



DATA ACCESS TUTORIAL 2012 OCB PI SUMMER WORKSHOP

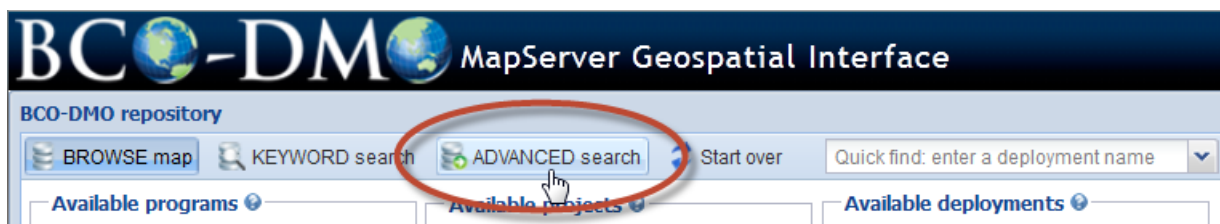
DATA ACCESS: MAP SEMANTIC SEARCH

SCENARIO 5: you have an idea what you are looking for, but you do not know the Program, Project, or Deployment name.

The MapServer has a semantically-enabled search that is useful for finding data in a way that makes sense to you. A beta version of this interface was released in May 2012, and is available as the ADVANCED search option from the MapServer.

The MapServer ADVANCED search is a semantically-enabled, faceted search. It provides access to the same data as the "BROWSE map" and "KEYWORD search" modes, but is intended to help visitors find data of interest without having to know much about how the data are organized by BCO-DMO. The facets (search categories) can be arranged in any order, and you can choose which search facets are active.

Try it out ...



Click on the ADVANCED search button. The map interface will reinitialize (clearing any selections you have made previously; click OK in the pop up dialog box to confirm this and continue). If necessary, slide the vertical panel divider between "Available datasets" and "Mapped datasets" (upper right of the browser window) so those two panels are more equal in width in your browser display. The browser display should look like the screen shot on the next page.

The screenshot displays the BCO-DMO MapServer Geospatial Interface (beta) in a web browser. The interface is divided into several sections:

- Search Panel (Left):** Contains three search categories: Programs, Projects, and People. Each category has a list of items with counts and a search input field.
 - Programs:** Arctic Summer Cloud Ocean Study (2), Census of Marine Life (449), Comparative Analysis of Marine Ecosystem Organization (1), Emerging Topics in Biogeochemical Cycles (10), Gulf of Mexico - Deepwater Horizon Oil Spill (43), Historical datasets (11).
 - Projects:** Algal Communities in Distress: Impacts and Consequences (3), Amazon influence on the Atlantic: CarbOn export from Nitrogen fixation by DiAtom Symbioses (9), Anammox, denitrification and nitrogen fixation in the Black Sea (3), Aplysina Red Band Syndrome: Investigating the etiology, pathogenesis, and ecology of (1).
 - People:** Abbott, Mark (4), Abraham, Edward (12), Ainley, David (6), Albert, Daniel (5), Alexander, Lendell (2), Alldredge, Alice (2), Allen, Andrew (3), Allen, John (3).
- Results Panel (Top Right):** Shows available datasets grouped by deployment. It includes a table with columns for Dataset and Deployment.

Dataset	Deployment
Deployment: AA8704	
EcoMon Plankton 100m^3	AA8704
EcoMon Plankton 10m^2	AA8704
Deployment: AB_63_1	
iioe_chaetognaths	AB_63_1
iioe_copepods	AB_63_1
iioe_decapods	AB_63_1

 Below the table, it indicates "Page 1 of 618" and "1 - 25 of 617".
- Mapped datasets Panel (Top Right):** A section for mapped datasets, currently showing "No datasets have been mapped. Click the plus icon next to a dataset to begin." with a "Remove all" button.
- Visible deployments Panel (Bottom Left):** A list of visible deployments with color-coded squares next to them: AB_63_1, AB_63_2, AB_63_3, AB_63_4A, AB_63_A, AB_64_5, AB_64_6, AB_64_7, AB_64_8, ACIDIC-CMEN..., ACIDIC-FCKX..., ACIDIC-SHLX..., and AE-X0908.
- Map Panel (Bottom Right):** A map of the Arctic region showing various data points and lines. It includes a zoom in button, a pan button, and a query button.

