**Florida Water Resources Monitoring Council (FWRMC)**
**Public Meeting of Monitoring Catalog Workgroup – December 5, 2012**

**Attendees:**

| Agency | Name |
|---|---|
| Florida Department of Environmental Protection (DEP) | Kate Muldoon |
| DEP | Stephanie Sunderman |
| DEP | Gail Sloane |
| DEP | Shel McGuire |
| Florida Department of Health (DOH) | Charlie Donahue |
| St. Johns River Water Management District (SJRWMD) | Aiša Cerić |
| SJRWMD | Tom Mirti |
| Southwest Florida Water Management District (SWFWMD) | Catherine Wolden |
| South Florida Water Management District (SFWMD) | Brian Turcotte |
| Miami-Dade County Department of Regulatory and Economic Resources | Susan Markley |
| University of South FL (USF) | Jim Griffin |
| USF | Jan Allyn |
| USF | Keith Bornhorst |
| USF | Jason Scolaro |

## Introduction

Kate Muldoon (KM) welcomed everyone to the Monitoring Catalog Workgroup (MCW) meeting and reviewed the agenda. The discussion will focus on continuing to develop and finalize a recommendation for the list of metadata elements to be included in the FL Catalog of Monitoring Programs.

KM clarified that the MCW is using the term "metadata" in a slightly different way than it is typically used in the GIS community. For the purpose of MCW meetings, metadata refers to the general elements or fields used to describe all critical information about a sampling program and its associated stations, but does not include the data (analytical results) themselves. KM suggested that the MCW may want to come up with a new term for the information, to avoid future confusion with other uses of the term "metadata".

## Discussion of Metadata Elements to be Included in FWRMC Catalog Effort

KM began by reviewed the metadata elements that were discussed at the last MWC meeting, and summarized any comments that have been received since the last MCW meeting.
At the 11/14/2012 MCW meeting, recommendations were completed for the following two elements:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| UNIQUE_ STATION_ID | Y | Primary key identifier for stations within the Monitoring Catalog database. Automatically generated by Water Atlas staff – likely to be a combination of the ORGANIZATION_NAME and STATION_ID. | VARCHAR2 | N | 150 | Water Atlas Staff | Station |
| ORGANIZATION _NAME | Y | Organization overseeing monitoring project collecting information at the station. | VARCHAR2 | Y | 100 | Data Providers | Project |

Additional suggestions were received after the 11/14/2012 MCW meeting for the following element:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| PROJECT_ NAME | Y | Monitoring project collecting information at the station. A station can belong to many projects overseen by the same organization or other projects overseen by other organizations. | VARCHAR2 | N | 250 | Data Providers | Project |

Comments received:

- Change definition to include: "A station can belong to many projects overseen by the same organization or other projects overseen by other organizations".
- Is it desirable to associate stations with projects from multiple organizations? If so, is this feasible? Many databases (e.g. FL STORET) are setup so that a "station" can only be associated with a single organization and its projects. If the same geographic location is monitored by two different organizations, then the database treats it as two separate "stations" (because the station's unique ID is a combination of the organization code and the station name / ID).

Discussion:

- KM indicated that the conclusion at the last MCW meeting was to develop a many-to-many relationship between projects and stations, so that a station can belong to many projects.
- KM indicated that there is interest in having the ability to associate a station to projects overseen by multiple organizations, and asked the Water Atlas staff if this is feasible.
- Keith Bornhorst (KB) indicated that this should not present any limitations, as it should be feasible to design a database with any type of relationship that is required.

Conclusion:

Change definition to include: "A station can belong to many projects overseen by the same organization or other projects overseen by other organizations".

