

# Data Discovery Studio

## Metadata Editor

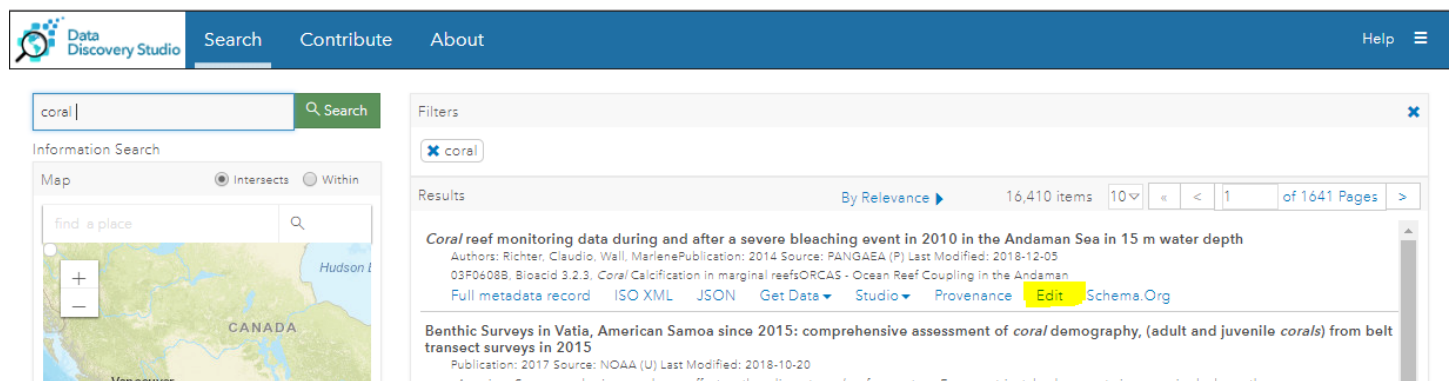
### Usage & Guidelines

Gary Hudman  
Dec 7, 2018

The Data Discovery Studio ingests metadata records from a variety of earth science repositories and enhances these records for improved discoverability with use of community specific ontologies, semantic annotation, and faceted searches. The metadata editor provides curators and stakeholders with additional ability to refine and improve search results of these registered datasets.

### Getting Started

To access the metadata editor from the Search page, click on the Edit link for a search result:



When opening the editor these are some of the key elements to get familiar with:

**Sign In** – Users must log in to be able to save their edits to the DDStudio database and search index. The user can perform all operations locally but cannot effect changes on the server until logging in.

**Curator** – Allows editors to optionally tag their work by name, in addition to the login.

**Validate** – Allows editors to identify and save records that have been reviewed and verified as appropriate.

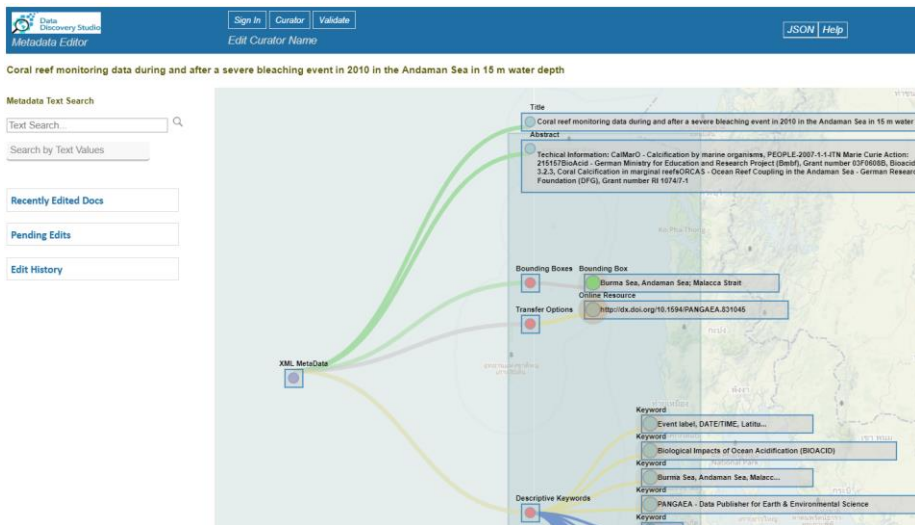
**Graphic Visual Editor** – in the right content panel, allows users to quickly identify key areas of the metadata and modify. The diagrammatic tree represents a subset of the structure of the metadata record.

**Meta Text Search** – Allows users to identify and modify any occurrence of specific text within the record that can be impacting its search rank.

**Recently Edited Docs** – keeps a local history of the metadata records edited by the user. These are links that allow users to retrieve these records.

