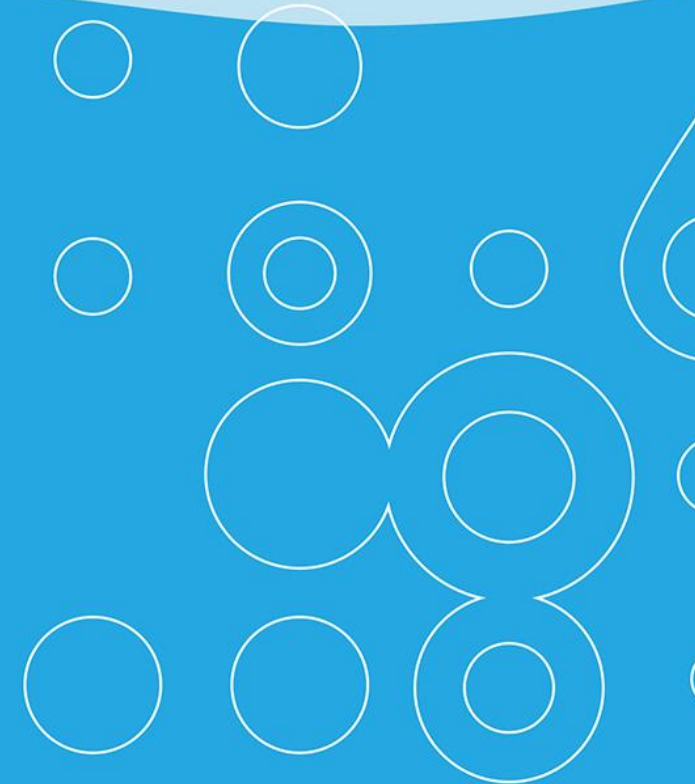




Fall 2020

# WPDx Ingestion Engine: Step-by-step guide for uploading data



# Before you start...

- Review your data to determine which columns of your dataset match with the [WPDx Data Standard](#). You may need to make some minor edits or combinations of columns within your dataset, depending on how your data is organized.
- The “[Ingestion Notes](#)” file may be useful to use to prepare for mapping your organization’s data to the WPDx Data Standard.
- Have your data file or weblink to the dataset available.

# Request Access:

- [upload.waterpointdata.org](https://upload.waterpointdata.org)
- Select “Login to the System”
- Requires a Google account



Login to the System



# Uploading data: Options

There are two options for uploading data:

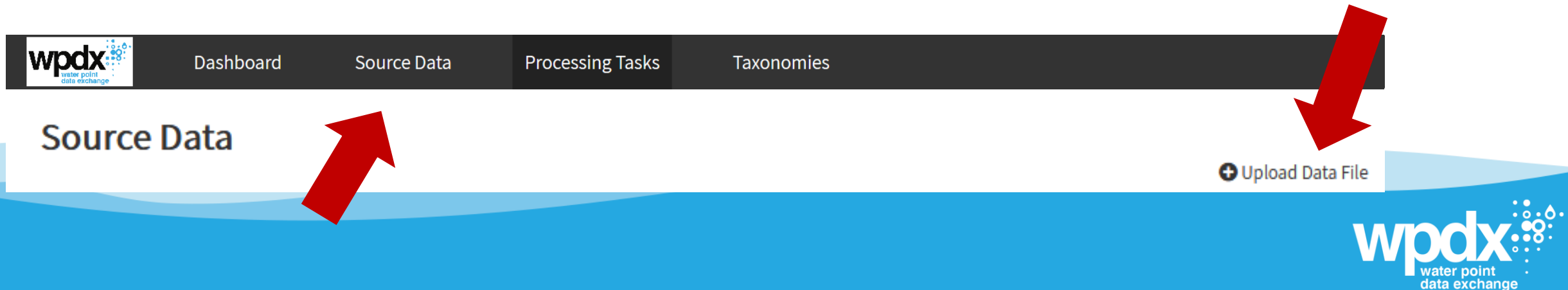
- Upload a physical file (.xlsx, .xls, .csv) from your computer
- Provide a web link to an API endpoint, Google Sheet, Dropbox or other online system

# Uploading data: Web link

- Must provide web link with permissions
  - For **Akvo**, request an API endpoint from your program manager. For more details, please see [here](#).
  - For **mWater**, create a [datagrid](#) formatted per the WPDx standard. This creates a permanent URL. Click on “Download as XLSX” and copy the download link. Use this in the direct URL box at the beginning of a new processing task. For more details, please see [here](#).
  - For **Dropbox**, copy the download link and not the sharing link to use in the WPDx ingestion engine.
  - For **Google Sheets**, ensure that the document is shared publicly.
  - For custom data platforms, please [contact us](#) to determine how we can best connect.
- If the web link does not work, an alternative method is to simply download the source file and upload via the ingestion engine.
- Skip to slide 7