At the 11/14/2012 MCW meeting, recommendations were completed for the following eight elements:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| PROJECT_URL | N | Address (URL) for a web page describing the project. | VARCHAR2 | N | 250 | Data Providers | Project |
| PROJECT_ DESIGN_ METHODOLOGY | N | Strategy to meet the goals and objectives of the project. Domain: *Deterministic; Probabilistic; Other; Unknown; Not Reported* | VARCHAR2 | Y | 25 | Data Providers | Project |
| PROJECT_ ABSTRACT | Y | Brief summary of the project. | VARCHAR2 | N | 4000 | Data Providers | Project |
| PROJECT_ PURPOSE | Y | Reason for information collection. Includes data quality objectives (DQOs). | VARCHAR2 | N | 4000 | Data Providers | Project |
| PROJECT_ DATA_ QUALITY_ OBJECTIVES | Y | Conveys whether documented data quality objectives (DQOs) are available. Domain: *Yes; No; Not Reported* | VARCHAR2 | Y | 12 | Data Providers | Project |
| STATION_NAME | Y | Full name of the site or station. | VARCHAR2 | N | 50 | Data Providers | Station |
| STATION_ID | Y | Code or identifier used as the station identifier within the data provider's internal database. | VARCHAR2 | N | 50 | Data Providers | Station |
| STATION_URL | N | Address (URL) for a web page describing the station. | VARCHAR2 | N | 250 | Data Providers | Station |

Additional suggestions were received after the 11/14/2012 MCW meeting for the following element:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| STATION_ DESCRIPTION | N | Description of the site or station. Information such as the address or parcel number associated with a station may be entered here. | VARCHAR2 | N | 250 | Data Providers | Station |

Comments received:

- Change definition to include: "Information such as the address or parcel number associated with a station may be entered here."

Discussion:

- KM clarified that including the suggested additions to the definition would help to clarify that this field may be used to convey supplemental information about the station's location. Information such as the address or parcel number associated with a station may be particularly helpful for locating groundwater (GW) stations.
- Tom Mirti (TM) indicated that it is not desirable to include information such as address or parcel number for most surface water stations. KM clarified that the proposed STATION_DESCRIPTION field is not required, and may be populated with any information that helps to describe the location or other features of the station.

Conclusion:

- Change definition to include: "Information such as the address or parcel number associated with a station may be entered here."

At the 11/14/2012 meeting, it was requested that a new straw-man proposal be developed for the metadata elements representing geographic location. The new proposal includes three options:

- **Option 1** – only allow points, and include general disclaimer for all stations that points may not represent exact sampling locations, and in some cases represent a larger sampling unit such as a polygon or transect.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| LATITUDE | Y | The decimal degree latitude. | FLOAT | N | 14,8 | Data Providers | Station |
| LONGITUDE | Y | The decimal degree longitude. | FLOAT | N | 14,8 | Data Providers | Station |

- **Option 2** – only allow points, but add an element called PROXIMITY_TO_SAMPLING_ACTIVITY

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| LATITUDE | Y | The decimal degree latitude. | FLOAT | N | 14,8 | Data Providers | Station |
| LONGITUDE | Y | The decimal degree longitude. | FLOAT | N | 14,8 | Data Providers | Station |
| PROXIMITY_TO_SAMPLING_ACTIVITY | Y | How close the LATITUDE and LONGITUDE are to the actual point where sampling activity occurred.<br>Domain:<br>*EXACT - Exact sampling location;*<br>*APPRX - Approximate sampling location;*<br>*VICIN - Vicinity of sampling location;*<br>*OFFST - Offset to exact sampling location;*<br>*START – Beginning vertex of a linear sampling location (e.g. transect);*<br>*BNDRY – One of a set of points used to describe the boundary of a polygon sampling location*<br>*CENTR - Center of polygon containing sampling location(s);*<br>*ADMIN - Administrative center of site or facility containing sampling location;*<br>*OTHER – Coordinates given represent a scenario not listed;*<br>*UNKNO – Relationship between the coordinates given and the sampling location is unknown;*<br>*NOREP – Information is not reported regarding the relationship between the coordinates given and the sampling location.* | VARCHAR2 | Y | 5 | Data Providers | Station |

- **Option 3** – Replace LATITUDE and LONGITUDE with two new fields. One field would contain the geometry type of the sampling location (point, line, or polygon), and another field would contain the location information, entered using a standard format (e.g. well-known text). This information could be stored as three feature classes, one for each geometry type.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| GEOMETRY_ TYPE | Y | The geometry type of the station location. Domain: *Point; Line; Polygon* | VARCHAR2 | Y | 7 | Data Providers | Station |
| LOCATION | Y | The geographic coordinates of the station, entered in well-known text format (or another standard format that the group agrees upon). Transferring information for lines and polygons by uploading ESRI shapefiles may eliminate the need to type this information into a field. Examples: POINT (27.91 -82.64) LINESTRING (27.91 -82.66, 27.92 -82.65, 27.91 -82.64) POLYGON ((27.94 -82.54, 27.93 -82.54, 27.92 -82.54, 27.93 -82.55, 27.94 -82.55)) | VARCHAR2 | N | 4000 | Data Providers | Station |

