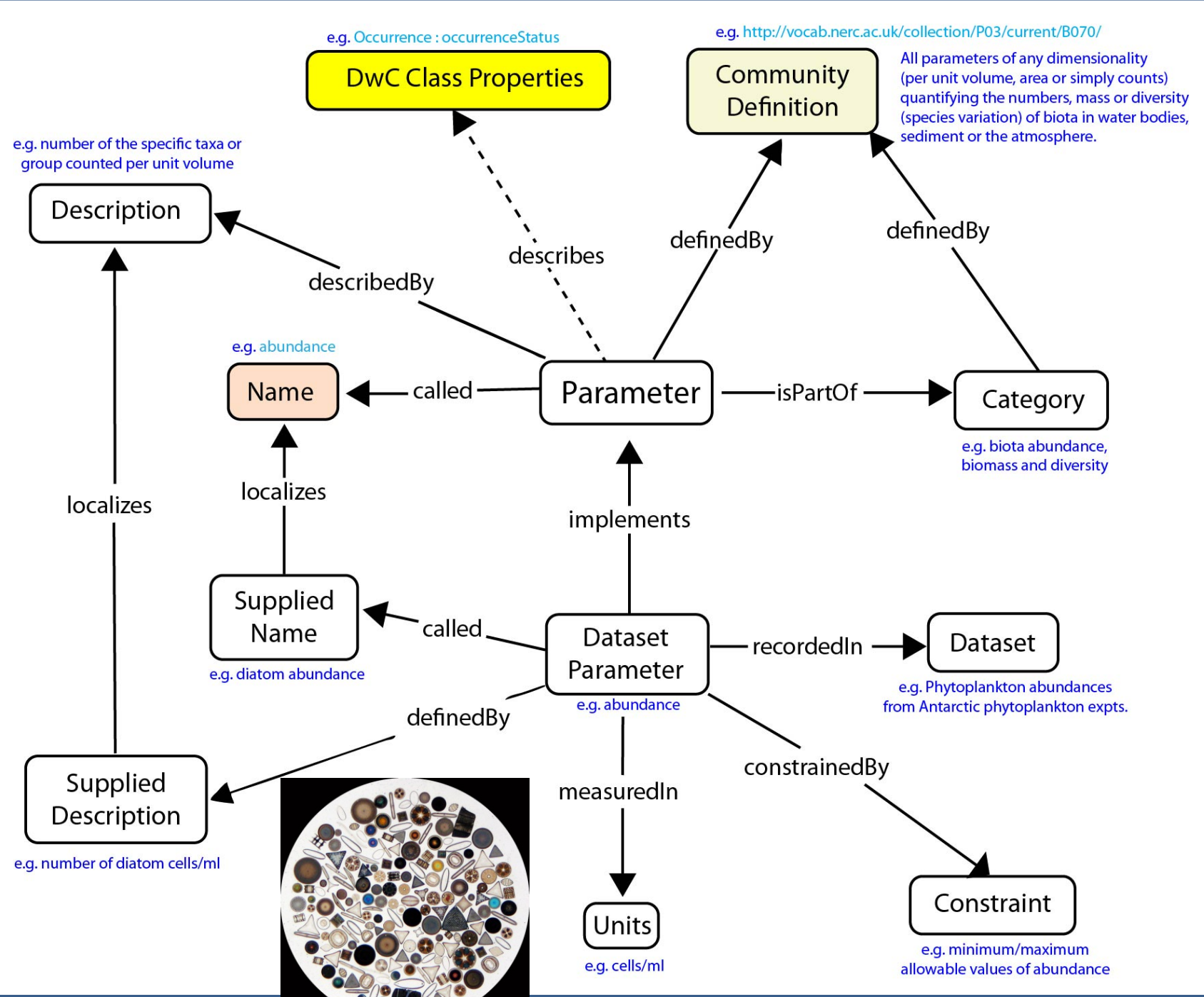
Annual Conference 2013, Florence, Italy 28 Oct – 1 Nov