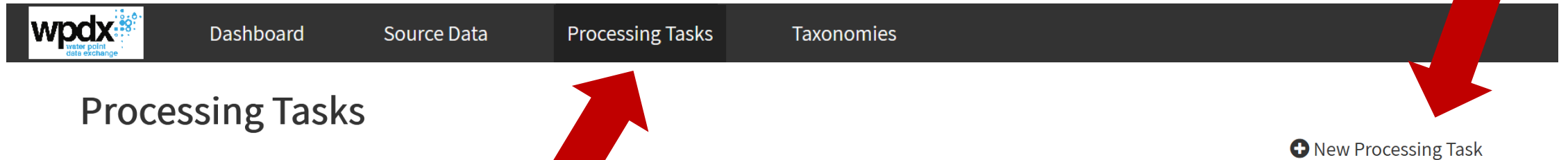
# Uploading data: Source Data tab

- Select “Source Data” tab
- Before you upload the file, to help keep the repository organized, please rename the file using the following format:
  - OrgName\_Country\_MonthYear
  - i.e., GWC\_Uganda\_Jan2020
- Select “+Upload Data File” and browse to datafile and click “Open”
- “File Upload Successful” message will appear at top of screen




# Uploading data: Processing Tasks tab

- Select “Processing Tasks” tab
- Select “+ New Processing Task”



# Processing Task: Task Name and Description

- Enter **Task Name** in the following format:
  - OrgName\_Country/Region\_Month/Year of data
  - For example, Global Water Challenge\_Global\_2019
- Processing tasks can be edited if additional information becomes available
- Provide the main purpose for the collected data under **Description**

 Dashboard Source Data Processing Tasks Taxonomies

## Edit Task

Task Name

Global Water Challenge\_Global\_2020

Description

Describe the main purpose for the collected data; examples may include: routine monitoring, project evaluation, regional inventory, national WASH inventory, water quality testing, etc.



# Processing Task: Metadata

- Complete the metadata prompts to provide a detailed overview of the data within your dataset.
- The metadata will be visible on the data page for your dataset within the WPDx data catalog.

About the Data

Data Source

Name of the organization that owns the data

Geographical scope

or 'sampling frame' of data collection (e.g. particular region, whole country etc.)

Types of water sources included

e.g. improved water sources only, handpumps only, water systems of a certain age, functional water points only, etc.

Sampling strategy

e.g. comprehensive census, simple random sampling, stratified random sampling, cluster random sampling, purposive sampling, convenience sampling, quota sampling etc.

Frequency of data collection

e.g., annually, monthly, one-time, etc.

# Processing Task: Point of Contact

- Complete Point of Contact details for dataset.
- To protect privacy, one option is to use an organizational level email (i.e., [data@name.org](mailto:data@name.org)) which can be forwarded by your organization to relevant contacts.

## Point of Contact

Point-of-contact information to be displayed on the dataset page  
We recommend using organizational level contact information and not personal.

Full Name

Email Address

# Processing Task: Agree to Data Sharing Terms

- Check box to agree to Data Sharing Terms
- Leave visibility as “Only Visible to Me”
- Select: Save & go to Workbench


Check this box




Agree to data sharing terms ☐

Visibility

Only Visible to Me

 Save & go to Workbench

 Save Task

Then press  
this button



# Data Import Workbench

- Select source file from dropdown
- Allow data to process (this may take a few minutes). The Direct URL and format boxes will auto-populate.
- If there are multiple sheets in your file, make sure the correct one is selected.
- Scroll down to continue (the “Data is Processing” message may still appear)
- If using a web address, enter directly in Direct URL text box and select the appropriate format option.
  - For JSON formats, be sure to leave the JSON Path field blank.

The screenshot displays the WPDx Data Import Workbench interface in two states. The top state shows the initial setup with a 'Please select a data source' message. The bottom state shows the 'Data is processing...' message. A blue arrow points to the 'Direct URL' field in the bottom state, and a red arrow points to the 'Uploaded Source File' dropdown in the top state.

**WPDx** Dashboard Source Data Processing Tasks Taxonomies

## Data Import Workbench

Save Please select a data source

**Source of Data**

Uploaded Source File:

Direct URL:

Format:

Save Data is processing...

**Source of Data**

Uploaded Source File:

Direct URL:

Format:

Sheet:

# Data Structure

- If your dataset is formatted to include only the column headers and the data, leave Skip Rows/Columns as “0”
- If there are additional rows or columns which should be skipped (i.e., additional headers or title cells) enter the number of rows/columns to skip.

	A	B	C	D	E	F	G	H	I
1	Data for WPDx								
2	Collected in 2020								
3	Latitude	Longitude	Is Water A	Date of Vis	Source Typ	Technolog	District	Subdistrict	Unique ID
4	12.34567	98.6543	Yes	1/1/2020	Borehole	Afridev ha	North	Northwest	B123
5	12.45678	98.5432	No	1/1/2020	Protected	Gravity sch	South	Southwest	PS321
6	12.45678	98.5432	No	1/1/2020	Protected	Tap stand	South	Southwest	T456
7	12.56789	98.4321	Yes	1/2/2020	Borehole	Handpump	West	Southwest	B234
8	12.6789	98.321	No	1/2/2020	Shallow w	Handpump	East	Northeast	SW123
9	12.7891	98.2109	Yes	1/2/2020	Borehole	Handpump	North	Northeast	B345
10	12.7892	98.2108	Yes	1/2/2020	Borehole	Tap Stand	North	Northeast	T123
11	12.7893	98.2107	Yes	1/2/2020	Borehole	Tap Stand	North	Northeast	T124
12	12.34598	98.98765	No	1/3/2020	Well	Handpump	South	Southwest	W123

## Data Structure

If the file headers are not located on the first row and column, set the following values to skip some rows or columns.

