ODM2: Developing a Community Information Model and Supporting Software to Extend Interoperability of Sensor and Sample Based Earth Observations

Jeffery S. Horsburgh

Anthony Aufdenkampe, Kerstin Lehnert, Emilio Mayorga, David Tarboton, Ilya Zaslavsky, David Valentine

Support: EAR 1224638
What is an “Information Model?”

• Representation of concepts, relationships, constraints, rules, and operations that specify the semantics of data for a domain of discourse
  – Defines the domain’s entity types and their attributes, relationships, and allowed operations on the entities
  – In a relational database implementation, entities become tables and their attributes become table columns

• Sharable, stable, and organized structure of information requirements for a domain context
  – Without constraining how that description is mapped to an actual implementation in software
  – There may be many physical implementations – e.g., relational database, XML schema, etc.

• Critically important to the effectiveness and interoperability of the cyberinfrastructure
An organization operates a network of monitoring sites
At each monitoring site a number of variables are measured
For each variable there is a time series of data values
Each data series is made up of individual, time-indexed values that are each characterized by location (where the observation was collected), time (when the observation was collected), and variable (what the observation represents)
Some Limitations of ODM 1.1.1

• ODM supports only point-based observations
• ODM doesn’t support sample-based data well
• Not all of the structure of ODM is required for each type of data
• Versioning and provenance of data can be difficult in ODM
• The central DataValues table doesn’t support every functional use case (e.g., metadata catalog)
Another Information Model Example

Overarching Goals for ODM2 Project

• Create an *information model* that is *integrative* and *extensible*
  – Accommodating a wide range of observational data
  – Aimed at achieving interoperability across multiple disciplines and systems that support publication of earth observations

• Allow a diverse range of geoscience observations to be consistently shared, discovered, accessed, and interpreted
Observations Information Model

- Samples Extension
- Sensor Extension
- Generic Extension
- Feature Model

Observations Core

Common Semantics for Earth Observations

- CUAHSI HIS
- EarthChem
- CZOData
- IOOS

Domain Cyberinfrastructures
ODM2 Functional Use Cases

- Archival Encoding
- Information Model
- Storage Encoding
- Catalog Encoding
- XML Schema Encoding
- Web Service Interface
ODM2 - Core

- Observations
- Features
- People
- Organizations
- Methods
- Variables
- QCLevels
- Units
ODM2 Observations

Open Geospatial Consortium Observations & Measurements – 2 Types of observations

1. Observations whose result is constant
2. Observations whose result varies with space or time

• Separate the concept of an “observation” and its “result”
ODM2 Observations

Observations whose result is constant
- Measurements
- Category observations
- Count observations
- Truth observations
- Temporal observations
- Geometry observations

Observations whose result varies in space and/or time
- Time series coverage
- Point coverage
- Profile coverage
ODM2 Sensor Extension

**Equipment**
- Model
- Serial number
- Owner
- Vendor
- Manufacturer
- Service history

**Site visits**
- Location
- Date
- People
- Conditions

**Field activities**
- Activity type
- Description
- Date

**Calibrations**
- Method
- Standard

**Deployments**
- Deployment type
- Description
- Dates
- Offsets

**Time Series Observations**
![Graph showing time series observations](image)
Managing Sensor Infrastructure: A Sensor Extension for ODM2
Amber Spackman Jones, Jeffery S. Horsburgh, Juan Caraballo, Maurier Ramirez
Utah Water Research Laboratory, Utah State University

**Background:** The Observations Data Model (ODM) is a relational data model for storage and management of environmental observations data designed to capture consistent descriptions for autonomous Interpretation among users. ODM 1.1 was developed for publication of point-based hydrologic observations using the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Hydrologic Information System (HIS). Because of its focus on data publication, ODM 1.1 does not provide all of the functionality needed to fully describe the data collection and management process for sensor and sample-based observations.

**Motivation:** Research sites conducting long term monitoring using in-situ sensors need the functionality to track equipment, deployments, calibrations, and other events related to the maintenance and to link this information to the observational data that they are collecting.

**ODM2 Core:** ODM2 consists of a core set of entities common to all types of observational data to describe location, observed variable, and time.

**ODM2 Extensions:** Additional metadata can be provided via extensions to the core to support particular types of observations.

**Extensions Under Development:**
- Sensors: describe monitoring equipment, deployments, and in situ time series.
- Samples: capture sample hierarchy and analytical methods.
- Features: detail the physical feature(s) from which an observation is made.
- Annotations: provide qualifiers or comments to variables or observations.
- Provenance: describe data versioning and processing.

**Sensors Extension:** The Sensors Extension permits information about individual sensors, deployments, and the associated time series to be recorded and managed.

**Table Descriptions:**
- Equipment: contains attributes of individual sensors and other pieces of field equipment while EquipmentModels describes attributes of equipment of the same type.
- Deployments: provides information about the sensor deployment in space and time.
- Site/Activities/Calibrations: record activities of field crews and sensor calibrations.
- Vendors: provides information on equipment manufacturers and suppliers and FactoryServiceEvents records sensor servicing.
- SensorOutputVariables: information on a specific variable being measured by a sensor, linked to the specific deployment through Deployment/MeasurementVariables.
- DataCollectionPrograms: tables allow for tracking of programs, files, and each variable measured.
- ControlledVocabularies: used for types of equipment, deployments, sites, vendors, activities, and calibration standards.

**Web Interface:** Field technicians and other researchers can use the web interface for the Sensors Extension to add and edit sensors, sites, and field activities such as calibrations and deployments. The interface can be used to address questions such as what sensors are deployed at a site, the deployment history or factory service history of a sensor, and the history of field activities performed at a site.

**Implementation:** The iUTAH (Innovative Urban Transitions and Arid Region Hydro- Sustainability) network of aquatic and terrestrial sensors is being used as a test case for the Sensors Extension. The ODM2 Sensors Extension and associated tools will be useful for similar large scale and long term monitoring networks. See data.iutahsource.org.
ODM2 Samples Extension

Sample collection

- Sample type
- Location / sampled feature
- Collection method
- Date
- People
- Preservation
- Identifiers
- Sampled medium

Sample splitting

- Sample hierarchy
- Parent/child relationships

Sample processing

- Preparation method
- Processing method
- People

Sample analysis

- Analytical method
- Instrument
- People
- Date

Sample analyses result in observations
ODM2 Samples Extension
ODM2 Features Extension
Linking Observations to the Geo-Environment

Geographic Features

<table>
<thead>
<tr>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

Sampling Feature

1. Stream Gage

Feature of Interest

<table>
<thead>
<tr>
<th>Feature of Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
Provenance and Annotations Extensions

- Better support for storing provenance of observational data
General Extensibility

• Ability to add extension properties to any entity
Impact

• Better support for sample based data in CUAHSI HIS, HydroShare, and related tools
• Support for and integration of sensor and sample-based data in the Critical Zone Observatory Integrated Data Management System (CZOData)
• Potential adoption of ODM within EarthChem
• Providing context and guidance for future WaterML 2.0 development (water quality samples)
Questions?