IOOS Vocabulary and Ontology Strategy for Observed Properties

Sara Haines
Department of Marine Sciences
University of North Carolina
Chapel Hill, NC

Vembu Subramanian
SECOORA
University of South Carolina
Columbia, SC 29208

Emilio Mayorga
NANOOS
University of Washington
Seattle, WA

Derrick Snowden and Rob Ragsdale
IOOS Program Office
Silver Spring, MD

Carlos Rueda
Monterey Bay Aquarium Research Institute
Moss Landing, CA

Matthew Howard
Department of Oceanography
Texas A&M University
College State, TX

***Abstract*—With the rapid growth of coastal ocean observations becoming available for integration by US Integrated Ocean Observing System (IOOS) Regional Associations and federal data assembly centers, there is a need for the establishment of IOOS Parameter Vocabulary strategy. Currently, different data naming conventions are being used by existing regional and sub-regional coastal ocean observing systems. This makes things complicated for the discovery, access and proper usage of the valuable data. To eliminate the misuse and misinterpretation of the data being made available and to facilitate the discovery and proper use of in data scientific research and other management applications, the authors have presented the development of IOOS Parameter Vocabulary and recommended a strategy to move this forward with ocean observing community engagement.**

***Index Terms*—Vocabulary, ontology, semantic web, ocean, observing.**

I. Introduction

Every discipline has its own terminology for data that they collect. Names and codes are assigned to variables (also known as observed properties, parameters, phenomena, etc) that are associated with the data obtained from multiple methods of observing our environment. The methods include data from direct measurement, sample collection, analysis, and numerical modeling. The terminology used to describe the data is collectively referred to as a vocabulary. The naming of variables and their variable descriptions (measurement procedure, medium, units, location and time) differ across disciplines and subgroups. These differences lead to difficulties in the use of data for scientific research, and complicate interpretation, data discovery, and integration of data from multiple sources. The problem is exacerbated when data providers use different names for the same measurement or use the same name based on different assumptions.

As data are integrated from organizations joined together to create regional and national ocean observing networks, a need for data naming conventions arises. However, because of the complexity and sheer number of possible names, no one controlled vocabulary can encompass all the terms needed for such a large enterprise.

Climate and Forecast (CF) Standard Name Table currently holds over 2000 unique terms for ocean and atmosphere states [1]. The United States Geological Survey (USGS) National Water Information System (NWIS) has over 18000 codes for associated hydrologic variables [2]. The Global Climate Change Master Directory (GCMD) contains hundreds of scientific key words [3]. The British Oceanographic Data Centre (BODC) Parameter Usage Vocabulary contains over 17,000 entries [4]. Other sources that can be identified are taxonomies and chemical substance registries. Each of these controlled vocabularies has been developed to meet specific requirements of each domain with terms that may or may not be relevant to the marine environment.

The problem of diversity in data naming is not new and solutions are emerging in the oceanographic community. Graybeal *et al.* [5] have built an ontology hosting and semantic mediation framework to address the problem associated with diversity in data naming, originally intended for the marine community. Similar semantic challenges have been addressed in the hydrological [6,7] and ecological [8] communities and BODC's National Environmental Research Council (NERC) DataGrid [4].

Copyright 2012 IEEE

There are two basic needs for data naming conventions. The first is providing names that data providers can understand as they relate to their data to collection processes or goals of the research and analysis. The second is providing names to concepts for integrated data delivery systems that are machine readable and available for automated discovery and delivery.

Network Common Data Format (netCDF) and Sensor Model Language (SensorML) are two data format standards and interoperable information technologies that need uniquely defined standard names and observed properties, respectively.

Hankin *et al.* [9] outlines the use of netCDF files created according to the CF Metadata Conventions and made available over networks using Data Access Protocol (DAP). The three standards (netCDF-CF-DAP) bring a high level of interoperability to many disparate oceanographic, atmospheric, and climate datasets [9]. The goal of the CF Standard Name Table, just one component of the entire CF Metadata Conventions, is to ensure a unique, well defined name for each variable stored in netCDF files that conform to the CF Metadata Conventions. One not so obvious consequence is that the data behind the CF Standard Name can become discoverable. For example, translation software exists that can extract essential metadata from CF compliant netCDF files and create ISO 19115-2 compliant metadata files (https://geo-ide.noaa.gov/wiki/index.php?title=NcISO). These ISO files can be loaded into service registries or catalogs to provide a data discovery service that complements the DAP data access service. The combination of data discovery services and multiple data access services begin to fill out an entire data distribution network. As the size of the network grows, a small number of controlled vocabularies combined with semantic mediation frameworks to relate these vocabularies to each other, become key enabling technologies that allow client software to be interoperable across all elements of a data distribution network.

Sensor Model Language (SensorML) is an Open Geospatial Consortium (OGC) standard and XML encoding for describing sensor measurements and processes that can be mined by OGC Web Services to provide access to distributed sensor and data networks [10,11]. SensorML encoding requires unique definitions, including ones for the parameter or property that is being sensed or described within the XML schema. A set of controlled vocabularies for observed properties is of utmost importance to assist data publishers with labeling data names with consistent interpretation.

The need for guidance and tools for vocabulary usage is three-fold. First, it should assist/help data providers with consistency and guidance in labeling ocean and atmospheric variables in data that they provide using standard services such as OGC Web Services and netCDF-CF-DAP. Second, this guidance should reduce the ambiguity inherent in human language where the same concept is often given two different names and provides relationships to broader concepts through mapping of these relationships. Third, it should promote discovery of terms or terminology in multiple domains and disciplines.