Discussion:

- KM summarized the benefits and limitations to using each option:
  - Option 1 is simplest, but users may not have complete information about the sampling location's geometry and accuracy.
  - Option 2 provides additional information about the given location's accuracy.
  - Option 3 is the most specific and provides the most information.
- Jim Griffin (JG) indicated that one of the first things that will be done is to associate that station with a Water Resource, which will provide fairly detailed information about the monitoring location.
- Brian Turcotte (BT) indicated that Option 3 allows the database to do SQL operations with GIS capability without actually having certain types of GIS software. This structure may make creating the GIS application to search the data repository easier because it would be easier to represent features using more than just a single coordinate pair.
- BT indicated that a coordinate pair and some descriptive information about that coordinate pair may be most useful to the greatest number of data providers. Option 2 is close to this, but perhaps there should be the ability to add a text description if "Other" is selected in the PROXIMITY_TO_SAMPLING_ ACTIVITY field.
- Charlie Donahue (CD) indicated that the Catalog may not need a high level of detail related to geographic location. Users could contact the data providers for more detailed information about the station location, since the final repository for the monitoring data (analytical results) will be with the data providers (and not part of the Catalog). BT agreed and suggested that the Catalog should have just enough information to make users want to take the next step and contact the data providers to find out more. Susan Markley (SM) agreed and suggested making the Catalog simple enough to encourage broad participation from data providers. Data providers will be less likely to update information if the process is too complex or time consuming.
- KM indicated that it seems like the MCW is favoring a simpler approach, and suggested using Option 1, and then adding information in the STATION_DESCRIPTION field if the station represents a larger sampling unit such as the centroid of a stratum in a probabilistic network.
- TM suggested using Option 2, as long as it would not be too onerous to populate the PROXIMITY_TO_SAMPLING_ ACTIVITY field. This would add valuable information about the uncertainty associated with the reported LATITUDE and LONGITUDE, without requiring users to find this information hidden among other comments in the STATION_DESCRIPTION field.
- SM indicated that her previous concern was that it is not practical to input every individual sampling point for probabilistic sampling programs. Instead it may be desirable to use the centroid of the polygon representing the stratum from which the random locations are selected. Similarly it may be desirable to use the beginning point of a transect, to represent the entire transect. Option 2 would address these scenarios.
- TM added that Option 2 also addresses scenarios such as the uncertainty associated with sampling at an approximate location (e.g. within 50m of a reported coordinate pair in a lake).
- Aiša Cerić (AC) suggested using Option 3 because it is the most flexible. Option 3 adds the possibility to include other types of data such as vegetation data collected along transects. KM indicated that while Option 2 may not include a representation of the entire transect, there are options for indicating the data was collected along a transect. TM and AC agreed that Option 2 does provide some additional information about transects, and may serve as a starting point for a future scenario in which transect data are represented in greater detail.

Conclusion:

Use elements described in Option 2.

At the 11/14/2012 MCW meeting, recommendations were completed for the following element:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| XY_DATUM | Y | The horizontal reference for measuring locations on the Earth's surface.<br>Domain:<br>*CAC – Cape Canaveral;,*<br>*HARN – High Accuracy Reference Network,*<br>*HPGN – High Precision GIS Network / High Precision Geodetic Reference;*<br>*NAD27 – North American Datum of 1927;*<br>*NAD83 – North American Datum of 1983;*<br>*WGS84 – World Geodetic Survey of 1984* | VARCHAR2 | Y | 10 | Data Providers | Station |

