Public Health Environmental Surveillance Open Data Model (PHES-ODM) Documentation

Version 2.0 RC2

2022-12-13

Table of contents

Table of contents

Table of contents

Introduction

High-level overview of the ODM

What is the ODM

Wastewater and environmental surveillance

Lorem ipsum

Values and priciples of the ODM

Open data, FAIR principles, etc.

ODM scope

- compartments - Water (including wastewater), surface, air
- domains - biologic, physical
- classes - alleles, variants, mutations, sequences, proteins
- measures -

Contributing, using, and citing

Lorem Ipsum

Acknowledgement

Getting to know the ODM

What are "Parts" and "Part Types"?

Within the Public Health and Environmental Surveillance Open Data Model (PHES-ODM, hereafter referred to as simply "the ODM"), every piece or component of the model structure is called a "part". Regardless of function, all of the tables, headers, and categorical inputs are all parts. This creates a foundational term so that the ODM and its structure can be discussed with greater ease by users and designers.

Because the term "parts" refers to every component of the ODM, "part types" are a way to differentiate between parts which have different functions and structures. The three most important part types are:

- Method Part Type
- Measure Part Type
- Attribute Part Type

Please see the parts reference document for more details, but essentially methods allow the recording of methods in environmental surveillance and testing, while measures allow the recording of observations and results, and attributes capture metadata.

The remaining 20 part types (for a total of 23 part types) are:

- Aggregation Part Type
- Aggregation Scale Part Type
- Aggregation Set Part Type

- Category Part Type
- Category Set Part Type
- Class Part Type
- Compartment Part Type
- Compartment Set Part Type
- Dictionary Part Type
- Domain Part Type
- Group Part Type
- Missingness Part Type
- Nomenclature Part Type
- Quality Part Type
- Quality Set Part Type
- Specimen Part Type
- Specimen Set Part Type
- Table Part Type
- Unit Part Type
- Unit Set Part Type

For information and details about all the part types, please see the parts reference document.

"Sets" and the rationale behind sets

"Sets" are a way of grouping together different possible categorical inputs within the ODM. The purpose behind sets is to group options together so that users only need to check a few options from a drop-down menu, rather than scroll through a long list. Sets are also designed so that a single part can be used in multiple sets, which avoids the need to create multiple versions of this part for these different use contexts. For example, if you're taking a measure of concentration, that measure will likely populate the unit choice drop-down list with units from the "standard concentration unit set". Units in this set include Milligrams per litre, parts per million, and Percent Primary Sludge. Similarly, if one were measuring the amount

of oxygen in a wastewater sample they would be offered units from the "dissolved gas unit set". The units in this set include parts per million, and Milligrams per litre. You see here how Milligrams per litre is used twice across the sets, but exists still as a single part in the parts list. Unit sets are only one type of set, with the others being:

- Aggregation sets
- Compartment sets
- Quality sets
- Specimen sets

Details about these sets are their similar - but unique - applications can be found in the parts reference document. Importantly, "category sets" are not considered a "set" like these others. See the section below for an explanation of this difference.

Why category sets are treated differently

One of the reasons why "sets" are set up and organized the way that they are, is to allow for the re-use of parts in multiple sets. Category sets are, however, an exception to this rule because the categories that make up a category set are used in only one set. Furthermore, the general sets are needed for almost every measure or method (which require unit, aggregation, quality, and compartment information), while category sets are used only in specific circumstances for specific fields. These are things such as collection metadata (example: Sample collection category set), dictionary metadata and model architecture (example: Data Type Category Set), or more detailed types of certain methods (example: Nucleic Acid Extraction Category Set). Category sets are also a distinct part type, and more information about them can be found in the parts reference document.

The "what" and "why" of Specimen IDs

Specimen IDs indicate the level at which a measure is being done. This can be: a site measure, ie. a measure at a site such as temperature or weather; a sample measure, ie. a laboratory measurement on a sample, such as gene copies of SARS-CoV-2 per mL; a person measure, ie. a measure of something at the level of an individual person, such as a blood pressure reading; or a population measure, ie. an aggregate measure of a population, such as the number of confirmed cases of a given illness. While there is currently nothing in version 2.0 of the ODM with uses the person specimen ID explicitly, it has been included to allow for the possibility of storing individual-level data in a future version. The reason specimen IDs were created was so that there could be a single manner in which measures and methods are recorded, regardless of the level at which they were performed, while still maintaining the possibility to collect site, sample, and population measures as a distinct types of metadata.

