

Appendix G: Oregon Water Data Portal Onboarding

Self-Service Integration Documentation Package

Compiled by the Internet of Water Consortium, Center for Geospatial Solutions, June 29, 2025

Appendix G: Oregon Water Data Portal Onboarding	1
Self-Service Integration Documentation Package	1
Part 1: Executive Summary	4
OWDP Pilot and Live Resources	4
Part 2: Data Verification Process Guide	5
2.1 PRE-SUBMISSION QUALIFICATION	5
Self-Assessment Questionnaire	5
Data Suitability Criteria Checklist	6
Exit Points for Inappropriate Data	6
2.2 EXISTING DATA SEARCH METHODOLOGY	7
Step-by-Step OWDP Search Process	7
Search Term Optimization Guide	7
Advanced Search Techniques	7
Search Result Interpretation	8
2.3 DUPLICATE DETECTION WORKFLOW	8
Comparison Criteria for Similar Datasets	9
2.4 CONFLICT RESOLUTION PROCEDURES	9
When Similar Data Exists (Decision Matrix)	9
Collaboration with Existing Data Providers	9
Data Integration vs. Separate Submission	10
Escalation Procedures	10
2.5 VERIFICATION CHECKLIST AND TOOLS	10
Quality Assessment Worksheet	11
Final Approval Criteria	11
2.6 COMMON SCENARIOS AND TROUBLESHOOTING	11
Part 3: Data Type Classification Guide	12
3.1 CLASSIFICATION DECISION SYSTEM	12
Classification Decision Tree Structure (Description)	13
Preliminary Classification Assessment	13
3.2 DETAILED DATA TYPE CATEGORIES	14
Water Quality Monitoring Data	14

Hydrological Data	15
Groundwater Data	15
Stream Flow and Surface Water	16
Weather and Climate Data	16
Infrastructure and Management Data	16
3.3 CLASSIFICATION TOOLS AND GUIDES	17
Data Type Quick Reference	17
3.4 EDGE CASES AND COMPLEX DATA	17
Part 4: Dataset Metadata Application Guide	18
4.1 OREGON OPEN DATA STANDARD COMPLIANCE	18
Complete Field Mapping	18
Required vs. Recommended vs. Optional Fields	19
Compliance Verification Procedures	19
4.2 FIELD-BY-FIELD METADATA GUIDE	19
Site Data Sheet Fields	19
1. siteID	19
2. uri	20
3. name	20
4. organization	20
5. featureOfInterest	21
Metadata Sheet Fields	21
1. metadataID	21
2. siteID	22
3. observedPropertyName	22
4. unitName	22
5. analyticalMethodName	22
4.3 METADATA QUALITY STANDARDS	23
Completeness Scoring System	23
Quality Assessment Criteria	23
4.4 DATA TYPE-SPECIFIC METADATA	24
4.5 BULK AND AUTOMATED METADATA	24
4.6 METADATA TOOLS AND TEMPLATES	24
4.7 TROUBLESHOOTING AND SUPPORT	25
Part 5: Data Formatting and Template Guide	25
5.1 TEMPLATE-BASED SUBMISSION SYSTEM	26
Template Selection Decision Tree	26

IOW-OWDP-TABULAR-TEMPLATE.xlsx Detailed Breakdown	26
Required vs. Optional Columns	27
5.2 EXCEL/CSV FORMATTING REQUIREMENTS	27
5.3 DATA PREPARATION PROCEDURES	27
5.4 TEMPLATE COMPLETION GUIDE	28
Sheet-by-Sheet Instructions	28
Cell Format and Formula Handling	28
5.5 FILE ORGANIZATION AND NAMING	29
5.6 UPLOAD PROCESS AND VALIDATION	29
5.7 QUALITY CONTROL CHECKLIST	30
5.8 TROUBLESHOOTING GUIDE	30
Part 6: API Integration Documentation	31
6.1 API OVERVIEW AND ARCHITECTURE	31
OWDP API Description	31
Endpoint: POST /v1/Things	32
Endpoint: POST /v1/Datastreams	33
Endpoint: POST /v1/Observations	33
6.2 COMPLETE CODE EXAMPLE	34
Python (requests library)	34
6.3 TESTING AND VALIDATION PROCEDURES	36
Unit Testing	36
6.4 PRODUCTION DEPLOYMENT CHECKLIST	37
Part 7: FAQs	38
Category: Getting Started	38
Category: Data Preparation	39
Category: Submission Process	39
Contact Information	40

Part 1: Executive Summary

The Oregon Water Data Portal (OWDP) is a critical resource for aggregating and sharing water data from diverse entities across the state. This documentation package provides a comprehensive, self-service framework for data providers to integrate their water-related data into the portal efficiently and accurately. The primary objective is to empower external stakeholders, including state agencies, local water authorities and research institutions, to onboard their data independently through either template-based or API-driven methods.

This initiative addresses the key challenge of standardizing disparate datasets to ensure interoperability and high data quality. By providing clear guidelines on data verification, classification, metadata application, and formatting, this documentation ensures all contributions align with state and federal standards. The expected outcome is a streamlined, scalable data pipeline that enriches the OWDP's repository with minimal direct intervention from the development team, fostering a more robust and collaborative water data ecosystem for Oregon.

OWDP Pilot and Live Resources

The current OWDP Pilot integration provides a live, functioning example of the data portal's capabilities. Data has been successfully integrated from several key sources, including the **Oregon Water Resources Department's** [Near-Real Time Hydrographics Data](#), and [Groundwater Information System](#), Oregon's [AWQMS database](#), and **XLSX templates** uploaded directly to the portal.

You can access the live pilot system and related resources at the following links:

- **User Interface (UI):** <https://owdp-pilot.internetofwater.app>
- **SensorThings API (v1.1):** <https://owdp-pilot.internetofwater.app/FROST-Server/v1.1>
- **OGC-API Endpoint:** <https://owdp-pilot.internetofwater.app/oapi/>
- **ETL Integration Code:** <https://github.com/cgs-earth/OregonWaterDataPortal-ETL>

These resources provide a transparent view into the portal's architecture and can serve as a valuable reference for developers and data managers.

The 3 Steps of Data Submission:

1. **Verify:** Check if your data already exists. (Part 2)
2. **Classify & Format:** Put your data into the standard template. (Parts 3, 4, 5)
3. **Submit:** Upload the template file via the portal's upload form or use the API. (Parts 5, 6)

Part 2: Data Verification Process Guide

Objective: To ensure that data submitted to the Oregon Water Data Portal (OWDP) is unique, valuable, and correctly formatted, preventing duplication and maintaining high data quality.

Audience: All potential data providers, including technical data managers, non-technical staff, third-party vendors, and researchers.

2.1 PRE-SUBMISSION QUALIFICATION

Before investing time in formatting and preparation, it's crucial to determine if your data is a good fit for the OWDP. This section provides a quick self-assessment to qualify your dataset.

Self-Assessment Questionnaire

Answer the following questions about your dataset. If you answer "No" to any of questions 1-5, your data may not be suitable for the OWDP at this time.

1. **Is your data related to water in Oregon?** (e.g., hydrology, water quality, water rights, infrastructure) Yes / No
2. **Do you have the legal right to publish this data?** (i.e., you are the owner or have explicit permission) Yes / No
3. **Is the data observational or modeled?** (i.e., it represents measurements, samples, or model outputs, not just plans or reports) Yes / No
4. **Can you provide specific location information for your data?** (e.g., latitude/longitude, site ID, well number) Yes / No
5. **Does your dataset contain time-series or discrete sample data?** (i.e., values associated with a specific date/time) Yes / No
6. **Is your data already available in a national repository like the Water Quality Portal (WQP) or USGS NWIS?** Yes / No
7. **Have you identified the primary topic for your data within the OWDP Topic Hierarchy (see Section 5.2)?** Yes / No
8. **Do you have methodological metadata describing how the data was collected and analyzed?** Yes / No

Data Suitability Criteria Checklist

Review this checklist. If your data does not meet the "Required" criteria, it cannot be accepted.

Criteria	Status	Details
Relevance to Oregon Water	Required	Data must pertain to water resources, management, or infrastructure within the state of Oregon.
Geospatial Reference	Required	Each observation must be tied to a specific, verifiable location (e.g., coordinates, site URI).
Temporal Reference	Required	Each observation must have a timestamp (date and time).
Observational Values	Required	The dataset must contain measured or modeled values (e.g., pH, streamflow in cfs, contaminant concentration).
Data Ownership	Required	You must be the originator or designated steward of the data with rights to distribute it publicly.
Structured Format	Preferred	Data is in a machine-readable, tabular format (e.g., CSV, XLSX). Unstructured data (e.g., PDFs, Word documents) is not suitable for direct integration.