Skip Rows:

Skip Columns:

For sample data shown,  
enter “2” in Skip Rows. Leave  
Skip Columns at “0”

# Data Taxonomy

- Leave Taxonomy as WPDx
- For “Empty Values”
  - If your dataset includes terms for blank/unknown values which should be ignored (i.e., Unknown, N/A, etc.), please enter those terms in the text box.
  - Use a comma as a separator between terms. Do **NOT** include any blank spaces between commas and terms.
  - For example: “unknown,Unknown,N/A,0,null,blank”

## Data Taxonomy

The detected data taxonomy for this dataset -

Taxonomy:

By default, only empty cells will be considered as empty.  
If other values need to be ignored as well, add them here in a comma separated list.

Empty Values:



# Data Mapping

- This section allows you to map your organization's data to the WPDx Data Standard.
- The **WPDx Data Standard parameters** are in bold on the left
- The **dropdown menus** contains your organization's data column headers from the source file.

**Column Mapping**

**Mandatory Fields**

<b>#lat_deg - Latitude</b> Provide the decimal value of the latitude in WGS 1984. At least four decimals should be included. North and East should be positive numbers. Take value from... <a href="#">more settings...</a>	#lat_deg
<b>#lon_deg - Longitude</b> Provide the decimal value of the longitude in WGS 1984. At least four decimals should be included. North and East should be positive numbers. Take value from... <a href="#">more settings...</a>	#lon_deg
<b>#status_id - Presence of Water when Assessed</b> Identify if any water is available on the day of the visit, recognizing that it may be a limited flow Take value from... <a href="#">more settings...</a>	#status_id
<b>#report_date - Date of Data Inventory</b> Provide the date that the data was collected on using ISO 8601. Time and time zone designator are optional Take value from... <a href="#">more settings...</a>	Date
<b>#source - Data Source</b> Provide the name of the organization collecting the data record Take value from... <a href="#">more settings...</a>	Constant...
<b>#water_source - Water Source</b> Describe the water source (e.g. shallow well, spring, borehole, river, pond, etc.). Take value from... <a href="#">more settings...</a>	No value for this field

Constant Value:

# Data Mapping: Getting Started

There are two methods to complete the data mapping process:

## Primary method..

- Using the dropdown menu, scroll to select the column header from your dataset which matches the WPDx standard.
- Some parameters may pre-populate, especially if your dataset is labeled with the WPDx #titles. Verify these selections.
- Note: you cannot map the same column to two different standard parameters.

## Optional method..

If there is a parameter which is not in your dataset, but for which a common value can be applied to all datapoints, Select “Constant...” from the dropdown.

- Examples
  - #source - Data Source → Constant: Name of Org
  - #country\_id - Country → Constant: “UG” or “GH”
  - #orig\_lnk - Public Data Source URL → Constant: URL

#lat\_deg - Latitude

Provide the decimal value of the latitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.

Take value from... Latitude

more settings...

#lon\_deg - Longitude

Provide the decimal value of the longitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.

Take value from... Longitude

more settings...

#status\_id - Presence

Identify if any water is available at the point of collection (e.g. Afridev, gravity scheme, etc.).

Take value from... Is Water Available?

more settings...

#water\_source - Water Source

Describe the water source (e.g. Afridev, gravity scheme, etc.).

Take value from... Source Type

more settings...

#water\_tech - Water Technology

Describe the system being used to transport water to the point of collection (e.g. Afridev, gravity scheme, etc.).

Take value from... Technology Type

more settings...

#country\_id - Country

Provide the country code (e.g. UG, GH, etc.).

Take value from... Country

more settings...

#orig\_lnk - Public Data Source URL

Provide the URL of the public data source (e.g. Afridev, gravity scheme, etc.).

Take value from... Public Data Source URL

more settings...

#lat\_deg - Latitude

Provide the decimal value of the latitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.

Take value from... Latitude

more settings...

#lon\_deg - Longitude

Provide the decimal value of the longitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.

Take value from... Longitude

more settings...

#status\_id - Presence

Identify if any water is available at the point of collection (e.g. Afridev, gravity scheme, etc.).

Take value from... Is Water Available?

more settings...

#water\_source - Water Source

Describe the water source (e.g. Afridev, gravity scheme, etc.).

Take value from... Source Type

more settings...

#water\_tech - Water Technology

Describe the system being used to transport water to the point of collection (e.g. Afridev, gravity scheme, etc.).

Take value from... Technology Type

more settings...

#country\_id - Country

Provide the country code (e.g. UG, GH, etc.).

Take value from... Country

more settings...

#orig\_lnk - Public Data Source URL

Provide the URL of the public data source (e.g. Afridev, gravity scheme, etc.).

Take value from... Public Data Source URL

more settings...

#source - Data Source

Provide the name of the organization collecting the data record

Take value from... Constant...

more settings...

#water\_source - Water Source

Describe the water source (e.g. Afridev, gravity scheme, etc.).

Take value from... Source Columns

more settings...

#water\_tech - Water Technology

Describe the system being used to transport water to the point of collection (e.g. Afridev, gravity scheme, etc.).

Take value from... Source Columns

more settings...

#country\_id - Country

Provide the country code (e.g. UG, GH, etc.).

Take value from... Constant...

more settings...

#orig\_lnk - Public Data Source URL

Provide the URL of the public data source (e.g. Afridev, gravity scheme, etc.).

Take value from... Constant...

more settings...