The "what" and "why" of Groups and Classes

Groups and classes, similar to sets, are ways of grouping together different measures within the ODM. Given that the ODM aims to be as robust as possible and has a very long list of possible measures, groups and classes were designed to give shorter lists of measures in the drop down lists by specifying details about the kind of measure that a user is recording. Groups and classes can work together to further specify what kinds of measures are being reported. For example, the group `sarsCov2` contains many measures, but by specifying that the class is an `allele` or `variant`, the list of possible measures is pared down. Alternatively, class can be said to be non-applicable, paring down the measures in the `sarsCov2` group to only be unspecified measures of the quantity of the virus.

Tables and table types within the ODM

Within the ODM there are three types of tables: program description tables, results tables, and look-up tables. These table types exist to differentiate the function of these tables and to highlight these differences to users. The different tables take different types of inputs and maintenance from users, so understanding the differences can be important.

Program description tables

Program description tables (represented in yellow in the ERD) are tables used to record metadata on the organizations, locations, methods, and appurtenance. These tables help to describe surveillance and testing programs, and are intended to be updated infrequently.

Results tables

Results tables (represented in blue in the ERD) are the tables used to record details on samples and measures. These tables record the main outcomes data and are updated daily, if not more frequently.

Look-up tables

Look-up tables (represented in green in the ERD) are the tables that are pre-programmed and pre-populated in the ODM. These hold information on sets, all parts, languages, and translation abilities. These are only updated by the ODM team in version updates.

Measures, Methods, and Attributes: Key parts

While there are many part types, there are three main part types users should be most familiar with: measures, methods, and attributes. These three have parallels with the three table types, and the differences and details of these parts are useful to understand.

Measures

Measures are actually types of measures that can be performed. These can range from temperature to the number of gene copies in a sample. The measure, or `measureID`, can be selected from a drop down in the templates and it specifies the kind of measure you intend to record. The actual value of the measure is then recorded in the `value` field of the `measures` table, with units and aggregation specified in the `unitID` and `aggregationID` fields.

Methods

Similar to measures, methods are types of methods that can be performed to accomplish a measurement. These can be diverse, ranging from incubation, qPCR, or nucleic acid extraction. The method itself (`methodID`) is selected from a drop down in the templates, specifying in general terms the kind of method the user wishes to record. From there, the `value` field of the `methodSteps` table can be populated by one of the inputs from the category set associated with that `methodID`. This provides a higher level of detail for a given method. For example, the `methodID` might be `solidSep` for solid separation of a sample. The value field might then be populated with `cent` for centrifugation, indicating more details about how the sample settling was done.

Attributes

Attributes are the largest category of part type, as these refer to any most of the fields in the ODM. Attributes are fields for metadata within the ODM and range greatly in the use. They comprise everything from collection dates, to sampling period, to names.

Data quality and reportability

Within the ODM there is a `qualityFlag` field present in both the `samples` and `measures` tables. This field serves to highlight whether or not there is a quality issues with the sample or the measure. The quality flag also allows a user to specify to type of quality issue. This is managed through the use of quality sets, which are the sets that contain the possible quality flags for a given measure or for samples. The idea is that this provides data on any issues with a sample and an indication of the nature of that issue. Having data about the nature and presence of a quality issue is often not sufficient for decision makers who are trying to use and interpret the data. As such, there is also a `reportable` field with is a Boolean indicator of whether or not data can or should be reported or included in final reports and decisions.

Time periods for samples and measures

Fo measures, there is a field `aDateStart` and `aDateEnd` which specify the date and time that an analysis was begun and finished. This allows for the recording of greater detail around timelines for especially long, multi-day analyses. For shorter analyses, which will likely make up the bulk of reported measures, the same dat can input in both fields. The idea is to have the reporting tables for measures be as robust as possible to allow for various kinds of timelines.

For samples, there is `collDT`, `collDTStart`, and `collDTEnd`. The first field is for the collection date and time of a single grab sample, so start and end are not necessary and the other two fields can be left blank. For composite or pooled sample, the start and end date and time for collection is crucial information to know. When these two fields are populated, the `collDT`field can be left blank.

Dates within the ODM

There are a number of date fields within the ODM which all serve different purposes. Collection datetime (`collDT`) is the date a sample was collected, used only for grab samples. This field is left blank if the collection datetime start and end fields (`collDTstart` and `collDTEnd`) have been populated instead. Inversely, collection datetime start and end fields should be left blank if the collection datetime filed has been populated. Similarly, analysis datetime start and analysis datetime end (`aDateStart` and `aDateEnd`) are used to report the date and or timeline of an analysis for a measure.