**Pending Edits** - the list of edits performed by the user that have completed but not yet saved.

**Edit History** – The list of changes performed on a specific record.



Note: All saves within the page only save data locally, to save changes to the server, you must login and press **Publish**.

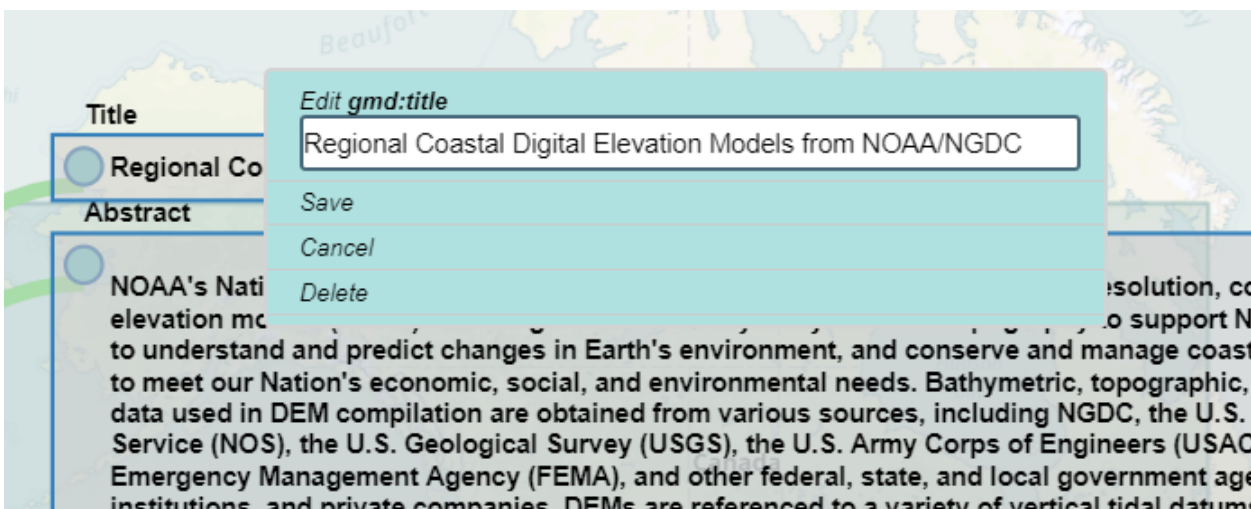
## Visual Graphic Tree Editor Basics

The graphic representation of the metadata provides the user with a quick view of the most important elements of the record (relevant to the search criteria). These important elements include:

- Title
- Abstract
- Extents
- Resource Links
- Keywords

Other elements can be represented within the graphic tree, but the diagram can be harder to view and understand if too many elements are displayed at once. This displays only elements considered the most relevant at the current time. However, the text search tool allows you to quickly access all elements in the metadata.

Each element in the diagram can be edited according to its data type and the constraints defined by the relevant metadata standard. For instance, right click on the title, a text editor appears:



Make the desired change and click save, to save to the local environment.

## Validation of Metadata with Visual Elements

The graphic tree is designed to allow users to quickly assess the quality of the metadata records. Some of these features include:

**Link Colors** - The color of each the links in the diagram identifies the elements status. This information can provide a quick view of the records status. The link colors represent:

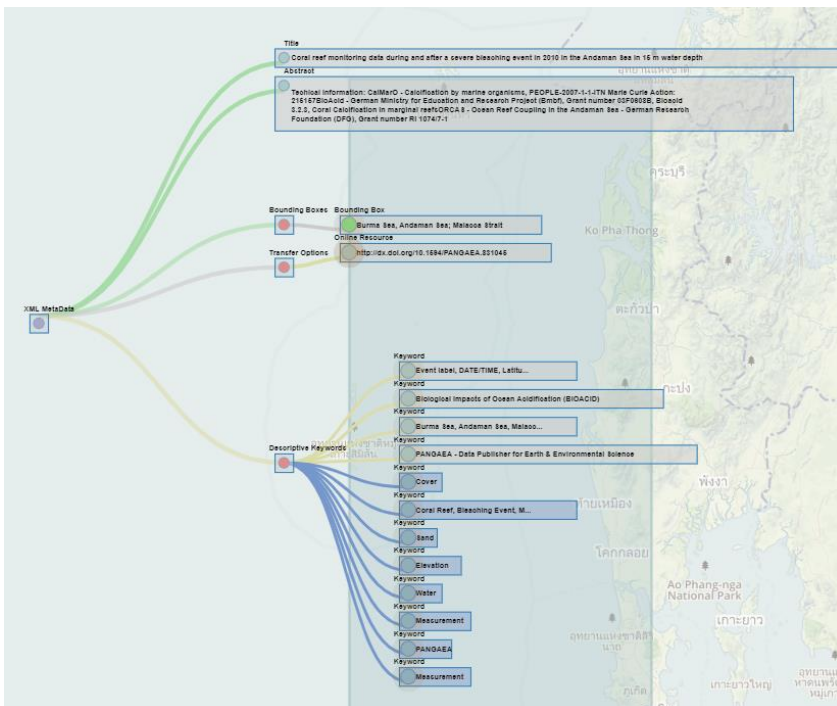
Green – required element is present and valid