Constant Value: Global Water Challenge



# Data Mapping: Required Fields

There are 6 mandatory parameters:

- #lat\_deg - Latitude
- #lon\_deg - Longitude
- #status\_id - Presence of Water when Assessed
- #report\_date - Date of Data Inventory
- #source - Organization providing data
- #water\_source - Water Source **AND/OR**
- #water\_tech - Water Point Technology

**Mandatory Fields**

**#lat\_deg - Latitude**  
Provide the decimal value of the latitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.  
Take value from...

**#lon\_deg - Longitude**  
Provide the decimal value of the longitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.  
Take value from...

**#status\_id - Presence of Water when Assessed**  
Identify if any water is available on the day of the visit, recognizing that it may be a limited flow  
Take value from...

**#report\_date - Date of Data Inventory**  
Provide the date that the data was collected on using ISO 8601. Time and time zone designator are optional  
Take value from...   
[more settings...](#)

**#source - Data Source**  
Provide the name of the organization collecting the data record  
Take value from...  Constant Value:

**#water\_source - Water Source**  
Describe the water source (e.g. shallow well, spring, borehole, river, pond, etc.).  
Take value from...

**#water\_tech - Water Point Technology**  
Describe the system being used to transport the water from the source to the point of collection (e.g. Afridev, gravity scheme, Maida, India Mark II, gravity scheme, etc.)  
Take value from...

# Data Mapping: #lat\_deg and #lon\_deg

- Latitude and longitude must be in decimal degrees in WGS84.
- Select the appropriate column header which matches with #lat\_deg.
- Go the next dropdown and make the selection to match #lon\_deg

## Column Mapping

### Mandatory Fields

#### #lat\_deg - Latitude

Provide the decimal value of the latitude in WGS 1984. At least four decimal positive numbers.

Take value from...

[more settings...](#)

#### #lon\_deg - Longitude

Provide the decimal value of the longitude in WGS 1984. At least four decimal positive numbers.

Take value from...

[more settings...](#)

#### #lat\_deg - Latitude

Provide the decimal value of the latitude in WGS 1984. At least four decimals should be included. North and East should be noted as positive numbers.

Take value from...

[more settings...](#)

#### #lon\_deg - Longitude

Provide the decimal value of the longitude in WGS 1984. At least four decimal positive numbers.

Take value from...

[more settings...](#)

#### #status\_id - Presence

Identify if any water is available.

Take value from...

[more settings...](#)

- Latitude
- Longitude
- Is Water Available?
- Date of Visit
- Source Type
- Technology Type
- District
- Subdistrict
- Unique ID
- Year of Installation
- Management Structure
- Payment
- Installer
- Condition

# Data Mapping: #status\_id

- Select the appropriate column header from the dropdown
- Default values include Yes/No. “Unknown” values (see slide 14) will be converted to a blank cell in the WPDx Global Data Repository
- If your dataset does not include Yes/No, but instead terms such as “Functional/Partial/Non-functional” select “more settings..” and enter those terms.
  - True Values = terms which indicate the water point IS functional
  - False Values = terms which indicate the water point is NOT functional
  - Do not leave any spaces between terms, just a comma (i.e., Yes,functional)

## #status\_id - Presence of Water when Assessed

Identify if any water is available on the day of the visit, recognizing that it may be a limited flow

Take value from...

Is Water Available?

[more settings...](#)

## #status\_id - Presence of Water when Assessed

Identify if any water is available on the day of the visit, recognizing that it may be a limited flow

Take value from...

Is Water Available?

Title:

status\_id

True Values:

Functional,Partial

# Data Mapping: #report\_date

- Select the appropriate column header from the dropdown
- The system will automatically detect the format of the dates in your dataset
- If there are errors indicated, select “more settings...” and choose a specific format. (This should only be an issue in rare circumstance)

## #report\_date - Date of Data Inventory

Provide the date that the data was collected on using ISO 8601. Time and time zone designator are optional

Take value from...

Date

[more settings...](#)

## #report\_date - Date of Data Inventory

Provide the date that the data was collected on using ISO 8601. Time and time zone designator are optional

Take value from...

Date

Title:

#report\_date - Date of D

Date Format:

Automatic

Custom Format:

More information on custom  
[less settings...](#)

Automatic - YMD  
Automatic - YDM  
Automatic - DMY  
Automatic - MDY  
MM/DD/YY - 04/22/14  
MM/DD/YYYY - 04/22/2014  
DD/MM/YY - 22/04/14  
DD/MM/YYYY - 22/04/2014  
YYYY-MM-DD - 2014-04-14  
YYYY/MM/DD - 2014/04/14  
DD-MMM-YY - 22-Apr-14  
MMM DD YYYY - Apr 22 2014  
YYYY-MM-DDTHH:MM:SS - 2014-04-22T00:00:00  
YYYY-MM-DDTHH:MM:SSZ - 2014-04-22T00:00:00Z

# Data Mapping: #source

- Provide the name of the organization providing the data.
- If your dataset includes data from multiple sources, please map the parameter to the appropriate column header.
- Otherwise, the entry for Data Source in the About the Data section will be applied to all uploaded records.

## #source - Data Source

Provide the name of the organization collecting the data record. If your dataset includes data from multiple sources, please map the parameter to the appropriate column header here. Otherwise, the entry for Data Source in the About the Data section will be applied to all uploaded records.

Take value from...