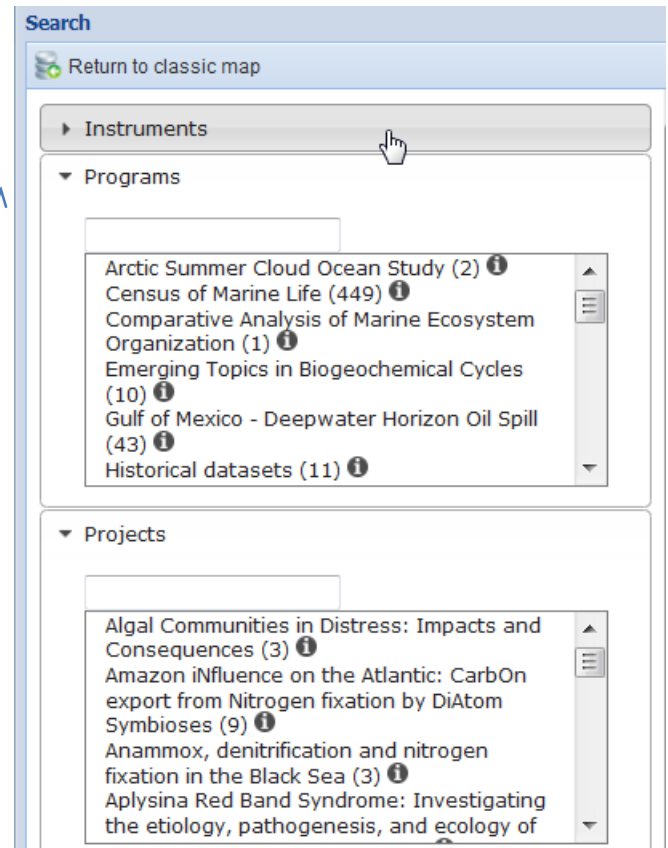
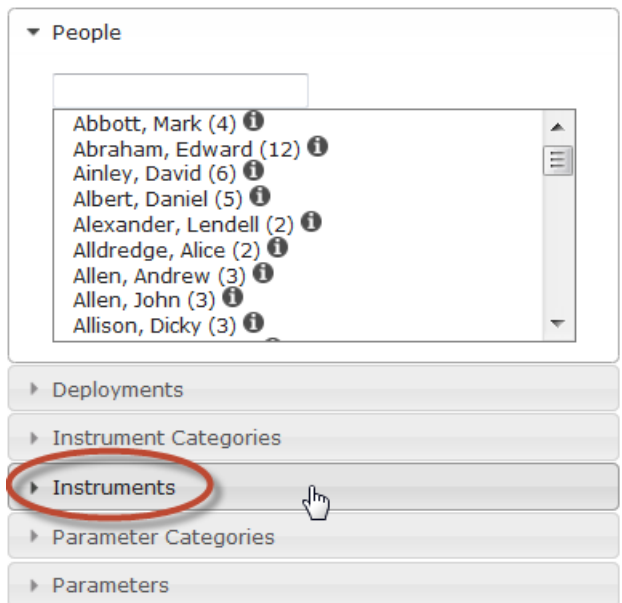
The elements of the ADVANCED (semantic) search interface are similar to the MapServer layout you have seen in the previous scenarios in this tutorial, but those elements have been rearranged. The search panel is now on the left, the map on the lower right, and the dataset panel is above right.

The Search panel initializes with the Programs, Projects and People search boxes open and active.

For this exercise, let's assume you know the type of instrument that collects the data you want to find. For example, you have read a paper about some data collected with a MOCNESS system, and the paper says the data are available from BCO-DMO.

In the search panel on the left, scroll down so you can see the "Instruments" bar.