Below, we outline the motivation and lineage of the Integrated Ocean Observing System (IOOS) Parameter Vocabulary and work done by IOOS Regional Associations (RAs) towards more integrated terminology. We describe the current work on mapping efforts between IOOS Core Variables and IOOS Parameter Vocabulary and CF Standard Names and provide a few query examples that open the semantic knowledge and understanding of concepts to details of data. By describing this work and summary of preliminary results, we conclude by making suggestions for future steps to meet the growing semantic needs of IOOS nationally and regionally.

II. Marine Metadata Initiative (MMI) Ontology Registry and Repository (ORR)

The Marine Metadata Interoperability (MMI) Project was first funded in 2004 by the National Science Foundation. The project goal was to improve data exchange in the marine community by introducing marine scientists and data managers to concepts and practices that would enable intelligent agents (i.e., computers) to discover data with less human oversight [5]. Although it is no longer funded, MMI and its web presence persist through the efforts of volunteers and benevolent organizations who host the MMI websites and tools. MMI's efforts are in concert with the concepts of the Semantic Web.

The Semantic Web is a collaborative effort to develop agreements on standards and common data formats used to name and tag data objects and express relationships between related terms. The reason for doing this is so intelligent agents can perform efficient searches with a high-probability of locating all of the desired information while returning little of the similar sounding but irrelevant information. For example, the system should correctly return information on ocean currents when an oceanographer searches for "currents" and not return information on electrical currents. An intelligent search agent will know that conductivity is related to salinity and will display data on the conductivity of seawater when asked for salinity data if the user wants related terms to be included in the return. The system will understand that the English word salinity, the Spanish word "salinidad", and the German word "salzgehalt" all refer to the same thing. Using any of these terms in a search should return the same result. The approach being taken by information architects to achieve this effect is to capture and encode metadata plus semantics (word meanings) and ontologies (concepts and relationships within a community's domain of use e.g., oceanography) in a form computers can process into efficient queries. Thus, the scope of work to achieve the Semantic Web concept includes creating software tools for encoding, decoding, and translating semantics and ontologies together with many lines of Boolean logic code needed to correctly traverse the chains of relationships.

The MMI Ontology Registry and Repository (ORR) (http://mmisw.org/) is a web application designed for human subject matter experts to create lists of terms and their meanings (vocabularies), to express relationships between terms (ontology) and to publish this information for broader use [12]. Once created, ontologies can be registered with the

Copyright 2012 IEEE

ORR then re-used inside other ontologies. Borrowing from existing ontologies is a common and recommended practice. For example, the meanings and relationships between the terms “station”, “platform”, “instrument”, and “sensor” may be the same in several domains of data collection. If the spellings and definitions of terms in an existing ontology are suitable for an application the user need only list the term or terms and reference the existing ontology.

The MMI ORR is a community site where ontology providers and users can collaborate in building, annotating and mapping vocabularies. MMI complements the ORR with many guidance documents for both data providers and data users. The MMI Guides on “Vocabularies, Dictionaries, Ontologies and More” provide helpful tips and tools on how to construct, and use different types of vocabularies, within the context of marine metadata [13]. At the time of this writing, the ORR contains 192 registered entries. Some entries are mappings, while others are complete definitions. Mappings are statements about the logical relationships between terms in different vocabularies or ontologies. Relationships between two terms can be set to “exact match”, “approximate match”, “broader than”, “narrower than”, or “related”. Semantic reasoning engines can use these relationships to match terms referenced in different ontologies [5]. Some of the mappings can be quite mundane. They may simply equate terms written in lower case to the same term written with leading letter capitalization or camel case. In other cases, ontologies include definitions, units, and references (e.g., IOOS Parameter Vocabulary, http://mmisw.org/ont/ioos/parameter). The ORR website includes tools to create or upload definitions and to publish these definitions in the standards-based Resource Description Framework (RDF) encodings. RDF encodings take the form of subject-predicate-object such as “temperature” “has units of” “degrees Centigrade”. RDF lists can be processed into the OWL Web Ontology Language which is the form machines use to process the information contained in the ontologies.

III. CF Standard Name Table

Atmospheric and oceanic forecast modelers were among the first to have the need to move large datasets across the network on a routine basis. Atmospheric model forecast outputs are used to force oceanic circulation models. High-resolution small-domain models are nested within low-resolution large-domain models. Coupled ocean-atmosphere models move model output in both directions. The modeling community quickly embraced the netCDF Application Programmer Interface (API) which uses a file format that is self-describing and allows for efficient direct access to subsets of multi-dimensional model output. While variable dimensions, variable names and numeric data types were described in the file, file authors were free to use whatever variable names, parameters units and array packing order they wanted. This meant the file subscriber had to write custom code for each source of model output received. Over time conventions were established which greatly simplified reading and using foreign files.

The Cooperative Ocean/Atmosphere Research Data Service (COARDS) conventions were first introduced in 1995. The conventions standardized parameter unit names, file name extensions (.nc), and array packing order for common 4-dimensional data (time, height, lat, lon) and constrained other aspects such as variable name construction. The CF conventions for metadata, introduced in 2003, built upon the COARDS conventions. The most significant new convention was the requirement to use standardized parameter names for value of the standard_name attribute in a COARDS/CF compliant netCDF file. The CF Standard Names list has all of the desirable attributes of a good controlled-vocabulary. These attributes include: broad adoption by a large active community, the vocabulary “spans the space” meaning it has a term for everything the community wishes to express, it has governance and is actively maintained, and there is a process for adding new terms through working groups. The CF Standard Name Table [1] is the authoritative source of the proper spelling of each parameter together with a definition and canonical units. A typical name is “upward_sea_water_velocity”. No capitalization is used and words are joined by the underscore character. Terms are grouped by category such as Atmospheric Dynamics, Ocean Dynamics, Cloud, Surface and others to make it easy to search for existing terms that might not be known ahead of time. Since 2001 the CF Standard Name Table [1] has been revised 19 times. Previous versions are available on the web. No terms have been dropped since v1 was established in 2003 (though this has led to ambiguities and changes in meaning [5]). The website contains guidance for the construction of CF Standard Names and the status of pending proposals for new standard names.