Abstract

The Biological and Chemical Oceanography Data Management Office (BCO-DMO) is funded to serve the data management requirements of investigators funded by the U.S. National Science Foundation (NSF) Biological and Chemical Oceanography Sections (OCE) and the Division of Polar Programs (PLR) Antarctic Sciences (ANT) Organisms & Ecosystems Program. Some of the 6500 datasets served by BCO-DMO include species and taxon names, abundance, counts and size information. Scientists supply names for the various types of data they are submitting such as ‘temperature’, ‘lat’, ‘species name’, ‘stage’ and ‘count’. These variable names are then matched to our internal standard names (parameters) which are in turn mapped, if possible, to the terms used by the SeaDataNet community, served by the Natural Environment Research Council (NERC) Vocabulary Server (NVS) at the British Oceanographic Data Centre (BODC). BCO-DMO is currently exploring ways to also incorporate or map other controlled vocabularies, including the Darwin Core terms, into the dataset metadata and our evolving ontology. BCO-DMO makes use of the dataset metadata to determine how best to visualize the data in order to facilitate data discovery and support data re-use. This presentation illustrates the vocabulary mapping effort and shows some examples of visualizing occurrence data.

BCO-DMO Parameter Concept Map

This concept map displays the relationships of several aspects of the term ‘parameter’ (which can be thought of as observations and measurements) as it is currently viewed by BCO-DMO. The example of diatom abundance is shown in blue text; concepts in colored boxes are further represented in the table, right. The dotted arrow toward DwC Class Properties indicates the connection has not yet been implemented.



Vocabulary Matching & Mapping

The process of matching and mapping supports the use of: 1) local vocabulary terms, familiar to the originating investigator; 2) intermediate, consistent terms managed by repository custodians (e.g. BCO-DMO); and 3) closest match terms shared by the larger community (e.g. SeaDataNet). Multi-level matching and mapping enables retention of important information while improving interoperability of data systems. Colored cells correspond to the examples in the concept map, left.