Red – required element is missing or invalid

Yellow – non-required element is present

Grey – non-required element is missing or empty

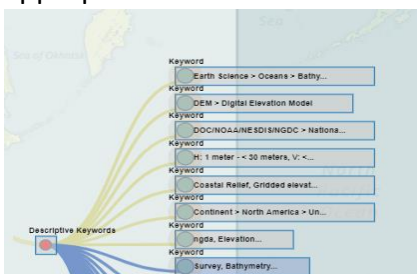
Blue – A DDStudio semantically enhanced element is present.



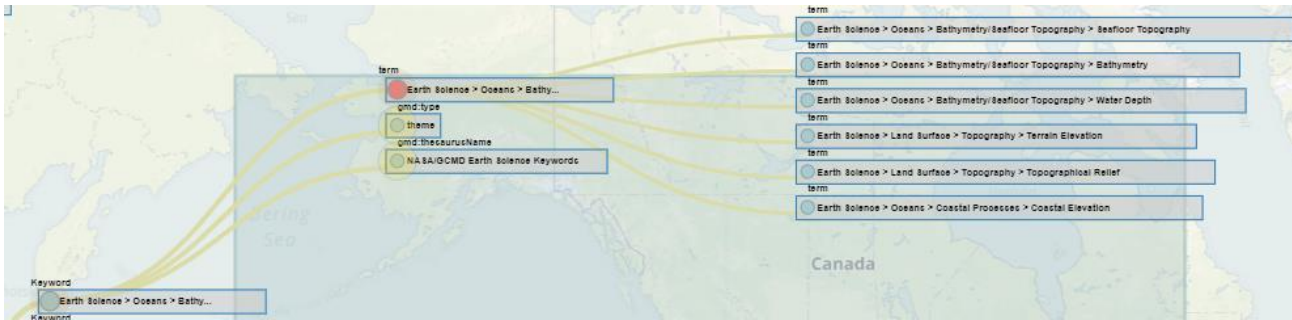
**Bounding Box Status** - If the records extent is valid, it will appear as a background image on the page. If a valid extent is not found the background will appear empty.

## Navigating the Tree

The editor tree can be expanded or collapsed as needed by clicking on any of the Rectangle-Dots. If there are sub elements, the tree will be expanded. Click again and the tree will collapse. If you are at the endpoint, the appropriate editor element will appear.



For example the first keyword element is expanded, which includes a nested sub list of keyword terms ..



Terms use a typeahead lookup from the DDStudio vocabulary. Using the standardize terms is highly recommended to ensure consistent search terminology.

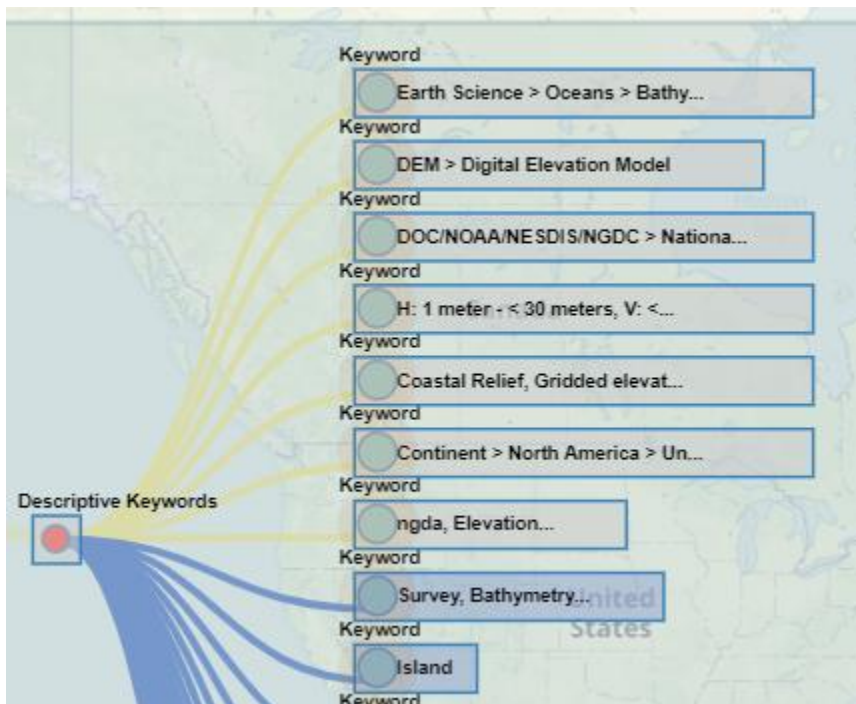
The keyword type is current a selectable list of standard types:

- DataCenter
- Discipline
- Instrument
- Place
- Stratum
- Theme
- Temporal

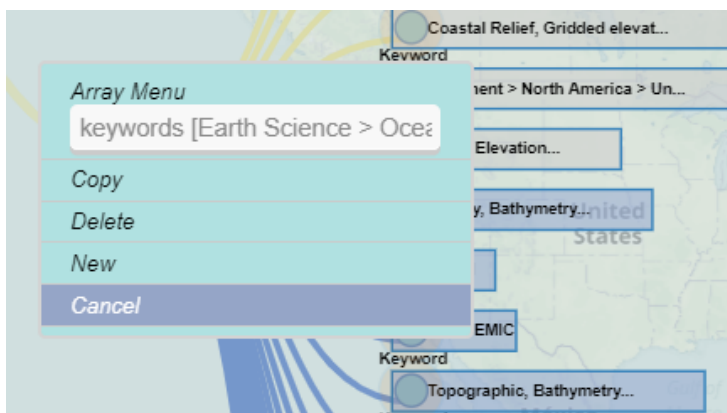
The thesaurus name is currently just a text entry, but future improvements will include a selectable list of domain relevant vocabulary sources.

## Add or Removing Data Elements

In metadata, there are specific type of elements that can have multiple occurrence. For instance, keywords, bounding boxes, and distribution links can show up 0, 1, or many times in a record. In the editor, these type of elements are represented as a Red-Dot-Rectangle.



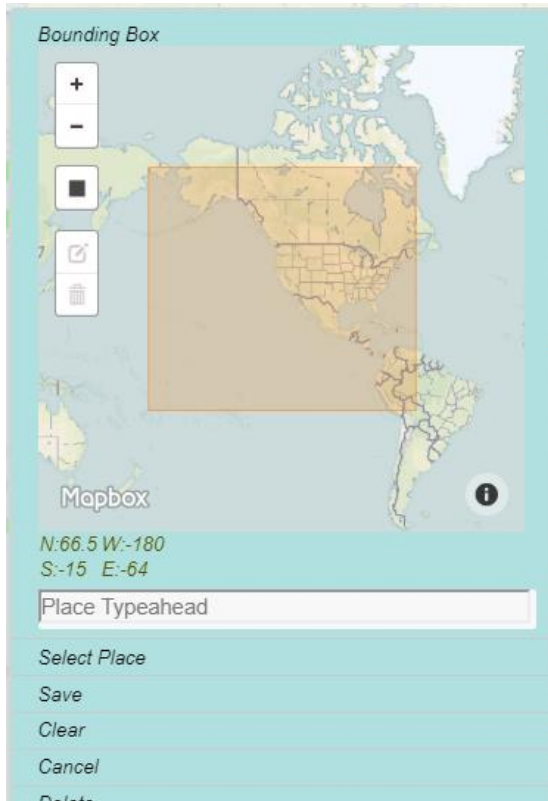
To make additions, right click on the Red-Dot and select New or Copy to add an additional item. The drop-down box allows you to select the pattern of the new element based on the selected existing one (since these sub elements can often have different data structures). The **Copy** command will make a copy keeping the values of the source record. The **New** command copies the structure but leaves the values empty.



You will see the new element branch appear, which you can then edit as desired.

## Bounding Box Edits

Bounding boxes have a widget that allows users to enter place names or map tools to set the extent for a metadata record. It is possible to have more than one bounding box for a record. The Red-Dot-Rectangle menu (right-click) can be used to copy or create a new one. To edit a bounding box - click on the bounding box element to bring up the map widget.



Map – standard pan and zoom functions

Rectangle Tool – select and drag a rectangle to set a new extent.

Place Name text – provides a text lookup for standard places from the ESRI gazetteer, then **Select place** will draw the place as an extent in the map.

Save – will save the new extent info locally

Clear – will remove existing extents from the map.

Cancel – cancel edits

Delete – delete this extent.