IV. IOOS Parameter Vocabulary

Vocabularies, glossaries, dictionaries and code lists are developed to meet needed semantic knowledge and consistent naming of concepts. The IOOS Parameter Vocabulary provides a list of marine terms, their definitions and units. This vocabulary was compiled to capture important observed parameters being integrated regionally by one or more of the 11 RA’s and potentially available for national integration by the IOOS Program Office. Figure 1 shows the IOOS Parameter Vocabulary as it is registered with MMI-ORR and viewable through the web interface provided by the MMI-ORR.

A little background on how the IOOS Parameter Vocabulary has evolved to its current status and content may provide some understanding of its similarities and differences to other parameter vocabularies specially that of CF Standard Name Table.

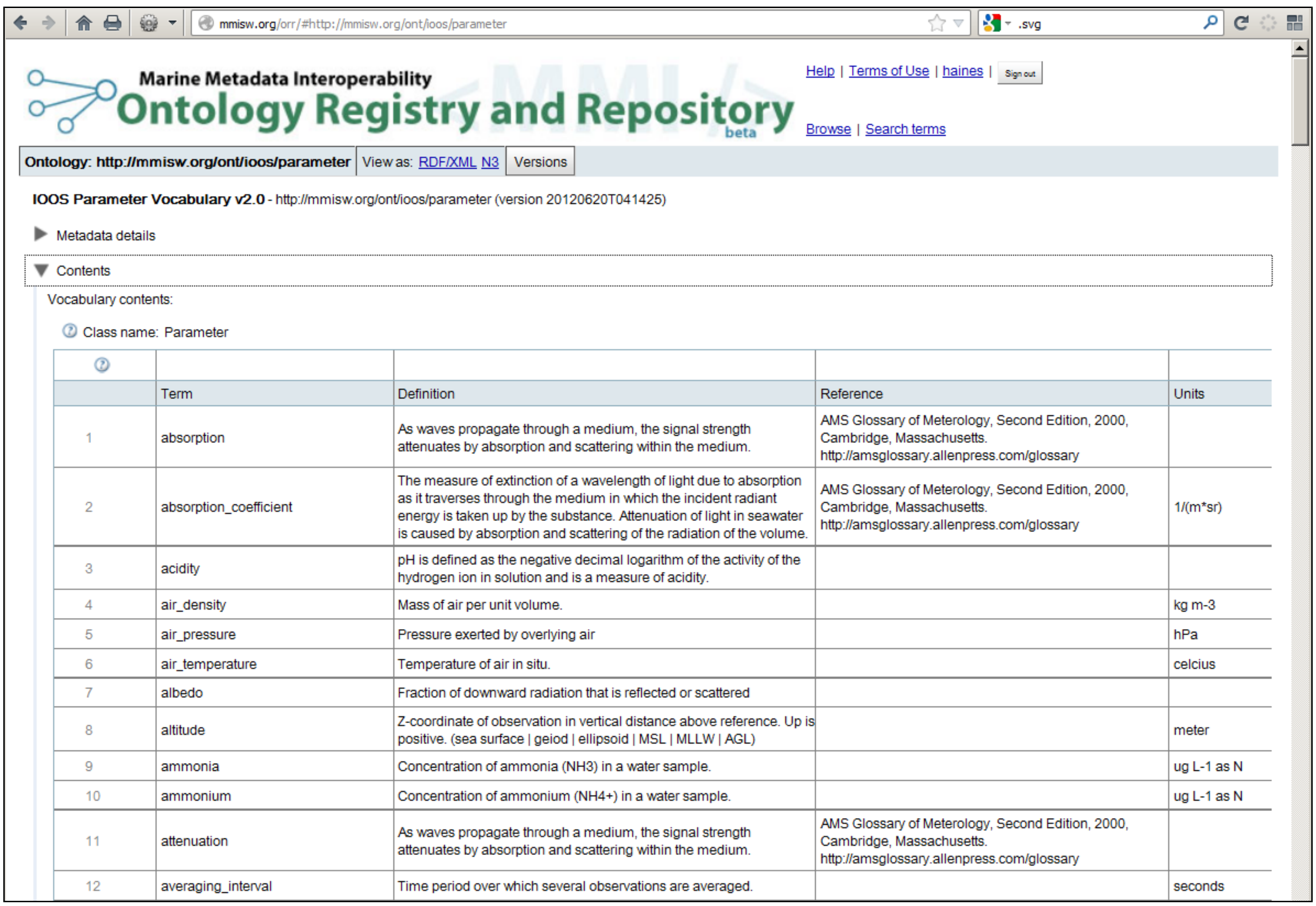
The Southeast Atlantic Coastal Ocean Observation System (SEACOOS), a consortium of universities with established coastal ocean observing programs worked together between 2002 and 2007 to construct a functional regional observing system. In order to share data a minimum set of standards were established to support data description, expression, and transport. SEACOOS relied heavily upon the CF Standard Name Table v1.0 and adapted the format where needed to

Copyright 2012 IEEE

support many different types of observations. At the time, the CF Standard Name Table v1.0 provided many atmospheric terms in support of atmospheric forecast modeling but did not include terms for most types of oceanographic measurements. The SEACOOS Data Dictionary was developed to specify the naming and referencing of the many types of marine observations in both the ocean and atmosphere to provide consistent naming of variables being shared and integrated by SEACOOS [14]. This was accomplished most efficiently by extending the CF Standard Name Table v1.0. A primary goal of extending the CF Standard Name Table was to provide names to parameters that can be applied across environments and observing networks. More specifically, unlike CF standard names, the SEACOOS Data Dictionary did not specify sea water as a substance where fresher water was the sample location, and did not specify location interfaces such as sea surface or cloud base when the depth or height defining surface or level varies with application and field measurement issues. When the Southeast Coastal Ocean Observing Regional Association (SECOORA) became incorporated in 2007, the SEACOOS Data Dictionary was adopted as the SECOORA Data Dictionary.