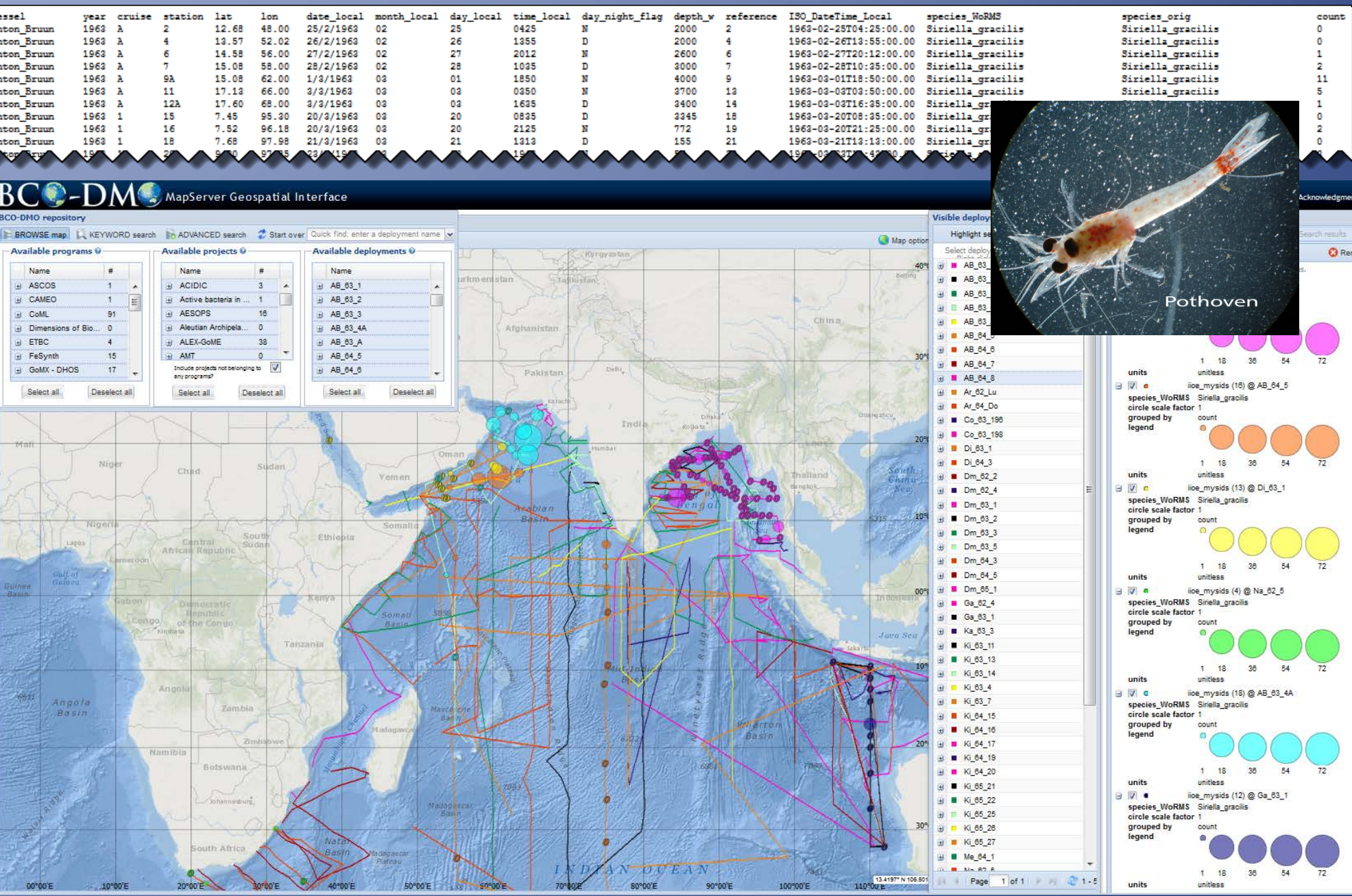
| BCO-DMO | SeaDataNet (NERC) | TDWG | TDWG |
|-------------------------------------|--|---|---------------------------------------|
| Consistent terms managed by BCO-DMO | Closest match shared by the larger Community | Darwin Core Concepts Class | Darwin Core Concepts Class Properties |
| abundance type terms | | | |
| abundance | BiotaAbundBiomassDivers /P03/current/B070/ | dwc:Occurrence http://rs.tdwg.org/dwc/terms /Occurrence | dwc:occurrenceStatus |
| count | BiotaAbundBiomassDivers /P03/current/B070/ | " | dwc:individualCount |
| biomass | BiotaAbundBiomassDivers /P03/current/B070/ | ? | ? |
| species/taxonomic terms | | | |
| taxon_code | none | Taxon | dwc:taxonID |
| species (genus+species) | none | " | dwc:scientificName |
| species_epithet | none | " | dwc:specificEpithet |
| family | none | " | dwc:family |
| phylum | none | " | dwc:phylum |
| stage | Animal development | " | dwc:lifeStage |
| sex | none | " | dwc:sex |
| common_name | none | " | dwc:vernacularName |
| other common terms | | | |
| lat | Latitude /P09/current/LATX | dcterms:Location | dwc:decimalLatitude |
| lon | Longitude /P09/current/LONX | " | dwc:decimalLongitude |
| depth | vertical spatial coordinate /P02/current/AHGT/ | " | dwc:verbatimDepth |
| ISODatetime | Date and time /P02/current/AYMD/ | Event | dwc:eventDate |
| temperature | Water column temperature /P03/current/D025 | none | none |

| | | | |
|-------------------|-------------------------------|---|--|
| BCO-DMO Terms | Abundance | http://osprey.bco-dmo.org/parameter.cfm?flag=view&id=6 | originates with principal investigator |
| NVS Terms | Biota Abundance | http://lod.bco-dmo.org/data/odo/parameter/6.rdf | URI resolves to RDF/XML (dereferencable) |
| Darwin Core Terms | Occurrence | http://vocab.nerc.ac.uk/collection/P03/current/B070/ | URI resolves to RDF/XML (dereferencable) |
| | Occurrence : occurrenceStatus | http://rs.tdwg.org/dwc/terms/ | URI resolves to html page |
| | | http://kos.gbif.org/wiki/Special:ExportRDF?pages=dwc:occurrenceStatus | ? |