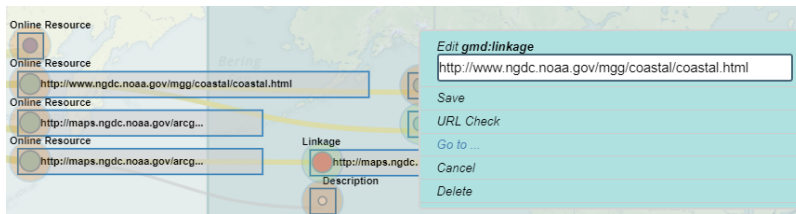
## Managing Resource Links

Resource links are some of the most common changes, and there are additional features for managing and validating links. These are identified in the Transfer Options Branch.



You can add or delete links by right clicking on the Red-Dot-Rectangle and selecting the appropriate option.

The data structure of the element can vary depending on the standard and user input. Click to expand down to the linkage element (the URL) and right click:



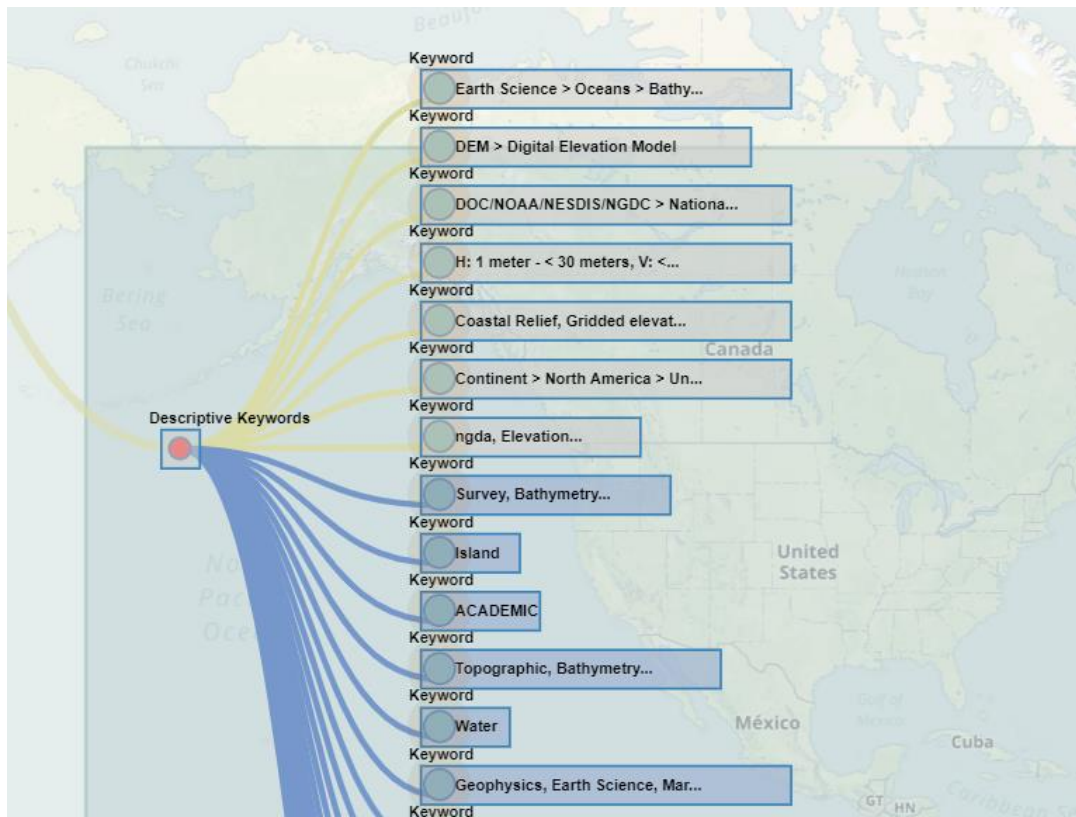
In the menu, you may

- Make text changes to the URL and Click Save
- Click URL to check if the link is working or not (turns element Red or Green)
- Go To .. will send you to the resource page in another tab.

