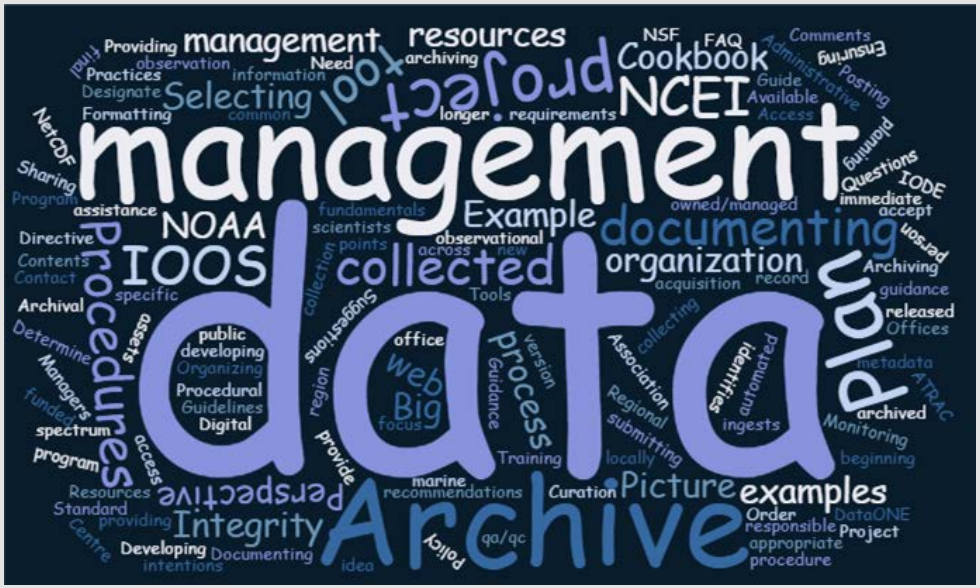
NCEI-IOOS Regional Archiving



http://www.wordclouds.com/

Mathew Biddle

Mathew.Biddle@noaa.gov

NCEI.IOOS@noaa.gov

June 2nd, 2016

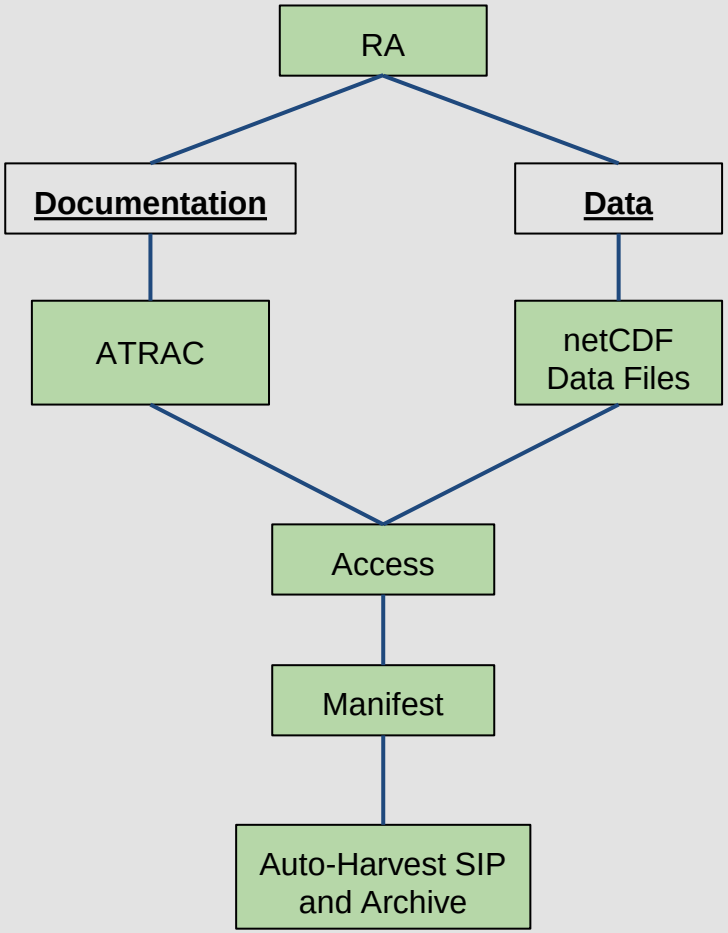
IOOS DMAC Meeting, Silver Spring, MD

Outline

- Implementation process
- Regional archiving status
- Tools
  - NCEI Archiving Cookbook walkthrough
  - ATRAC
  - Mapping SOS to NCEI netCDF Template
- NCEI netCDF template changes (v1.1 to v2.0)
- NCEI IOOS Archive Data Portal

Implementation Process

1. Select non-Federal assets to be archived.
2. Format the data to conform to the NCEI netCDF templates.
3. Provide direct access to data files.
4. Provide checksums for data files to be archived.
5. Review and approve the ATRAC documentation.