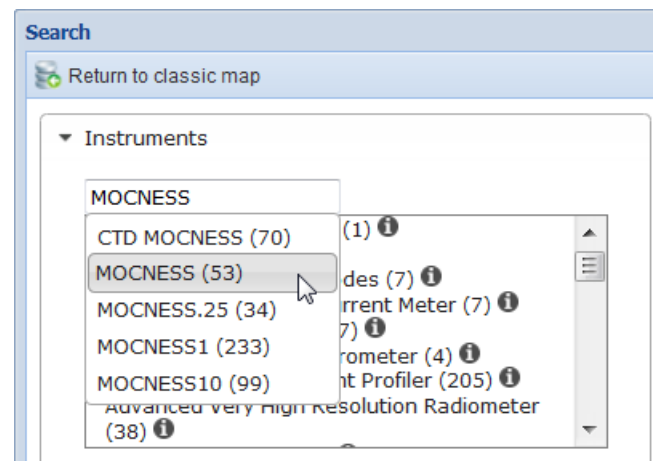
Click anywhere in the Instruments bar and drag it up to the top of the Search panel and drop it above the “Programs” box.



Now, click the triangle at the left end of the Instruments bar to open that search box.

Enter an instrument name to search for; for example, in the text entry box under the label “Instruments”, type “MOCNESS”. Select “MOCNESS” from the list of matches that pop up.

Observe how the lists change in the open search boxes (Programs, Projects, People) below Instruments and the list of “Available datasets” and “Visible deployments” (the panels on the right of the browser window) also get updated to reflect that you are only interested in data from a MOCNESS tow.



At this point you could further narrow the search by selecting a deployment on the map, or by selecting a dataset that looks interesting, but let’s try choosing a deployment (cruise ID) from the “Deployments” search box. In the Search panel on the left, scroll down and slide the Deployments bar up above “Programs” (below “Instruments”). Open the “Deployments” search box (click on the triangle at the left end of the Deployments bar). Choose EN307 from the list (second in the list).

Observe how the panels on the right update to display only results from deployment EN307.

If you scroll down in the Search panel, you can learn that the EN307 MOCNESS data were reported by Peter Wiebe as part of the US GLOBEC Georges Bank project. Your browser window should look similar to the screen shot below.

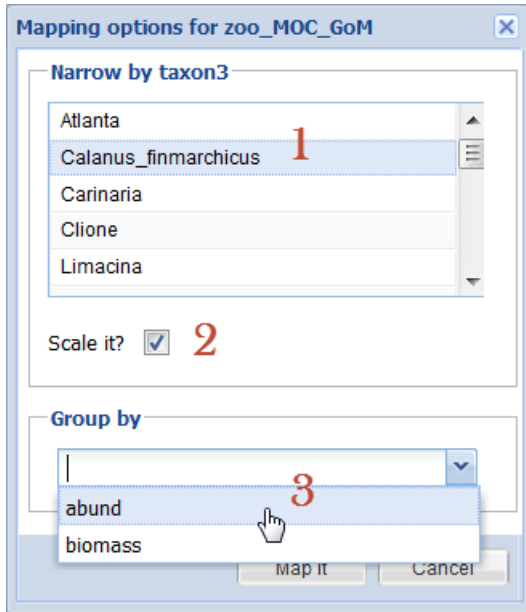
The screenshot displays the BCO-DMO MapServer Geospatial Interface (beta). The interface is divided into several panels:

- Search Panel:** Contains a "Return to classic map" link and a list of search results categorized by Programs, Projects, and People. The "Programs" section shows "U.S. GLOBal ocean ECosystems dynamics (1)". The "Projects" section shows "U.S. GLOBEC Georges Bank (1)". The "People" section shows "Wiebe, Peter (1)".
- Results Panel:** Displays "Available datasets" grouped by deployment. The "Deployment: EN307" section shows a single dataset: "zoo_MOC_GoM".
- Mapped datasets Panel:** Shows a message: "No datasets have been mapped. Click the plus icon next to a dataset to begin."
- Visible deployments Panel:** Shows a list of visible deployments, including "EN307".
- Map Panel:** Displays a map of the Georges Bank area. The map includes a coordinate grid and labels for "Boston", "Georges Bank", and "Bay of". The coordinates shown are 44°00'N, 42°00'N, 72°00'W, 70°00'W, 68°00'W, and 66°00'W. A specific location is marked with coordinates 42.6838° N 72.2103° W.

This close-up view of the Results panel shows the "Available datasets" section. The "Group by:" dropdown is set to "deployment". The "Deployment: EN307" section displays a table with two columns: "Dataset" and "Deployment". The table contains one entry: "zoo_MOC_GoM" under the "Dataset" column and "EN307" under the "Deployment" column. A green plus icon is visible next to the "zoo_MOC_GoM" dataset name.