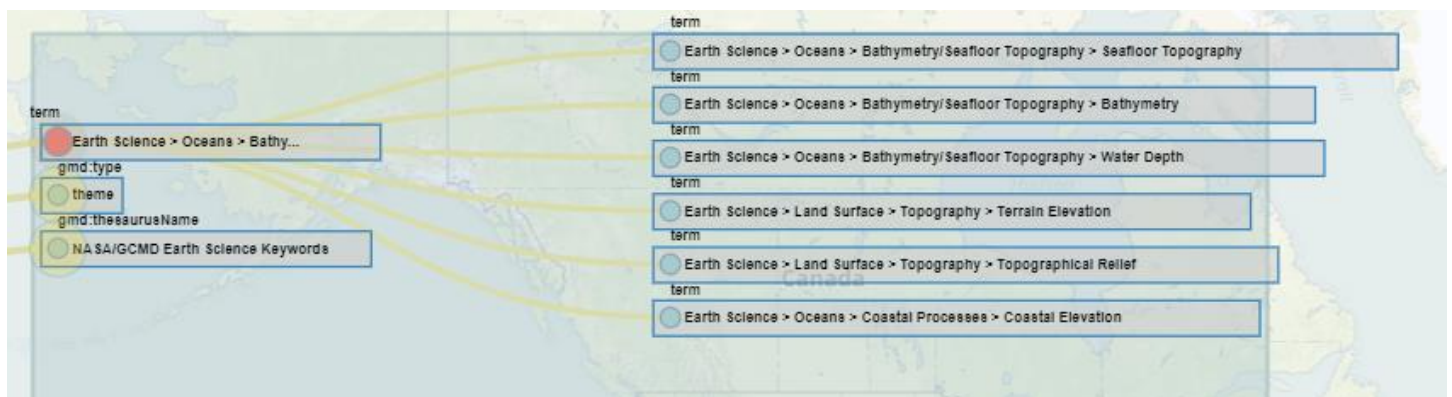


## Managing Keywords

Keywords are perhaps the most important and complex set of elements in determining search relevance. There are two main categories of keywords, the keywords included in the original source document, and keywords generated by the DDStudio enhancement algorithms. The enhanced elements are identified as elements and links in blue:



Keywords sets or groupings are characterized by type (gmd:type), and dictionary (thesaurus) that defines the vocabulary. Keyword sets are often grouped by the these identifying characteristics:

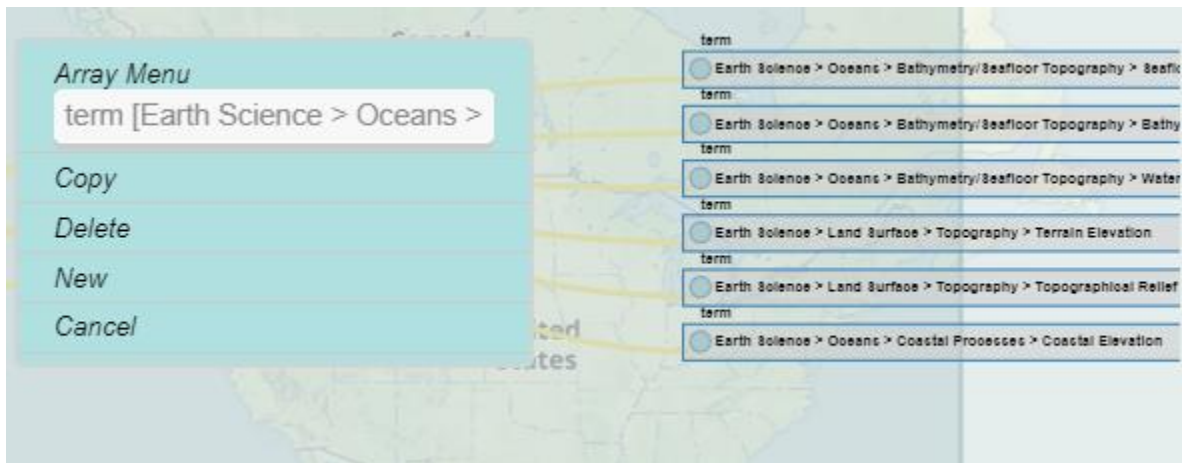


So in modifying keywords, it is important to keep in mind the type and thesaurus Name of the keyword set you are editing. If there isn't an existing branch that has the characteristics you need, you may want to start a new keyword branch and establish the correct type and dictionary for your new keywords.