Additional suggestions were received after the 11/14/2012 MCW meeting for the following element:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| LOCATION_ COLLECTION_ METHOD | Y | The method or mechanism used to derive the location measurements.<br>Domain:<br>*ADDM – Address Matching,*<br>*AGPS – Autonomous GPS,*<br>*CALC – Calculated by GIS Software,*<br>*CSUR – Cadastral Survey,*<br>*DGPS – Differentially Corrected GPS (USGS coordinate datum code "D"),*<br>*DMAP – Digital Map Interpolation (USGS coordinate datum code "N"),*<br>*DPHO – Digital Aerial Photography With Ground Control,*<br>*GGPS – Geodetic Quality GPS,*<br>*LORN – LORAN-C Navigational Device (USGS coordinate datum code "L"),*<br>*MMAP – Manual Map Interpolation (USGS coordinate datum code "M"),*<br>*MPHO – Manual Aerial Photography With Ground Control,*<br>*OTHR – A Method Not Listed,*<br>*PARC – Parcel Matching*<br>*SATI – Satellite Imagery With Ground Control,*<br>*T- Transit, theodolite, or other surveying method,*<br>*TOWN – Township / Range Matching*<br>*WGPS – GPS with Wide-Area Augmentation Service Correction,*<br>*UNKN – Unknown Method,*<br>*ZIP2 – ZIP Code + 2 Segment Centroid,*<br>*ZIP4 – ZIP Code + 4 Segment Centroid,*<br>*ZIPC – Zip Code Centroid* | VARCHAR2 | Y | 4 | Data Providers | Station |

Comments received:

- Add additional domain values, including those from USGS NWIS (National Water Information System).

Discussion:

- KM clarified that the suggested additions include expanding the definitions for DGPS, DMAP, LORN, and MMAP as indicated above, and adding the following values: PARC, T, and TOWN.
- CD asked if the domain could be expanded in the future, if other values are desired. JG indicated that the domains can be easily expanded during the development phase, and likely expanded with some effort required at later phases of the project.
- KM clarified that these proposed additions do not include storing associated information (parcels, zip codes, etc.).

Conclusion:

Add all suggested domain values.

At the 11/14/2012 MCW meeting, recommendations were completed for the following two elements:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| COUNTY | Y | The name of the FL county station is located in.<br>Domain will include the names of all 67 counties and "Not Applicable". | VARCHAR2 | Y | 50 | Data Providers | Station |
| DRAINAGE_ BASIN | Y | The name of the 8-digit USGS HUC station is located in.<br>Domain will include all 8-digit USGS HUCs in FL and "Not Applicable". | VARCHAR2 | Y | 14 | Water Atlas Staff | Station |

Additional suggestions were received after the 11/14/2012 MCW meeting for the following element:

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| WATER_ RESOURCE_ NAME | Y | The name of the waterbody that the station is located in. For groundwater wells, this is the name of the aquifer tapped. (Note: "Unknown" is permitted if aquifer tapped is not known.) | VARCHAR2 | N | 100 | Data Providers | Station |

Comments received:

- Change definition to: "The name of the waterbody that the station is located in. For groundwater wells, this is the name of the aquifer tapped. (Note: "Unknown" is permitted if aquifer tapped is not known.)"

Discussion:

- KM clarified that the suggested additions include listing the aquifer tapped for GW wells as the WATER_RESOURCE_NAME. "Unknown" could be entered in this field if the aquifer tapped was not known.
- TM asked if there are any other fields that include information about the aquifer tapped. KM indicated that there are not. TM and BT indicated that this would be desirable because it would allow users to search by the aquifer tapped, and will make the GW resource data more useful.
- CD asked why WATER_RESOURCE_NAME is a required field for surface water resources because it seems like information that could be derived once the geographic location (LATITUDE and LONGITUDE) is known. BT indicated that WATER_RESOURCE_NAME information will be helpful for quality assurance purposes, especially in situations where the accuracy of the geographic location is not high.
- BT suggested using a list of suggested domain values, but also allowing data providers to enter other values if the desired value is not present. This may make it easier to allow users to search by WATER_RESOURCE_NAME. KB indicated that the Water Atlas has a list of water bodies for areas currently covered by the Water Atlas, and suggested leaving the WATER_RESOURCE_NAME field as free-text, with a list of suggested values available (but not required to be used).

Conclusion:

Change definition to: "The name of the waterbody that the station is located in. For groundwater wells, this is the name of the aquifer tapped. (Note: "Unknown" is permitted if aquifer tapped is not known.)" Have a list of suggested domain values available, but allow data providers to enter values that are not on the list.