In 2010, an effort spearheaded by the Gulf of Mexico Coastal Ocean Observing System (GCOOS) RA collected input from all of the 11 RA's to review and suggest terms that span the domain of what is measured in each region. The SECOORA Data Dictionary was the starting point. Additional terms were added mostly encompassing wave parameters, water quality measurements and a few chemical constituent terms. This list was registered on MMI-ORR as the IOOS Parameter Vocabulary v1.0. Comparisons to equivalent or related CF Standard Names v13 were made but no changes to parameter names in the regional list were made as it was recommended that a mapping of associated parameters could be made.

In 2011, SECOORA funded work to ensure all terms from the IOOS Parameter Vocabulary v1.0 were given unique definitions and researched for consistency and continuity. All term names and definitions were reviewed and references and units provided. This vocabulary was registered on the MMI ORR in 2012 as the IOOS Parameter Vocabulary v2.0 (Figure 1). At the time of this writing, it contains 185 terms covering



	Term	Definition	Reference	Units
1	absorption	As waves propagate through a medium, the signal strength attenuates by absorption and scattering within the medium.	AMS Glossary of Meterology, Second Edition, 2000, Cambridge, Massachusetts. http://amsglossary.allenpress.com/glossary	
2	absorption_coefficient	The measure of extinction of a wavelength of light due to absorption as it traverses through the medium in which the incident radiant energy is taken up by the substance. Attenuation of light in seawater is caused by absorption and scattering of the radiation of the volume.	AMS Glossary of Meterology, Second Edition, 2000, Cambridge, Massachusetts. http://amsglossary.allenpress.com/glossary	1/(m*sr)
3	acidity	pH is defined as the negative decimal logarithm of the activity of the hydrogen ion in solution and is a measure of acidity.		
4	air_density	Mass of air per unit volume.		kg m-3
5	air_pressure	Pressure exerted by overlying air		hPa
6	air_temperature	Temperature of air in situ.		celcius
7	albedo	Fraction of downward radiation that is reflected or scattered		
8	altitude	Z-coordinate of observation in vertical distance above reference. Up is positive. (sea surface | geiod | ellipsoid | MSL | MLLW | AGL)		meter
9	ammonia	Concentration of ammonia (NH3) in a water sample.		ug L-1 as N
10	ammonium	Concentration of ammonium (NH4+) in a water sample.		ug L-1 as N
11	attenuation	As waves propagate through a medium, the signal strength attenuates by absorption and scattering within the medium.	AMS Glossary of Meterology, Second Edition, 2000, Cambridge, Massachusetts. http://amsglossary.allenpress.com/glossary	
12	averaging_interval	Time period over which several observations are averaged.		seconds

Fig. 1. The IOOS Parameter Vocabulary as viewed within MMI-ORR user interface

water quality, water chemistry, water properties, data quality, ocean currents, waves, and atmospheric properties. However, numerous terms on fish and zooplankton species and abundance and chemical substances were not included in this version of the vocabulary.

The IOOS Program Office and IOOS partners have joined efforts to develop a consistent, well described observed-properties vocabulary and ontology strategy based on these regionally led efforts and previous vocabulary activities coordinated nationally. The next iteration of work on the IOOS Parameter Vocabulary expands the relationships or mapping of terms to other controlled vocabularies, mainly that of CF Standard Names. In addition, terms are associated with IOOS Core Variables and Areas of Societal Benefit [15] and the general categories mentioned above to create some hierarchies for subsequent searches. This work is summarized in the following sections.

V. Mapping Terms

The term mapping tool provided by MMI ORR allows users to establish relationships between terms across multiple vocabularies using standards-based properties. Each IOOS term was pitted against CF terms that matched in content and then a relationship applied between them if applicable. Table 1 lists the five types of relationships available in MMI ORR and associated symbol used in the figures and the tool. Figure 2 shows how the IOOS term for "salinity" is an exactMatch (=) with CF term "sea_water_salinity", and has a narrowMatch (>) with the CF term "sea_surface_salinity" as it modifies the parameter to a specific location. A narrowMatch (>) relationship from IOOS term to CF term was used for other modifiers such as "_threshold" or "_anomaly." A relatedMatch (~) could have been used in a multitude of comparisons but the decision was made to not make the relationship to reduce confusion.

Also in Figure 2a, we show how the IOOS Parameter term salinity relates to the IOOS Core Variable "Salinity." The Core Variable is a broader concept to "salinity" and includes "conductivity" and "water_temperature" and "water_pressure" parameters as salinity is dependent on these. Finally, "weather and climate" is mapped as an even broader concept of which Core Variable "Salinity" is one of many narrower concepts to weather and climate. Figure 2a also shows an alternative hierarchy using Category "Water Property" to provide a different context for the IOOS and CF terms. Figure 2b is a summary diagram showing the number of relationships created in this process. Mappings such as these are the missing link necessary to begin building the semantic web for discovery of data and information about the data.

TABLE I. Definition of Simple Knowledge Organization System (SKOS) relationships used in one-to-one mappings.