Date fields that are more related to data processing and labratory infrastructure are: the last edited date (`lastEdited`) which indicates the last time a table, or a measure or sample details, were last edited or updated; the sent date (`sentDate`), or date that a sample was sent to the lab from the field; the received date (`recDate`), or the date the sample was recieved in the lab; and the report date (`repDate`), or the date that the analysis results or measures were reported.These give a greater indication of how up to date data is, but also on the speed and efficiency of the pipeline between sampling and reported results. Dates should be reported in day/month/year format to accommodate the most popular global convention.

Translation and language capabilities

The default language of the ODM is English, but French translations of all descriptive elements of the data model dictionary are also available. As additional nation states and partners adopt the ODM, we anticipate that these fields will be translated into other languages as well. The translation capacities are managed through the language look-up table (`languageLUs`), the translation look-up table (`translationLUs`), and the parts table (`partsLUs`). When a translation for a given part is not available, the dictionary will default to the English term.The language look-up table stores linguistic and classification codes for spoken human languages, with the most recent ISO639 code being the language ID (`languageID`). In the translation table, the language ID is paired with every part ID (`partID`), along with the label (`partLabel`), description (`partDesc`), and instructions (`partInstr`) translated to that language. The part IDs are linked to the full parts list which otherwise contains metadata that is coded in variables and requires no further translation.

Questions and online community

If any users have additional questions or issues with the ODM, we invite them to check out our Discourse Page for discussion boards and community support. For larger issues, and to ask the ODM team to add additional fields or variables into the model, we encourage users to visit the project's GitHub repository and to create and issue there so that a team member can respond.

How to use the ODM

1) How to use templates

The main purpose of templates it to provide you with an easy way to enter in your wastewater data in the PHES-ODM format. Below you will find a series of Guides that will walk you through this process. To begin with, in these Guides, "field" refers to a column in the template, and "entry" refers to a row of information that you entered. While these Guides will also provide definitions and explanations for some terms, most can be found here ..

The templates contain four main types of fields and, in the Guides, they have been associated with a colour for instructional purposes:

- **Identifier, Unique Identifier and Composite Unique Identifier (Red):** In these types of fields you will create a value that will be used to identify something. For instance, "Sample ID" is used to identify a particular Sample. In an Identifier, these values can be repeated between entries, while in a Unique Identifier, they cannot. There is usually only one Unique Identifier per template. A Composite Unique Identifier is a type of Unique Identifier formed by two or more fields. An example can be found in colums A and B below, where while there is repetition in A and B, there is no repetition in the combination of the two.

- **Drop-down and Conditional Drop-down Menus (Green):** These fields have a drop-down menu which contain the values that

21

	A	B
1	Frog	Chair
2	Frog	Table
3	Cat	Sofa
4	Dog	Sofa

can be entered. These valuess can be selected **or** typed in. In conditional drop-down menus, the valuess depend on what you entered in a previous field.

- **Free form (Blue):** In these fields, there are no restrictions on how to enter your data.

2) How to report methods:

In this guide you will learn how to enter Methods into the Method Sets and Methods Step templates. As described in , a Method is "A procedure for collecting a sample or performing a measure". A Method Set is essentially a Method . In the Method Sets template, each entry is a step in a Method Set (a step will be defined shortly).

A Method Set is made up of steps, called Method Steps. In the Method Steps template, each entry represents one of these steps. A key thing to note is that a step can be a Method or a Measure. An individual step can also be used in multiple Method Sets.

For instance, you might have a Method Set for an extraction procedure. This set will have multiple steps that go through this process.

For further information regarding terms and the colour-coding of columns that you will see, please refer to the beginning of this guide (1) How to use templates) and

Quick Start

Below you will find the fields that are mandatory for the Method Steps and Method Sets template. You will also find some examples of entries of data, with the fields filled and colour-coded. A more detailed description of some of the key concepts can be found in the following section, "Detailed Description".

Method Steps Template

1. **Mandatory fields**

Method Step ID	Method (if entering a Method)
Measure (if entering a Measure)	Name
Value (if entering a Measure)	Unit (if entering a Measure)
Aggregation (if entering a Measure)	

2. **Definitions**

- Method Step ID: The unique identifier for a specific method step.

- Method: A procedure for collecting a sample or performing a measure.

- Measure: A measurement or observation of any substance including a biological, physical or chemical substance.

- Name: Name of the entity.

- Value: Value of a measure, observation or attribute.

- Unit: The units of a measurement.

- Aggregation: Statistical measures used to report a measure. Each aggregation has a corresponding value.

