

Michigan State University Superfund Research Center MIFlowCyt_1.0.0 Contextual Data Curation

Contextual Data (Metadata) Curation

Overview

The purpose for this template is to harmonize the metadata collection for flow cytometry analyses and support the submission to the International Society for Advancement of Cytometry (ISAC) FLOWRepository.

Metadata will be provided by the data generators who are responsible for its accuracy at time of collection.

This MIFlowCyt template developed by the Michigan Sate University Superfund Research Center Data Management and Analysis Core is intended to be used with *DataHarmonizer*:

```
ſĠ
@article{,
   author = {Gill, I. S., Griffiths, E. J., Dooley, D., Cameron, R., Savic
Kallesoe, S., John, N. S., Sehar, A., Gosal, G., Alexander, D., Chapel, M.,
Croxen, M. A., Delisle, B., Di Tullio, R., Gaston, D., Duggan, A., Guthrie, J.
L., Horsman, M., Joshi, E., Kearny, L., Knox, N., Lau, L., LeBlanc, J. J., Li,
V., Lyons, P., MacKenzie, K., McArthur, A. G., Panousis, E. M., Palmer, J.,
Prystajecky, N., Smith, K. N., Tanner, J., Townend, C., Tyler, A., Van
Domselaar, G. and Hsiao, W. W. L.},
   title = {The DataHarmonizer: a tool for faster data harmonization,
validation, aggregation and analysis of pathogen genomics contextual
information.},
   journal = {Microb Genom},
   pages = \{\},
   DOI = \{10.1099/mgen.0.000908\},
   url = \{\},
   year = \{2023\},
   type = {Journal Article}
}
```

Table of Contents

- 1. Usage
 - a. Getting to know DataHarmonizer
 - b. Collecting investigation and study level metadata
 - c. Loading the MIFloCyt templates
 - d. Before adding metadata
 - e. Populating the templates
 - f. Missing metadata
 - g. MIATE organization
- 2. FAQ
- 3. Record of Changes

Usage

1. Familiarize yourself with *DataHarmonizer*

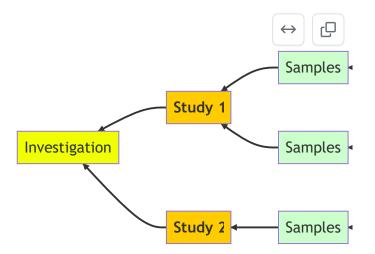
Before beginning to collect your (meta)data it is strongly recommended that you familiarize yourself with *DataHarmonizer* created by the Centre for Infectious Disease Genomics and One Health (CIDGOH), at Simon Fraser University. See the following video for an overview of *DataHarmonizer*.

Click here for DataHarmonizer video

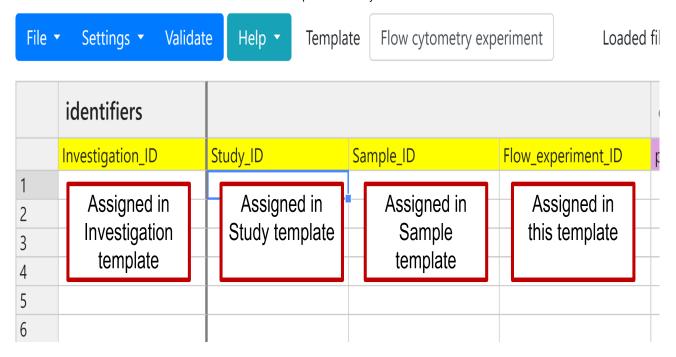
Notes:

- DataHarmonizer will not save your progress. It is good practice to save your work and upload the output next time you want to work on (meta)data collection.
- If you decide to fill some of the data offline using Excel, ensure that you meet the expectations for individual fields. These can be verified by uploading and validating.