Symbol	SKOS Relation	Definition
=	skos:exactMatch	indicates a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications.
≈	skos:closeMatch	sufficiently similar that they can be used interchangeably in some information retrieval applications.
>	skos:narrowMatch	used to state a hierarchical mapping link between two concepts.
<	skos:broadMatch	used to state a hierarchical mapping link between two concepts.
~	skos:relatedMatch	used to state an associative mapping link between two concepts

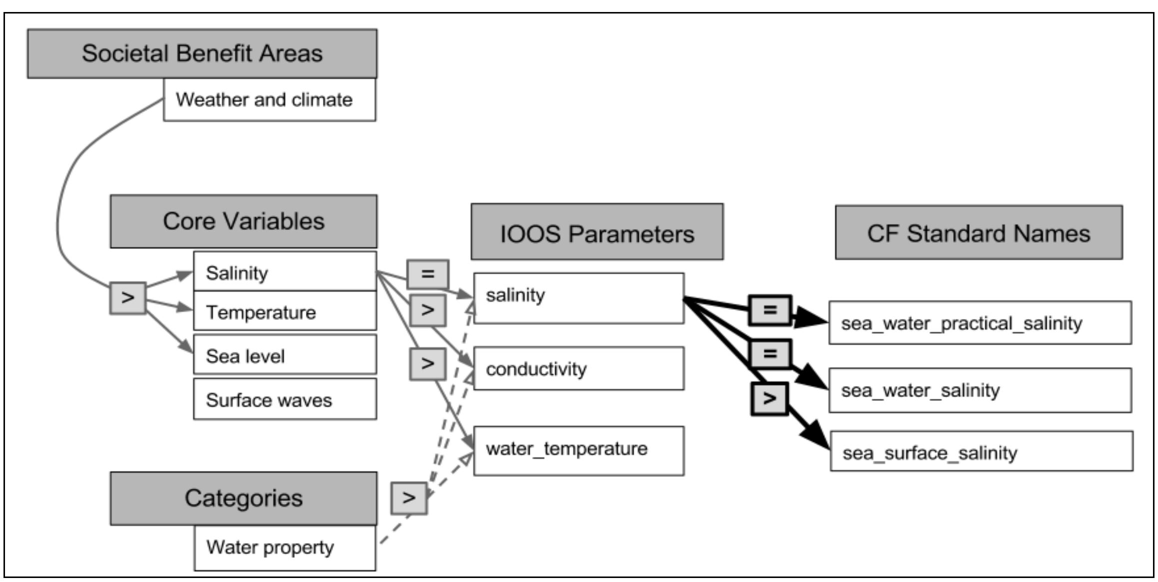


Fig. 2a. Concept mapping between salinity terms in CF Standard Names and IOOS Parameters provides the relationship between the two terms that have the same meaning but different labels (thick black arrows). Also, when mappings are provided in context of a knowledge hierarchy such as Core Variables that apply to societal concerns (gray arrows) or with alternative categories (dashed arrows), more meaning is brought to each concept.

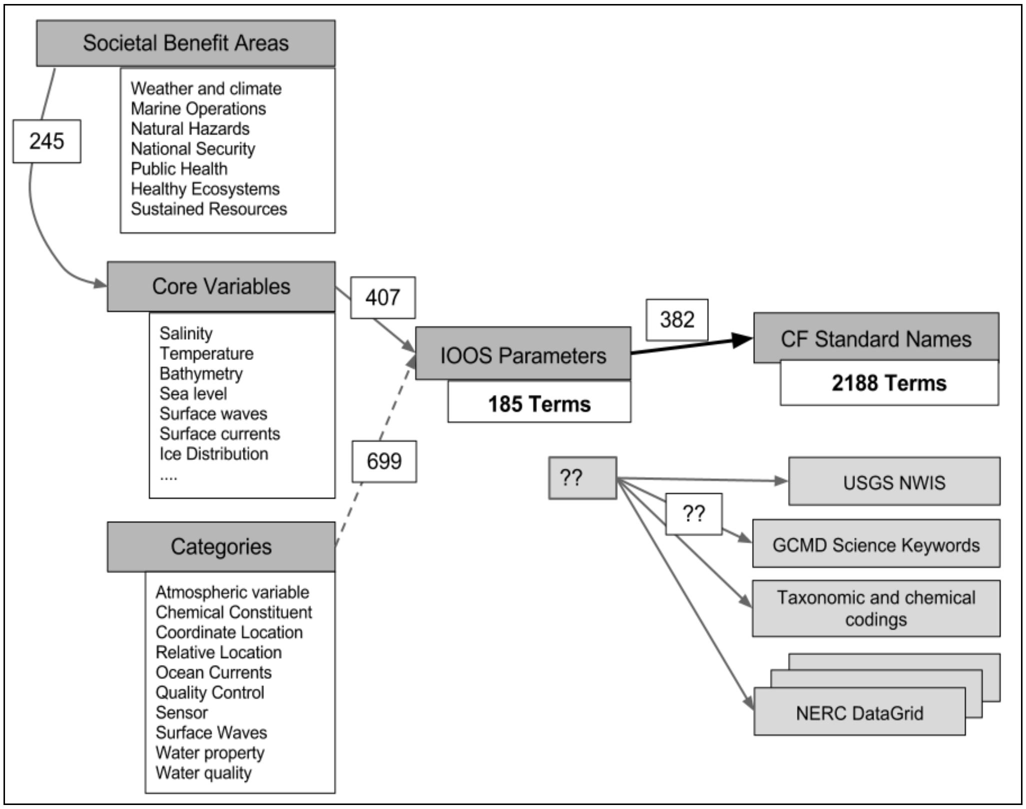


Figure 2b. The number of relationships mapped between IOOS Parameter Vocabulary and CF Standard Names (382) and hierarchical relationships to IOOS Parameter Vocabulary from Core Variables (407), Categories (699) and from Societal Areas to Core Variables (245) based on [8]. As more mappings are created between chosen vocabularies, many, many relations (??) and inferred data names [4] can be searched and extracted through the MMI-OOR [4].