Exit Points for Inappropriate Data

If you determined your data is not a good fit, here are some potential next steps:

- **Narrative Reports/Plans:** Consider publishing these on your organization's website and providing a link to the OWDP team for potential inclusion in the "Resources" section.
- **Data from Outside Oregon:** We recommend seeking a more appropriate state or regional data portal.
- **No Geospatial Reference:** If locations cannot be identified, the data cannot be integrated. Efforts should be made to georeference the data before proceeding.

2.2 EXISTING DATA SEARCH METHODOLOGY

A thorough search of the OWDP is essential to avoid duplicating existing datasets.

Step-by-Step OWDP Search Process

1. **Navigate to the OWDP Catalog:** Access the main data portal at <https://www.oregonwaterdata.org/> and find the "Catalog" or "Search" page.
2. **Start with Broad Keywords:** Begin with general terms related to your data's topic and location.
 - *Example:* Willamette River water quality
3. **Use Filters to Refine:** Use the available filters on the search results page. Prioritize filtering by:
 - **Topic Hierarchy:** (e.g., Water Quality > Physical/Chemical)
 - **Geographic Area:** Use the map interface to draw a bounding box around your area of interest.
 - **Organization:** Filter by your own organization or known partner organizations.
4. **Inspect Potential Matches:** Click on any datasets that appear relevant. Pay close attention to the dataset's description, date range, and parameters measured.
5. **Document Your Search:** Keep a simple log of the search terms you used and the results you found.

Search Term Optimization Guide

If you have...	Try searching for...	Example
A Site ID	The full Site ID and partial IDs.	USGS-14211720, 14211720
A Waterbody Name	The official name and common local names.	Tualatin River, Tualatin
A Specific Parameter	The parameter name and common abbreviations.	Dissolved Oxygen, DO
An Organization Name	The full name and acronyms.	Oregon Department of Environmental Quality, DEQ
A Geoconnex URI	The full URI.	https://geoconnex.us/ref/mainsteams/1431511

Advanced Search Techniques

- **Boolean Operators:** Use AND, OR, NOT to combine terms. (e.g., Clackamas River AND temperature NOT turbidity)
- **Wildcard Searches:** Use an asterisk * to find variations of a word. (e.g., hydro* would find hydrology, hydrographic)
- **Phrase Searching:** Enclose phrases in double quotes " for exact matches. (e.g., "North Santiam River")

Search Result Interpretation

- **Look at Metadata:** The metadata is key. Check the **organization**, **siteID**, **featureOfInterest**, and **date range**.
- **Check Parameters:** Does the existing dataset measure the same parameters (observedPropertyName) as yours?
- **Review Data Quality:** Note any data quality flags (dataStatus, dataQuality) that might differentiate your dataset.

2.3 DUPLICATE DETECTION WORKFLOW

This workflow helps you determine if a similar dataset you found is a true duplicate of yours.

1. **Start:** "Potential Match Found in OWDP"
2. **Decision 1:** "Is the organization name the same as yours?"
 - **Yes:** Proceed to Decision 2.
 - **No:** Proceed to Decision 3.
3. **Decision 2 (Same Org):** "Is the siteID and date range identical?"
 - **Yes:** This is likely a **Duplicate**. Do not upload. Contact your organization's data manager.
 - **No:** Proceed to Decision 4.
4. **Decision 3 (Different Org):** "Does the dataset monitor the exact same geographic feature (featureOfInterest URI or lat/lon)?"
 - **Yes:** Proceed to Decision 4.
 - **No:** This is likely **Not a Duplicate**. The data represents a different location. Proceed with submission.
5. **Decision 4 (Same Feature/Site):** "Do the measured parameters (observedPropertyName) and time periods overlap?"
 - **Yes:** This is a **Potential Conflict**. Proceed to Section 2.4 Conflict Resolution."
 - **No:** This is likely **Complementary Data**. Proceed with submission, but note the existing dataset in your comments."

Comparison Criteria for Similar Datasets

Use the IOW-OWDP-TABULAR-TEMPLATE.xlsx as your reference.

Field	What to Compare	Importance
organization	The entity responsible for data collection.	High
siteID / uri	The unique identifier for the monitoring location.	High
Geometry	The latitude and longitude coordinates.	High
featureOfInterest	The URI of the specific feature being monitored.	High
sample Timestamp	The start and end dates of the data collection period.	High
observedPropertyName	The specific parameters measured (e.g., pH, Turbidity).	Medium
samplingMethodName	The method used to collect the data.	Low

2.4 CONFLICT RESOLUTION PROCEDURES

This section guides you on how to proceed when you find a dataset that is not an identical duplicate but has significant overlap.

When Similar Data Exists (Decision Matrix)

Overlap Type	Action
Temporal Overlap (Same location, same parameters, different times)	Integrate. Your data represents a new time period for an existing monitoring effort. Proceed with submission.
Parameter Overlap (Same location, same time, different parameters)	Integrate. Your data adds new measured parameters to an existing site. Proceed with submission.
Methodological Overlap (Same location, time, parameters, but different method)	Collaborate. Contact the existing data provider to discuss linking to their data in each other's metadata..

Collaboration with Existing Data Providers

1. Find the contact information for the organization listed in the existing dataset's metadata.
2. Draft a professional email explaining your dataset and the overlap you've identified.
3. Propose a short meeting to discuss how your datasets can complement each other.

4. Be prepared to share your metadata and a sample of your data.

Data Integration vs. Separate Submission

- **Integration is Preferred:** If your data fills gaps (temporal or parameter) at an existing site, it should be integrated. This means using the same siteID and featureOfInterest URI.
- **Separate Submission:** Only submit as a completely new dataset if it represents a new monitoring location or a fundamentally different effort that cannot be reconciled with existing data.

Escalation Procedures

If you cannot resolve a conflict through collaboration or if you identify a critical data conflict (e.g., two different values for the same parameter, place, and time), you must escalate the issue.

1. Contact OWDP support at OWDP@DEQ.oregon.gov
2. Provide the siteID or uri of both datasets.
3. Clearly describe the conflict.
4. Attach your completed Verification Checklist (Section 2.5).
5. An OWDP data curator will investigate and provide guidance within 5-7 business days.

2.5 VERIFICATION CHECKLIST AND TOOLS

Check #	Task	Notes
1	Completed Pre-Submission Self-Assessment (2.1).	
2	Confirmed data is suitable for OWDP (2.1).	
3	Performed a thorough search of the OWDP Catalog (2.2).	<i>Search terms used:</i>
4	Documented all potential matches found.	<i>Matching Site IDs:</i>
5	Completed the Duplicate Detection Workflow (2.3).	<i>Result:</i>
7	All potential conflicts identified and resolved (2.4).	<i>Resolution path:</i>
8	Contacted existing data providers if necessary.	<i>Contacted:</i>
9	Confirmed the final submission path (Integrate vs. Separate).	
Final Step	Ready to Proceed to Data Formatting	

Quality Assessment Worksheet

This is a self-evaluation to help you think about your data quality.

- **Completeness:** Are there missing values in your key fields (siteID, Timestamp, dataValue)?
- **Accuracy:** Have you performed QA/QC on your data? Can you provide accuracy bounds?
- **Consistency:** Are the units (unitName) consistent for each parameter throughout the dataset?
- **Documentation:** Is your methodological metadata (samplingMethodName, analyticalMethodName) complete?

Final Approval Criteria

You are ready to proceed to the next stage of data submission when you can confidently state:

1. My data is unique and does not duplicate an existing OWDP dataset.
2. I have resolved any potential overlaps or conflicts with existing data providers or OWDP staff.
3. I have all the necessary metadata required by the [IOW-OWDP-TABULAR-TEMPLATE.xlsx](#).

2.6 COMMON SCENARIOS AND TROUBLESHOOTING

Here are realistic examples to help you navigate the verification process.

1. **Scenario:** I have water quality data from a well. The well has a state-assigned ID, but I can't find it in the OWDP.
 - **Solution:** Your well is likely a new monitoring site. Search by the well's coordinates to be certain no one else is monitoring it under a different ID. If no match is found, proceed with submitting it as a new site.
2. **Scenario:** I found a dataset from my own agency at the same location, but it only contains data up to 2020. My data is from 2021-2023.
 - **Solution:** This is a perfect case for **integration**. Your data is a temporal extension. Use the exact same siteID and featureOfInterest as the existing dataset when you format your template.
3. **Scenario:** Another organization is monitoring the same river reach, but their site is 2 miles upstream.
 - **Solution:** This is **not a duplicate**. It is a separate monitoring location. Proceed with your submission. You may want to use the same featureOfInterest URI for the river reach itself to create a link between the datasets.
4. **Scenario:** My search for DEQ returns too many results.

