EMN Metadata Template Instructions

# Purpose

The purpose of the templates are to incorporate advanced metadata and data into the datahub. With the ultimate goal to be able to search and sort across private (within HydroGEN projects) and public databases (HydroGen Consortium + Literature + Archival Data+ Materials Project).

## What is metadata

Metadata is like a digital notebook, it is information associated with an experiment, material test. Something you would put in an experimental section of a paper, or in your lab notebook.

# Metadata

Metadata template consists of **Metadata Fields** on left and **Metadata Values** on the right. The **Fields** will be searchable “fields” returning the **Values**.

Ex: search “laboratory” returns “LBNL, NREL, *institution name*”

Some fields are already part of the DataHub during upload of data. We are making this part of the process easier. Others that we add are added in levels, think about going deeper into describing the data and how it was acquired.

## Metadata Levels

Metadata is organized as a hierarchy of levels. This helps clarify the meaning of the metadata for both the producer and consumer (e.g., during search).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Lab (Dataset Level Metadata)** | | | |  |  |
| * Where was the experiment done, by whom, etc. * Comment * Hardware/Experimental setup | | | |  |  |
|  | **Experiment:Fixed (Resource Level Metadata)** | | |  |  |
|  | * Fixed experimental variables * What hardware was used | | |  |  |
|  |  | **Experiment:Variable** | |  |  |
|  |  | * Experimental conditions being varied * Specific to the experiment performed and technology   + Temperature   + Pressure | |  |  |
|  |  |  | **Sample** |  |  |
|  |  |  | * Specific to the sample that is used and characterized |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

### Tags

For all the metadata levels, you can have more than one section by adding a *tag* after the name of the level, separated by a colon, e.g. Sample:Material. Thus, for example, if you were performing an electrolyzer test and needed to characterize the material, MEA, and cell, you could use three metadata sections for the Sample level: Sample:Material ; Sample:MEA; Sample:Cell.

# Data

Contain two distinct areas. While data can be attached and uploaded separate of the metadata file, it can be included in columns, with the header **Data:**”name”. The second type is the searchable data “SearchData” these are figures of merit for the data that you want to come up in a search, for instance, maximum power density, a current at voltage, surface area value.

Searchable Data (SearchData)

* What is the main result that you want reported along with the sample name?
  + “My super catalysts” : current @ voltage
  + “Ultra membrane” : ASR

Tabular data

* Columns with headers and units, e.g. like in a spreadsheet

Raw data

* Files of arbitrary information, e.g., images or videos

# Making your own template

While templates exist for common experiments, you can create your own and create a standard that you’d like to follow. Here is the procedure.

1. Add field/value pairs for each metadata level (Lab, Experiment, Measurement, or Sample). We will call the metadata for a given level, a metadata "section".
   1. Create a "header" row that has the name of the metadata level (with a tag, if desired), and a blank cell to the right. The capitalization of the metadata level name is not important ("experiment" or "EXPERIMENT" would also work).

|  |  |
| --- | --- |
| Experiment |  |

* 1. Below the header row, add rows with the metadata. Put the name of the field in the first column, and the value in the second column. Any spaces, tabs, or linebreaks in the name of the field will be converted to underscores, so it is recommended to just use underscores from the start to avoid confusion.

|  |  |
| --- | --- |
| Experiment |  |
| data\_source | LBNL LTE test stand |
| environment | Indoor experiment |

* 1. Repeat this for each metadata section. In a spreadsheet, you can put metadata sections in columns and/or in rows. There must be one row of blank cells between sections (otherwise the second section will be seen as just more fields). You can do this in a CSV file as well, but it may be much more convenient to merely stack the sections up in rows.

|  |  |  |  |
| --- | --- | --- | --- |
| Experiment |  | Sample:MEA |  |
| data\_source | LBNL LTE test stand | type | CCM |
| environment | Indoor experiment |  |  |
|  |  |  |  |
| Measurement |  | Sample:Material |  |
| Direction | x | anode\_catalyst\_material | Pt/C |
| Current\_or\_voltage\_step | 1 | anode\_catalyst\_particle\_size(nm) | 5 |

1. Add SearchData section(s)
   1. A SearchData section is just like a metadata section, with a header and fields.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Experiment |  | Sample:MEA |  | SearchData |  |
| data\_source | LBNL LTE test stand | type | CCM | Potential\_at\_2 A/cm2\_(V) | 1.94 |
| environment | Indoor experiment |  |  | Current\_at\_1.6V\_(mA/cm2) | 1.87 |
|  |  |  |  |  |  |
| Measurement |  | Sample:Material |  |  |  |
| Direction | x | anode\_catalyst\_material | Pt/C |  |  |
| Current\_or\_voltage\_step | 1 | anode\_catalyst\_particle\_size(nm) | 5 |  |  |

* 1. You can use tags, just like with metadata, to have multiple SearchData sections. However, each tag should match a corresponding tag on a Data section. The simplest case of this is that there is only one Data section, with no tag, and the SearchData section also does not have a tag.

1. Add Data section(s)
   1. Add a "Data" header row (like for a metadata section)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Experiment |  | Sample:MEA |  | SearchData |  | Data |  |
| data\_source | LBNL LTE test stand | type | CCM | Potential\_at\_2 A/cm2\_(V) | 1.94 |  |  |
| environment | Indoor experiment |  |  | Current\_at\_1.6V\_(mA/cm2) | 1.87 |  |  |
|  |  |  |  |  |  |  |  |
| Measurement |  | Sample:Material |  |  |  |  |  |
| Direction | x | anode\_catalyst\_material | Pt/C |  |  |  |  |
| Current\_or\_voltage\_step | 1 | anode\_catalyst\_particle\_size(nm) | 5 |  |  |  |  |

* 1. Add a second header row with names for each column (variable), including units in parentheses. You can have as many data columns as you wish. The first blank cell in that row is the end of the data columns.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Experiment |  | Sample:MEA |  | SearchData |  | Data |  |
| data\_source | LBNL LTE test stand | type | CCM | Potential\_at\_2 A/cm2\_(V) | 1.94 | X (J / mA/cm) | Y (Ewe / V) |
| environment | Indoor experiment |  |  | Current\_at\_1.6V\_(mA/cm2) | 1.87 |  |  |
|  |  |  |  |  |  |  |  |
| Measurement |  | Sample:Material |  |  |  |  |  |
| Direction | x | anode\_catalyst\_material | Pt/C |  |  |  |  |
| Current\_or\_voltage\_step | 1 | anode\_catalyst\_particle\_size(nm) | 5 |  |  |  |  |

* 1. Fill in the data for each column

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Experiment |  | Sample:MEA |  | SearchData |  | Data |  |
| data\_source | LBNL LTE test stand | type | CCM | Potential\_at\_2 A/cm2\_(V) | 1.94 | X (J / mA/cm) | Y (Ewe / V) |
| environment | Indoor experiment |  |  | Current\_at\_1.6V\_(mA/cm2) | 1.87 | 2300 | 2.1 |
|  |  |  |  |  |  | 2219.1 | 2.4 |
| Measurement |  | Sample:Material |  |  |  |  |  |
| Direction | x | anode\_catalyst\_material | Pt/C |  |  |  |  |
| Current\_or\_voltage\_step | 1 | anode\_catalyst\_particle\_size(nm) | 5 |  |  |  |  |