SIP - Submission Information Package

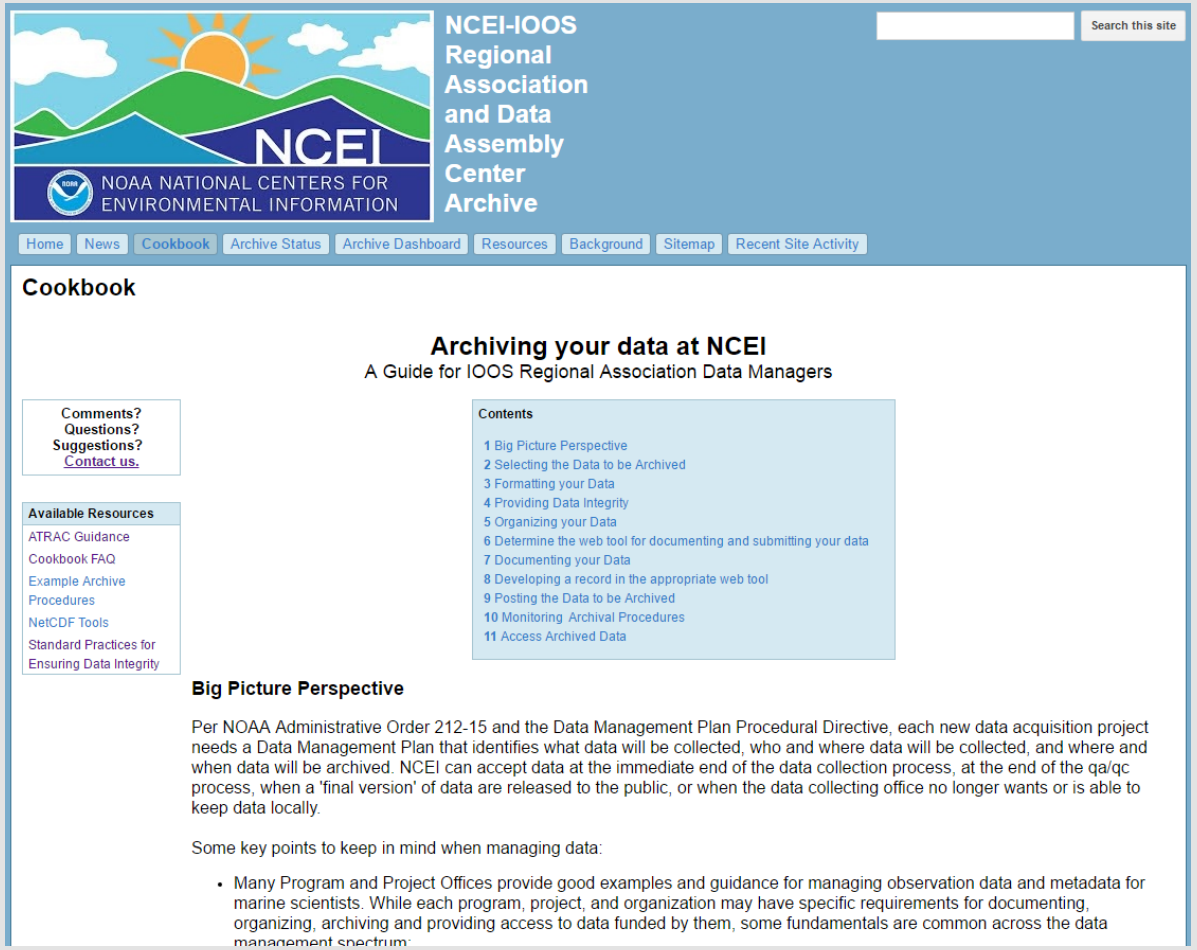
Regional Archiving Status

Regional Association	Conversation started?	ATRAC started?	Request to Archive?	ISO?	Submission Agreement ?	Data Ready?	Data Access Available?	Automation Initialized?	non-Federal Obs. Data Archive Status	NCEI HDT Number	ATRAC Project ID	percent complete (%)
SECOORA	Y	Y	Y	Y	Y	Y	Y	Y	complete	19108	6396	100
GLOS	Y	Y	Y	Y	Y	Y	Y	Y	complete	19615	na	100
SCCOOS	Y	Y	Y	Y	Y	N	Y	N	incomplete	23127	7918	75
GCOOS	Y	Y	Y	Y	Y	N	Y	N	incomplete	23126	7909	75
CariCOOS	Y	Y	Y	N	N	N	Y	N	incomplete	22684	7667	50
NANOOS	Y	Y	Y	N	N	N	N	N	incomplete	21940	6901	37.5
CeNCOOS	Y	Y	N	N	N	N	Y	N	incomplete	23128	7950	37.5
PacIOOS	Y	Y	N	N	N	N	Y	N	incomplete	23129	7907	37.5
MARACOOS	Y	Y	N	N	N	N	N	N	incomplete	20964	5702	25
AOOS	Y	Y	N	N	N	N	N	N	incomplete	23125	7947	25
NERACOOS	Y	N	N	N	N	N	N	N	incomplete	22684		12.5

Tools

NCEI Archiving Cookbook

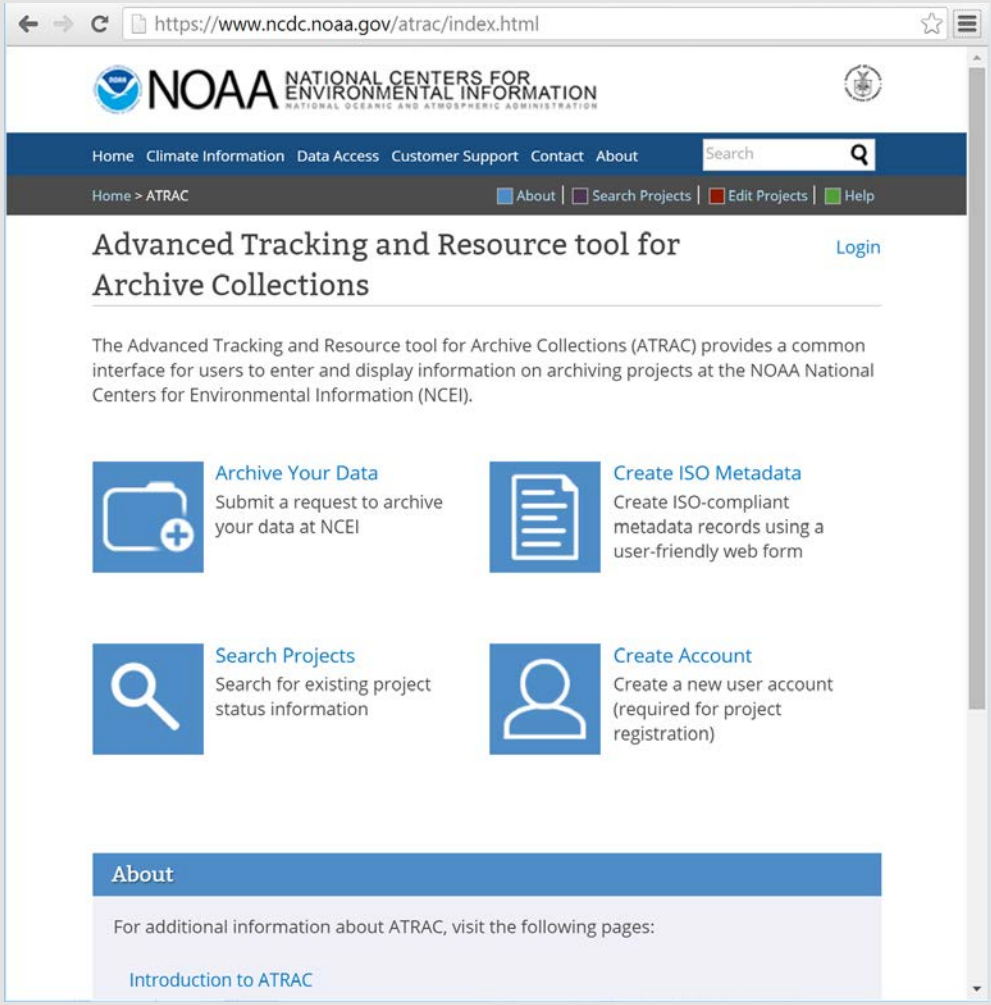
https://sites.google.com/a/noaa.gov/ncei-ioos-archive/cookbook