- **Solution:** Use advanced search techniques. Combine DEQ with a geographic area on the map and a more specific topic, like Infrastructure > Wastewater.
- 5. **Scenario:** The upload_form.jpg shows a simple upload. Why is this process so detailed?
 - **Solution:** The form is the final step. This verification process ensures that the data being uploaded via that form is high-quality and doesn't corrupt or duplicate existing portal data.
- 6. **Scenario:** I have data on fish populations. Is that "water data"?
 - **Solution:** Yes. According to the Water Data Standards Guidebook, this falls under the Water Quality > Ecological topic hierarchy.
- 7. **Scenario:** My data is in a proprietary format from my monitoring equipment.
 - **Solution:** You must first export this data into a tabular format (like CSV) and then transfer it to the IOW-OWDP-TABULAR-TEMPLATE.xlsx. The portal cannot process proprietary formats directly.
- 8. **Scenario:** I am a researcher and my data is for a one-time study.
 - **Solution:** Your data is still valuable. The key is to provide very detailed metadata about your study's methodology so others can understand the context. It will be submitted as a new dataset.
- 9. **Scenario:** I found a dataset with the same site and parameters, but the values are slightly different.
 - **Solution:** This is a **potential conflict**. It could be due to different calibration, methods, or a simple error. Contact the original data provider to understand the discrepancy. If that fails, escalate to OWDP support.
- 10. **Scenario:** I don't have a URI for my featureOfInterest.
 - **Solution:** This is common. Provide the best possible name and description. The OWDP curation team can assist in assigning or finding the correct URI during the ingestion process.

Part 3: Data Type Classification Guide

Objective: To provide data providers with a clear and systematic process for accurately classifying their datasets according to the Oregon Water Data Portal (OWDP) topic hierarchy and data standards.

Audience: All data providers, particularly those new to the OWDP or those submitting complex datasets.

3.1 CLASSIFICATION DECISION SYSTEM

Correctly classifying your data is the most critical step to ensure it is discoverable and interoperable. This section provides a structured approach to guide you through the

classification process.

Classification Decision Tree Structure (Description)

The classification process follows a decision tree structure to guide you from a general understanding of your data to a specific OWDP classification. The process starts with a single broad question, which then branches into more specific questions based on your answers.

- **Level 1 (Origin):** The first question is: "Does your data measure or model a **natural** feature OR an **anthropogenic** (human-made) feature?"
- **Level 2 (Natural Features Branch):** If "Natural," the next question is: "Is the feature **Surface Water** OR **Groundwater**?"
 - **Level 3 (Surface Water Branch):** If "Surface Water," subsequent questions distinguish between data about the *features themselves* (e.g., a stream network, Hydrography) versus data *collected from* the features (e.g., water samples, Hydrology or Water Quality). Questions will include: "Are you measuring the flow, level, or volume?" (leading to Hydrology) or "Are you measuring the chemical, physical, or biological constituents?" (leading to Water Quality).
 - **Level 3 (Groundwater Branch):** If "Groundwater," questions will differentiate between aquifer characteristics (Groundwater Features), well observations (Groundwater Sites), and water quality samples taken from wells.
- **Level 2 (Anthropogenic Features Branch):** If "Anthropogenic," the next question is: "Is the feature **Infrastructure**, a **Boundary**, or a **Monitoring System**?"
 - **Level 3 (Infrastructure Branch):** This is the most complex branch. It splits into: Water Supply, Wastewater, and Stormwater. Further decision points will ask: "Does your data concern pipes/conduits?", "Does it concern a treatment plant?", "Does it concern a storage tank or reservoir?". This will lead to very specific sub-classifications.
 - **Level 3 (Boundaries Branch):** Questions will distinguish between Administrative Boundaries (e.g., service areas, government jurisdictions) and Statistical Units (e.g., census tracts).

This branching logic, with over 15 decision points, ensures that by the end of the process, you have narrowed your data down to one or more specific categories from the OWDP Topic Hierarchy.

Preliminary Classification Assessment

Answer these questions to get a preliminary classification for your dataset.

1. **What is the primary subject of your measurement?**
 - A) The chemical/physical/biological properties of a water sample. -> **Likely Water Quality**

- B) The volume, flow rate, or level of water in a river, lake, or aquifer. -> **Likely Hydrology/Stream Flow**
- C) The physical characteristics of a man-made object like a pipe, dam, or treatment plant. -> **Likely Infrastructure**
- D) Precipitation, temperature, or wind speed. -> **Likely Weather and Climate**
- 2. **Where was the data collected?**
 - A) From a river, stream, or lake. -> **Likely Surface Water**
 - B) From a well or spring. -> **Likely Groundwater**
 - C) At a weather station. -> **Likely Weather and Climate**
 - D) Within a water treatment plant, distribution system, or sewer system. -> **Likely Infrastructure**
- 3. **Is your data a direct measurement or a model output?**
 - Both are acceptable, but if it is from a model, it must be classified under the relevant topic (e.g., a Hydrology > Model).

3.2 DETAILED DATA TYPE CATEGORIES

Water Quality Monitoring Data

- **Definition and Scope:** This category includes data on the chemical, physical, biological, and ecological characteristics of water. It typically involves discrete samples taken at a specific point in time or continuous monitoring from a sensor (sonde).
- **Specific Parameter Examples:**
 1. pH
 2. Turbidity (NTU)
 3. Dissolved Oxygen (mg/L)
 4. Temperature (°C)
 5. Nitrate/Nitrite Concentration (mg/L as N)
 6. E. coli (count/100mL)
 7. Pesticide concentrations (e.g., Atrazine in µg/L)
 8. Heavy metals (e.g., Lead, Mercury in µg/L)
 9. Benthic macroinvertebrate index scores
 10. Chlorophyll-a concentration (µg/L)
- **Required Metadata Fields (from XLSX Template):** siteID, organization, sampleTimestamp, observedPropertyName (e.g., "pH"), dataValue, unitName (e.g., "pH units"), analyticalMethodName.
- **Submission Requirements:** For regulatory compliance data, follow the WQX standard. For continuous sensor data, ODM2 is recommended. Both must be augmented with URIs where possible.
- **Quality Standards:** Data should include QA/QC information, such as detectionLimits

and `accuracyBounds`. Data status must be marked as `provisional` or `final`.

- **Common Errors and Solutions:**

- **Error:** Inconsistent units for the same parameter.
- **Solution:** Standardize all measurements for a given parameter to a single unit before submission.
- **Error:** Missing `analyticalMethodName`.
- **Solution:** This is critical for data interpretation. If the method is unknown, it must be stated as such, but every effort should be made to find this information.

Hydrological Data

- **Definition and Scope:** This category focuses on the quantity, distribution, and movement of water. It includes data about water in all its forms (liquid, solid) and its interaction with the land.
- **Measurement Types:**
 1. Streamflow / Discharge (cfs or m³/s)
 2. Water Level / Stage / Gage Height (feet or meters)
 3. Precipitation (inches or mm)
 4. Evapotranspiration (inches or mm)
 5. Lake/Reservoir Storage Volume (acre-feet)
 6. Snow Water Equivalent (SWE) (inches)
 7. Soil Moisture (%)
 8. Infiltration Rate (in/hr)
- **Temporal Requirements:** This data is often time-series, ranging from 15-minute intervals (for streamflow) to daily (for precipitation). The `sampleTimestamp` is crucial.
- **Spatial Considerations:** Data must be linked to a specific monitoring site. For surface water features, this should include the `featureOfInterest` URI, preferably the `mainstem_uri` from the 3DHP.
- **Integration with Existing Networks:** If the data is from a USGS gauge, the `siteID` must match the official USGS site number to ensure proper integration.

Groundwater Data

- **Definition and Scope:** This includes data on water found beneath the Earth's surface, including well characteristics, water levels, and quality.
- **Well Monitoring Specifications:** Datasets should include well depth, screen intervals, and construction details where available. The `GroundwaterML2` standard is recommended for describing hydrogeologic units.
- **Water Level vs. Quality Data:**
 - **Level Data:** A type of hydrological data, measuring the depth to water level in a well over time.
 - **Quality Data:** Water quality samples taken from a well. These are classified as `Water Quality` data, but the `featureOfInterest` will be the well or aquifer.

- **Reporting Frequencies:** Can range from continuous to monthly, quarterly, or annually.
- **Regulatory Compliance Requirements:** Data submitted for regulatory purposes must meet state requirements for accuracy and reporting. The `aquifer_uri` is a key field for linking data to the correct groundwater body.