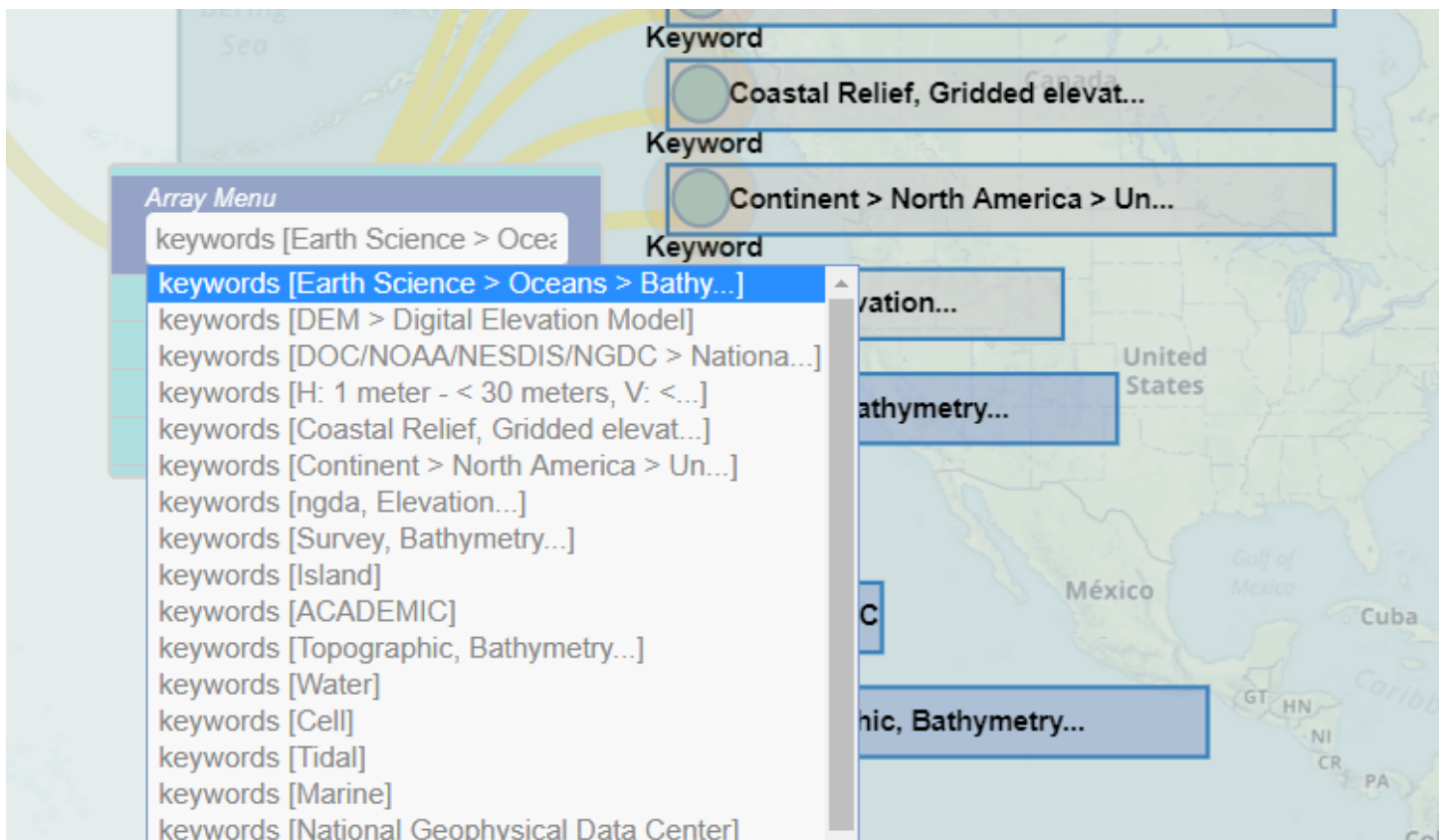
1 - To edit an existing keyword – just click, edit and save.



2 - Adding a keyword to an existing set – right click on the Red-Dot-Rectangle in the set, and in the menu add appropriately:

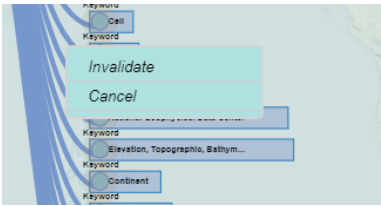


3 - Adding a new top level keyword set - right click on the **Descriptive Keywords** Red-Dot-Rectangle and make a new top level branch using an existing as a template:



### Special Instructions for DDStudio Enhanced Keyword Sets

Since the DDStudio generated keywords are algorithmically generated, we recommend that direct edits not be performed on these, and if a keyword is inappropriate for a search simply mark as Invalidate. This allow the change to be processed within the automated pipeline for future improvement.



### Using the Text Search Tool

Occasionally it occurs that in editing metadata records, that the text affecting the search, is somewhere embedded in the record and not appearing in the graphic tree. Since the metadata record tree can be quite large and complex, including every element in the tree can make finding the key elements more difficult than necessary. To quickly find these hidden elements enter the text and click search. Using the drop-down box below the search box, these are searches can optionally be performed by value, by the metadata element name, or can optional return a tree of all elements:

#### Regional Coastal Digital Elevation Models from NOAA/NGDC

Metadata Text Search

Search by Text Values

Recently Edited Docs

Pending Edits

Edit History

Search Results for water

Replace All Close

Metadata

identificationInfo

DataIdentification

abstract

descriptiveKeywords

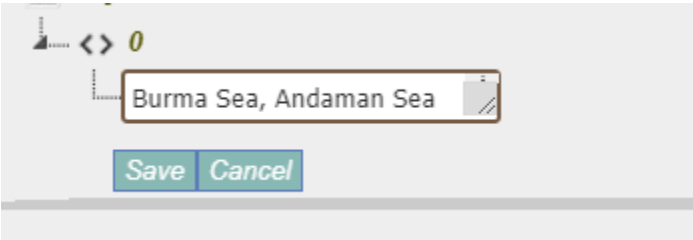
Keywords

keyword

NOAA's National Geophysical Data Center (NGDC) builds and distributes high-resolution, coastal digital elevation models (DEMs) that integrate ocean bathymetry and land topography to support NOAA's mission to understand and predict changes in Earth's environment, and conserve and manage coastal and marine resources to meet our Nation's economic, social, and environmental needs. Bathymetric, topographic, and shoreline data used in DEM compilation are obtained from various sources, including NGDC, the U.S. National Ocean Service (NOS), the U.S. Geological Survey (USGS), the U.S. Army Corps of Engineers (USACE), the Federal Emergency Management Agency (FEMA), and other federal, state, and local government agencies, academic institutions, and private companies. DEMs are referenced to a variety of vertical tidal datums, generally Mean High Water, and horizontal datum of World Geodetic System 1984 (WGS84). Cell size is based on user requirements and generally range from 1 arc-second (~30 meters) to 24 arc-seconds.

Earth Science > Oceans > Bathymetry/Seafloor Topography > Water Depth

These search items have a similar edit feature, essentially allowing you to modify any metadata element. Click on the element, change and save (locally) as needed.



## Saving Your Changes to the Server

Once you have complete the changes, you can view the list of pending edits in the Right Nav Bar. Additionally, you may look at the json representation of the data by clicking on the JSON button in the top right corner. If the changes appear correct, to save to the server you must:

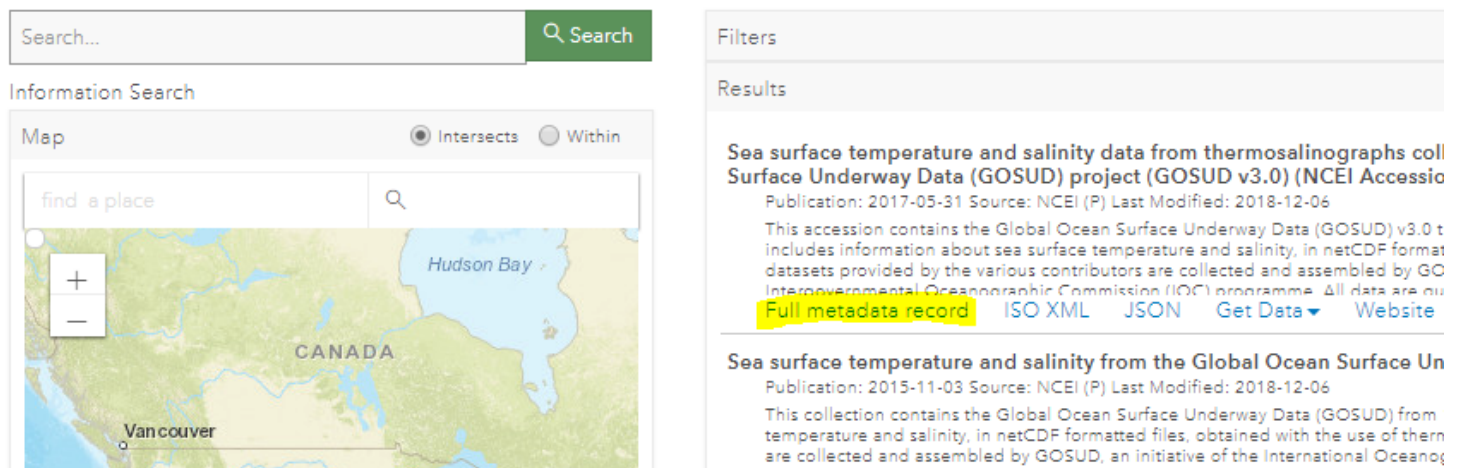
- 1 – Login with an authoritative login.
- 2 – Enter your name as the curator.



When successfully logged in you will see the Publish button in the upper right. Click publish to save to the Server. This will accomplish the following:

- 1 – Update the DDStudio database, recording the changes with modification and new version info.
- 2 – Update the geoportal index (the data that is used in the actual searches).

Once you have completed a change, you can re-execute the search to verify the changes you have made are in fact modifying the search as desired. Within the DDStudio, you can verify that your changes appear correctly by clicking on the **Full metatadata record** in the search results pane:



The process of correcting records to match the desired search results may take several iterations, as the search engine within geoportal performs a complex ranking system to determine which records are most relevant.

Use Case #1

TBD