Fundamental goal is to enable discovery

Developing a common vocabulary is the key. This requires semantic mediation/mapping and the difficulty of this varies between group. For example some benefit from existing semantic lists (USGS/EPA). Example of a mature system is the Cuashi which scrapes metadata of point time series {Richard Hooper, Cuashi), that scrapes the metadata physical, chemical point time series. Federal STORET, Ocean-DMAC, ASOS. Scraping the data as a Web Services strategy, based basic physical variables time, space, and community specific granularity. Hurdle is often at the registration of the data set (The costs of the data registration, SQL-dependent).

Some legacy systems (LTER) has a meta-data catalog, need to develop the community wide control vocabulary, example everyone calls the stuff a different name. So it is organizational issue that establishes a standard; need a flexible bottom-up driven system. **Can we establish a semantic standards approach that (SEEK program) is driven by functional needs**? **Knowledge encoding needs to evolve, so standardizing the approaches for developing agreed upon methods to evolve environmental observatory ontology's**. Method, bottom up vocabulary work (GRDDL), use case approach to allow the discovery, inventory, and detail. First, discovery provides the scientific context in a domain that is semantically linked to other the domains (ESIP.org). **We must engage lots existing community expertise. For example the ESIP community and provide a give and take but at the dawn of the GEOSS era that is largely the ISO compliant.**

We need a group, or body, to represent the observatory the groups as a whole. This is even required to even the share the controlled vocabulary the group, and we have the developed the means to mediate between the groups. **We should develop the means to mediate the development the control vocabulary through an informed ontologeny. This should be a hierarchal and taxonomic attributes. Probably best developed with a group summit focused on some use case networks allowing the CI engineers, scientists, and is thus rooted in reality. Advantage is you can then map the ontology, and the data gaps can be inferred, and this lowers the registration hurdle. There is a level difficulty, and it will take time and be organic. Data curation would be great after the community has seen the demonstration of the power.**

How would ensure the data quality? By creating an engine to reach the NSF Environmental observatory data portals. There are governance issues for security that would suggest that there would be a common type authentication. **Need to define a data publishing policy, and these policies can be built into the software infrastructure and outlines the shared principles (the rules of the road). This publishing policy cuts across all observatories, and that will evolve in time. This could be also tackled by the EO board.**

Simple versioning of databases is often not sufficient. Raising the issues of defining the community wide standards. But this is the data publishing issue. Need to keep track to of who accesses the data, and the community. We need to quantify the metrics and then use as a basis for provenance.