3. **Examples**

	A	B	C	D	E	F	G
1	Method Step ID	Method ID	Measure ID	Name	Summary	Method Step Version	Reference link
2	dougVol2		Extraction Volume of Sample	aliquot volume	NA	1	NA
3	dougSpikeBSRV1	spikeTarget		brsv culture Spike Target	WW solids w brsv spike, add to eppendorf tube	1	NA
4	dougSpikeBSRV2		BRSV-N	brsv quantity	NA	1	NA
5	dougSpikeRef	spikeMat		Recovery Efficiency reference sample - Raw wastewater Sample	NA	1	NA

	H	I	J	K	L	M	N	O
1	Organization	Contact ID	Instrument ID	Value	Unit	Aggregation ID	Update date - Last Edited	Notes
2	tmu	carl1	NA	40	ml	single	28-04-2022	NA
3	tmu	carl1	Pipette used	NA	NA	NA	28-04-2022	NA
4	tmu	carl1	NA	200	Gene Copies per L	NA	28-04-2022	NA
5	UNZ	susan4	Type of tube used	NA	NA	NA	28-04-2022	NA

Method Sets Template

1. **Mandatory fields**

Method Set Report ID	Method Set ID
Method Step ID	Method Step Index

2. **Definitions**

- Method Set Report ID: Unique identifier for a method set.

- Method Set ID: A unique identifier for a given method set.

- Method Step ID: The unique identifier for a specific method step.

- Method Step Index: Specifies the index of a method step within a method set, i.e. the order of the steps in a set.

3. **Examples**

	A	B	C	D
1	**Method Set Report ID**	**Method Set ID**	**Method Step ID**	**Method Step Index**
2	SARSQuantificationWW-0	SARSQuantificationWW	description	0
3	SARSQuantificationWW-2	SARSQuantificationWW	dougConc	1
4	SARSQuantificationWW-2.1	SARSQuantificationWW	dougCentri1	1.1
5	SARSQuantificationWW2-0	SARSQuantificationWW2	description	0

	E	F	G	H
	Method Set Version	**Method Set provance ID**	**Update date - Last Edited**	**Notes**
	1	NA		
	1	NA		
	1	NA		
	2	SARSQuantificationWW		

Detailed Description

Method Steps Template

1. **Columns A to G**

	A	B	C	D	E	F	G
1	Method Step ID	Method ID	Measure ID	Name	Summary	Method Step Version	Reference link
2							
3							

   i) Column A:

- Method Step ID (column A) is the Unique Identifier field for this template and cannot be repeated between entries. You can think of each Method Step ID value as representing a unique Method Step.

ii) Columns B to C:

- These are Drop-down Menu fields.
- You only need to enter a value into Method ID (column B) or Measure ID (column C). The former is for when you are entering a Method and the latter is for a Measure. Leave the other field **blank**.

iii) Columns D to G:

- These are Free Form fields and you can enter the relevant information in any format that you wish.
- Name (column D) refers to
- Summary (column E) is a short description of the Method Step that you are entering
- Method Step Version (column F) refers to whether or not you have multiple version of this Step. If you do, then you can differentiate the different versions with "1", "2", etc.

2. **Columns H to O**

	H	I	J	K	L	M	N	O
1	Organization	Contact ID	Instrument ID	Value	Unit	Aggregation ID	Update date - Last Edited	Notes
2								
3								

i) Columns H to J:

- These are Identifier fields.
- You can repeat values between entries. For instance, if you are entering two different Method Steps from the same Organization, then "Organization" (column H) will be the same.

- You may have already created these Identifiers in another template. For instance you may have created "Organization" in the Organizations Template.

ii) Column K:

- A Free Form field in which you will enter the Value of the Measure. If a Method is being entered, then you can put "NA" here.

iii) Columns L and M:

- These are Conditional Drop-down Menu fields.
- Unit (column L) depends on what you entered in Measure ID (column C). If you entered a Method, then you can put "NA" in this field.
- Aggregation (column M) depends on what you entered in Unit (column L). If you entered a Method, then you can put "NA" in this field.

iv) Columns N and O:

- These are Free Form fields where you will enter values for the indicated fields.

Method Sets Template

1. **Columns A to D**

A	B	C	D
Method Set Report ID	Method Set ID	Method Step ID	Method Step Index

i) Columns A to D:

- Method Set Report ID (column A) is the Composite Unique Identifier field for this template.
  - This value is automatically generated by the template by combining Method Set ID (column B) and Method Step Index (column D). It is important that you ensure that any combination of these two fields are not repeated.
  - Each Method Set Report ID value represents a Step within a Method Set.
- Method Set ID (column B) and and Method Step ID (column C) are Identifier fields.
- Method Step Index (column D) is a Free Form field and indicates the position of the Method Step in its Method Set. For instance, if the Method Step being entered is the third step in the Method Set, then this value would be "3".