At the 11/14/2012 meeting, it was requested that a new straw-man proposal be developed for STATION_TYPE.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| STATION_ TYPE | Y | The type of station that data collection occurs at / in.<br>Domain:<br>(*See options from new straw-man proposal below*) | VARCHAR2 | Y | 25 | Data Providers | Station |

- **Option 1** – (from FL STORET and current Water Atlas)

*Bay*
*Canal*
*Cave*
*Channelized Stream*
*Combined Sewer*
*Constructed Wetland*
*Estuary*
*Facility*
*Gallery*
*Great Lake*
*Lake*
*Land*
*Land Runoff*
*Landfill*
*Mine / Mine Discharge*
*Ocean*
*Reservoir*
*River / Stream*
*Riverine Impoundment*
*Spring*
*Storm Sewer*
*Tidal Fresh*
*Tidal Brackish*
*Waste Pit*
*Waste Sewer*
*Well*
*Wetland*
*Other*
*Not Applicable*

- **Option 2** – (from USGS NWIS, definitions available at http://help.waterdata.usgs.gov/site_tp_cd)

*Atmosphere*
*Ocean*
*Ocean - Coastal*
*Estuary*
*Lake*
*Stream*
*Stream - Canal*
*Stream - Ditch*
*Stream - Tidal stream*
*Spring*
*Well*
*Well - Collector or Ranney type well*
*Well - Extensometer well*
*Well - Hyporheic-zone well*
*Well – Interconnected wells*
*Well – Multiple wells*
*Well – Test hole*
*Other subsurface*
*Other subsurface – cave*
*Other subsurface – groundwater drain*
*Other subsurface – Tunnel, shaft, or mine*
*Other subsurface – Unsaturated Zone*
*Wetland*
*Land*
*Land – Excavation*
*Land – Outcrop*
*Land – Sinkhole*
*Land – Soil hole*
*Facility*
*Facility – Cistern*
*Facility – Combined Sewer*
*Facility – Diversion*
*Facility – Field, Orchard, or Nursery*
*Facility – Golf course*
*Facility – Hydroelectric plant*
*Facility – Laboratory or sample-prep area*
*Facility – Landfill*
*Facility – Outfall*
*Facility – Pavement*
*Facility – Septic system*
*Facility – Storm sewer*
*Facility – Thermoelectric plant*
*Facility – Waste-injection well*
*Facility – Wastewater sewer*
*Facility – Wastewater disposal*
*Facility – Wastewater-treatment plant*
*Facility – Water-supply treatment plant*
*Facility – Water-use establishment*
*Aggregate water-use establishment*
*Aggregate groundwater use*
*Aggregate surface-water-use*
*Other*
*Not applicable*

- **Option 3** – (include some / all of the entries from options 1 and 2, plus the following, as requested by workgroup members)

*Well – Potable Well*
*Well – Non-potable Well*
*Well - Monitoring Well*
*Well - Irrigation Well*
*Well - Production Well*
Well – Drainage Well (DW)
Well – Monitoring - Ground Water Quality Monitoring Well (IW)
Well – Observation - Ground Water Observation Well (OW)
Well – Other Well (XW)
Well – Unknown (UNK)
Well – Recharge Well (RW)
Well – Water Supply / Production - Agricultural Supply Well (AS)
Well – Water Supply / Production - Industrial Supply Well (IS)
Well – Water Supply / Production - Private Drinking Water Well (VW)
Well – Water Supply / Production - Public Drinking Water Well (BW)

Discussion:
- KM clarified that the desired list of domain values could include any combination of items from the lists above.
- CD, TM, and Catherine Wolden (CW) suggested using the values from Option 1. This list is simpler than the list from Option 2. If users need more detailed information they can contact the data providers.
- TM indicated that Option 3 may be too detailed, and problems could arise in cases where a station could be classified as multiple types (e.g. "Well – Drainage Well" and "Well – Monitoring").
- KM indicated that some values from Option 1 may want to be removed because they refer to types of stations not likely to be found in FL (e.g. "Great Lake" and "Gallery").
- CW suggested adding "Atmosphere" to the list of values from Option 1.
- CD suggested replacing the value "Well" with two new values: "Well – Potable Well" and "Well – Non-potable Well". All monitoring wells would be classified as "Well – Non-potable Well".

Conclusion:
Use domain values from Option 1, but remove "Great Lake" and "Gallery". Add "Atmosphere" to domain. Replace "Well" with "Well – Potable Well" and "Well – Non-potable Well".

As of the 11/14/2012 MCW meeting, no recommendations have been completed for the remaining elements. Details associated with most of these elements have not yet been discussed at workgroup meetings.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| STATION_ PHOTOGRAPH | N | Link to photo of station. (Data providers upload photograph(s), and Water Atlas Staff populate URL. | VARCHAR2 | N | 250 | Data Providers & Water Atlas Staff | Station |

Comments received:
- New element.