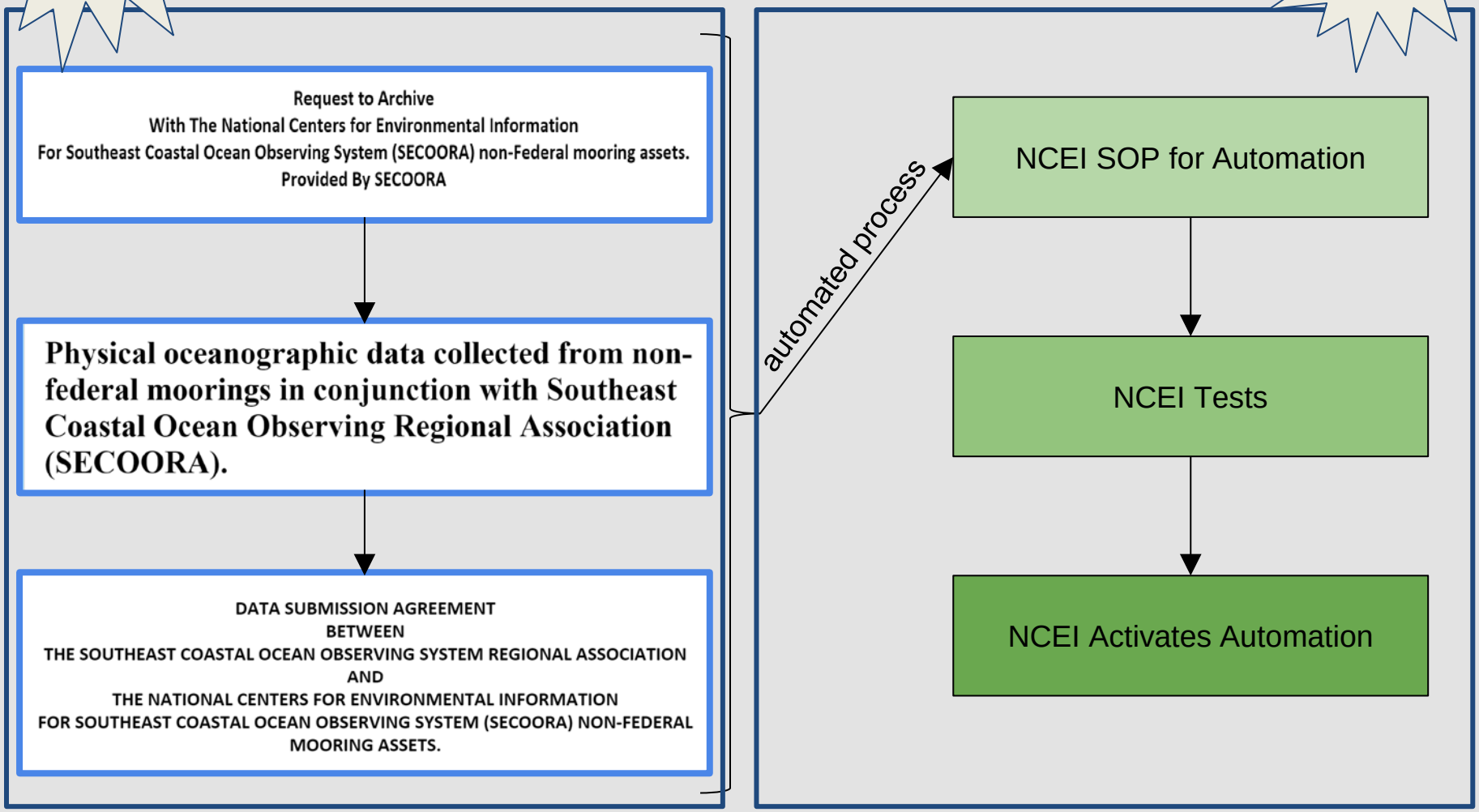
ATRAC

https://www.ncdc.noaa.gov/atrac/index.html

- 'Heart' of an automated archive process.
- Provides a common interface for users to enter and display information on archiving projects at NCEI.
- Allows multiple users to add and update the documentation for an archive project.
- Generates ISO 19115-2 metadata record.
- Information collected in ATRAC is directly transferred to an automation framework.