To map a dataset, select (click on) the "zoo_MOC_GoM" dataset from the "Available datasets" panel in the upper right.

A dialog box pops up with "Mapping options" for this dataset. This particular type of dataset has some custom mapping options.

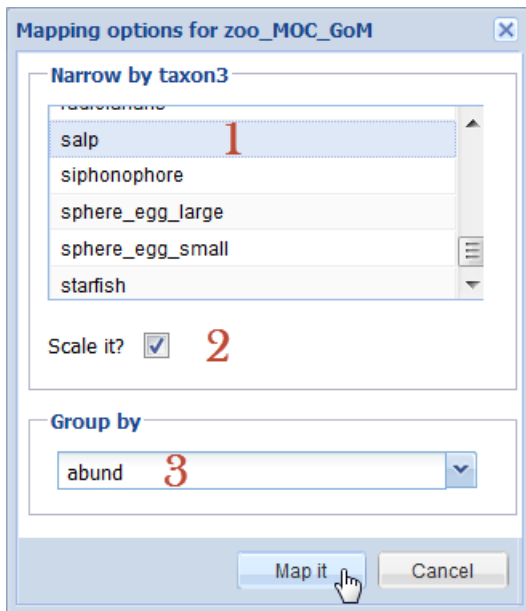


Select some options in the dialog box:

1. Select something from the list, for example: *Calanus finmarchicus*
2. Select the "Scale it" option to display the results using symbol sizes scaled to indicate higher or lower data values.
3. Finally, select "abund" (abundance) as the measurement by which to group the data.

Click the "Map it" button to see the results.

The relative size of the circles indicates differences in the abundance of *Calanus finmarchicus* at the MOCNESS sampling locations on this cruise.



One can display multiple taxa from a MOCNESS dataset. Click on the zoo_MOC_GoM dataset once again (up in the "Available datasets" panel). In the pop up dialog box, scroll to the bottom of the taxon list and select (1) salp, (2) "Scale it" and (3) abund.

Click the "Map it" button to put the scaled symbols on the map indicating abundance of salps from the MOCNESS tow data.

As with the other MapServer search modes, clicking or right-clicking on different elements in the display pops up dialog boxes that display either information about the item or additional options for that item (e.g. access to that selection in the database, or the option to do 'quick view' plots).

The final browser window looks something like this if you have clicked on the northern most sampling location, and clicked on the EN307 cruise ID in the “Visible deployments” list.

The screenshot displays the BCO-DMO MapServer Geospatial Interface (beta) in a Firefox browser window. The interface is divided into several sections:

- Search:** Includes a search bar and a list of results. The results list includes EN307 (1), EN330 (1), EN331 (1), KIWI6 (2), KIWI8 (2), KM0414 (1), LMG0104 (1), and LMG0203 (1). Below this is a section for Programs, showing "U.S. GLOBal ocean ECosystems dynamics (1)".
- Results:** Contains a table of available datasets. The table has columns for Dataset and Deployment. The dataset "zoo_MOC_GoM" is listed under the deployment "EN307".
- Mapped datasets:** Shows a list of mapped datasets, including "zoo_MOC_GoM (4) @ EN..." and "zoo_MOC_GoM (4) @ EN...".
- Visible deployments:** A list of visible deployments, with "EN307" selected. A red arrow points to the text "click the cruise ID to view cruise information.".
- Map:** A map showing the study area, including the Bay of Fundy and Georges Bank. A yellow line indicates the cruise track. A tooltip for EN307 provides details about the deployment, including the platform name (R/V Endeavor), start date (1997-10-08 00:00:00), end date (1997-10-17 00:00:00), location (Gulf of Maine and Georges Bank), deployment report URL, and chief scientist (Charles Greene).
- Pop-up window:** A window titled "zoo_MOC_GoM @ EN307" showing the location (43.5157° N 67.8867° W) and available on-line data at this point. It includes a link to the full dataset, the total count (0.7007), and the cast number (4). The X-axis is labeled "abund" and the Y-axis is labeled "abund". A link "view and get data" is also present.