Discussion:
- JG indicated that it would be possible to include links to photographs of stations, but expressed concern that this would be a large quantity of information to store. JG asked the data providers if STATION_PHOTOGRAPH information will be useful enough to include.
- CD suggested changing the definition of STATION_PHOTOGRAPH to: "Conveys whether photographs of the station are available." The domain could be limited to: "*Yes", "No", and "Not Reported".*

- CW suggested that STATION_PHOTOGRAPH information should not be include in the Catalog. If users want to obtain this type of information, they can visit the websites associated with the projects / stations / data retrieval.
- Gail Sloane (GS) indicated that this information is useful for samplers trying to locate stations for re-sampling, but it may not be desirable to include photographs because data providers will need to obtain permission to display the photographs, if they include private property.
- TM, BT, and CW indicated that including any information about photographs in the catalog my require too much effort, especially since this information is already available through internal resources (and could be made available to others by contacting the data provider).

Conclusion:
Do not include STATION_PHOTOGRAPH in the Catalog.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| START_DATE | Y | Date monitoring activity started at the station. | DATE | N | | Data Providers | Station? |
| END_DATE | N | Date the monitoring activity ended or is anticipated to end at the station (use 99/99/9999 for ongoing monitoring with no anticipated end date) | DATE | N | | Data Providers | Station? |

Comments received:
- Is it necessary for this to be a required field?
- Is the monitoring activity relative to the start and end dates of the data collection in relation to the project information as filled in above or is it to be associated with the start and end date of data collection in relation to the site itself?

Discussion:
- BT and TM indicated that it is preferable to associate the start / end date with a station. Associating the start / end date with a project may be misleading because the data collected for one project that has ended may be used for another ongoing or future project with a larger period of record.
- CD asked if start / end date information is needed for the Catalog because users may just go to a location of interest and download all information available. SM indicated that users may want to know how old the information available is.
- TM indicated that START_DATE is desirable, and would be easy to populate because it would never change.
- CD asked why START_DATE is necessary. KM indicated that it is helpful to allows users to see when the earliest information was collected. It is helpful to provide some information about the period of record to users before they contact the data providers for more information. GS indicated that this will be helpful for identifying stations with long periods of records that may be useful for trend analysis.
- TM and SM indicated that it would not be desirable to include start and stop dates associated with every project at a given station.
- TM and BT indicated that it is not desirable (requires too much effort form data providers) to populate the END_DATE with the most recent sampling date for ongoing projects.
- SM and GS indicated that it is useful to include END_DATE information to help identify gaps in monitoring coverage and to help users understand when data is no longer being collected (e.g. some users may not want data from stations that have not been sampled in the last 10 years).
- BT suggested adding a field to identify stations as “active” or “inactive”. TM indicated that not including the date when monitoring ended will not capture enough detail related to the final date of information collection (for gap analysis, etc).
- KM suggested adding “present” or “active” as an option for END_DATE. An end date would not be required for active stations, but would be required for inactive stations.
- CW indicated that it is helpful to have this information in the Catalog, but it is not desirable to make this information required because a lot of effort is needed to maintain it.
- JG indicated that the Water Atlas updates the date last sampled information as the data is uploaded. However, the Catalog may be different because there may not be a need to update information about each station every time that it is sampled.

Conclusion:
Keep START_DATE as described in the straw-man proposal. For END_DATE, change definition to: "Date the monitoring activity ended or is anticipated to end at the station (enter "active" for stations with ongoing monitoring and no anticipated end date)".

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| SAMPLING_ FREQUENCY | Y | The time interval between sampling events at the station. | VARCHAR2 | N | 250 | Data Providers | Station? |

Comments received:
- Suggest a list to be used but do not enforce it. Will help for consistency.
- Suggested Domain: Sampled Every 5 minutes / Sampled Every 15 minutes / Sampled Every 30 minutes / Sampled Hourly / Sampled Quarterly / Sampled Semi-Annually / Sampled Annually / Triggered / Other