ATRAC Record to Automation



NCEI netCDF Template Update

https://www.nodc.noaa.gov/data/formats/netcdf/v2.0/

- Content Changes (**3**):
  - NODC → NCEI
  - Conventions
  - standard_name_vocabulary
- Additional attributes:
  - Variable (**2**):
    - missing_value
    - coverage_content_type
  - Global (**18**):
    - publisher_name
    - geospatial_bounds
    - geospatial_bounds_crs
    - geospatial_bounds_vertical_crs
    - time_coverage_duration
    - creator_type
    - creator_institution
    - publisher_type
    - program
    - geospatial_vertical_units
    - date_issued
    - date_metadata_modified
    - product_version
    - platform_vocabulary
    - instrument_vocabulary
    - metadata_link

Recent Changes		
v2.0	2015-12-07	New template for swath data Updated attributes to conform to ACDD 1.3 Updated guidance table Minor corrections and improvements
v1.1	2014-05-12	Minor corrections

NetCDF to SOS

- The primary goal of the mapping is to bring NCEI netCDF metadata into line with IOOS SOS Templates – GetCapabilities, DescribeSensor, and GetObservation.
- The secondary goal was to reuse NCEI Templates' attributes as much as possible in the course of reaching the primary goal.
- The most important gaps between NCEI netCDF and IOOS SOS metadata turned to be:
  - no short names in the NCEI Templates vs. universal use of them in the IOOS Templates;
  - lack of asset identification that matches IOOS SOS Templates;
  - issues with mapping multiple platforms/instruments.
- The crosswalk closes the identified gaps by augmenting the NCEI Templates with required 'ioos_code' (preferred) or 'short_name' variable attributes:
  - the former is an exact copy of the complete IOOS asset identifier;
  - the latter allows direct mapping to the essential fields of the IOOS asset identifiers (<label> for platforms, and <component> for instruments).
- The implementation of the 'ioos_code' shall be coordinated with the revision of the IOOS Convention for Asset Identification (see Kyle Wilcox's presentation).

Crosswalk from netCDF to SOS: rules for platform/instrument identifiers

- IOOS Asset Identifier is a URN ***urn:ioos:<asset_type>:<authority>:<label>:<component>***, where

	Value for platform	Value for sensor
asset_type	Fixed “station”	Fixed “sensor”
authority	naming_authority	naming_authority
label	platform_variable:short_name	platform_variable:short_name
component	N/A	geophysical_variable:standard_name

- It is strongly recommended (a.k.a. required) that a description of each geophysical variable in netCDF file provides reference to the platform & instrument variables, e.g.

  sea_water_temperature:platform = “platform_1 steady_as_rock_platform quick_as_thought_platform" ;

  sea_water_temperature:instrument = “instrument_1 fantastic_sensor odd_tool whimsical_gadget" ;

- For a platform, the following rules apply (in order of preference):
  1. ‘ioos_code’ attribute (similar to existing attributes ‘wmo_code’, ‘nodc_code’, ‘imo_code’, etc.) is strongly recommended for each ‘platform_variable’ with the value identical to an IOOS Asset Identifier;
  2. alternatively, a value of ‘platform_variable:short_name’ attribute (e.g. platform_variable_1:short_name = “carquinez”) should be used for the <label> field of the IOOS Asset Identifier**;**
  3. if neither 1) nor 2) is available, the value of a variable with a 'cf_role' attribute should be used for the <label> field;
  4. very reluctantly, and only if neither of the above options is available, the value of a global 'id’ attribute should be used for the <label> field.
- For a sensor, the following rules apply (in order of preference):
  1. it is required/highly recommended that each 'instrument#' variable should have the 'ioos_code' attribute; the value of the 'ioos_code' serves as an IOOS Asset Identifier;
  2. alternatively, the <label> field is defined as described for a platform, and the value of a 'geophysical_variable:standard_name' attribute (e.g. "water_temperature”) can be used for the <component> field of the IOOS Asset Identifier;

NetCDF to SOS

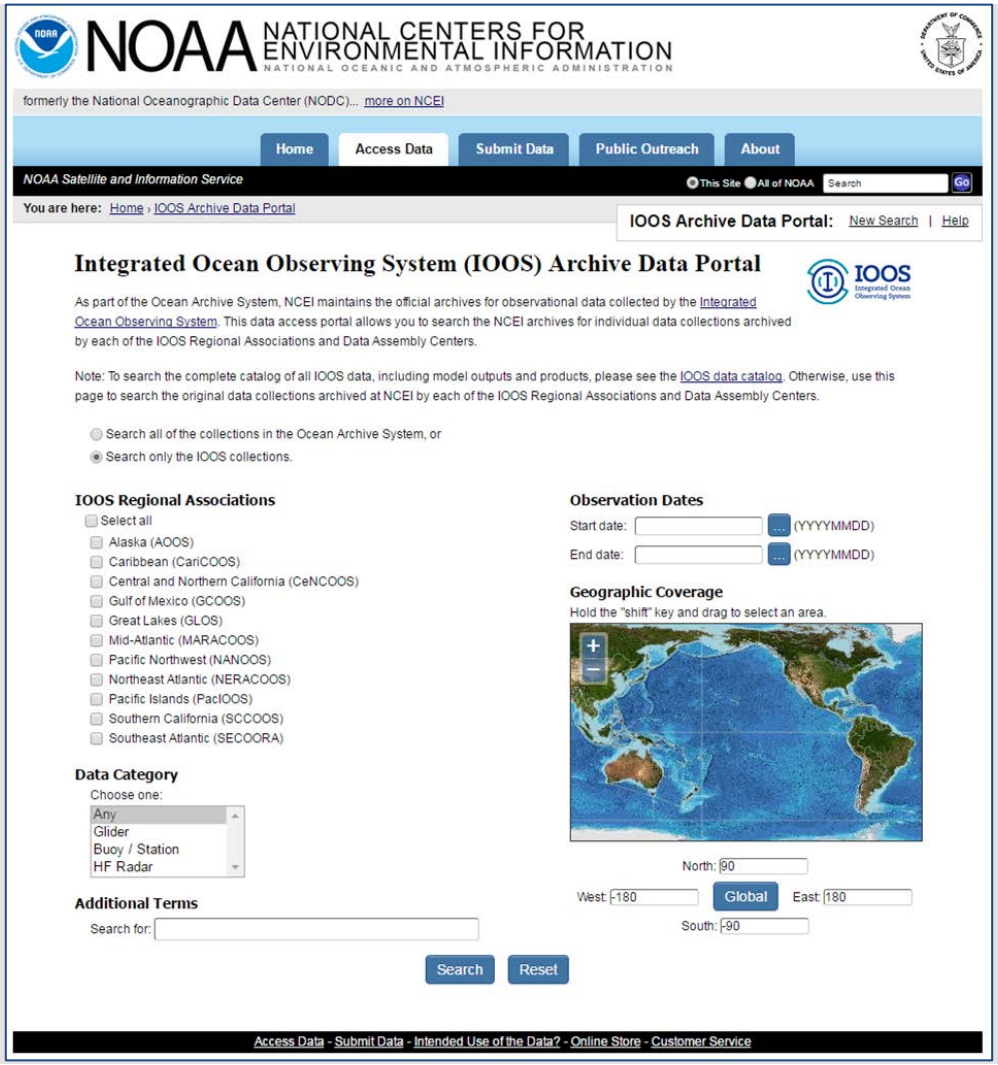
The following sample illustrates the use of ‘ioos_code’ attribute in netCDF metadata:

char instrument1 ;
instrument1:long_name = "Acoustic Doppler Current Profiler (ADCP)" ;
instrument1:comment = "Environmental/oceanographic probes and sensors" ;
instrument1:i**oos_code = "urn:ioos:sensor:GCOOS:sea_water_speed1" ;**
instrument1:calibration_date = "" ;
instrument1:serial_number = "" ;
instrument1:make_model = "" ;
char platform ;
platform:long_name = "DISL Dauphin Island, AL" ;
Platform:**ioos_code = “urn:ioos:station:DISL”**
platform:comment = "" ;
platform:call_sign = "" ;
platform:ncei_code = "147F, 3614" ;
platform:wmo_code = "" ;
platform:imo_code = "" ;

double sea_water_speed(timeSeries) ;
sea_water_speed:_FillValue = -999. ;
sea_water_speed:long_name = "sea water speed" ;
sea_water_speed:standard_name = "sea_water_speed" ;
sea_water_speed:ncei_name = "sea_water_speed" ;
sea_water_speed:units = "cm s-1" ;
sea_water_speed:scale_factor = 1. ;
sea_water_speed:add_offset = 0. ;
sea_water_speed:missing_value = -999. ;
sea_water_speed:valid_min = 0. ;
sea_water_speed:valid_max = 30. ;
sea_water_speed:coordinates = "time lat lon z" ;
sea_water_speed:coverage_content_type = "physicalMeasurement" ;
sea_water_speed:grid_mapping = "crs" ;
sea_water_speed:source = "GCOOS LDN upload/SOS." ;
sea_water_speed:references = "http://www.mymobilebay.com/stationdata/StationInfo.asp?jday=&property=&chartyear=&StationID=628" ;
sea_water_speed:cell_methods = "time: point lat: point lon: point z: point" ;
sea_water_speed:ancillary_variables = "instrument platform" ;
sea_water_speed:**platform = "platform" ;**
sea_water_speed:**instrument = "instrument1" ;**
sea_water_speed:comment = "" ;

NCEI IOOS Archive Data Portal

http://www.nodc.noaa.gov/ioos/



What's Next?

- ➔ Complete documentation for all RA non-Federal station data.
- ➔ Continue non-Federal RA station data archive implementation.
- ➔ Develop netCDF "Gold Standard" examples.
- ➔ Release 1-Page Checklist for archive.

Thanks!

Questions?

NCEI.IOOS@noaa.gov

NCEI netCDF Template Update

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v1.1.cdl, Top line: 1

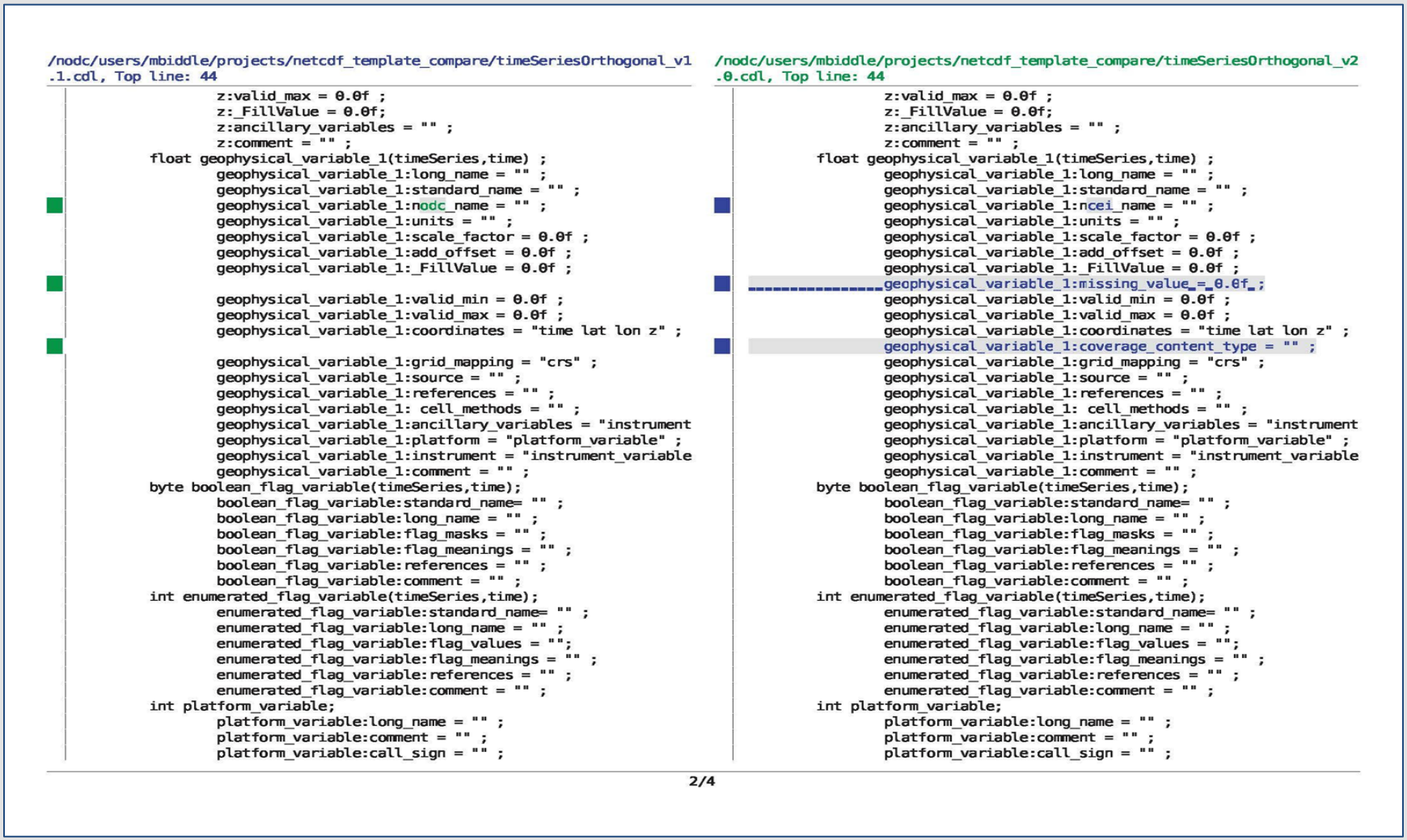
```
netcdf NODC_timeseries_orthogonal {
dimensions:
        time = < dim1 >;
        timeSeries = <dim2>;
variables:
         int timeSeries(timeSeries);
                 timeSeries:long_name = "Unique identifier for each featu
                 timeSeries:cf_role = "timeseries_id";
         double time(time) ;
                 time:long_name = "" ;
                 time:standard_name = "time" ;
                 time:units = "seconds since 1970-01-01 00:00:00 0:00" ;
                 time:calendar = "julian" ;
                 time:axis = "T" ;
                 time:ancillary_variables = "" ;
                 time:comment = "" ;
         float lat(timeSeries) ;
                 lat:long_name = "" ;
                 lat:standard_name = "latitude" ;
                 lat:units = "degrees_north" ;
                 lat:axis = "Y" ;
                 lat:valid_min = 0.0f ;
                 lat:valid_max = 0.0f ;
                 lat:_FillValue = 0.0f;
                 lat:ancillary_variables = "" ;
                 lat:comment = "" ;
        float lon(timeSeries) ;
                 lon:long_name = "" ;
                 lon:standard_name = "longitude" ;
                 lon:units = "degrees_east" ;
                 lon:axis = "X" ;
                 lon:valid_min = 0.0f ;
                 lon:valid_max = 0.0f ;
                 lon:_FillValue = 0.0f;
                 lon:ancillary_variables = "" ;
                 lon:comment = "" ;
         float z(timeSeries) ;
                 z:long_name = "" ;
                 z:standard_name = "" ;
                 z:units = "" ;
                 z:axis = "Z" ;
                 z:positive = "" ;
                 z:valid_min = 0.0f ;
```

