

**NSF EarthCube Integrative Activities:
Cross-Domain Observational Metadata Environmental Sensing
Network (X-DOMES)**

September 2015 – June 2019

**Focusing on Sensor Metadata Creation; Data Quality Assessment;
Sensor Metadata Interoperability; Content Management**

using community adopted standards

OGC: SWE; SOS; SensorML;

W3C: Semantic Web; RDF; OWL



PI: Janet Fredericks, Woods Hole Oceanographic Institution

Co: Pis Krzysztof Janowicz (UCSB), Mike Botts, Carlos Rueda (MBARI), Felimon Gayanillo (TAMUCC) & consulting with John Graybeal (MMI)

***ASIC – Oakland, CA
September 2018***

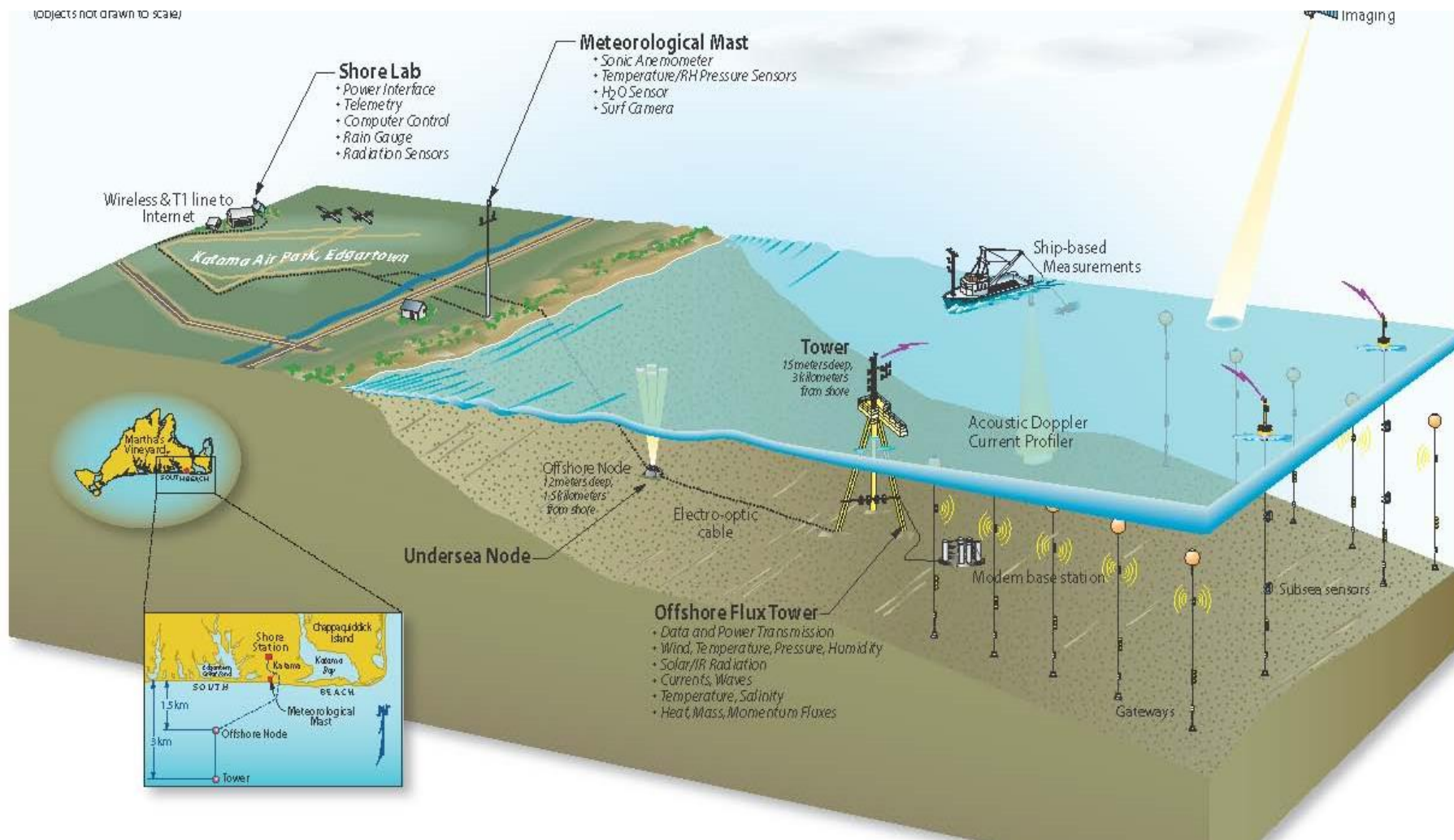


Martha's Vineyard Coastal Observatory (In Ocean Testbed for the CMR)

