Step-by-step tutorial: generating metadata and submitting to LIMS

Each set of data submitted to the LIMS requires a metadata file. The LIMS system provides tools to generate these metadata files. If you are generating metadata for less than three (< 3) samples, it's recommended that you use the metadata form, as this will be the fastest method to produce your metadata files. If you are generating metadata for more than three (> 3) samples, it's recommended that you use the metadata batch processor to generate your metadata files. This tutorial provides step-by-step instructions for both methods.

Metadata Form (recommended for < 3 samples)

1. Go to the LIMS metadata form.
2. Enter information into the form fields where necessary.
    1. **Required fields summary**:
        1. LIMS Account: The LIMS account is your account name: you can see this at the top right corner of the page. If you enter a non-existent account name, the system will tell you the account is invalid when you submit the form.
        2. Experimental ID: The experimental ID is the serial number analog for your sample. If you previously characterized your sample in the LIMS and are now submitting additional characterization information (e.g. you previously submitted an NMR spectrum and now you're submitting a mass spectrum), you **must** use the same experimental ID. If this is a new sample, then click the “Generate ID” button to generate a new unique identifier.
        3. Component: The simplest kind of molecule/chemical species as understood by the LIMS. Specify your component by common name or PubChem CID number (the latter is preferred!) and mark its component class: monomer, solvent, catalyst, or additional reagent (i.e. other).
        4. Experiment, Synthesis Dates: The date the experiment (characterization) was performed and the date the sample was synthesized. Recommendation: use the calendar tool over manual entry.
        5. Exp. Synopsis: A brief synopsis, or summary, of what's going on in this experiment. Recommendation: include a truncated version of your methods and instrument parameters here.
        6. Synthesis Platform: Indicate where/how the sample was synthesized.
        7. Instrument: Indicate the instrument from which the data you are submitting was acquired.
        8. Data File(s): Indicate which instrument data files are included as part of this metadata submission.
            1. ***New as of 2023/04/14***: When using the batch processor only, you may specify a data file folder name as the value for your data_files field in the template. This allows you to process dozens of files while entering only one value in the field (the name of the folder). You must ensure this folder exists in the LIMS Transfer folder alongside the metadata files.
        9. Data Category: Indicate the manner in which the data should be interpreted (i.e., is this a spectrum, or a chromatogram, or a mean squared displacement measurement?). Data categories are used to distinguish between different data representations that might be available from a singular raw data source file.
        10. Data Integrity: Indicate whether the identity or identities of the

component(s)/polymer(s)/composite is confirmed. In other words: have you confirmed that your sample is as you've described above? If so, mark it as verified. If not, mark it as unverified.

11. Data Availability: Indicate whether the data is accessible to all qualified BioPACIFIC MIP users or whether it should be restricted. For the LIMS beta phase, all data should be marked as available to all users.

2. **Conditional fields summary**:
    1. Parent Exp. ID (***New as of 2023/05/04***): The parent experimental ID is the serial number analog associated with the sample from which your current sample was derived. This is useful in cases where you e.g. collected fractions from an earlier synthesis and isolated a material from that fractionation. The new material, being derived from some previous material, has an atypical origin. We can capture this information to some degree by collecting the experimental ID of the parent from which the sample originated. This is an optional field.
    2. Polymer: If your sample is a polymer (i.e. an assembly of multiple components), then you will provide the common name or abbreviation (or, if available, PubChem CID number) in this field. You must also provide information concerning the following:
        1. *Polymer components*: Indicate which components belong to/were used to construct the specified polymer. You will do so by separating comments with the hash-tag symbol, #.
        2. *Component ratios*: Indicate the ratios of these components, expressed as either mass or mole ratios.
        3. *Chemical formula*: Provide the chemical formula, with repeating units indicated appropriately.
        4. *Molecular weight*: Provide the molecular weight in g/mol.
        5. *Polymerization mechanism(s)*: Indicate which mechanism(s) was/were employed to construct the polymer.
    3. Composite: If your sample consists of multiple polymers as part of a complex macromolecular species, then you must provide the common name your composite and include the following information:
        1. *Constituent ratios*: Indicate the ratios of the polymers used to build the composite, expressed as either mass or mole ratios.
3. **Fully optional fields summary**
    1. Some fields are always optional. These are highlighted in a bright, cyan blue color and can be found by clicking the document-shaped icon beside polymer and composite. These fields include:
        1. Polymer
            1. *Tg*: The glass-phase transition temperature, in degrees C.
            2. *Polydispersity*: The polydispersity index, PDI.
            3. *Isomeric SMILES*: The isomeric SMILES string
            4. *bigSMILES*: The bigSMILES string
        2. Composite
            1. *Polydispersity*: The polydispersity index, PDI
        3. Context Files
            1. Context files are optional file attachments. These are files that provide contextual information that would help another researcher better understand the unorthodox or unique analytical interpretations you have implemented for the data you have provided. Think of these files as “How-to” guides to help researchers understand how to understand