#source



# Data Mapping: #water\_source & #water\_tech

- At least one of #water\_source or #water\_tech must be mapped for the upload to proceed.
- Select the appropriate column header/s from the dropdown
- If the information is constant for all values, you can instead select “Constant..” and enter in the appropriate value in the text box.

## #water\_source - Water Source

Describe the water source (e.g. shallow well, spring, borehole, river, pond, etc.).

Take value from...

Source Type

[more settings...](#)

## #water\_tech - Water Point Technology

Describe the system being used to transport the water from the source to the point of collection (e.g. Afridev, gravity scheme, Malda, India Mark II, gravity scheme, etc.)

Take value from...

Technology Type

# Data Mapping: Optional Fields

- The “Optional Fields” are not required, but they do help to provide a more robust dataset for understanding the status of the local water sector. **Please map as many of the WPDx parameters as possible.**
- For any parameters which do not align with your dataset, you can select “No value for this field” (this is the default selection) and go on to the next parameter.
- For example, if your dataset does not include any information on payment:

## #pay - Payment for Water

Provide the payment amount and basis (e.g. monthly, per jerry can, when broken, etc.). If no amount is provided, the basis can be provided alone. An amount without a payment basis cannot be included.

Take value from...

No value for this field

# Data Mapping: #adm1 and #adm2

- #adm1 and #adm2 are official administrative division designations
  - If you have questions, look at GADM.org (see tutorial on next slides) or statoids.com to determine the appropriate designations.

## #adm1 - Primary Administrative Division

Provide the name of the primary administrative division. The correct unit can be found at <http://www.statoids.com>. This corresponds to "First Order" and "First Level" administrative units at <http://Geonames.org> and <http://www.gadm.org> respectively

Take value from...

[more settings...](#)



## #adm2 - Secondary Administrative Division

Provide the name of the secondary administrative division. The correct unit can be found at <http://www.statoids.com>. This corresponds to "Second Order" and "Second Level" administrative units at <http://Geonames.org> and <http://www.gadm.org> respectively

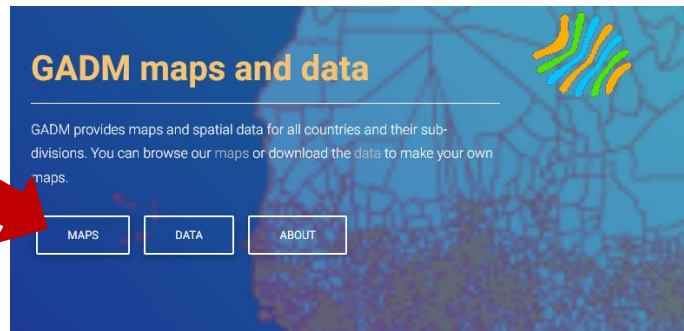
Take value from...