<http://www.whoi.edu/mvco>

Owned & Operated by WHOI: Heidi Sosik (Chief Sci)

Janet Fredericks, Jay Sisson, Hugh Popenoe and Steve Faluotico +++



Talk Overview

Discussion of **Agents** in sensor and data management
(knowledge, needs and roles)

Presentation of **Enabling Technologies**
(Standards and X-DOMES Implementation Model)

How **WE** can promote adoption and implementation



AGENTS – needs and knowledge



EARTH CUBE
TRANSFORMING GEOSCIENCES RESEARCH



Sensor Manufacturer

- * best understands instrument capabilities and limitations*
- * builds specific instruments and can describe and uniquely identify them, as built*
- * is interested in how and where their sensors are used*
- * needs to be able to target data sets that may be affected by manufacturing defects or limitations, when discovered*



Sensor Owner/Operator

Sensor Management:

Where IS that sensor? When was it deployed? Was it damaged?
When was it recovered? When was it serviced? Was it repaired? If so, when?

How was it was configured, which can change sensor capabilities.

This knowledge is needed to manage sensor networks and needs to be able to be associated with data to enable data quality assesment.



Data Manager

- * Needs to know the limitations of a sensor (operational limits, accuracy and precision) to accurately handle the data.*
- * Knows what quality control steps were taken (spike elimination, gap processing etc)*
- * needs to be able to associate data with the sensor used and the processing steps taken to enable quality assessment by potential users*



Data Facility

* Needs access to FAIR data ([doi:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)) Findable, Accessible, Interoperable and Reusable

-- to automate ingestion of data and metadata if access to standards-based systems are implemented ...

--- reducing loss of knowledge and potential errors

(does the data facility really understand what information is important for this sensor type and processing stream?)

* Can serve as a broker for discoverability and interoperability



Data Consumer

- * Needs to be able to determine if discovered data are suitable for use/reuse

 - What were the operational ranges;
 - what processes were used in quality control, sampling frequency, reporting frequency, etc.

- * Must be able to discover the sensor type used in data collection and be able to associate it with authoritative literature on sensor performance.

- * Would benefit from knowledge of the sensor ID to enable discovery of potential sensor issues



X-DOMES Model

(with an eye toward data quality assessment)

Use Role-Based content creation (SensorML) with links to registered terms (enabling Semantic Web functionality).

This project is funded by the National Science Foundation (NSF) as an EarthCube Integrative Activity. EarthCube is a collaboration between the Division of Advanced Cyberinfrastructure (ACI) and the Geosciences Directorate (GEO) of the US National Science Foundation (NSF). For official NSF EarthCube content, please see: <http://www.nsf.gov/geo/earthcube/>. (Award No. 1541008)



Community Adopted Standards:

Semantic Interoperability: Semantic Web Technologies (W3C)

Syntactic Interoperability: Sensor Web Enablement (OGC)

SensorML describes

*Input/Output/Capabilities/Characteristics/Systems ... how an
observation came to be
using machine-actionable encoding.*



WHY SEMANTIC WEB? The ORR creates the content upon which meaning is captured but the power of the W3C-Semantic Web can enable ontologies to be developed that can provide better results:

oceanTemperature SAME-AS seaWaterTemp

Map QC tests to be the same if they use the same processing and inputs
(rmoutlier ... spiketest)

and map QC flagging across conventions and values



AGENTS – roles in creating content



EARTH CUBE
TRANSFORMING GEOSCIENCES RESEARCH



Domain Specialists (mfgs/users/researchers)