how to interpret your data. Context files can be of any file format, including raw code (e.g. *.py or *.m files).

3. After completing all of the required fields for your sample, click “Generate Metadata File”.
4. If there are no validation errors, you should see the something similar to the following at the top of the metadata form:

5. Click on the “Download Metadata” button. This will ask you to download your formatted metadata JSON file (to either a location of your choosing, your Downloads folder, or another location – whichever is specified by your browser settings).
6. Move/copy/save the JSON file to the LIMS Transfer folder location on the instrument. Consult with the project scientist and/or laboratory technician(s) overseeing the lab for more information.
7. Copy the instrument data files to the LIMS Transfer folder. The name of each file must match exactly with the names of the data and context (if provided) files in the metadata file. If the file names do not match, they will not be transferred to the LIMS.
8. At this point, you're finished! The system utilities of the instrument computer will take care of the rest, sending your metadata, data, and context (if provided) files to the LIMS. After a successful transfer, the files will be archived in a separate “archive” folder inside the LIMS Transfer folder.

Metadata Batch (recommended for > 3 samples)

1. Go to the LIMS metadata batch processor.
2. Open up the metadata batch template spreadsheet (the link to which is at the top of the page and can also be found here.
3. Make a copy of the spreadsheet (File → Make a copy) or, if Excel is available locally on your machine, download it (File → Download → Microsoft Excel).
4. Fill out the required cells. See the above “Required fields summary” and its counterparts for more information. Additionally, each column header should include a tool-tip description when you hover your cursor over it. There is only one unfamiliar field column in this spreadsheet, called “group_number”, and it is described below.
    1. group_number: Related sample information is “grouped” together in the spreadsheet via the use of the group_number column. All the contents of a particular sample, e.g. Sample JKL, should share the same group number, e.g. 4. All of the rows with group number 4 will be treated as a unit and parsed accordingly to produce a metadata file for that sample.
5. Enter your account name in the “LIMS Account” form field.
6. Browse for and select your metadata batch spreadsheet file from the Batch File form field.
7. Click Process Batch File.
8. If successful, you should receive a message like the following at the top of the form:

9. Click the “Download Batch Zip” button. As with “Download Metadata File”, it will either ask for the location of or automatically download to the location of your browser's specified Downloads folder. The difference compared to the single submission form is that, instead of downloading a single JSON file, you'll be downloading a zipped folder of JSON files (with a JSON file for each unique group in your metadata spreadsheet).
10. Unzip the folder's contents and put the metadata JSON files into the LIMS Transfer folder. It should look something like this:

11. Copy your data and context (if provided) files to the LIMS Transfer folder.

12. You're done – celebrate! The transfer scripts will take care of the rest and archive metadata files after successful transfer.

From:
https://bpm-wiki.cnsi.ucsb.edu/ - **NSF BioPACIFIC MIP Wiki**

Permanent link:
https://bpm-wiki.cnsi.ucsb.edu/doku.php?id=lims-guide-data-prep-tutorial&rev=1683222458

Last update: **2023/05/04 17:47**