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v2.0.cdl, Top line: 1

```
netcdf NCEI_TimeSeries_Orthogonal {
dimensions:
        time = < dim1 >;
        timeSeries = <dim2>;
variables:
         int timeSeries(timeSeries);
                 timeSeries:long_name = "Unique identifier for each featu
                 timeSeries:cf_role = "timeseries_id";
         double time(time) ;
                 time:long_name = "" ;
                 time:standard_name = "time" ;
                 time:units = "seconds since 1970-01-01 00:00:00 0:00" ;
                 time:calendar = "julian" ;
                 time:axis = "T" ;
                 time:ancillary_variables = "" ;
                 time:comment = "" ;
         float lat(timeSeries) ;
                 lat:long_name = "" ;
                 lat:standard_name = "latitude" ;
                 lat:units = "degrees_north" ;
                 lat:axis = "Y" ;
                 lat:valid_min = 0.0f ;
                 lat:valid_max = 0.0f ;
                 lat:_FillValue = 0.0f;
                 lat:ancillary_variables = "" ;
                 lat:comment = "" ;
        float lon(timeSeries) ;
                 lon:long_name = "" ;
                 lon:standard_name = "longitude" ;
                 lon:units = "degrees_east" ;
                 lon:axis = "X" ;
                 lon:valid_min = 0.0f ;
                 lon:valid_max = 0.0f ;
                 lon:_FillValue = 0.0f;
                 lon:ancillary_variables = "" ;
                 lon:comment = "" ;
         float z(timeSeries) ;
                 z:long_name = "" ;
                 z:standard_name = "" ;
                 z:units = "" ;
                 z:axis = "Z" ;
                 z:positive = "" ;
                 z:valid_min = 0.0f ;
```

1/4

NCEI netCDF Template Update cont.