Stream Flow and Surface Water

- **Definition and Scope:** While a subset of Hydrological Data, this focuses specifically on flowing water in natural channels (rivers, streams).
- **Gauge Data Requirements:** Data must include the gauge location (`Geometry`), `siteID`, and the measured parameter (e.g., "Discharge", "Gage height").
- **Flow Measurement Standards:** Measurements should adhere to standards set by agencies like the USGS.
- **Rating Curve Documentation:** If streamflow is calculated from gage height via a rating curve, this documentation should be referenced in the metadata. The OWDP does not store rating curves themselves, but knowing they were used is important context.

Weather and Climate Data

- **Definition and Scope:** Data related to atmospheric conditions. This is crucial as it represents the primary input for the hydrologic cycle.
- **Meteorological Parameters:**
 1. Air Temperature
 2. Precipitation (Rain, Snow)
 3. Wind Speed and Direction
 4. Relative Humidity
 5. Barometric Pressure
 6. Solar Radiation
- **Station Metadata Requirements:** Must include the station's name, ID, and precise location. `facilityType` should be specified (e.g., "weather station").
- **Quality Control Procedures:** Describe any QC procedures applied, such as outlier removal or sensor drift correction.
- **Climate vs. Weather Distinctions:**
 - **Weather:** Short-term atmospheric conditions (e.g., daily rainfall). Most submitted data will be weather data.
 - **Climate:** Long-term patterns (e.g., 30-year average rainfall). This is typically derived data.

Infrastructure and Management Data

- **Definition and Scope:** Data about the physical, human-made systems used to manage and distribute water. This category is broad and uses specific identifiers.

- **System Capacity Information:**
 - *Example:* For a Water Treatment Plant (wtp), wtpDesignCapacity (MGD).
 - *Example:* For a Storage Tank, storageTankCapacity (Gallons).
- **Operational Data Requirements:**
 - *Example:* For a Water Supply intake, surfaceSourceAverageWithdrawal (MGD).
 - *Example:* For a Wastewater Treatment Plant (wwtp), wwtpAverageDailyFlow (MGD).
- **Asset Management Integration:** This data allows for linking observational data to specific assets. For example, a water quality sample can be linked to a specific dwPipeID (Drinking Water Pipe ID) to monitor water age or disinfectant residuals.
- **Regulatory Reporting:** Many fields are tied to regulatory reporting, such as pwsid (Public Water System ID) and npdesPermitID (NPDES Permit ID).

3.3 CLASSIFICATION TOOLS AND GUIDES

Determine the most accurate classification for your data.

Data Type Quick Reference

Data Type	Definition	Key Identifiers	Primary Parameter Examples
Water Quality	Chemical, physical, biological condition of water.	siteID, analyticalMethodName	pH, Temp, Turbidity, E. coli
Hydrology	Quantity, distribution, and movement of water.	siteID, mainstem_uri	Streamflow, Water Level, SWE
Groundwater	Water beneath the Earth's surface.	wellID, aquifer_uri	Water Level, Well Depth
Infrastructure	Man-made water management systems.	pwsid, dwPipeID, npdesPermitID	Capacity, Flow, Pressure
Weather/Climate	Atmospheric conditions.	stationID	Precipitation, Air Temp

3.4 EDGE CASES AND COMPLEX DATA

- **Multi-type Datasets:** It is common for a dataset to cross multiple topics. The [Water Data Standards Guidebook](#) gives the example of "a database of drinking water quality samples" which should be tagged with **both** Water Quality-Physical/Chemical and Infrastructure-Drinking Water-Distribution. When submitting, choose the primary topic for classification but provide all relevant metadata.
- **Derived Data Classifications:** If you have calculated a new value from observed data (e.g., calculating nutrient load from concentration and flow), it is considered Modeled Data. The classification should match the subject (e.g., Water Quality > Modeled), and the methodological metadata must clearly describe the calculation used.

- **Modeled vs. Observed Data:** The OWDP template is designed to handle both.
 - **Observed:** A direct measurement from a sensor or lab analysis.
 - **Modeled:** An output from a computer model (e.g., a rainfall-runoff model or a climate model). In the template, the `samplingMethodName` should reflect the model used (e.g., "PRMS Model Output").
- **Historical Data Considerations:** Historical data is extremely valuable but may lack complete metadata. Submit what you have. The most critical fields are `Location`, `Date`, `Parameter`, and `Value`. In the metadata fields, it is acceptable to state "Unknown" if the information is truly lost, but this should be a last resort.

Part 4: Dataset Metadata Application Guide

Objective: To provide data providers with a definitive guide for applying complete, accurate, and standardized metadata to all datasets submitted to the Oregon Water Data Portal (OWDP), ensuring compliance with state and federal recommendations.

Audience: All data providers, with a focus on data managers and technical staff responsible for data formatting and submission.

4.1 OREGON OPEN DATA STANDARD COMPLIANCE

Compliance with metadata standards is not optional; it is essential for data interoperability, discovery, and long-term value. The OWDP's requirements are based on the principles of the Oregon Open Data program and best practices recommended by the Internet of Water (IoW), as materialized in the IOW-OWDP-TABULAR-TEMPLATE.xlsx.

Complete Field Mapping

The OWDP standard is a profile of established data models like WQX and ODM2. The provided XLSX template represents the direct implementation of this standard for tabular data submissions.

- **Conceptual Standard:** Internet of Water (IoW) Recommended Elements, as described in the Water Data Standards Guidebook.
- **Implementation Standard:** The columns and data models defined in the three tabs of the IOW-OWDP-TABULAR-TEMPLATE.xlsx:
 1. **Site Data:** Describes the physical location and context of the monitoring site.
 2. **Metadata:** Describes the "what" and "how" of the measurements—the parameters, methods, and units.
 3. **Data:** Contains the actual time-series observations and their corresponding values.

Required vs. Recommended vs. Optional Fields

This classification dictates the level of effort required for a successful submission.

Level	Definition	Consequence if Missing	Example Fields
Required	The field must contain a valid value for the record to be accepted.	Submission will fail automated validation.	siteID, organization, sampleTimestamp, observedPropertyName, dataValue
Recommended	The field provides critical context and is expected for high-quality datasets.	Submission may be flagged for manual review; data will be less discoverable and useful.	unit_uri, analyticalMethod_uri, featureOfInterest
Optional	The field provides useful, but non-essential, supplementary information.	No direct consequence, but completeness is encouraged.	comments, accuracyBounds

Compliance Verification Procedures

1. **Automated Check:** Upon submission (via form or API), an automated validator will check for the presence of all **Required** fields and verify the data type and format of each entry.
2. **Manual Review:** Datasets that pass the automated check but are missing a high number of **Recommended** fields may be flagged for a manual review by an OWDP data curator.
3. **Self-Validation:** Before submitting, use the checklists in Section 4.6 to validate your metadata completeness and quality.

4.2 FIELD-BY-FIELD METADATA GUIDE

This section provides a detailed breakdown of the key fields in the **Site Data** and **Metadata** tabs of the IOW-OWDP-TABULAR-TEMPLATE.xlsx.

Site Data Sheet Fields

1. siteID

- **Technical Identifier:** siteID
- **Purpose:** A unique alphanumeric string that you, the provider, assign to identify a specific monitoring location. This ID is the primary key linking your site information to your metadata and observation data.