# GADM.org: Check administrative divisions

1. Go to GADM.org and Select “Maps”

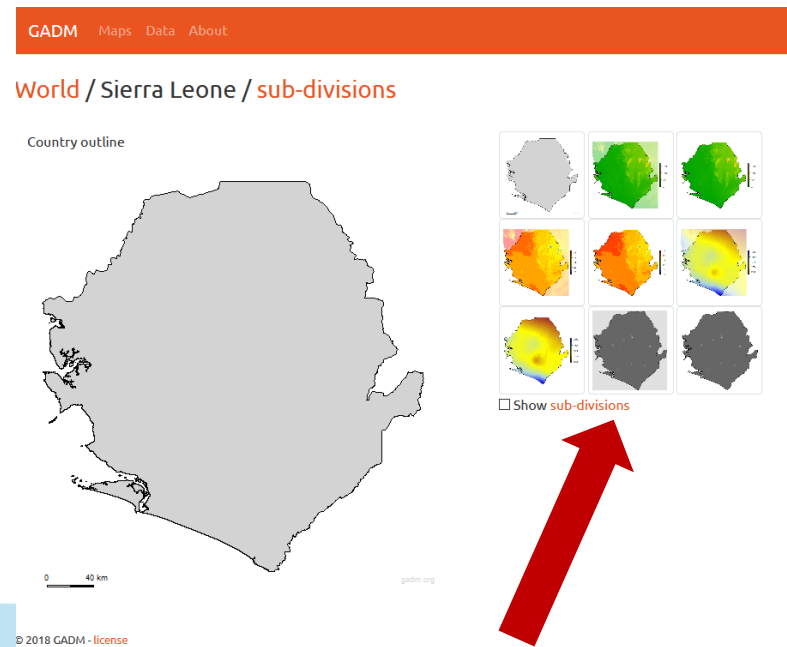


2. Click on country of interest



# GADM.org: Check administrative divisions

3. Select “Show sub-divisions”



4. This creates a map and a list of first-level subdivisions

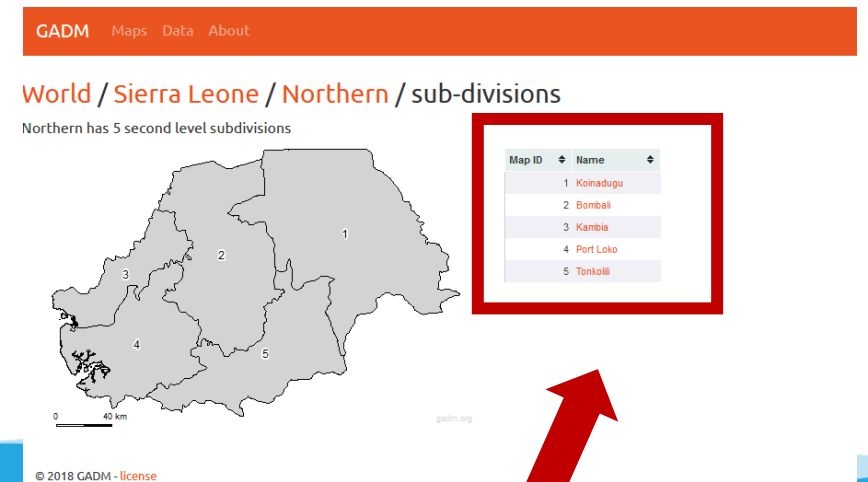
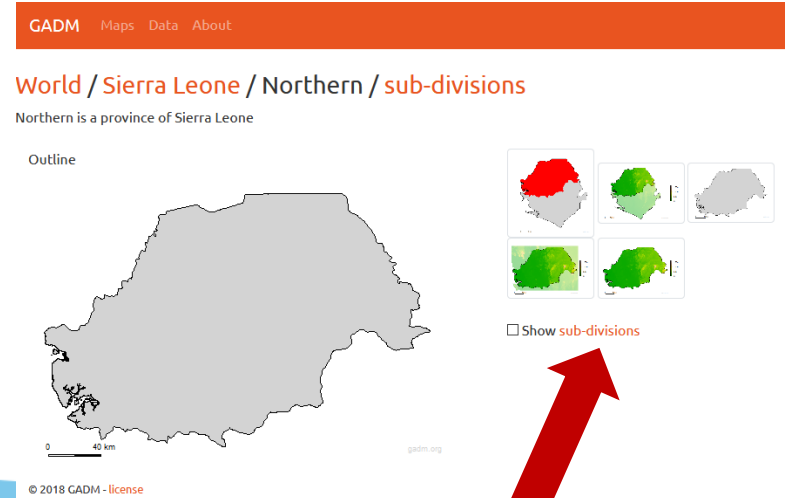
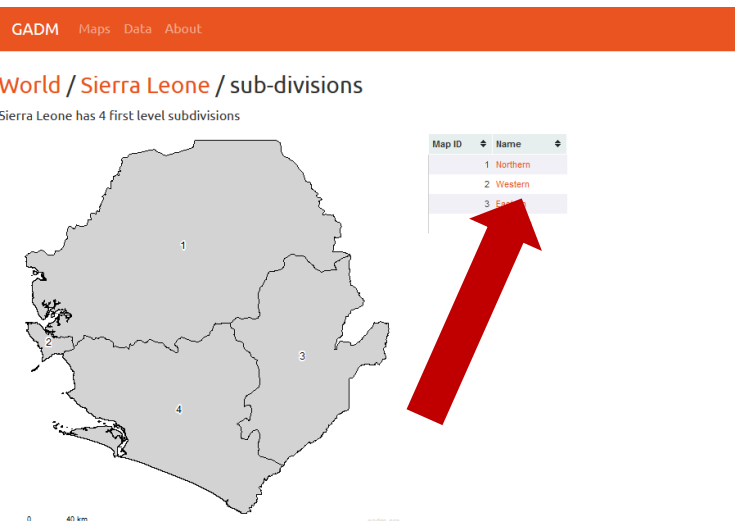


# GADM.org: Check administrative divisions

5. Click on one of the first level sub-divisions

6. Click on “Show sub-divisions

7. This creates a map and list of second level subdivisions



# Data Mapping: #activity\_id

- Select the appropriate column header from the dropdown
  - If a locally or globally recognized standardized identification number exists (i.e., a physical well ID number or barcode) within your dataset, please use that column, **OR**
  - If your organization has a unique id system which would allow water points to be matched within your organization over time, please use that column

## #activity\_id - Water Point ID

Provide the Unique ID for the specific water point infrastructure, as reported by data collector. This could be a physical ID on the water point or an internal system ID.

Take value from...



# Data Mapping: #country\_id

Select the ISO two letter country classification code, selected from a list of [all ISO country codes](#).

- If your dataset includes entries from different countries, this information should be included in your data file. Select the appropriate column header from the dropdown menu.
- If your dataset only includes entries from a single file, you can select “Constant..” and enter a value to be applied to all rows.

## #country\_id - Country

Select the ISO two letter country classification code, selected from a list of all ISO country codes.

Take value from...

Constant...



Constant Value:

# Data Mapping: #install\_year

- Select the appropriate column header from the dropdown
- Note that this field accepts only a four-digit year, and not a full installation date.

■ #install\_year - Installation Year

Provide the 4 digit installation year

Take value from...

Year of Installation



# Data Mapping: #management

- Select appropriate column header from dropdown.
- Select the management classification of the entity that directly manages the water point. Example management types include:
  - Direct Government Operation
  - Private Operator/Delegated Management
  - Community Management
  - School
  - Healthcare Facility
  - Other Institutional Management
  - Other

## #management - Management Structure

Select the classification of the entity that directly manages the water point.

Take value from...

Management Strucutre

# Data Mapping: #pay

- Select appropriate column header from dropdown.

## #pay - Payment for Water

Provide the payment amount and basis (e.g. monthly, per jerry can, when broken, etc.). If no amount is provided, the basis can be provided alone. An amount without a payment basis cannot be included.

Take value from...