2. **Columns E to H**

E	F	G	H
Method Set Version	Method Set provance ID	Update date - Last Edited	Notes

i) Columns E to F:
- Method Set Version (column E) is a Free Form field where you will indicate the version of the Method Set. For instance, if the current Method Set that you are entering is an updated version of the original one, you would enter "2".
- Method Set Provenance ID (column F) is the Method Set that the one that is being entered is based upon. Thus,

28

you would enter "NA" here unless the value of Method Set Version is 2 or greater.

ii) Columns G to H:

- These are free form fields and you can enter the relevant information in any format that you wish.

You have now entered your data in the Method Steps and Methods Sets Template, congratulations!

3) How to report measures:

In this guide you will learn how to enter Measures into the Measures and Measures Set template. Each entry in this template represents a Measure. As described in , a Measure is a "measurement or observation of any substance including a biological, physical or chemical substance". You will enter individual Measures into the Measures Template. You can then group certain Measures together and this will be entered into the Measure Sets template. For further information regarding terms and the colour-coding of columns that you will see, please refer to the beginning of this guide (1) How to use templates) and

Quick Start

Below you will find the fields that are mandatory for these templates. You will also find some defintion of terms and examples of entries of data, with the fields filled and colour-coded. A more detailed description of some of the key concepts can be found in the following section, "Detailed Description".

Measures Template

1. **Mandatory fields**

Measure Report ID	Sample ID
Site ID	Analysis Date End Value
Specimen	Fraction analyzed (only if Specimen ID = "Sample")
Measure	Value
Unit	Aggregation

2. **Definitions**

- Measure Report ID: Unique identifier for a measure.

- Sample ID: Unique identifier for a sample.

- Site ID: Unique identifier for the location where a sample was taken.

- Analysis Date End: Date the measurement or analysis was completed.

- Specimen: The substance or thing upon which the observation was made.

- Fraction analyzed: Fraction of the sample that is analyzed.

- Measure: A measurement or observation of any substance including a biological, physical or chemical substance.

- Value: Value of a measure, observation or attribute.

- Unit: The units of a measurement.

- Aggregation: Statistical measures used to report a measure. Each aggregation has a corresponding value.

3. Examples

	A	B	C	D	E	F	G	H
1	Measure report ID	Method set report ID	Sample ID	Purpose	Polygon ID	Site ID	Dataset ID	Measure set report ID
2	rw12345	SARSQuantWW2	rwsp0831	Regular	gta	tmu	queens0830	MSet2
3	rw12346	SARSQuantWW2	rwsp0901	Education	gta	tmu	queens0830	MSet2
4	rop12345	temp1	ropec082922ex	Provisional report	ottGat	delatolla	ropDelEnv	MSet9

	I	J	K	L	M
1	Analysis date start	Analysis date end	Measure Report Date	Specimen ID	Fraction analyzed
2	01/09/2022	01/09/2022	02/09/2022	Sample	Mixed/homogenized sample
3	02/09/2022	02/09/2022	02/09/2022	Sample	Mixed/homogenized sample
4	30/08/2022	31/08/2022	02/09/2022	Site	

	N	O	P	Q	R	S
1	Group	Class ID	Measure ID	Value	Unit	Aggregation ID
2	Miscellaneous viruses group	Alleles class	PMMoV-CP	134	gene copies per mL	Arithmetic Mean
3	SARS-CoV-2 Group	Alleles class	SARS-CoV-2-N1	238	gene copies per mL	Arithmetic Mean
4	Site features group		Environmental Temperature	27	Degrees Celcius	Median

Measure Sets Template

1. Mandatory fields

Report Set ID	Method Set Report ID

2. Definitions

	T	U	V	W	X
1	Index	Quality Flag	Reportable data flag	Update date - Last Edited	Notes
2	1	No quality concerns	TRUE	02/09/2022	Lorem ipsum
3	2	No quality concerns	TRUE	02/09/2022	
4	1	Quality concerns	FALSE	02/09/2022	Lorem ipsum

- Report Set ID: Unique identifier that links together a group of related measures.

- Method Set Report ID: Unique identifier for a method set.

3. **Examples**

   Image will be added.

Detailed Description

Measures Template

1. **Columns A to H**

	A	B	C	D	E	F	G	H
1	Measure report ID	Method set report ID	Sample ID	Purpose	Polygon ID	Site ID	Dataset ID	Measure set report ID
2								
3								
4								

   i) Columns A-C, and E-H

   - These are all Identifier fields.