VI. Term Discovery

Finally the goal of data discovery is to find data with specific labels within metadata and to find other data with related concepts. Term querying discovers related terms and related concepts that are the labels used for data naming. A query interface provided by ORR is based on the Simple Protocol and RDF Query Language (SPARQL, pronounced as "sparkle"), a query language built for the RDF file formats supported on ORR to perform such queries. Demonstration of increased semantics is provided through these types of queries.

One such question was: how many and which CF Standard Names are exactly the same or narrower in definition than terms in the IOOS Parameter Vocabulary. Figure 3 and 4 demonstrate how SPARQL can be used to address such

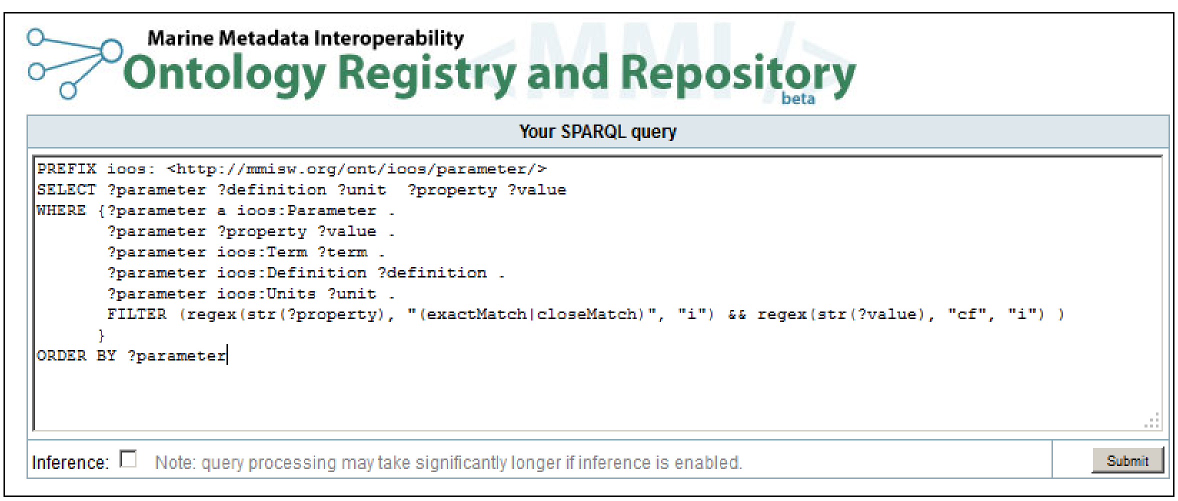


Fig. 3. Screen capture of MMI functional form for testing SPARQL queries. In this example, the query performs a search of all the IOOS Parameter terms to find the exactMach or closeMatch mappings to CF Standard Names. The output is displayed in Figure 4.

Copyright 2012 IEEE

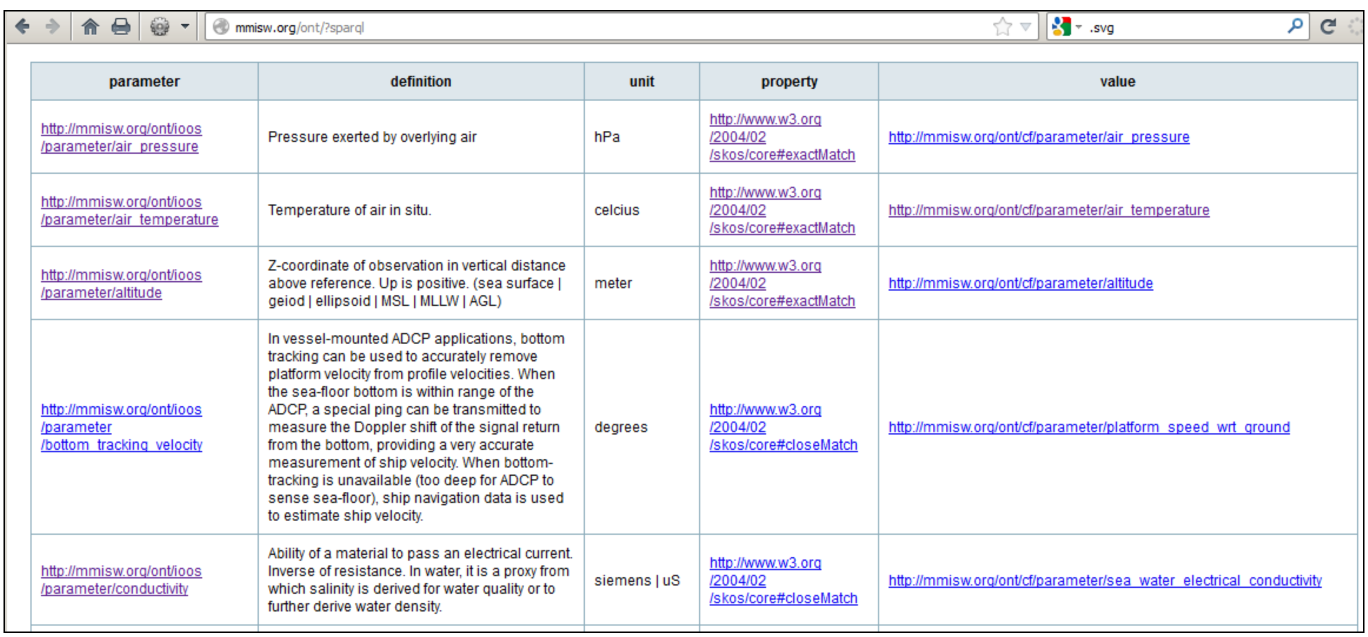
questions.