/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v1.1.cdl, Top line: 44

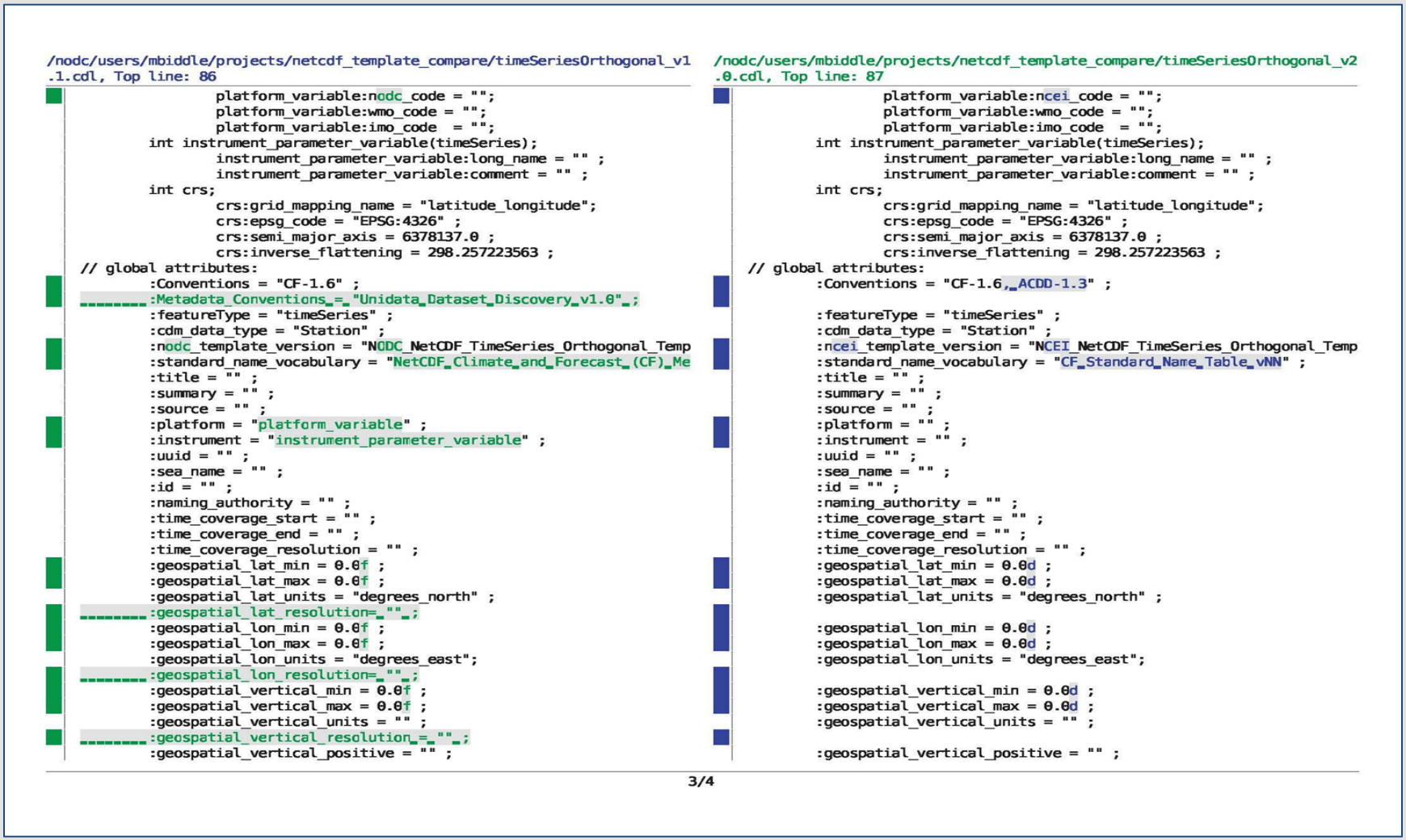
```
                z:valid_max = 0.0f ;
                z:_FillValue = 0.0f;
                z:ancillary_variables = "" ;
                z:comment = "" ;
        float geophysical_variable_1(timeSeries,time) ;
                geophysical_variable_1:long_name = "" ;
                geophysical_variable_1:standard_name = "" ;
                geophysical_variable_1:nodc_name = "" ;
                geophysical_variable_1:units = "" ;
                geophysical_variable_1:scale_factor = 0.0f ;
                geophysical_variable_1:add_offset = 0.0f ;
                geophysical_variable_1:_FillValue = 0.0f ;

                geophysical_variable_1:valid_min = 0.0f ;
                geophysical_variable_1:valid_max = 0.0f ;
                geophysical_variable_1:coordinates = "time lat lon z" ;

                geophysical_variable_1:grid_mapping = "crs" ;
                geophysical_variable_1:source = "" ;
                geophysical_variable_1:references = "" ;
                geophysical_variable_1: cell_methods = "" ;
                geophysical_variable_1:ancillary_variables = "instrument
                geophysical_variable_1:platform = "platform_variable" ;
                geophysical_variable_1:instrument = "instrument_variable
                geophysical_variable_1:comment = "" ;
        byte boolean_flag_variable(timeSeries,time);
                boolean_flag_variable:standard_name= "" ;
                boolean_flag_variable:long_name = "" ;
                boolean_flag_variable:flag_masks = "" ;
                boolean_flag_variable:flag_meanings = "" ;
                boolean_flag_variable:references = "" ;
                boolean_flag_variable:comment = "" ;
        int enumerated_flag_variable(timeSeries,time);
                enumerated_flag_variable:standard_name= "" ;
                enumerated_flag_variable:long_name = "" ;
                enumerated_flag_variable:flag_values = "";
                enumerated_flag_variable:flag_meanings = "" ;
                enumerated_flag_variable:references = "" ;
                enumerated_flag_variable:comment = "" ;
        int platform_variable;
                platform_variable:long_name = "" ;
                platform_variable:comment = "" ;
                platform_variable:call_sign = "" ;
```

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v2.0.cdl, Top line: 44

```
                z:valid_max = 0.0f ;
                z:_FillValue = 0.0f;
                z:ancillary_variables = "" ;
                z:comment = "" ;
        float geophysical_variable_1(timeSeries,time) ;
                geophysical_variable_1:long_name = "" ;
                geophysical_variable_1:standard_name = "" ;
                geophysical_variable_1:ncei_name = "" ;
                geophysical_variable_1:units = "" ;
                geophysical_variable_1:scale_factor = 0.0f ;
                geophysical_variable_1:add_offset = 0.0f ;
                geophysical_variable_1:_FillValue = 0.0f ;
                geophysical_variable_1:missing_value = 0.0f ;
                geophysical_variable_1:valid_min = 0.0f ;
                geophysical_variable_1:valid_max = 0.0f ;
                geophysical_variable_1:coordinates = "time lat lon z" ;
                geophysical_variable_1:coverage_content_type = "" ;
                geophysical_variable_1:grid_mapping = "crs" ;
                geophysical_variable_1:source = "" ;
                geophysical_variable_1:references = "" ;
                geophysical_variable_1: cell_methods = "" ;
                geophysical_variable_1:ancillary_variables = "instrument
                geophysical_variable_1:platform = "platform_variable" ;
                geophysical_variable_1:instrument = "instrument_variable
                geophysical_variable_1:comment = "" ;
        byte boolean_flag_variable(timeSeries,time);
                boolean_flag_variable:standard_name= "" ;
                boolean_flag_variable:long_name = "" ;
                boolean_flag_variable:flag_masks = "" ;
                boolean_flag_variable:flag_meanings = "" ;
                boolean_flag_variable:references = "" ;
                boolean_flag_variable:comment = "" ;
        int enumerated_flag_variable(timeSeries,time);
                enumerated_flag_variable:standard_name= "" ;
                enumerated_flag_variable:long_name = "" ;
                enumerated_flag_variable:flag_values = "";
                enumerated_flag_variable:flag_meanings = "" ;
                enumerated_flag_variable:references = "" ;
                enumerated_flag_variable:comment = "" ;
        int platform_variable;
                platform_variable:long_name = "" ;
                platform_variable:comment = "" ;
                platform_variable:call_sign = "" ;
```

2/4

NCEI netCDF Template Update cont.