- **Data Type:** Text
- **Constraints:** Must be unique within your organization's submissions. No spaces or special characters other than hyphens (-) or underscores (_). Max length: 50 characters.
- **Requirement: Required**
- **Good Examples:** TUALATIN-RIVER-GAUGE-01, DEQ-WELL-45B, CITY-OF-EUGENE-WWTP-OUTFALL
- **Bad Examples:**
 - My Site (Too generic, contains a space)
 - 123 (Not descriptive, could easily conflict with other providers)
 - Tualatin@River (Contains an invalid special character)

2. uri

- **Technical Identifier:** uri
- **Purpose:** A globally unique, persistent web identifier for the monitoring site itself. This enables linked data applications and prevents ambiguity across different data systems.
- **Data Type:** Text (URL format)
- **Constraints:** Must be a valid URI. If you don't have one, this can be constructed by OWDP upon ingestion. Format: <https://geoconnex.us/owdp/sites/<your-siteID>>
- **Requirement: Recommended**
- **Good Examples:** <https://geoconnex.us/owdp/sites/DEQ-WELL-45B>, <https://waterdata.usgs.gov/monitoring-location/444508122171000>
- **Bad Examples:**
 - www.mywebsite.com/site1 (Not a persistent identifier)
 - DEQ-WELL-45B (This is an ID, not a URI)

3. name

- **Technical Identifier:** name
- **Purpose:** The official, human-readable name of the monitoring site.
- **Data Type:** Text
- **Constraints:** Max length: 255 characters.
- **Requirement: Required**
- **Good Examples:** Tualatin River at West Linn, DEQ Monitoring Well #45B near Salem, City of Eugene Wastewater Treatment Plant Outfall 001
- **Bad Examples:**
 - Site 1 (Not descriptive)
 - TRWL (Overly abbreviated, not human-readable)

4. organization

- **Technical Identifier:** organization

- **Purpose:** The full name of the entity responsible for operating the site and collecting the data.
- **Data Type:** Text
- **Constraints:** Use a consistent name across all submissions.
- **Requirement: Required**
- **Good Examples:** Oregon Department of Environmental Quality, City of Portland Bureau of Environmental Services, Tualatin Riverkeepers
- **Bad Examples:**
 - DEQ (Use the full name)
 - John Smith (Use the organization, not the individual, unless it's a private citizen project)

5. featureOfInterest

- **Technical Identifier:** featureOfInterest
- **Purpose:** The specific, real-world feature that the monitoring site is observing. This is a critical field for linking data.
- **Data Type:** Text (preferably a URI)
- **Constraints:** Should be a persistent URI from a recognized system like Geoconnex. If a URI is not available, a descriptive name can be used.
- **Requirement: Recommended**
- **Good Examples:**
 - <https://geoconnex.us/ref/mainstems/1431511> (URI for the North Santiam River)
 - <https://geoconnex.us/ref/dams/1075935> (URI for a specific dam)
 - Willamette River Mainstem (Descriptive name if URI is unknown)
- **Bad Examples:**
 - River (Too generic)
 - My monitoring location (Describes the site, not the feature it's monitoring)

Metadata Sheet Fields

1. metadataID

- **Technical Identifier:** metadataID
- **Purpose:** A unique ID you assign to a specific combination of parameter, method, and unit. This allows you to define a measurement type once and reuse it for many observations.
- **Data Type:** Text
- **Constraints:** Must be unique within your submission. Max length: 50 characters.
- **Requirement: Required**
- **Good Examples:** PH-YSI-PROBE-LAB, TURBIDITY-HACH-2100Q, DISCHARGE-ACOUSTIC-DOPPLER
- **Bad Examples:** Metadata1 (Not descriptive), pH (Too generic, will conflict if you have multiple methods for measuring pH)

2. sitelD

- **Technical Identifier:** sitelD
- **Purpose:** Links this metadata record to a specific monitoring site defined in your Site Data sheet.
- **Data Type:** Text
- **Dependencies:** The value **must** exactly match a sitelD in the Site Data sheet.
- **Requirement: Required**
- **Good Examples:** TUALATIN-RIVER-GAUGE-01 (if it exists in Site Data)
- **Bad Examples:** Tualatin-River-Gauge-01 (Case sensitivity matters), My Site (if it doesn't exist in Site Data)

3. observedPropertyName

- **Technical Identifier:** observedPropertyName
- **Purpose:** The specific parameter or property being measured.
- **Data Type:** Text
- **Constraints:** Use a consistent, preferably standard, name. Refer to controlled vocabularies like the [ODM2 VariableName](#) vocabulary.
- **Requirement: Required**
- **Good Examples:** pH, Turbidity, Dissolved oxygen, Water temperature, Streamflow
- **Bad Examples:** Temp (Ambiguous, use "Water temperature" or "Air temperature"), Flow (Use "Streamflow"), ph (Use consistent capitalization "pH")

4. unitName

- **Technical Identifier:** unitName
- **Purpose:** The unit of measurement for the observedPropertyName.
- **Data Type:** Text
- **Constraints:** Use standard abbreviations. Refer to controlled vocabularies like the ODM2 Units vocabulary.
- **Requirement: Required**
- **Good Examples:** pH units, NTU, mg/L, deg C, cfs
- **Bad Examples:** Degrees (Ambiguous, use "deg C" or "deg F"), milligrams per liter (Use abbreviation mg/L), none (For pH, use "pH units")

5. analyticalMethodName

- **Technical Identifier:** analyticalMethodName
- **Purpose:** The name of the method, instrument, or procedure used to determine the value.
- **Data Type:** Text
- **Constraints:** Be as specific as possible. If using a standard lab method, include its number.
- **Requirement: Recommended**

- **Good Examples:** YSI ProDSS Sonde, Hach 2100Q Turbidimeter, EPA 300.0, USGS Price AA meter
- **Bad Examples:** Probe (Not specific enough), Lab analysis (Which analysis?), My own method (Provide a description or reference)

4.3 METADATA QUALITY STANDARDS

Completeness Scoring System

Score	Rating	Description
A	Excellent	All Required and all Recommended fields are present and valid.
B	Good	All Required fields are present; >75% of Recommended fields are present.
C	Sufficient	All Required fields are present; <75% of Recommended fields are present. Dataset may have limited discoverability.
F	Insufficient	One or more Required fields are missing or invalid. Dataset will be rejected.

Quality Assessment Criteria

- **Accuracy:** Are the names, IDs, and methods spelled correctly?
- **Consistency:** Is the same organization name and siteID format used for all related records?
- **Clarity:** Is the name of the site human-readable and unambiguous?
- **Linkability:** Are URIs used for uri and featureOfInterest where possible?

4.4 DATA TYPE-SPECIFIC METADATA

Data Type	Critical Metadata Fields to Prioritize
Water Quality	analyticalMethodName, samplingMethodName, detectionLimits, observedProperty_uri
Hydrological	featureOfInterest (with mainstem URI), unitName (ensure it's correct, e.g., cfs vs cms)
Groundwater	featureOfInterest (with aquifer URI), facilityType (e.g., "well")
Infrastructure	facilityType, featureOfInterest (linked to the specific asset URI)
Historical	comments (to describe data origin and any known issues), organization (original collector)
Real-Time Streams	accuracyBounds, dataStatus (to indicate "provisional")

4.5 BULK AND AUTOMATED METADATA

- **Batch Metadata Procedures:** For multiple sites or parameters, populate the Site Data and Metadata sheets completely first. Use spreadsheet features like fill-down to apply the same organization or analyticalMethodName to many records at once.
- **Template-Based Completion:** The IOW-OWDP-TABULAR-TEMPLATE.xlsx is your primary tool. A single metadataID can be created on the Metadata sheet and then referenced by thousands of records on the Data sheet, saving significant effort.
- **API Metadata Submission:** When submitting via API, the system will expect a JSON object that includes the content of the XLSX template within the SensorThings API data model. \
- **Update and Versioning Procedures:** To update metadata for an existing dataset, re-submit the Site Data or Metadata sheet with the corrected information, using the same siteID and/or metadataID. The system will overwrite the previous entries.

4.6 METADATA TOOLS AND TEMPLATES

- **Metadata Completion Template:** The IOW-OWDP-TABULAR-TEMPLATE.xlsx is the primary template.
- **Validation Checklist:**
 1. ☐ Have all **Required** fields been filled?
 2. ☐ Does every siteID on the Metadata and Data sheets have a matching entry on the Site Data sheet?
 3. ☐ Does every metadataID on the Data sheet have a matching entry on the Metadata sheet?

4. [] Are all date/time values in ISO 8601 format (YYYY-MM-DDTHH:MM:SSZ)?
5. [] Are all URIs properly formatted?
6. [] Have units been checked for consistency for each parameter?

4.7 TROUBLESHOOTING AND SUPPORT

- **Common Metadata Errors (Examples):**
 1. **Error:** Using a comma in a text field when the file is CSV.
 - **Solution:** Remove the comma or enclose the text in double quotes.
 2. **Error:** Incorrect date format.
 - **Solution:** Reformat all dates to YYYY-MM-DDTHH:MM:SSZ.
 3. **Error:** siteID mismatch between sheets.
 - **Solution:** Use VLOOKUP or other spreadsheet tools to find and fix mismatches.
 4. **Error:** Non-unique siteID or metadataID.
 - **Solution:** Append numbers or more descriptive text to create unique IDs.
 5. **Error:** Leaving unitName blank for pH.
 - **Solution:** Use the standard value "unitless".
- **Validation Failure Solutions:** The automated validator will provide an error report indicating the specific row and field that caused the failure. Correct the indicated error in your template and resubmit.
- **Contact Procedures:** If you encounter a complex metadata issue that is not covered in this guide, or if a validation error seems incorrect, contact OWDP support. Please include your organization name and the siteID in question.