Discussion:
- BT indicated that this information is not stored in most data providers' databases, but can be derived (requires effort from data providers). However, this information is frequently requested, so it will be helpful for Catalog users. GS agreed, and indicated that sampling frequency information is desired by a variety of users, including DEP.
- SM indicated that sampling frequency at a station varies by parameter, but indicated that data providers may want to derive a broad summary of SAMPLING_FREQUENCY information at a station (e.g. seasonal information is available).
- KM indicated that SAMPLING_FREQUENCY was entered for each individual parameter for the GOMA Catalog of Monitoring Program.
- SM suggested adding "Monthly" to the list of suggested domain values.
- BT indicated that "Other" may be selected very frequently because many stations are sampled at variable frequencies or the sampling frequency has changed over time. KM suggested adding an additional free-text field to be populated if "Other" was selected. BT indicated that populating this level of detail may require too much effort from data providers.
- BT suggested lumping all sampling that occurs more frequently than daily into "Continuous", and then listing others that are sampled less frequently as "Discrete". KB suggested including additional sub-options for "Discrete", such as weekly, monthly, quarterly, etc.
- CW asked what options will be available for updating / maintaining information in the catalog, and indicated that it is desirable to have options for quickly updating information, such as SAMPLING_FREQUENCY, for a list of stations.

Conclusion:
Make this a domain limited field. The domain should include the following values: *Continuous (sampled more frequently than daily); Discrete - Sampled Daily; Discrete - Sampled Weekly; Discrete - Sampled Monthly; Discrete - Sampled Quarterly; Discrete - Sampled Semi-Annually; Discrete - Sampled Annually; Discrete – Triggered; Discrete – Varied; Discrete – Other*.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| SAMPLING_ TYPE | Y | The method in which the data was collected.<br>Domain:<br>*Observation;*<br>*Measurement;*<br>*Direct Grab;*<br>*Remotely Sensed;*<br>*Other.* | VARCHAR2 | Y | 25 | Data Providers | Station? |

Comments received:
- New element.

Discussion:
- CD suggested adding "Composite" to the domain.
- BT, AC, and CW indicated that the information level for SAMPLING_TYPE should be a single sampling event, not a station or project, therefore it may be difficult to include in the catalog.

- SM and TM indicated that it would be difficult to populate this information because some parameters collected at a single station, as part of a single sampling event, are measured using different sampling methods (e.g. salinity is measured in situ and phosphorous is analyzed from a direct grab sample).
- BT and CW indicated that this information would be helpful to users trying to locate stations with specific types of sampling (e.g. autosamplers), so perhaps this element should be included but not required.
- KM suggested designing the field to allow multiple domain values to be selected for a single station (e.g. "direct grab" and "observation"). TM indicated that designing the database to allow multiple selections for a single element may complicate data transfer.
- TM indicated that sampling type may also change throughout time, so the catalog would be limited to only capturing the most recent sampling type.
- TM, BT, and CW suggested not including this element in the Catalog.

Conclusion:
Do not include SAMPLING_TYPE in the Catalog.

| METADATA ELEMENT NAME | REQUIRED (Y/N) | DEFINITION | DATA TYPE | DOMAIN LIMITED (Y/N) | DATA SIZE | POPULATED BY | INFORMATION LEVEL |
|---|---|---|---|---|---|---|---|
| PARAMETERS | Y | List of parameters / variables / information collected at the station, including STORET/WQX codes when possible. | VARCHAR2 | N | 500 | Data Providers | Station? |

Comments received:
- Increase data size…..proposed change to maximum allowable
- Is this a list of parameters across the period of record or is it only those for the project listed above? Would high level data collection type values be accepted – i.e. "water quality", "water level", "atmospheric", etc?

Discussion:
- ***KM indicated that the group will need to identify the desired information level*** - project / station / sampling event / date range.
- ***KM indicated that the group will need to identify the desired level of detail*** - groups of analytes (e.g. nutrients) / individual analytes (e.g. dissolved oxygen).
- BT indicated if the group decides to target a higher level of detail such as groups of analytes, then the taxonomy for grouping individual analytes will need to be developed and agreed upon. It may be possible to find a taxonomy for grouping that is already in use by another effort.

Conclusion:
***Continue discussion at next meeting. Jason Scolaro, BT, KM, and SS will draft a straw-man proposal for parameter information to discuss at the next meeting.***

**Next Meeting**
***KM suggested that the workgroup meet again via webinar during the week of December 17-21*** to continue discussion of the remaining elements from the straw-man proposal and address the outstanding issues mentioned above.