- "Report ID" (column A) is the Unique Identifier for this template, and cannot be repeated between entries. In essence, each value of "Report ID" represents a unique Measure.
- For the rest, you can repeat the Identifer between entries (if needed). For instance, if you are entering two different Measures from the same Sample, then the "Sample ID" (column C) will be the same.
- You may have already created these Identifiers in another template. For instance you may have created "Sample ID" in the Samples Template.

ii) Column D

- This is a Drop-down Menu field in which you can select a value representing the Purpose of your Measure from the menu.

2. **Columns I to M**

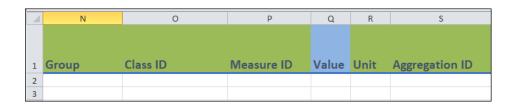|	I	J	K	L	M
1	Analysis date start	Analysis date end	Measure Report Date	Specimen ID	Fraction analyzed
2					
3					
4					

i) Columns I to K

- These are all Free Form fields where you can enter various dates associated with your Measure.

ii) Columns L and M

- These are Drop-down Menu fields where you can select a value that represents the Specimen, and the Fraction of the Samples that was analyzed for your Measure.

3. **Columns N to S**

	N	O	P	Q	R	S
1	Group	Class ID	Measure ID	Value	Unit	Aggregation ID
2						
3						

i) Columns N and O

- These are Drop-down Menu fields which contain the possible values for Group and Class, respectively, of the Measure that you are entering.
- Group and Class are not mandatory, so if you do not know them, leave these fields **blank**. .

ii) Column P

- This is a Conditional Drop-down Menu fields where you will enter the Measure of your entry.
- If you chose to enter in a "Class ID" (column O), the drop down menu will contain the Measures found in that Class. If you did not enter a value in column O, the menu will contain all possible Measures.

iii) Column Q

- This is a Free Form field, where you will enter in the Value of your Measure.
- For instance, if your Measure is a temperature of 20 °C, you would enter "20".

iv) Column R

- This is a Conditional Drop-down Menu field where you will choose the Units of your Value.
- Using the temperature example in the previous point, it would be "Celcius". What appears in this menu depends on what you entered in the "Measure ID" field (column P).

v) Column S

- This is a Conditional Drop-down Menu field which contains the types of Aggregations of your value. For example does your value represent a mean, median, etc.
- What appears in this menu depends on what you entered in the Unit field (column Q).

4. **Columns T and W**

	T	U	V	W	X
1	Index	Quality Flag	Reportable data flag	Update date - Last Edited	Notes
2					
3					

i) Columns T, W and X

- These are Free Form fields in which you can enter the indicated information.
- Index (column T) is used if you have multiple entries with the exact same values in the other fields. If this is the case you could use "1", "2", etc. to differentiate them.

ii) Columns U and V

- These columns consist of a Conditional Drop-down Menu field (Quality Flag) and a Drop-down Menu field (Reportable Data Flag).
- "Quality Flag" depends on the value of "Measure" (column P). You can use this to indicate it there are any quality issues with the Measure.
- "Reportable" refers to whether or not your entry is of good enough quality that it can be used by others.

Measure Sets Template

1. **Columns A to H**

	A	B	C	D	E	F	G	H
1	Measure report ID	Method set report ID	Sample ID	Purpose	Polygon ID	Site ID	Dataset ID	Measure set report ID
2								
3								
4								

i) Columns A, B, E and F

- These are all Identifier fields.
- "Set Report ID" (column A) is the Unique Identifier for this template, and cannot be repeated between entries. In essence, each value of "Set Report ID" represents a unique Measure Set.
- For the rest, you can repeat the Identifer between entries (if needed). For instance, if you are entering two different Measure Sets that came from the same Organization, then the "Organization ID" (column E) will be the same.
- You may have already created these Identifiers in another template. For instance you may have created "Organization ID" in the Organizations Template.

ii) Column D

- This is a Drop-down Menu field in which you will choose from the menu the type of Measure Set that you are entering.

iii) Columns C, G and H

- These are all Free Form fields in which you can enter in information in any format that you wish.
- "Name" (Column C) refers to .

You have now entered your data in the Measures and Measure Sets Template, congratulations!

4) How to report samples:

In this guide you will learn how to enter Samples through the Samples template. So, each entry in this template represents a Sample. As described in , a Sample is " " .

For further information regarding terms and the colour-coding of columns that you will see, please refer to the beginning of this guide (1) How to use templates) and

Quick Start

Below you will find the fields that are mandatory for the Samples template. You will also find some examples of entries of data, with the fields filled and colour-coded. A more detailed description of some of the key concepts can be found in the following section, "Detailed Description".

