Data Management Plan

Types of data

As part of this project, many types of data will be collected. This includes environmental resource data gathered from sensors including temperature, humidity, wind direction and speed, and solar irradiance. It also includes electrical grid data including voltage, current, power, and reactance. Finally, it includes power and energy data from distributed generation devices such as photovoltaic systems.

All of the above data will be experimental measurement data obtained through physical harvesting of data from the surrounding environment and from energy production and storage devices. The data will be captured using standard environmental sensors such as anemometers, power meters, pyrometers, etc. It will be stored in a custom-built database, as discussed next.

Data and metadata standards

While there are many systems that have been developed to store environmental sensor data, such as GeoCENS, Pachube, and the Berkeley Sensor Database, there are no commonly recognized standards for the formatting, storage, or transmission of the data we will be collecting, storing, and analyzing. Because of this, we plan to store our data using a custom, but publicly documented, set of database schemas in an internet-accessible server running a standard open source LAMP (Linux, Apache, MySQL, Python) stack. We also plan to design and implement a REST protocol for web service oriented access and manipulation of the data.

Policies for access and sharing and provisions for appropriate protection and privacy

As noted above, we plan to design and implement a web service that will enable external access to the data by interested researchers. The environmental data that we will collect, store, analyze, and publicize will not reveal personal characteristics of users. We will restrict real-time access to power and/or energy consumption data, or aggregate this data as required, in order to prevent users from being able to gain behavioral insight from patterns of energy usage.

Policies and provisions for re-use and re-distribution

We plan to make the data collected in this research freely available under a Creative Commons Attribution CC BY license. This license will let others use, distribute, and analyze the data we collect without restriction as long as they credit us as the original creators of the data.

We believe that the data we collect will be of interest to others developing ”smart microgrid” systems as it will provide, at a minimum, baseline data for environmental conditions we experienced during the course of our research.

Plans for archiving and preservation of access

During the course of this research, we will be storing the data in a web accessible database as discussed above. Archiving the data past the conclusion of the study will be done via data archiving services provided by the University of Hawaii. All research reports and collateral documents created in response to the data will also be permanently archived through the University of Hawaii technical report services. The University of Hawaii implements standard best practices for data storage backup and retrieval, including off-site storage, redundant power supplies, RAID disk storage, and so forth.