No value for this field





# Data Mapping: #installer

- Select appropriate header from dropdown

## #installer - Installer

Provide the name of the entity or entities that installed the water system. This should be the entities that completed or was directly responsible for the construction, rather than a donor or other involved stakeholder. If a system is rehabilitated, both the initial installer and rehabilitation entity can be listed if data is available.

Take value from...



# Data Mapping: #status

- Select appropriate header from dropdown

## #status - Condition

Provide any descriptive status regarding the condition of the water point.

Take value from...



# Data Mapping: #orig\_ink

- If the data is available via a public link, select 'Constant' from the dropdown and enter it so that it can be applied to all rows.
- If there is to a public link, leave as 'No value for this field'

## #orig\_ink - Public Data Source URL

Provide the public link to the data record for a specific water point or full data set, including any non-standard compliant data.

Take value from...

Constant...



Constant Value:

http://www.mydatalink.

[more settings...](#)



# Data Mapping: #photo\_ink

- Select appropriate column header from dropdown

## #photo\_ink - Photograph

Provide the URL of a photograph of the water system. Multiple URLs can be included, with each URL separated by a comma.

Take value from...

# Data Mapping: #converted

- The #converted parameter is for internal use by WPDx to indicate when data has been interpreted by WPDx staff to fit within the standard.
- This only applies when WPDx staff have uploaded data on the behalf of another organization.

# Data Mapping: #fecal\_coliform\_presence

- Select appropriate column header from the dropdown
- Default values include Present/Presence and Absent/Absence. If your dataset include other terms, select 'more settings...' and enter the terms into the True Value and False Value text boxes.
- Separate terms with a comma but do not include any spaces.
- Complete associated metadata questions

## #fecal\_coliform\_presence - Fecal Coliform – Presence or Absence

Results of E. coli or thermotolerant coliform water quality test in a 100ml sample. Total coliform should not be included in this attribute. If thermotolerant, must be noted in the metadata. The water should be taken directly from the water point.

Take value from...

No value for this field

[more settings...](#)

## #fecal\_coliform\_presence - Fecal Coliform – Presence or Absence

Results of E. coli or thermotolerant coliform water quality test in a 100ml sample. Total coliform should not be included in this attribute. If thermotolerant, must be noted in the metadata. The water should be taken directly from the water point.

Take value from...

No value for this field

Title:

fecal\_coliform\_presence

True Values:

Present, Presence, presen

False Values:

Absent, Absence, absent,

[less settings](#)

# Data Mapping: #fecal\_coliform\_value

- Select appropriate column header from dropdown
- Complete associated metadata questions

## #fecal\_coliform\_value - Fecal Coliform – Value

Results of e. coli or thermotolerant coliform water quality test in a 100ml sample. Total coliform should not be included in this attribute. If thermotolerant, must be noted in the metadata. Value should represent the most probable number or colony forming units in 100ml. The water should be taken directly from the water point.

Take value from...

No value for this field



# Data Mapping: #subjective\_quality

- Select the appropriate column header from dropdown

## ■ #subjective\_quality - Subjective Quality

Information regarding the perceived quality of the water including taste, appearance, and/or odor.

Take value from...

Overall Water Quality





# Data Mapping: #scheme\_id

- Select the appropriate column header from dropdown

## #scheme\_id - Scheme Identification

The identifier for a small piped scheme that connects multiple individual water points.

Take value from...

No value for this field



# Data Mapping: #notes

- Select appropriate column from header or apply Constant value is appropriate.
- The #notes parameter can be used to enter custom data which the host country government or organization has selected.
  - For example, some organizations want to track seasonality, additional administrative districts, or some combination.
  - Multiple parameters can be included by creating a column that includes the parameters of interest, separated by a “;” or “...” delimiter.

## #notes - Notes

This field can be used to incorporate any additional information not already part of the WPDx standard that is useful to the data provider.

Take value from...

No value for this field



[more settings...](#)

# Water Quality and Notes Metadata

- If you mapped the #fecal\_coliform\_presence, #fecal\_coliform\_value or #notes columns, please complete the additional metadata question section.

## Water Quality Data

If your dataset includes water quality data, please indicate below the methodology used and the coliform type measured

Measurement unit

Sampling Methodology

Coliform Type

Coliform Incubation Method

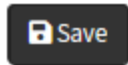
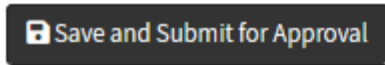
## #notes Metadata

If you mapped the #notes parameter, please include a brief description of what information is included. (e.g., Seasonality - months when water point is dry. Multiple parameters can be included and separated by "...", e.g. Param1...Param2...Param3)

#notes Field Contents

# Once mapping is complete

- Select “Save” or “Save and Submit for Approval”
  - Select Save and Submit for Approval when your data has been fully mapped and is ready for upload
- The status in the Processing Tasks tab will now show as “Pending”
- An administrator will be notified and will complete the uploading process
- Once approved, an email will be sent to the uploader’s email address
- If the mapping was not successful, you will see an error message indicating which parameter was not mapped and explanation of why. Once the error has been fixed, you can submit the processing task for approval.