1. **Mandatory fields**
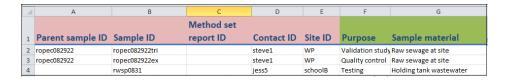
Sample ID	Site ID
Sample Material	Sample Collection
Collection Period	Collection Number
Collection Date Time	

2. **Definitions**

- Sample ID: Unique identifier for a sample.

- Site ID: Unique identifier for the location where a wastewater sample was taken.

- Sample Material: Type of sample.

- Sample Collection: Method used to collect the sample.

- Collection Period: Value of a measure, observation or attribute.

- Collection Number: The number of subsamples that were combined to create the sample. Use NA for continuous, proportional or passive sampling.

- Collection Date Time: The date, time and timezone the sample was taken.
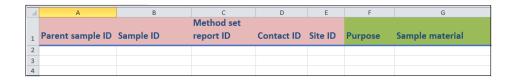
3. **Examples**

	A	B	C	D	E	F	G
			Method set				
1	**Parent sample ID**	**Sample ID**	**report ID**	**Contact ID**	**Site ID**	**Purpose**	**Sample material**
2	ropec082922	ropec082922tri		steve1	WP	Validation study	Raw sewage at site
3	ropec082922	ropec082922ex		steve1	WP	Quality control	Raw sewage at site
4		rwsp0831		jess5	schoolB	Testing	Holding tank wastewater

	H	I	J	K	L	M	N
1	**Dataset ID**	**Sample origin**	**Replicate Type**	**Sample collection**	**Collection Period**	**Collection Number**	**Pooled**
2	delatolla	Derived sample	Laboratory Duplicate	Grab sample		3	1 N
3	delatolla	Synthetic sample	Laboratory Duplicate	Grab sample		3	1 N
4	hsnri	Field sample	Unique Sample	COSCa ball		8	3 Y

	O	P	Q	R	S	T	U
	Collection date	**Collection date**	**Collection date**	**Date sample was**		**Update date - Last**	
1	**time**	**time start**	**time end**	**sent**	**Quality Flag**	**Edited**	**Notes**
2	29/08/2022			01/09/2022		01/09/2022	lorem ipsum
3	29/08/2022			01/09/2022		01/09/2022	lorem ipsum
4		10:00PM 8/31/2022	6:00AM 9/1/2022	02/09/2022	Sample spilled	02/09/2022	lorem ipsum

Detailed Description

1. **Columns A to G**

   i) Columns A to E:

      - These are Identifier fields.

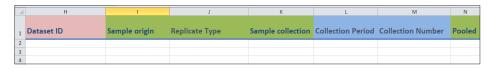|	A	B	C	D	E	F	G
1	Parent sample ID	Sample ID	Method set report ID	Contact ID	Site ID	Purpose	Sample material
2							
3							
4							

- Sample ID (column A) is the Unique Identifier for this template and cannot be repeated between entries. You can think of each Sample ID value as representing a unique Sample.

- For the rest of the fields, you can repeat values between entries. For instance, if you are entering two different Samples from the same Site, then the "Site ID" (column E) will be the same.

- You may have already created these Identifiers in another template. For instance you may have created "Site ID" in the Sites Template.

ii) Columns F to G:

- These are Drop-down Menu fields where you can enter information regarding the Purpose and Sample Material of your Sample.

2. **Columns H to N**

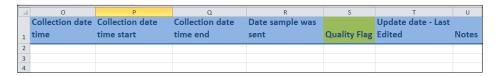|	H	I	J	K	L	M	N
1	Dataset ID	Sample origin	Replicate Type	Sample collection	Collection Period	Collection Number	Pooled
2							
3							
4							

i) Columns H:

- These is an Identifier field for the Dataset that your Sample is associated with.

ii) Columns I to K, N:

- These are Drop-down Menu fields related to various characteristics of your Sample.
- Sample collection (column K) refers to the collection technique used to obtain the Sample that you are entering.
- Pooled (column N) refers to whether or not the Sample that you are entering is made up of multiple child Samples.

iii) Columns L and M:

- These are Free Form fields.
- Collection Period (column L) refers to the number of hours that it took to collect your Sample.
- Collection Number (column M) refers to the number of subsamples that were used to create Sample that you are entering.

3. **Columns O to U**

O	P	Q	R	S	T	U
Collection date time	Collection date time start	Collection date time end	Date sample was sent	Quality Flag	Update date - Last Edited	Notes

i) Columns O to R, T to U:

- These are Free Form fields.
- Most are related to various dates associated with your Sample.

ii) Column S:

- This is a Drop-down Menu field.
- You can use this field to indicate if there are any quality issues with the Sample.