VII. Guidance, Maintenance, and Governance

The collaboration between the IOOS Program Office and RA partners on observed-property vocabularies and ontologies is part of a new, broader IOOS strategy that addresses vocabulary needs across core components of distributed observation systems. This strategy encompasses additional vocabularies for observation platforms; organization names, types and roles; and other characteristics. All such vocabularies are being defined on the MMI ORR, and existing community vocabularies and ontologies will be adopted as appropriate. All references to terms will be done via MMI term URL's (e.g., http://mmisw.org/ont/ioos/parameter/ammonia). MMI-hosted IOOS vocabularies and mappings are available at http://mmisw.org/ont/ioos.

An initial IOOS Data Management and Communications (DMAC) prototype project (2007-2010) focused on 7 "core variables" defined via an XML schema hosted by IOOS [16]. As the number of core variables and the scope of the data integration efforts were expanded, the limitation and burden of this approach became more apparent. As a result, CF standard names were adopted starting in 2010 as accepted observed-parameter terms. At the same time, the development of the regionally led IOOS Parameter Vocabulary v1.0 made clear the need for addressing parameters not currently found in the CF standard names. The IOOS Parameter Vocabulary will serve as a vocabulary of terms known to be important to RA's, some of which may not be present among CF terms. At the same time, the IOOS Parameter Vocabulary incorporates many CF terms that are widely used among RA's and other IOOS partners. The MMI-hosted vocabulary mapping between IOOS Parameter and CF terms will enable the automatic linking of a term in one vocabulary to an appropriate term on the other vocabulary if such a match exists. Currently, CF standard names only define scalar properties. Previously, IOOS defined three composite or aggregate properties made up of two or more scalar properties: winds, waves and currents. These composite properties will be supported as terms in the IOOS Parameter Vocabulary, and will be mapped to appropriate scalar properties.

Creating a robust useful vocabulary is a difficult task, and keeping that vocabulary relevant as the body of knowledge in any one field matures requires a long term commitment. The community behind the CF conventions has demonstrated the ability to create a good vocabulary through rigorous science based input from an international body of participants. Further, they have demonstrated a long term commitment (19 version updates over ten years). For these reasons, IOOS strongly encourages the use of CF standard name terms when a relevant term exists in that vocabulary. This preference exists to support interoperability with the wider community using CF. When the desired term is not found on the CF vocabulary, an IOOS Parameter Vocabulary term should be used. The IOOS Parameter Vocabulary will be managed and maintained by the IOOS Program Office and IOOS partners. IOOS and RA partners will set up a clear, simple and reasonably fast governance process for requesting and adding new terms to the IOOS Parameter Vocabulary, when needed. In parallel, IOOS will maintain an engagement with the CF community to request adding parameters that are not currently present. Nonetheless, it



parameter	definition	unit	property	value
http://mmisw.org/ont/ioos/parameter/air_pressure	Pressure exerted by overlying air	hPa	http://www.w3.org/2004/02/skos/core#exactMatch	http://mmisw.org/ont/cf/parameter/air_pressure
http://mmisw.org/ont/ioos/parameter/air_temperature	Temperature of air in situ.	celcius	http://www.w3.org/2004/02/skos/core#exactMatch	http://mmisw.org/ont/cf/parameter/air_temperature
http://mmisw.org/ont/ioos/parameter/altitude	Z-coordinate of observation in vertical distance above reference. Up is positive. (sea surface | geiod | ellipsoid | MSL | MLLW | AGL)	meter	http://www.w3.org/2004/02/skos/core#exactMatch	http://mmisw.org/ont/cf/parameter/altitude
http://mmisw.org/ont/ioos/parameter/bottom_tracking_velocity	In vessel-mounted ADCP applications, bottom tracking can be used to accurately remove platform velocity from profile velocities. When the sea-floor bottom is within range of the ADCP, a special ping can be transmitted to measure the Doppler shift of the signal return from the bottom, providing a very accurate measurement of ship velocity. When bottom-tracking is unavailable (too deep for ADCP to sense sea-floor), ship navigation data is used to estimate ship velocity.	degrees	http://www.w3.org/2004/02/skos/core#closeMatch	http://mmisw.org/ont/cf/parameter/platform_speed_wrt_ground
http://mmisw.org/ont/ioos/parameter/conductivity	Ability of a material to pass an electrical current. Inverse of resistance. In water, it is a proxy from which salinity is derived for water quality or to further derive water density.	siemens | uS	http://www.w3.org/2004/02/skos/core#closeMatch	http://mmisw.org/ont/cf/parameter/sea_water_electrical_conductivity

Fig. 4. Example HTML table output from SPARQL query provided by the MMI SPARQL endpoint. This output shows all IOOS Parameters (parameter column) found that have either exactMatch and closeMatch (property column) mappings to CF Standard Names (value). Definition and unit were extracted and selected to be shown along with each returned triple. The links can be followed to show the definitions and units for each term.

Copyright 2012 IEEE

is recognized that there may be terms that may not be appropriate for the CF community or may be best drawn from a different vocabulary. Such terms will be hosted and maintained in the IOOS Parameter Vocabulary, until a broader policy for observed parameters is developed.

While the CF standard names and the current IOOS Parameter Vocabulary address many of the variables common in the physical disciplines (physical oceanography and meteorology), there is a need to embrace the multidisciplinary nature of IOOS. Increasingly water quality, chemistry and biology data are being added to the IOOS data distribution network. As previously mentioned these fields have well developed vocabularies and IOOS must determine how best to support and interact with the vocabulary communities in these disciplines. This may include using existing vocabularies, such as Darwin Core, mapping to them through semantic technologies like MMI, and eventually collaborating with them to evolve the vocabularies based on new needs.

