**iUTAH Data Collection Plan**

Please complete the following for each data collection effort with any funding support from iUTAH. ***All fields are required.*** Note that if there are substantive changes to the types of data, methods, data formats, products, or availability of data, an updated plan should be submitted.

**Your Name and Affiliation:**

Please provide your name and affiliation (iUTAH participating organization).

**Effort Name:**

Please provide a brief name for your data collection effort.

**Collaborators:**

Please list potential collaborators and affiliations associated with your data collection effort. This is the first step in planning for authorship of datasets and other publications resulting from this effort.

**Summary and Purpose:**

Please provide a brief description of your data collection effort. Include information on the purpose of and the research questions to be addressed by the data collection/creation. Include a description of how the effort fits into iUTAH goals and objectives.

**Data Collection Methods:**

Please thoroughly describe the methods you will use to collect/create your data.

**Location of Data Collection:**

Please describe where your data will be collected. How many sites?

**Timing of Data Collection:**

Please describe the timing of the data collection/creation. When will it begin and end? Is it periodic or continuous?

**Parties Responsible for Data Collection:**

Who will actually collect/create the data? Include names and affiliations if known.

**Individual(s) Responsible for iUTAH Metadata Completion:**

Who will draft the metadata describing the data for the iUTAH publication system (<http://repository.iutahepscor.org>). Include name and affiliation.

**Datasets Expected to be Generated:**
Please complete this table to describe each dataset(s)/data product(s) that will result from your data collection effort. Specific fields are described here. ***All fields are required.***

* Dataset: Please provide a separate entry for each dataset or data product expected to be generated from the effort.
* Authorship: Who do you anticipate will be the published authors of this product? Could be individuals or a group (e.g., iUTAH GAMUT Working Group).
* Data Type: Please indicate the Data Type (A, B, C, or D) using the typology from the iUTAH data policy. This is found below the table.
* Resulting Data Format: Please describe the file format(s) that will be used for the final data products.
* Data Storage: Please describe how the data will be stored and managed during data collection.
* Timeframe: Please indicate dates (or date range) when you anticipate that the data product will be released/published. Guidelines for timeframe for each data type are found below the table.
* Access During Collection: Please indicate any restrictions on data access during data collection. Names of individuals may be used.
* Access After Completion: Please indicate any restrictions on data access after data collection.
* Anonymization: If anonymization of the data product is required (i.e., per IRB) before sharing, please describe how this will be done.
* Publication: Will the data product be shared/published via the iUTAH Data Publication System (<http://repository.iutahepscor.org>)? If not (rare), what is your plan for making data broadly available?

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Authorship** | **Data Type\*** | **Resulting Data Format** | **Data Storage** | **Timeframe\*\*** | **Access During Collection** | **Access After Completion** | **Anonymization** | **Publication** |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

**\*Data Typology:**

**Type A** - Primary iUTAH datasets or research products. These include raw and QA/QC calibrated sensor data from iUTAH facilities, baseline sampling datasets across iUTAH facilities and sites, and general datasets collected by iUTAH for the community of iUTAH participants.

**Type B** - Datasets or other research products for which monetary or material support was provided by iUTAH, but that are created by a specific investigator, student, or iUTAH group to support a particular research question or goal.

**Type C** - Type A and Type B datasets or products that include personally identifiable information or information about human subjects/participants and are subject to IRB restrictions.

**Type D** – Datasets or other research products procured by iUTAH or iUTAH participants supported by iUTAH that are subject to licensing, copyright, or use restrictions/agreements from the data source that may prohibit general distribution of the data.

**\*\*Data Typology Timeframe**:

**Type A** - Automated data streams from iUTAH facilities will be streamed directly into live databases and will be made available online in near real time. Quality controlled and derived data products from iUTAH facilities will be published within six months of data collection. All other primary datasets will be published as they become available (e.g., as soon as results are created).

**Type B** - Finalized data will be submitted to the iUTAH MDF within one year of the completion of data creation activities. Students collecting data must submit their finalized data as a condition of their thesis/dissertation defense. For long running data creation activities (i.e., efforts that last longer than one year), the following will be required:

1. The initial metadata description will be reviewed and updated at least once per year.
2. Data will be submitted at least every 6 months. Data will not be released until the dataset is finalized by the data creator.
3. Finalized data will be submitted within one year of collection or by the end of the project, whichever comes first.

**Type C** - Type C datasets will be subject to time requirements described for Type A and Type B datasets. However, they may require the additional step of anonymization or aggregation with methods described in the Data Collection Plan.

**Type D** - Type D datasets will be published as soon as possible (within one month) and to the greatest extent allowable by the licensing, copyright, and/or data use agreements under which they were created/procured.  Some Type D datasets may be permanently restricted.