You have now entered your data in the Samples Template, congratulations!

How to report metadata:

In this guide you will learn how to enter Metadata through a series of 8 templates. Each entry in a template will usually represent the the title of the template. For instance, in the Organizations template each row will represent an Organization. Metadata is data that gives information about other data. In the context of the PHES-ODM, metadata is anything that gives additional information about Measures, Methods and Samples.

For further information regarding terms and the colour-coding of columns that you will see, please refer to the beginning of this guide (1) How to use templates) and

Quick Start

Below you will find the fields that are mandatory for the metadata-related templates. You will also find some examples of entries of data, with the fields filled and colour-coded. A more detailed description of some of the key concepts can be found in the following section, "Detailed Description".

Addresses Template

1. **Mandatory fields**

Address ID	Dataset ID
Address Line 1	City
State, Province or Region	Country

2. **Definitions**

- Address ID: A unique identifier for an address.

- Dataset ID: The name of the dataset that stores information for Measures, Samples and other reporting tables.

- Address Line 1: Line 1 (the street name, number and direction) for a given address.

- City: The city where a site or organization is located; part of the address.

- State, Province or Region: The state, province, or region where a site or organization is located; part of the address.

- Country: The country where a site or organization is located; part of the address.

3. **Examples**

Images will be added

Contacts

1. **Mandatory fields**

Contact ID	Organization ID
Contact email	

2. **Definitions**

- Contact ID: A unique identifier for a given contact person.

- Organization ID: A unique identifier for the organization to which the reporter is affiliated.
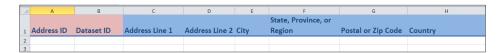
- Email: Contact e-mail address, for the lab.s.

3. **Examples**

Images will be added

Detailed Description

Addresses Template

1. **Columns A to H**

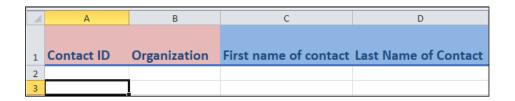| A	B	C	D	E	F	G	H
Address ID	Dataset ID	Address Line 1	Address Line 2	City	State, Province, or Region	Postal or Zip Code	Country

   i) Columns A and B:

- These are Identifier fields.

- "Address ID" (column A) is the Unique Identifier for this template, and cannot be repeated between entries. In essence, each value of "Address ID" represents a unique Address.

- For column B (Dataset ID), you can repeat the value between entries (if needed). For instance, if you are entering two different Addresses from the same DataSet, then this column will be the same.

- You may have already created the Dataset ID in the Datasets Template.

   ii) Columns C to H:

- These are Free-form fields in which you can enter various information regarding the Address.

43

	A	B	C	D
1	**Contact ID**	**Organization**	**First name of contact**	**Last Name of Contact**
2				
3				

Contacts Template

1. **Columns A to D**

    i) Columns A and B:

     - These are Identifier fields.

     - "Contact ID" (column A) is the Unique Identifier for this template, and cannot be repeated between entries. In essence, each value of "Contact ID" represents a unique Contact.

     - For column B (Organization ID), you can repeat the value between entries (if needed). For instance, if you are entering two different Contacts from the same Organization, then this column will be the same.

     - You may have already created the Organization ID in the Organizations Template.

    ii) Columns C to D:

     - These are Free-form fields in which you can enter the name of the Contact.

2. **Columns E to I**

    i) Columns E and I:

     - These are Free-form fields in which you can enter various information regarding the Contact.

44

Example data for each of the 11 report tables.

E	F	G	H	I
Contact email	Contact phone	Role of Contact	Update date - Last Edited	Notes

Example data for each of the 11 report tables.

Tutorials

If needed, we can have chapters with differen tutourials

Part I.

Reference guide

The book section has the parts and other lists. The chapters are organized so that there can be persistent links for the look-up tables (parts, sets, language, translations.)

Parts

The list of parts from alphabetical order of part ID.

Ctreat the list so that there is a persistent link for each part. i.e.

Sets

The list of sets in aphabetical order of setsID

Languages

The language table by langID.

Translations

The list of translations by languageID.

Tables

The list of tables in the three main groups (partType = tables, class = results, prgramDesc, lookup).

Report tables

Program Description tables

Look-up tables

partTypes

A list of partTypes, grouped by partType. Group by partType, group and class. ie. for list measures by group and classes within groups.

partTypes

Aggregations

Attributes

Categories

Category sets

Changes column

Classes

Compartments

Data types

Domains

Groups

Measures

group #1 of measures

class #1 of measues within group #1

Methods

Missingness

Nomeclature

Part types

Quality measures

Specimens

Tables