Conclusion

In recent years, with the sensor technological advancements in coastal ocean observing more and more physical, meteorological, chemical, biological and geological variables are being observed and made available. With the US Integrated Ocean Observing System (IOOS), the 11 Regional Associations currently engaged in aggregating and integrating such data collected and provided from various organizations, there is a need for guidance and tools for data naming conventions to facilitate the discovery, access, proper usage and archival of the data. By leveraging the previous and existing vocabulary efforts of MMI and IOOS RAs collaborative vocabulary work, we have outlined and described the path to the development, establishment of the IOOS Parameter Vocabulary V2.0. We have presented the results of an example mapping of IOOS Parameter Vocabulary with Climate Forecasting Terms and demonstrated the use of registering and mapping vocabularies with the tools such as MMI-ORR tools to facilitate discovery of data. We have recommended a community collaboration strategy between the IOOS Program Office and the Regional Associations for the oversight, maintenance and governance of the IOOS Parameter Vocabulary as we move forward. We also have echoed the need for broader community involvement (regional, national and global level) and engagement to provide input to the existing IOOS Parameter vocabulary framework towards the goal of achieving semantic interoperability.

References

[1] CF Metadata Group, 2012. “NetCDF Climate and Forecast (CF) Metadata Convention and CF Standard Name Table,” Version 19, 28 April 2012. http://cf-pcmdi.llnl.gov/documents/cf-standard-names. Accessed August 30, 2012.

[2] USGS 2011, “Technical Documentation of USGS Water-Quality Web Services,” http://qwwebservices.usgs.gov/technical-documentation.html. Accessed August 30, 2012.

[3] Olsen, L.M., G. Major, K. Shein, J. Scialdone, R. Vogel, S. Leicester, H. Weir, S. Ritz, T. Stevens, M. Meaux, C.Solomon, R. Bilodeau, M. Holland, T. Northcutt, R. A. Restrepo, 2007. “NASA/Global Change Master Directory (GCMD) Earth Science Keywords.” Version 6.0.0.0.0.

[4] R. Lowry, R. Cramer, M. Gutierrez, M. Hughes, K. Kleese van Dam, S. Kondapalli, S. Latham, B. Lawrence, K. O'Neil, A. Wolf, 2005. “Parameter Vocabularies in the NERC DataGrid” Project Proceedings of the UK e-Science All Hands Meeting 2005 19th - 22nd September, Nottingham UK. http://www.allhands.org.uk/2005/proceedings/papers/310.pdf. Accessed August 30, 2012.

[5] J. Graybeal, A. W. Isenor, C. Rueda. 2012. “Semantic mediation of vocabularies for ocean observing systems,” Computers & Geosciences 40: 120-131, doi:10.1016/j.cageo.2011.08.002

[6] Piasecki, M. and B. Beran. 2009. “A semantic annotation tool for hydrologic sciences,” Earth Science Informatics 2(3): 157-168, doi:10.1007/s12145-009-0031-x

[7] Yu, J., S. Cox, G. Walker, P. J. Box, and P. Sheahan. 2011. “Use of standard vocabulary services in validation of water resources data described in XML,” Earth Science Informatics 4(3): 125-137, doi:10.1007/s12145-011-0084-5

[8] Madin, J., S. Bowers, M. Schildhauer, S. Krivov, D. Pennington, and F. Villa. 2007. “An ontology for describing and synthesizing ecological observation data,” Ecological Informatics 2(3): 279-296, doi:10.1016/j.ecoinf.2007.05.004

[9] Hankin S. C., Blower J. D., Carval T., Casey K. S., Donlon C., Lauret O., Loubrieu T., Srinivasan A., Trinanes J., Godoy O., Mendelssohn R., Signell R., De La Beaujardiere J., Cornillon P., Blanc F., Rew R., Harlan J. (2010). “NETCDF-CF-OPENDAP : standards for ocean data interoperability and object lessons for community data standards processes,” Oceanobs 2009, Venice Convention Centre, 21-25 septembre 2009, Venice.

[10] Botts, M. 2007. OGC “Implementation Specification 07-000: OpenGIS Sensor Model Language (SensorML),” Open Geospatial Consortium, Wayland, MA, USA.

[11] Bröring, A., J. Echterhoff, S. Jirka, I. Simonis, T. Everding, C. Stasch, S. Liang, and R. Lemmens. 2011. “New generation sensor web enablement,” Sensors 11(3): 2652-2699, doi:10.3390/s110302652

[12] Rueda, C., Bermudez, L., Fredericks, J. (2009). “The MMI Ontology Registry and Repository: A Portal for Marine Metadata Interoperability,” Oceans 09 MTS/IEEE Biloxi.

[13] Neiswender, C., Isenor, A., Montgomery, E., Bermudez, L., Miller, S.P. 2011. "Vocabularies: Dictionaries, Ontologies, and More." In The MMI Guides: Navigating the World of Marine Metadata. http://marinemetadata.org/guides/vocabs. Accessed August 30, 2012.

[14] Fletcher, M., J. Cleary, J. Cothran, and D. Porter, 2008. Southeast “Atlantic Coastal Ocean Observation System (SEACOOS) Information Management: Evolution of a Distributed Community System,” MTS Journal, Volume 42, Number 3, Fall 2008 , pp. 28-34(7).

[15] Table 3-3, U.S. Integrated Ocean Observing System: A Blueprint for Full Capability Ver 1.0 (Nov 2010). http://www.ioos.gov/library/us_ioos_blueprint_ver1.pdf. Accessed August 28, 2012.

[16] US IOOS Program Office. Data Integration Framework (DIF) Final Assessment Report, November 22, 2010. http://www.ioos.gov/library/ioos_dif_assmnt_report_final.pdf.

Copyright 2012 IEEE

Copyright 2012 IEEE