Part 5: Data Formatting and Template Guide

Objective: To provide data providers with precise, step-by-step instructions for formatting their data using the official IOW-OWDP-TABULAR-TEMPLATE.xlsx, ensuring a successful and efficient submission process.

Audience: All data providers preparing tabular data for upload to the OWDP. This guide is essential for both technical and non-technical users.

Estimated Time to Complete: 2 - 6 hours, highly dependent on the initial state and complexity of the raw data.

5.1 TEMPLATE-BASED SUBMISSION SYSTEM

The foundation of tabular data submission to the OWDP is the standardized IOW-OWDP-TABULAR-TEMPLATE.xlsx. This system ensures that all data, regardless of its source, is structured consistently, making it interoperable and discoverable.

Template Selection Decision Tree

All non-API data submissions use the same comprehensive IOW-OWDP-TABULAR-TEMPLATE.xlsx. The key is to understand which parts of the template are relevant to your submission.

1. **Is this your first time submitting data for this set of monitoring sites?**
 - **Yes:** You must complete all three sheets: Site Data, Metadata, and Data.
 - **No:** Proceed to question 2.
2. **Are you submitting new observations for existing sites using previously defined measurement methods?**
 - **Yes:** You only need to complete the Data sheet. Ensure the **siteID** and **metadataID** values match your previous submissions.
 - **No:** Proceed to question 3.
3. **Are you adding a new type of measurement (e.g., a new parameter or method) to an existing site?**
 - **Yes:** You need to add new rows to the Metadata sheet to define the new measurement and then add your new observations to the Data sheet. You do not need to edit the Site Data sheet.
 - **No:** Contact support, as your scenario is unique.

IOW-OWDP-TABULAR-TEMPLATE.xlsx Detailed Breakdown

The template consists of three interconnected sheets (tabs). This relational structure separates static information from time-series data, which is highly efficient.

- **Site Data Sheet:** Describes the "where." Each row represents a single, unique monitoring location. This information is typically static.
- **Metadata Sheet:** Describes the "what" and "how." Each row defines a unique combination of a parameter, unit, and method for a specific site. This information is also relatively static.
- **Data Sheet:** Contains the "when" and the "value." This is where your time-series observations are recorded. It is the most dynamic part of the template and will contain the most rows.

How they link:

- The Data sheet links to the Metadata sheet using the **metadataID** column.
- The Metadata sheet links to the Site Data sheet using the **siteID** column.

Required vs. Optional Columns

Within the template, some columns are marked as **Required**. Any submission with missing data in a required column will fail validation. Refer to the *Dataset Metadata Application Guide* (Part 4) for a field-by-field breakdown of requirements.

5.2 EXCEL/CSV FORMATTING REQUIREMENTS

Strict adherence to these formatting rules is mandatory for a successful upload.

- **File Format:** You must submit your data in a Microsoft Excel file (.xlsx)
- **Column Structure:** Do not add, remove, or reorder any columns in the template. The structure must remain identical to the provided template.
- **Data Type Formatting:**
 - **Dates (sampleTimestamp, resultTimestamp):** Must be in ISO 8601 format: YYYY-MM-DDTHH:MM:SSZ. Example: 2023-10-26T14:30:00Z. Excel's default date formats are not compliant. You must format these cells as "Text" to prevent Excel from auto-formatting them incorrectly.
 - **Numbers (dataValue):** Must be formatted as a standard number with no thousands separators (commas). Use a period (.) for the decimal separator. Example: 1234.56.
 - **Text:** All other fields should be formatted as "Text" to avoid unintended conversions by Excel.
- **Special Character Handling:**
 - IDs (siteID, metadataID): Must **not** contain spaces or special characters other than hyphens (-) and underscores (_).
 - Other Text Fields: Avoid using commas, tabs, or pipe characters (|) unless absolutely necessary. If you are exporting to CSV, any field containing a comma must be enclosed in double quotes (").
- **Encoding:** All files must be saved with **UTF-8** encoding to ensure proper handling of all characters. When saving as a CSV from Excel, use the "CSV UTF-8 (Comma delimited)" option.

5.3 DATA PREPARATION PROCEDURES

Before populating the template, your raw data must be cleaned and standardized.

1. **Raw Data Transformation:** Export your data from its native format (e.g., logger software, lab database) into a basic spreadsheet.
2. **Quality Control Pre-checks:**
 - Scan for obviously erroneous values (e.g., a pH of 20, negative flow).
 - Check for data gaps and understand their cause (e.g., equipment failure).
3. **Missing Data Handling:**
 - If a value for a specific timestamp is missing, **do not** include that row in the Data sheet.

- Do not use placeholders like -9999, NA, or null. Simply omit the row.
- 4. **Outlier Identification and Treatment:** Identify statistical outliers. If an outlier is a valid measurement, it can be included. If it is an error, it should be removed. You can note the presence of outliers that were kept in the comments field.
- 5. **Unit Conversion:** Convert all measurements to the standard units required by the OWDP. For example, if you measured temperature in Fahrenheit, you must convert it to Celsius (deg C). All required units are specified in the *Dataset Metadata Application Guide*.

5.4 TEMPLATE COMPLETION GUIDE

Sheet-by-Sheet Instructions

1. **Site Data Sheet:**
 - For each unique monitoring location, create one row.
 - Fill in all required columns, such as siteID, name, organization, and Geometry (latitude and longitude in separate columns, in decimal degrees).
2. **Metadata Sheet:**
 - For each unique type of measurement at a given site, create one row.
 - For example, if you measure pH and Turbidity at SITE-01, you will create two rows in this sheet for SITE-01.
 - Create a unique, descriptive metadataID for each row.
 - Fill in the observedPropertyName, unitName, and other fields.
3. **Data Sheet:**
 - This sheet will contain the bulk of your records.
 - For every single observation, create one row.
 - In each row, enter the dataValue you measured and the sampleTimestamp when you measured it.
 - Crucially, fill in the metadataID that corresponds to the measurement you are recording. This single ID links the observation to the correct parameter, unit, method, and site.

Cell Format and Formula Handling

- **Cell Formats:** As stated in 5.2, format columns as 'Text' or 'Number' as appropriate. **Never** use 'Date' formats.
- **Formulas:** Your final submitted file must contain **only values**. Do not leave any Excel formulas (e.g., =SUM(A1:A5)) in the cells. Use "Paste Special > Values" to convert all calculated cells to static values before saving.

5.5 FILE ORGANIZATION AND NAMING

- **File Naming Conventions:** Use a clear and consistent naming convention.
 - **Recommendation:** ORGANIZATION_SITENAME_YYYY-MM-DD.xlsx
 - **Example:** TualatinRiverkeepers_GalesCreek_2023-10-26.xlsx
- **Directory Structure:** Keep your raw data, prepared templates, and submission receipts in organized folders on your local system.
- **Version Control:** If you need to submit a correction, save it as a new version.
 - **Example:** TualatinRiverkeepers_GalesCreek_2023-10-26_v2.xlsx

5.6 UPLOAD PROCESS AND VALIDATION

This section describes the process on the "[Oregon Water Data Portal Tabular Data Upload](#)" page.

1. **Step 1: Select File:** Navigate to the upload page. Drag and drop your completed .xlsx or .csv file into the upload box, or click to browse your local files.
2. **Step 2: Submit:** Click the "Submit" button.
3. **Step 3: Real-Time Validation:** The system will immediately perform the first level of validation. It will check for:
 - Correct file format (.xlsx).
 - File size (uploads may be limited to 500MB as per the example).
 - Correct column headers and order.
4. **Step 4: Progress and Confirmation:** If the initial validation passes, the file will be uploaded, and a background process will begin to validate the content. You will receive an on-screen confirmation of a successful upload.
5. **Resubmission:** If you receive a validation error report, correct the issues in your template file, save it (preferably as a new version), and repeat the upload process.

5.7 QUALITY CONTROL CHECKLIST

Use this checklist before every submission to catch common errors.

Check	Common Formatting Error Example & Solution
File	Error: File saved as .xls (old Excel format). Solution: Re-save as .xlsx.
Structure	Error: Deleted the comments column because it was empty. Solution: Re-add the column. Never change the template structure.
IDs	Error: siteID is City Well 1. Solution: Change to City-Well-1. No spaces allowed.
Dates	Error: Date is 10/26/2023 14:30. Solution: Convert to 2023-10-26T14:30:00Z. Format cell as Text.
Numbers	Error: Value is 1,234.56. Solution: Remove the comma: 1234.56.
Text	Error: organization is "City, County of Example". Solution: For CSV, enclose in quotes: "City, County of Example".
Links	Error: Mismatched metadataID between Data and Metadata sheets. Solution: Use VLOOKUP to find and correct mismatches.
Values	Error: Formulas like =A2*B2 left in cells. Solution: Copy and "Paste Special > Values" to convert to static numbers.
Blanks	Error: Used -9999 for a missing value. Solution: Delete the entire row for that missing observation.