/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v1.1.cdl, Top line: 86

```
                platform_variable:nodc_code = "";
                platform_variable:wmo_code = "";
                platform_variable:imo_code  = "";
        int instrument_parameter_variable(timeSeries);
                instrument_parameter_variable:long_name = "" ;
                instrument_parameter_variable:comment = "" ;
        int crs;
                crs:grid_mapping_name = "latitude_longitude";
                crs:epsg_code = "EPSG:4326" ;
                crs:semi_major_axis = 6378137.0 ;
                crs:inverse_flattening = 298.257223563 ;
// global attributes:
        :Conventions = "CF-1.6" ;
        :Metadata_Conventions = "Unidata_Dataset_Discovery_v1.0" ;
        :featureType = "timeSeries" ;
        :cdm_data_type = "Station" ;
        :nodc_template_version = "NODC_NetCDF_TimeSeries_Orthogonal_Temp
        :standard_name_vocabulary = "NetCDF_Climate_and_Forecast_(CF)_Me
        :title = "" ;
        :summary = "" ;
        :source = "" ;
        :platform = "platform_variable" ;
        :instrument = "instrument_parameter_variable" ;
        :uuid = "" ;
        :sea_name = "" ;
        :id = "" ;
        :naming_authority = "" ;
        :time_coverage_start = "" ;
        :time_coverage_end = "" ;
        :time_coverage_resolution = "" ;
        :geospatial_lat_min = 0.0f ;
        :geospatial_lat_max = 0.0f ;
        :geospatial_lat_units = "degrees_north" ;
        :geospatial_lat_resolution = "" ;
        :geospatial_lon_min = 0.0f ;
        :geospatial_lon_max = 0.0f ;
        :geospatial_lon_units = "degrees_east";
        :geospatial_lon_resolution = "" ;
        :geospatial_vertical_min = 0.0f ;
        :geospatial_vertical_max = 0.0f ;
        :geospatial_vertical_units = "" ;
        :geospatial_vertical_resolution = "" ;
        :geospatial_vertical_positive = "" ;
```

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v2.0.cdl, Top line: 87

```
                platform_variable:ncei_code = "";
                platform_variable:wmo_code = "";
                platform_variable:imo_code  = "";
        int instrument_parameter_variable(timeSeries);
                instrument_parameter_variable:long_name = "" ;
                instrument_parameter_variable:comment = "" ;
        int crs;
                crs:grid_mapping_name = "latitude_longitude";
                crs:epsg_code = "EPSG:4326" ;
                crs:semi_major_axis = 6378137.0 ;
                crs:inverse_flattening = 298.257223563 ;
// global attributes:
        :Conventions = "CF-1.6, ACDD-1.3" ;

        :featureType = "timeSeries" ;
        :cdm_data_type = "Station" ;
        :ncei_template_version = "NCEI_NetCDF_TimeSeries_Orthogonal_Temp
        :standard_name_vocabulary = "CF_Standard_Name_Table_vNN" ;
        :title = "" ;
        :summary = "" ;
        :source = "" ;
        :platform = "" ;
        :instrument = "" ;
        :uuid = "" ;
        :sea_name = "" ;
        :id = "" ;
        :naming_authority = "" ;
        :time_coverage_start = "" ;
        :time_coverage_end = "" ;
        :time_coverage_resolution = "" ;
        :geospatial_lat_min = 0.0d ;
        :geospatial_lat_max = 0.0d ;
        :geospatial_lat_units = "degrees_north" ;

        :geospatial_lon_min = 0.0d ;
        :geospatial_lon_max = 0.0d ;
        :geospatial_lon_units = "degrees_east";

        :geospatial_vertical_min = 0.0d ;
        :geospatial_vertical_max = 0.0d ;
        :geospatial_vertical_units = "" ;

        :geospatial_vertical_positive = "" ;
```

3/4

NCEI netCDF Template Update cont.

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v1.1.cdl, Top line: 129

```
        :institution = "" ;
        :creator_name = "" ;
        :creator_url = "" ;
        :creator_email = "" ;
        :project = "" ;
        :processing_level = "" ;
        :references = "" ;
        :keywords_vocabulary = "" ;
        :keywords = "" ;
        :acknowledgment = "" ;
        :comment = "" ;
        :contributor_name = "" ;
        :contributor_role = "" ;
        :date_created = "" ;
        :date_modified = "" ;
        :publisher_name = "" ;
        :publisher_email = "" ;
        :publisher_url = "" ;
        :history = "" ;
        :license = "" ;
        :metadata_link = "" ;
}
```

/nodc/users/mbiddle/projects/netcdf_template_compare/timeSeriesOrthogonal_v2.0.cdl, Top line: 130

```
        :institution = "" ;
        :creator_name = "" ;
        :creator_url = "" ;
        :creator_email = "" ;
        :project = "" ;
        :processing_level = "" ;
        :references = "" ;
        :keywords_vocabulary = "" ;
        :keywords = "" ;
        :acknowledgment = "" ;
        :comment = "" ;
        :contributor_name = "" ;
        :contributor_role = "" ;
        :date_created = "" ;
        :date_modified = "" ;
        :publisher_name = "" ;
        :publisher_email = "" ;
        :publisher_url = "" ;
        :history = "" ;
        :license = "" ;
        :metadata_link = "" ;
        :geospatial_bounds = "" ;
        :geospatial_bounds_crs = "" ;
        :geospatial_bounds_vertical_crs = "" ;
        :time_coverage_duration = "" ;
        :creator_type = "" ;
        :creator_institution = "" ;
        :publisher_type = "" ;
        :publisher_institution = "" ;
        :program = "" ;
        :date_issued = "" ;
        :date_metadata_modified = "" ;
        :product_version = "" ;
        :platform_vocabulary = "" ;
        :instrument_vocabulary = "" ;
}
```

4/4