What I am doing for CBI “pass-through” WOF service:

* I have a script called **build\_cbi\_cache.py** to make a local database cache for the CBI metadata, including Sites, Parameters, Units, and Series Catalogs
  + **Sites**: Sites come from the IOOS Reg file at <http://lighthouse.tamucc.edu/ioosobsreg.xml>
    - Parse the xml to extract site name, code, latitude, and longitude of sites
    - Not all Sites in the IOOS Reg file are contained in the CBI SOS Capabilities document available at <http://lighthouse.tamucc.edu/sos> so I am NOT including the sites that are not in both sources.
      * Is this ok?
      * If the Capabilities document and the Reg file are later brought into agreement, then all the sites will be added upon re-running build\_cbi\_cache.py
      * The site codes that are in IOOS Reg file but not in the Capabilities document are: 193, 167, 183, 165, 166, 155, 174, 156, 192, 173, 159.
    - The IOOS Reg file has vertical datums and vertical offsets, but they are at the “Series” (combination of site and variable) level, not the Site level. A single Site can have multiple vertical datums and offsets. Typically they are all 0m from MSL, except for “WaterLevel”. I am not sure how to handle this. Any ideas?
  + **Parameters**:
    - Valid parameter codes are extracted from the Capabilities document.
    - The “observationNames” in the IOOS Reg file do not match those in the GCOOS parameter file.
    - The parameter codes are cross-referenced to the GCOOS ontology file at <http://mmisw.org/ont?form=rdf&uri=http://mmisw.org/ont/gcoos/parameter> which (sort of) gives the **Units**.
      * The units are not really well described. See “Units from GCOOS” table in accompanying XLS file for a listing of all extracted units for CBI service. Should this be improved? How?
  + **SeriesCatalogs:**
    - Series are represented in the Capabilities document in the <sos:observationOffering> elements. From these elements I can parse the site and parameter codes and create SeriesCatalogs in the local cache
      * Date range of series is given in gml:beginPosition and gml:endPosition elements
        + I assume that if the <gml:endPosition> element is empty, then data collection is current.
        + My CbiDao class will use the current datetime of the server when a GetSiteInfo request is made to fill in the series EndDateTimeUTC. Is this ok?