5.8 TROUBLESHOOTING GUIDE

- **Upload Failures:**
 - **"Invalid File Format":** Ensure your file extension is .xlsx
 - **"File Too Large":** Split your Data sheet into multiple files, each under the size limit. The Site Data and Metadata sheets can be identical in each file.
- **Template Corruption:** If your Excel file becomes corrupted, try opening it in a different program (like Google Sheets or LibreOffice Calc) and re-saving it as a new .xlsx file. As a last resort, copy the data values into a fresh, blank template.
- **Performance Optimization:** For very large files (>100,000 rows), working in CSV format using a text editor or a data scripting language (like Python with pandas) can be much faster than using Excel.

Part 6: API Integration Documentation

Objective: To provide developers with a definitive technical reference for integrating their data submission systems with the Oregon Water Data Portal (OWDP) API.

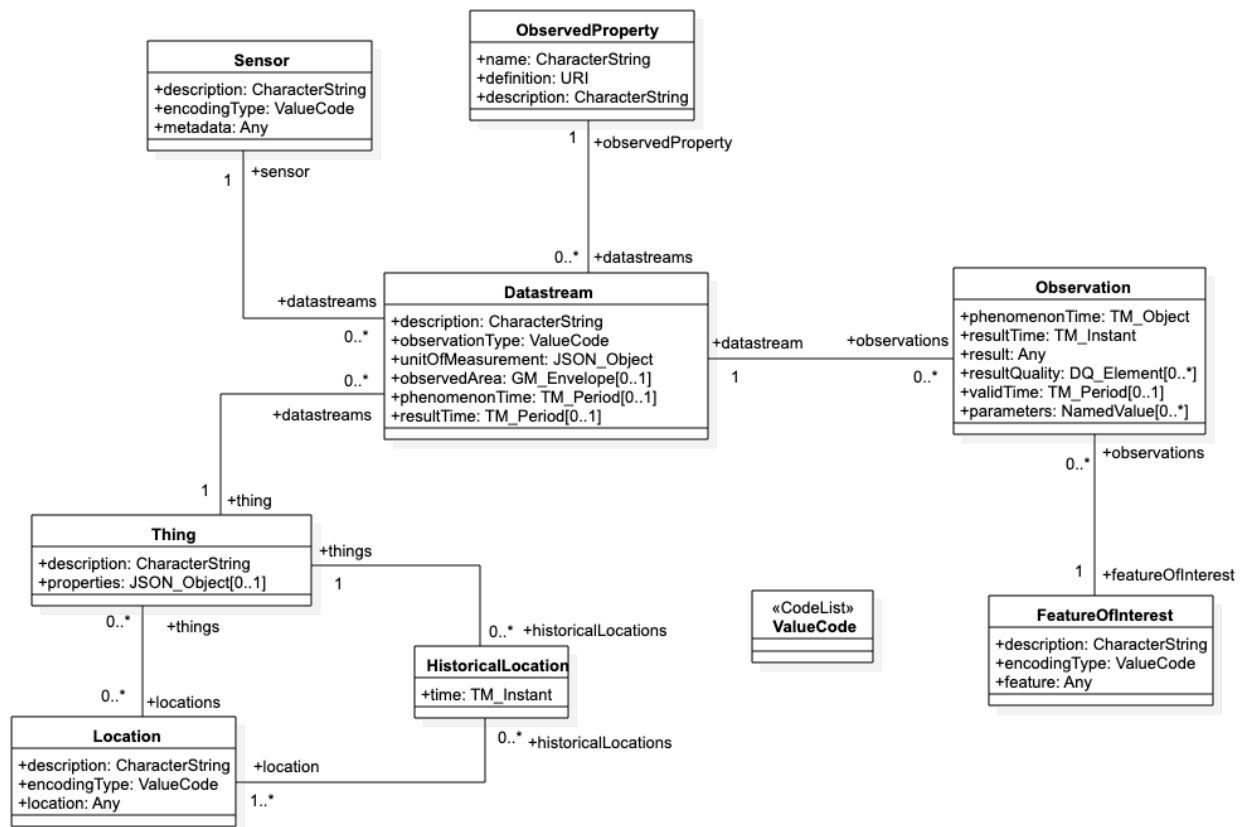
Audience: Software developers, data engineers, and technical data managers at partner organizations.

API Standard: The OWDP API is a compliant implementation of the **OGC SensorThings API Part 1: Sensing, Version 1.1**. Familiarity with this standard is beneficial but not required, as this guide provides all necessary information.

6.1 API OVERVIEW AND ARCHITECTURE

OWDP API Description

The SensorThings API Standard has the data model illustrated below:



Core Entities Mapping:

- **Thing:** Represents the object of interest. In OWDP, this is a **Monitoring Site**.
- **Location:** The physical location of the Thing.
- **Datastream:** A collection of Observations from a single Sensor measuring a specific ObservedProperty. This corresponds to a metadataID in the template, linking a site to a parameter.
- **Sensor:** The instrument used to generate the Observation. This corresponds to the method.
- **ObservedProperty:** The property being measured. This is the parameter (e.g., pH, Turbidity).
- **Observation:** A single data value recorded at a specific time.

Endpoint: POST /v1/Things

- **Purpose:** Create a new monitoring site (a Thing).
- **HTTP Method:** POST
- **Authentication:** Required.
- **Request Body Schema:** A JSON object representing the Thing and its associated Location.

```
{
  "name": "Tualatin River at West Linn",
  "description": "Monitoring station on the Tualatin River, managed by Tualatin Riverkeepers.",
  "properties": {
    "organization": "Tualatin Riverkeepers",
    "siteID": "TR-GAUGE-01"
  },
  "Locations": [{
    "name": "Tualatin River at West Linn",
    "description": "Primary location of the monitoring station.",
    "encodingType": "application/vnd.geo+json",
    "location": {
      "type": "Point",
      "coordinates": [-122.6141, 45.3736]
    }
  }]
}
```

- **Success Response:** 201 Created. The response will include a Location header with the URL of the newly created Thing (e.g., [https://owdp-pilot.internetofwater.app/FROST-Server/v1.1/Things\(123\)](https://owdp-pilot.internetofwater.app/FROST-Server/v1.1/Things(123))). The body will contain the full JSON of the created Thing.

- **Error Responses:**

- 400 Bad Request: The request body is malformed or missing required fields.
- 409 Conflict: A Thing with the same siteID under your organization already exists.

Endpoint: POST /v1/Datastreams

- **Purpose:** Create a new measurement definition (a Datastream). This links a site (Thing) to a parameter (ObservedProperty) and method (Sensor).
- **HTTP Method:** POST
- **Authentication:** Required.
- **Request Body Schema:** A JSON object defining the Datastream.

```
{
  "name": "Water Temperature - Tualatin River at West Linn",
  "description": "Water temperature measurements in Celsius.",
  "observationType": "http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement",
  "unitOfMeasurement": {
    "name": "Degree Celsius",
    "symbol": "deg C",
    "definition":
"http://www.qudt.org/qudt/owl/1.0.0/unit/Instances.html#DegreeCelsius"
  },
  "Thing": {"@iot.id": 123},
  "ObservedProperty": {"@iot.id": 1},
  "Sensor": {"@iot.id": 1}
}
```

(Note: You must first create or find the @iot.id for the Thing, ObservedProperty, and Sensor)

- **Success Response:** 201 Created.
- **Error Responses:**
 - 400 Bad Request: Malformed JSON or invalid @iot.id references.
 - 404 Not Found: The referenced Thing, ObservedProperty, or Sensor does not exist.

Endpoint: POST /v1/Observations

- **Purpose:** Submit a single observation record.
 - **HTTP Method:** POST
 - **Authentication:** Required.
 - **Request Body Schema:**
- ```
{
```

```

 "phenomenonTime": "2023-10-26T14:30:00Z",
 "result": 12.5,
 "DataStream": {"@iot.id": 456}
}

```

- **Success Response:** 201 Created.