2. Collecting investigation and study level metadata

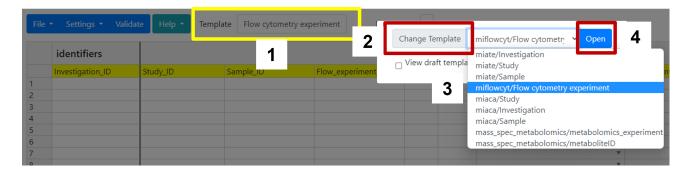


The MIFlowCyt template is intended to collect **Assay** level data and **must be complemented** with study level data. For study level data please see the *Minimum information about animal* toxicology experiments (MIATE) or Minimum information about cellular assays (MIACA) templates. This template will ask for identifiers which were assigned in these other templates as shown in the following image:

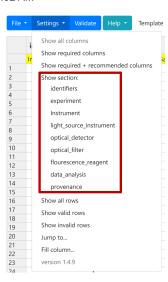


2. Loading the Minimum Information about Flow Cytometry Experiments (MIFlowCyt) template

Go to the Superfund Research Center_DataHarmonizer site and choose the MIFlowCyt template from the menu bar as shown below.



This will load the template into the browser. The template is broken down into **9 sections** which you can display at the same time or choose to display specific sections as shown here.



Here you can also choose to display only required, recommended, or required+recommended columns.

The sections broadly capture the following. For more details see section 6.

- Identifiers: Here you will enter all the relevant identifiers to map the data to individual samples/subjects, funding source, project details, etc. You will also assign a unique identifier to each row.
- **Experiment**: Here you will enter the general details about the experiment such as the purpose, contact person, QC procedures, and major conlusions. This is required for upload to FLOWrepository
- **Instrument**: Here you will enter the general details about the flow cytometer such as vendor, model, flow cell etc.
- Light source (instrument): Enter details about the flow cytometer light source.
- Optical detector (instrument): Enter details about the flow cytometer optical detector.
- Optical filter (instrument): Enter details about the flow cytometer optical filter.
- **Fluorescence reagent**: Enter details about the reagents used to detect the molecule or process of interest.
- **Data analysis**: Here you will include analysis procedures including raw and processed data file names.
- **Provenance**: This is autopopulated, please do not fill.

3. Before adding metadata

Before you begin to add your metadata:

• Review the fields in the template of the Validator application. For a summary overview you can select the Reference Guide from the Help menu.



- Review this protocol to make sure that you are using the template as intended.
- Review your dataset.

4. Populating the template

4A. Direct entry:

The template can be used as a typical excel template with some limited function. Simply type, copy & paste, or choose from dropdown menus to fill the relevant fields in the template. You can also copy from an excel document on your own computer.

For more information about the columns `double click` on the column name. This will describe the column and provide some examples.

To add multiple values (when allowed), seperate the values using the ';' character.

To delete an item from a dropdown menu, select it and hit the 'delete' button on your keyboard.

4B. (Meta)data import:

Using the `Save As...` option in the `File` menu you can export a template to use for populating. It will be missing some of the dropdown menus so you will have to confirm that your selections are valid. This can be done either by looking up the value or by uploading your template and choosing the `Validate` button.

4C. Hybrid:

In development

5. Unavailable or missing metadata

Every required field must have a value. If this information is missing then please select a null value from the dropdown menu indicating where it is:

- Not applicable
- Missing
- Not collected
- Not provided
- Restricted acces

If a term is missing in either a dropdown menu as a column in the *DataHarmonizer* template please submit a Github *issue* or contact Rance Nault

6. Template organization

Section	Required fields
Indentifiers	investigation_id, study_id, sample_id, flow_experiment_id
Experiment	experiment Variables, primary Contact, quality Control Measu
Instrument	
Light_source_instrument	
Optical_detector	
Optical_filter	
Fluorescence_reagent	analyte, detector, analyteReporter
Data_analyis	
Provenance	
→	

bold indicates a unique identifier to be used as key for database records.

Record of changes