**User Defined Analysis Simulation Data Preparation Guidance**

# Format and Structure of Simulation Data

* There are two parts of the simulation data: (1) meta data, and (2) sensor data.
* All data are stored in the directory: “simulation\_data/{technical\_route}/{building\_type\_or\_name}/{weather}/”, for example: “simulation\_data/general\_guidance/small\_commercial\_building/TN\_Knoxville/”.
* The user needs to provide data only when the “technical\_route” is “user\_defined\_analysis”. If “technical\_route” is “general\_guidance”, the analysis framework will automatically download relevant simulation data from the FRP case study.

# Meta Data

* The meta data file is a csv file that aggregates high-level simulation information from each fault simulation model (representing combinations of fault type and intensity). The meta data file name should match the value stored in the “simulation\_metadata\_name” field in the “config.json” file.
* An example of meta data is shown below:

Graphical user interface

Description automatically generated with low confidence

* Required columns include: (1) “id”, (2) “fault\_type”, (3) “unmet\_hours\_during\_occupied\_cooling”, (4) “unmet\_hours\_during\_occupied\_heating”. (3) and (4) are used for sensor cost analysis.
* Each id requires one simulation data csv file in the same folder.
* The meta data file must include a row entry with fault\_type set to “baseline”. This represents the non-fault (normal operation) simulation case, which is a key input to the analysis framework.

# Simulation Data

* Simulation data are simulation results in csv format (our FRP case study dataset is made up of EnergyPlus results, as an example). There is one simulation data file for each simulation.
* An example of simulation data is shown below:

Table

Description automatically generated

* The simulation data file name should follow the convention: “{simulation\_id}\_sensors.csv”, for example, “2ce70b53-365f-4214-8625-2dbcdad712e9\_sensors.csv”. And the “simulation\_id” should have a corresponding row entry in the meta data file.
* The first column should be the timestamp. There is no requirement for the column name of the timestamp or for the time format. The timestep should be 15 minutes. The ideal simulation duration is one year, resulting in 4 \* 8760 = 35040 rows in the simulation data file.
* To facilitate sensor cost analysis, the simulation data file should include columns named “electricity\_facility [W]”and “gas\_facility [W]”, which are common EnergyPlus outputs.