## 6.2 COMPLETE CODE EXAMPLE

### Python (requests library)

```

import requests
import json
import time
import os

--- Configuration ---
BASE_URL = "https://owdp-pilot.internetofwater.app/FROST-Server/v1.1/ "
API_KEY = os.environ.get("OWDP_API_KEY") # Recommended: load from
environment variable
HEADERS = {
 "x-api-key": API_KEY,
 "Content-Type": "application/json"
}

def create_monitoring_site(site_payload):
 """Creates a new monitoring site (Thing)."""
 try:
 response = requests.post(f"{BASE_URL}/Things", headers=HEADERS,
data=json.dumps(site_payload))
 response.raise_for_status() # Raises an HTTPError for bad responses (4xx or
5xx)
 print("Site created successfully!")
 return response.json()
 except requests.exceptions.HTTPError as err:
 print(f"HTTP Error creating site: {err}")
 print(f"Response body: {err.response.text}")
 except requests.exceptions.RequestException as err:
 print(f"Request Error creating site: {err}")
 return None

```

```

def submit_observation_batch(datastream_id, observations):
 """Submits a batch of observations for a given Datastream."""
 # The SensorThings standard allows for batch submission via a "CreateObservation"
 request in a batch request.
 # We will construct a payload that does this.
 payload_components = []
 for i, obs in enumerate(observations):
 component = {
 "id": f"obs-{i}",
 "method": "POST",
 "url": "/v1/Observations",
 "headers": {"Content-Type": "application/json"},
 "body": {
 "phenomenonTime": obs["timestamp"],
 "result": obs["value"],
 "Datastream": {"@iot.id": datastream_id}
 }
 }
 payload_components.append(component)

 batch_request_payload = {
 "requests": payload_components
 }

 try:
 response = requests.post(f"{BASE_URL}/$batch", headers=HEADERS,
data=json.dumps(batch_request_payload))
 response.raise_for_status()
 print("Batch submitted successfully!")
 return response.json()
 except requests.exceptions.HTTPError as err:
 print(f"HTTP Error submitting batch: {err}")
 print(f"Response body: {err.response.text}")
 return None

if __name__ == "__main__":
 if not API_KEY:
 print("Error: OWDP_API_KEY environment variable not set.")
 else:
 # --- Workflow Example ---
 # 1. Create a Site (Thing)

```

```

site_data = {
 "name": "Python Test Site",
 "description": "A test site created via Python.",
 "properties": {"organization": "My Python Script", "siteID": "PY-TEST-01"},
 "Locations": [{"name": "Python Test Site", "description": "Location",
"encodingType": "application/vnd.geo+json", "location": {"type": "Point", "coordinates": [-
122.7, 45.4]]}]
}
created_site = create_monitoring_site(site_data)

In a real workflow, you would now create the Sensor, ObservedProperty, and
Datastream.
For this example, we will assume a Datastream with ID 789 exists.
if created_site:
 datastream_id_for_temp = 789

2. Submit a batch of observations
observation_data = [
 {"timestamp": "2024-01-15T10:00:00Z", "value": 8.1},
 {"timestamp": "2024-01-15T10:15:00Z", "value": 8.2},
 {"timestamp": "2024-01-15T10:30:00Z", "value": 8.3}
]
submit_observation_batch(datastream_id_for_temp, observation_data)

```

## 6.3 TESTING AND VALIDATION PROCEDURES

### Unit Testing

When writing code that interacts with the API, your unit tests should mock the API endpoints. This allows you to test your application's logic (e.g., payload construction, response handling) without making actual network calls.

#### Conceptual Python Unit Test (using `pytest` and `requests-mock`):

```

In this example, 'my_app' contains the 'create_monitoring_site' function
from my_app import create_monitoring_site
def test_create_monitoring_site_success(requests_mock):
 """Test that the site creation function handles a 201 Created response correctly."""
 # Mock the API endpoint
 requests_mock.post("https://api-sandbox.owdp.oregon.gov/v1/Things",
 json={"@iot.id": 123, "name": "Test Site"},
 status_code=201)

```

```
Call the function being tested
test_payload = {"name": "Test Site", ...} # a valid payload
result = create_monitoring_site(test_payload)

Assert that the function returned the expected data
assert result is not None
assert result["@iot.id"] == 123
```

## 6.4 PRODUCTION DEPLOYMENT CHECKLIST

Use this checklist before, during, and after deploying your integration to the production environment.

| Phase                 | Status | Task                                                                                                                                                                                                  |
|-----------------------|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Pre-Deployment</b> | [ ]    | All unit and integration tests pass successfully against the <b>sandbox</b> environment.                                                                                                              |
|                       | [ ]    | Application is configured to use the <b>Production API URL</b> ( <a href="https://owdp-pilot.internetofwater.app/FROST-Server/v1.1/">https://owdp-pilot.internetofwater.app/FROST-Server/v1.1/</a> ). |
|                       | [ ]    | The production <b>API Key</b> is loaded from a secure source (e.g., environment variable, secrets manager), NOT hardcoded.                                                                            |
|                       | [ ]    | Robust error handling logic is implemented, including retries with exponential backoff for 429 and 5xx errors.                                                                                        |
|                       | [ ]    | Comprehensive logging is in place to capture request details, response status codes, and error messages.                                                                                              |
|                       | [ ]    | Alerting is configured to notify your team of sustained API failures or high error rates.                                                                                                             |
| <b>Go-Live</b>        | [ ]    | Deploy the application/code to your production environment.                                                                                                                                           |
|                       | [ ]    | Manually trigger a single, small data submission to the production API to verify connectivity and successful                                                                                          |

|                        |     |                                                                                                                        |
|------------------------|-----|------------------------------------------------------------------------------------------------------------------------|
|                        |     | processing.                                                                                                            |
|                        | [ ] | Confirm that the initial test record appears correctly in the OWDP portal or via a GET request.                        |
| <b>Post-Deployment</b> | [ ] | Closely monitor application logs and alerts for the first 24-48 hours.                                                 |
|                        | [ ] | Establish a performance baseline (e.g., average submission latency, success rate) for ongoing monitoring.              |
|                        | [ ] | Have a documented rollback plan to disable the integration or revert to a previous version in case of critical issues. |

## Part 7: FAQs

This section answers the most common questions. Find your question or browse by category.

### Category: Getting Started

1. Q: How do I get an account for the OWDP?

A: To get an account, please contact [OWDP@DEQ.oregon.gov](mailto:OWDP@DEQ.oregon.gov) .

2. Q: What are the system requirements to use the portal?

A: You will need a modern web browser (Chrome, Firefox, Edge, Safari) and spreadsheet software capable of opening .xlsx files (like Microsoft Excel, Google Sheets, or LibreOffice Calc). No other special software is required for template submissions.

3. Q: Is there a cost to submit data to the OWDP?

A: No, submitting data to the Oregon Water Data Portal is free of charge. We welcome contributions from all relevant data providers

4. Q: What kind of water data does the OWDP accept?

A: The portal accepts a wide range of water-related data for Oregon, including water quality, hydrology, groundwater levels, stream flow, and water infrastructure

data. For a complete list, please see the Data Type Classification Guide.

### **Category: Data Preparation**

1. Q: How do I know if my data is already in the portal?  
A: Before submitting, you must perform a thorough search of the OWDP catalog. Use keywords, location filters, and organization names to check for existing data.
2. Q: What is the difference between "Site Data," "Metadata," and "Data" in the template?  
A: "Site Data" describes your physical monitoring location (the "where"). "Metadata" describes what you are measuring and how (e.g., pH using a YSI probe). "Data" contains the actual measurement values and timestamps.
3. Q: My date/time values are not in ISO 8601 format. How do I fix this in Excel?  
A: First, format the column as "Text" to prevent Excel from auto-formatting. Then, use the formula =TEXT(A1, "yyyy-mm-ddThh:mm:ss")&"Z" (assuming your date is in cell A1) and copy the resulting values into your template.
4. Q: What do I do with missing values or gaps in my data?  
A: Do not use placeholders like -9999 or NA. Simply delete the entire row corresponding to the missing observation from the Data sheet of your template.

### **Category: Submission Process**

1. Q: I clicked "Submit." What happens now?  
A: Your file is first checked for basic formatting. If it passes, it's uploaded for asynchronous processing.
2. Q: The upload failed with an "Invalid Format" error. What's wrong?  
A: This usually means your file is not a valid .xlsx, or the internal structure (column headers, sheet names) has been altered. Please re-download a fresh template and carefully copy your data into it.
3. Q: How long does it take for my data to be processed?  
A: Processing time depends on file size and system load. Small files often process

in a few minutes. Very large files may take up to an hour. If you have not received a status email after 2 hours, please contact support.

4. Q: My submission failed. How do I fix the errors and resubmit?

A: The error report email will specify the sheet, row, and column for each error.

Correct these issues in your local file, save it (preferably as a new version), and re-upload it through the portal. You do not need to create a new dataset entry.

### Contact Information

- **Email:** For general, non-urgent inquiries, you can reach us at [OWDP@DEQ.oregon.gov](mailto:OWDP@DEQ.oregon.gov).