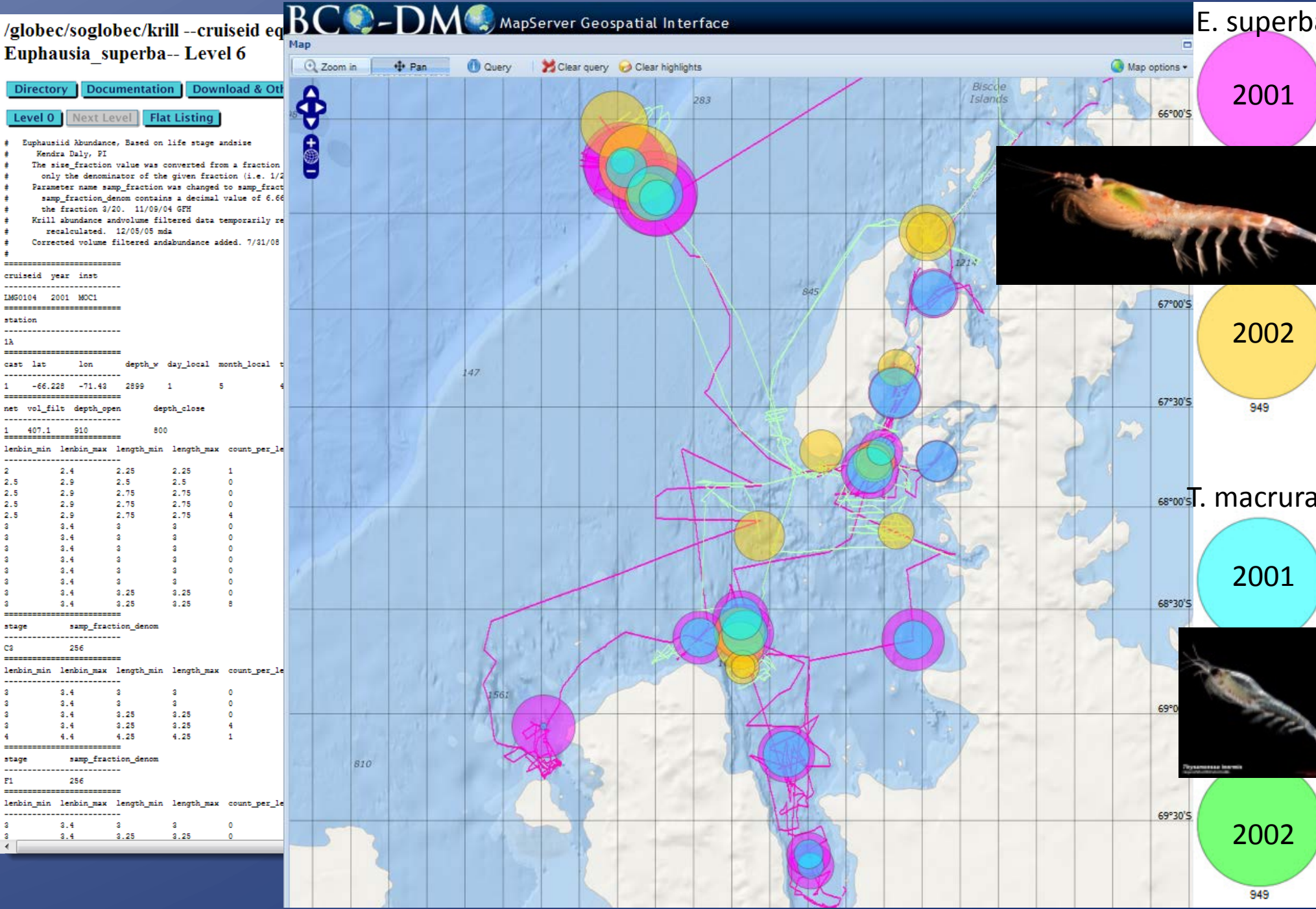


Species Data Visualization at BCO-DMO

Subset of the abundance data for the mysid species *Siriella gracilis* collected in the Indian Ocean in 1962-65. Results from 6 cruises are mapped.



Distribution of Antarctic krill, Euphausia superba and Thysannoessa macrura, winter of 2001 & 2002.



Acknowledgments

The BCO-DMO is funded by the National Science Foundation. We acknowledge the work done by all of the investigators who contribute data to BCO-DMO. The geospatial interface to the BCO-DMO data system was developed in collaboration with Charlton Galvarino (Second Creek Consulting, LLC).