- * create terms that are needed to describe content (sensor types, observable properties, capabilities; processes; characteristics) and register in the ORR
- * create standards-based profiles (RELAXNG) to guide manufacturers in the creation of common content

Sensor Manufacturers

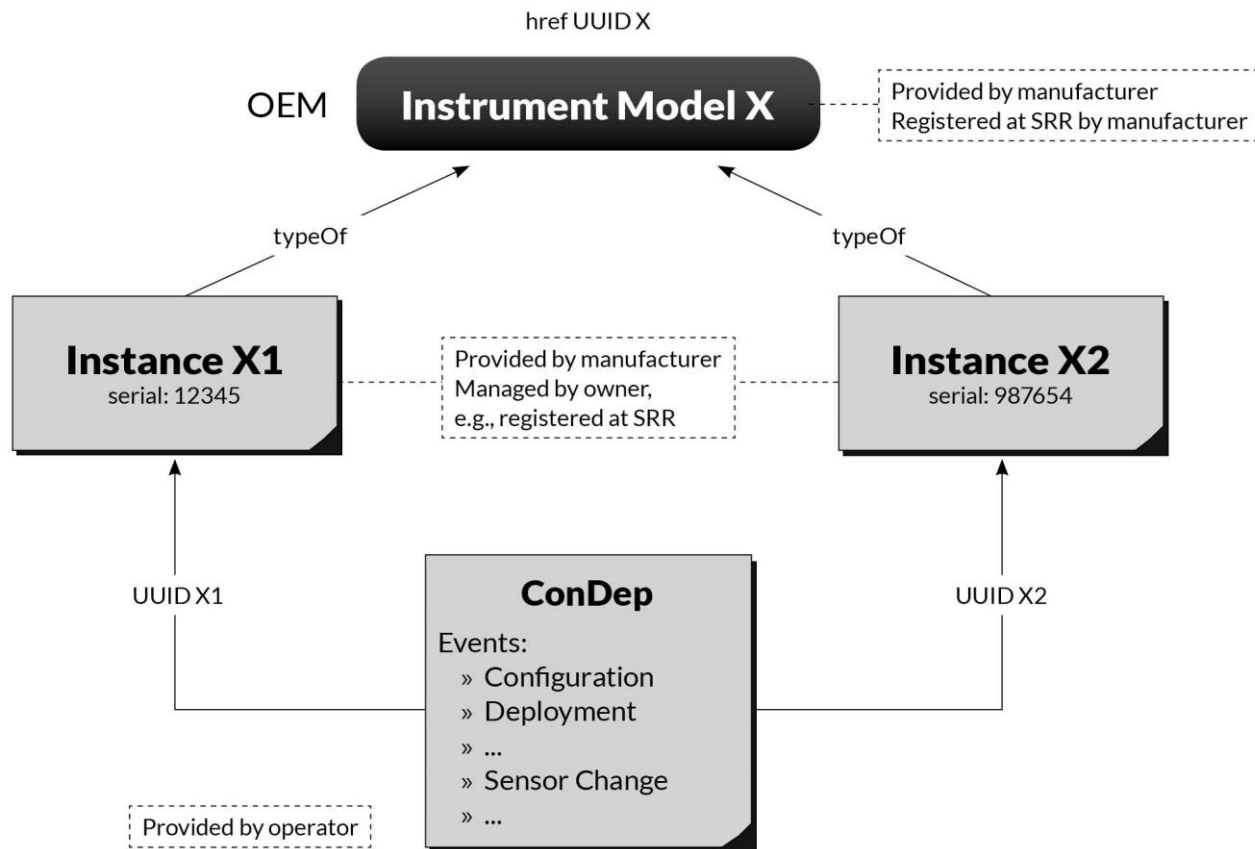
MODEL

- * create the SensorML description of the Sensor Model (OEM) - using profiles
- * register the OEM in the community registry/repository SRR

INSTANCE (as built)

- * create SensorML description of as-built sensor (Instance), including
 - definition a unique ID (URN)
 - a link to the OEM as a typeOf ... enabling inheritance
- * communicate the Instance SensorML to the sensor owner





Sensor Owners/Field Operators

- * create or have access to a community registry of SensorML Instance files and registers its Instance file
- * with each action that affects data (configuration/deployment/calibration), the Instance file is referenced. This also associates it to a sensor model (OEM) and its associated sensor type.