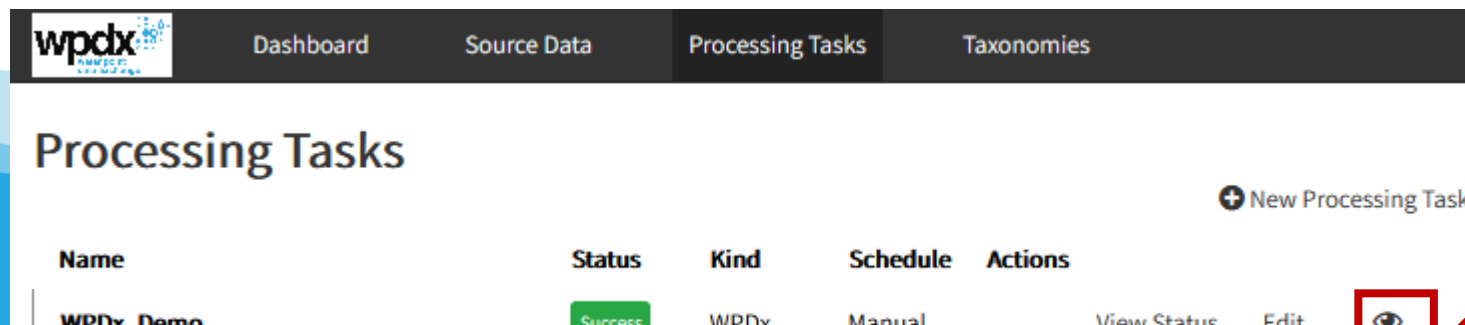
  Mapping completed successfully, click to submit data for approval

## Data Sample

✔ Original Data 9 rows			✔ Mapped Data 9 rows		✔ Transformed Data 9 rows						
#	Latitude	Longitude	Is Water Available?	Date of Visit	Source Type	Technology Type	District	Subdistrict	Unique ID	Year of Installation	Management Status
1	12.34567	98.6543	Yes	2020-01-01T00:00:00	Borehole	Afridev handpump	North	Northwest	B123	2000	Community Managed
2	12.45678	98.5432	No	2020-01-01T00:00:00	Protected spring	Gravity scheme	South	Southwest	PS321	2015	Community Managed
3	12.45678	98.5432	No	2020-01-01T00:00:00	Protected spring	Tap stand	South	Southwest	T456	2015	Community Managed
4	12.56789	98.4321	Yes	2020-01-02T00:00:00	Borehole	Handpump	West	Southwest	B234	1995	Private Operator
5	12.6789	98.321	No	2020-01-02T00:00:00	Shallow well	Handpump	East	Northeast	SW123	2010	None

# Successful upload

- Once the data upload has been completed by an administrator, the status in the Processing Task will be marked as “Success”
- You can view an overview of the dataset by clicking on the eye icon
- The overview page includes:
  - Metadata and contact details
  - Ingestion report - summary statistics of the number of rows uploaded and any errors encountered.
  - Link to download source file
- Data will be visible on the WPDx data repository within 24 hours



# Need to make changes?

- Users can edit their datasets and processing tasks to correct errors or make other additions (i.e., add a new column that was not previously mapped).
- To remove data from WPDx, please contact the administrator at [info@waterpointdata.org](mailto:info@waterpointdata.org) with “Request to remove data from WPDx” in the subject headline. Include the name of the source file and the reason for the removal request.

# Source Data: Update Contents or Delete

- If you realize you have made an error and need to edit or amend an existing dataset, select 'Update Contents' and upload a revised file.
- Once the file has been updated, go back to the associated Processing Task and check/edit the Processing Task content and data mapping and hit "Save and Submit" at the bottom of the Data Import Workbench page.
- Do not use 'Update Contents' to initiate a new dataset upload. Instead upload a new file and start a new Processing Task.

## Source Data

This screenshot shows the file management interface for WPDx. At the top right, there is a button labeled '+ Upload Data File'. Below it, there are three sorting options: '☒ Sort Alphabetically', '☐ Sort by Update Date', and '☐ Show only My Files'. The main area displays a single file entry for 'WPDx\_Test Data.xlsx', which is 11.10 KB, dated 9/14/2020, and owned by Katy Sill. A red box highlights the 'Update Contents' button next to the file name. A large red arrow points from the bottom center towards this button. In the bottom right corner, the WPDx logo is visible, consisting of the letters 'wpdx' in white and blue, followed by the text 'water point data exchange'.

# Editing a Processing Task

- If you want to add/edit the metadata for your dataset or make changes to the way that the data is mapped to the standard, select “Edit” from the Processing Task tab.
- Make any changes and hit “Save and Submit” at the bottom of the Data Import Workbench page.
- An admin will be alerted of your update and will process the upload.

The screenshot shows the WPDx (water point data exchange) interface. At the top is a navigation bar with the WPDx logo and links to Dashboard, Source Data, Processing Tasks, and Taxonomies. Below this is a section titled 'Processing Tasks' with a '+ New Processing Task' button. A table lists the tasks, with columns for Name, Status, Kind, Schedule, and Actions. The first task, 'WPDx\_test', has a status of 'Pending', kind of 'WPDx', and a schedule of 'Manual'. In the Actions column, there are links for 'View Status' and 'Edit'. The 'Edit' link is highlighted with a red box, and a red arrow points to it from the bottom right. The WPDx logo is also visible in the bottom right corner.

Name	Status	Kind	Schedule	Actions
WPDx_test	Pending	WPDx	Manual	<a href="#">View Status</a> <a href="#">Edit</a>





# Questions?

Please contact  
[info@waterpointdata.org](mailto:info@waterpointdata.org)