Data Managers

- * associates data to the Instance file
- * ULTIMATELY (beyond the scope of our current funding) ... the processing steps can be described using SensorML and SOS can reference these steps and the knowledge about the sensor (DescribeSensor) and the SOS associates a request for data (GetObservations) through the GetCapabilities ... <http://q2o.who.edu> -- NOAA funded project showing how this would work!



Create User-Friendly tools that build upon these rules to enable non-experts to provide consistent, complete content in describing how an observation came to be.



EARTH CUBE
TRANSFORMING GEOSCIENCES RESEARCH



X-DOMES Tools (<http://xdomes.org>)

The **Ontology Registry & Repository (ORR)** – Semantic Web based content is created by simple methods ... thus one has a persistent, authoritative reference to each term.

The **online SensorML Editor/Viewer** – enabling creation of rule-based (RELAXNG) profiles to guide in creating consistent, accurate SensorML files. Provides SPARQL query to the ORR to link to registered terms. **

The **SensorML Registry & Repository (SRR)** – to register SensorML document to create persistent, authoritative, referenceable content that is machine-actionable.

The ORR and SRR can be deployed by any community for its own use (e.g., the ESIP COR)



EARTH CUBE
TRANSFORMING GEOSCIENCES RESEARCH



UniqueID: urn:teledyne_rdi:workhorse_1200 ...

The Monitor is Teledyne RD Instruments most popular direct-reading Acoustic Doppler Current Profiler (ADCP). The unit is typically bottom frame-mounted and hard-wired to shore to provide real-time monitoring of coastal currents. The Monitor's high data accuracy and reliability make it a favorite for deployments in highvolume traffic areas such as ports and harbors, where the data is often integrated into a Vessel Traffic Monitoring System. In fact, the Monitor has been selected for most major port programs undertaken in the United States.

Keywords : oceanography waves currents acoustic Doppler profiler ...

▼ Identification

Long Name	Teledyne RDI Workhorse Monitor 1200 ...
Short Name	RDI Workhorse ...
Manufacturer	Teledyne RDInstruments ...
Model Number	Workhorse 1200 ...

▼ Classification

Intended Application	Ocean Waves ...
Intended Application	Ocean Currents ...
Sensor Type	Acoustic Doppler ...
Sensor Type	profiler ...

▼ capabilities (Doppler Properties)

Beam Frequency	1228800 Hz ...
Bandwidth	307200 Hz ...
Beam Angle	20 deg ...
Depth Cell	5 to 1600 cm ...
Expected Velocity Vector Accuracy	2 cm/s ...
Expected Direction Accuracy	3 deg ...
Velocity Vector Resolution	.1 cm/s ...
Direction Resolution	.1 deg ...
Speed Range	-5 to +5 m/s ...

▼ Contacts

Adoption and Implementation – We Need YOU!

- **Domain Specialists needed** to create a list of sensor types, observable properties and derived products with associated processes. You can help us create the profiles (exercising our editor) by defining what is needed for each sensor type. The group would ‘own’ the terms and the RELAXNG profiles. Using ORR and online SensorML editor.
- **Sensor Manufacturers needed** as Domain Specialists AND to test and assess the usability of the editor AND to actually create the OEM files as a first step. Using the editor and the SRR, to register, own, maintain the content.

SUSTAINING A Community to Work with Manufacturers

- Developing a community within **the ESIP Federation EnviroSensing Cluster** http://wiki.esipfed.org/index.php/EnviroSensing_Cluster and the ESIP Community Ontology Registry (COR) <http://cor.esipfed.org>
- Working with US National Science Foundation EarthCube Community <https://earthcube.org/group/x-domes>
- Aligning with the activities of the **W3C SSN/SOSA** (<http://w3c.github.io/sdw/ssn>)
- Working with the Marine Metadata Interoperability Project and utilizing its Ontology Repository & Registry: <http://marinemetadata.org> and <http://mmisw.org/orr>

janet@whoi.edu

Sensor Technical Descriptions out of human-readable content



Discoverable and machine-actionable content!

THANK YOU

janet@who.edu