1. **Introduction:**

This document is intended to serve as a road map for student work on the Carleton MySQL Energy Database and a vision for overall objectives and next steps beyond the 2011-2012 academic year. This project is focused on developing a MySQL database that uses Python and PHP to clean, compile, query and export data. It is intended to roll into a larger initiative to develop a permanent, web-accessible database that will compile Carleton’s energy data into a single repository , allow data entry via the web for a limited group of database administrators, and make Carleton energy data easily seen and accessed by the greater Carleton community through a simple web interface.

This project plan document is organized into the following sections:

* Project Stakeholders
* Communication Plan
* Project Scope(s) and Current Status
* Project Schedule
* Document Control

The master copy of this project plan will be maintained by Martha Larson as updates are required. The original draft and all revisions will be saved in:

**I:\Committees\Sustainability\Energy Information Database\MySQL Database**

1. **Project Stakeholders:** *(let me know if there are others we should mention here…)*
   1. **Primary** – this is the core project team and includes key project advisors, creators, and users who should be kept regularly informed of project progress and next steps:
      1. Martha Larson
      2. Paula Lackie
      3. Jie Lin
      4. Daniel Alabi
      5. Data-focused Sustainability Assistants: Max Timm, Ryan Noe
   2. **Secondary** – these are managers and specialized experts who can serve as resources and should be provided with high-level updates of project progress and activity:
      1. Steve Spehn, Facilities Director
      2. Mitch Miller, Maintenance Manager
      3. Joel Cooper, ITS Director
      4. Richard Goerwitz, Data Warehouse Administrator
      5. Rich Graves, ITS Senior Administrator
      6. Chris Dlugosz, ITS Senior Administrator
   3. **Tertiary –** these are future database users or administrators who should be provided with a demonstration and allowed to use the database once the it’s is in working order.
      1. Denise Gillen, facilities accountant
      2. Kathy Beckers, facilities administrative assistant
      3. Maintenance + steam plant staff
      4. Aaron Swoboda, Economics / ENTS Professor
      5. Jeff Ondich, Computer Science Professor
      6. David Liben-Nowell, Computer Science Professor
      7. Tun Myint, Political Science Professor
      8. Dan Hernandez, Biology / ENTS Professor
      9. Kim Smith, ENTS Director / Professor
      10. Sustainability Assistants: Courtney Dufford, Laura Henry, Alex Lai, Nina Whitney, Libby Nachman
      11. Chris Remley, Res Life Sustainability Committee Chair
      12. Environmental Advisory Committee (EAC)
2. **Communications Plan:**
   1. Primary stakeholders will schedule meetings or workshops on an as-needed basis to achieve project milestones per the schedule. Weekly email updates will be issued to all primary stakeholders by those working on specific tasks.
   2. Secondary stakeholders will be updated as follows:
      1. Martha Larson will update Fred Rogers, Steve Spehn in regular one-on-one management meetings
      2. Martha Larson will update Mitch Miller, Rich Goerwitz, Joel Cooper in weekly (Monday) Facilities / ITS task force meetings
      3. Paula Lackie will update Joel Cooper in regular management meetings [do you have those]?
      4. Paula Lackie will update Rich Graves and Chris Dlugosz as needed for areas that cross into their specific areas of expertise.
   3. Tertiary stakeholders will be updated by primary stakeholders and provided with database demonstrations or training at appropriate progress points.
3. **Project Scope(s) and Current Status:** *(Daniel / Jie, please complete the table below and any items in red that require your input.)*
   1. **Project #1: Existing Data Cleaning and Compiling**

**T**heprimary purpose of this project is to compile existing energy data from multiple sources and varied formats into a single database repository in a clean and consistent format. The database must be extensible, fully exportable and easily migrated to a new database system in the future (1-3 years).

A Summary of Project #1 scope and progress to date is as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Data Set** | **Data Source** | **Smallest Increment** | **Oldest Data Entry Date** | **Newest Data Entry Date** | **Program used to compile existing data (Python or PHP) + filename + file location** | **Program used to upload new data (Python or PHP)+ filename + file location** |
| Building Data (building names, dates online, square footage) | Eventually link to Colleague, for now a static table that administrators will manually update as needed. | N/A | N/A | N/A | *Moodle:* Martha's Energy Data Management Projects Siemens Point data and Building data This zip file contains the txt version of building data and siemens point data, which are used by the two php scripts - senddata.php and campusData.php.  campusData.php collates the siemens data excel files into a php array that senddata.php can handle. It calls the functions within senddata.php directly without writing the data into a temporary file.  senddata.php is the main php file we use to access the MySQL database and insert data into the database. |  |
| Population Data (number of staff or residents per building by date) | This will need to be a static table that is manually updated by administrators. | N/A | N/A | N/A |  |  |
| Steam Plant Data (all water) | Operator Logs | Daily | *08/01/1996* | *12/01/2011* | Filename: water\_gets\_intodb.php  File location: Temporary location in github  Data file: Everything in the same folder can be a data file. (Don’t put irrelevant files in the same folder) | Filename: water\_gets\_intodb.php  File location: Temporary location in github  Data file: Use the same format as steam plan logs after 2008. Follow the same naming convention for data files. |
| Steam Plant Data  (Steam, Gas) | Operator Logs | Daily | *01/01/2008* | *12/01/2011* | Filename: putInDB.php  File location: on one of the computers in Paula’s office. (Need uploading to Moodle)  Note: Many helper functions are inside senddata.php. | Filename: putInDB.php  File location: on one of the computers in Paula’s office. (Need uploading to Moodle)  Note: Many helper functions are inside senddata.php.  Data file: Use the same format as steam plan logs after 2008. Follow the same naming convention for data files. |
| Steam Plant Data  (Oil) | Operator Logs | Daily | *01/01/2009* | *12/01/2011* | Filename: putInDB.php  File location: on one of the computers in Paula’s office. (Need uploading to Moodle)  Note: Many helper functions are inside senddata.php. | Filename: putInDB.php  File location: on one of the computers in Paula’s office. (Need uploading to Moodle)  Note: Many helper functions are inside senddata.php.  Data file: Use the same format as steam plan logs after 2008. Follow the same naming convention for data files. |
| Siemens Steam Data: 17 bldgs | Carleton Meters | Daily | *07/01/2009* | *11/26/2011* | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addCampusSteam();  Data filename: CAMPUS STEAM.txt  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | If the new data are in the same format as “CAMPUS STEAM.txt”, use the same code listed on the left.  If not, no code exists yet. |
| Siemens Water Data:17 bldgs | Carleton Meters | Daily | *11/18/2009* | *11/26/2011* | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addCampusWater();  Data filename: CAMPUS WATER.txt  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | If the new data are in the same format as “CAMPUS WATER.txt”, use the same code listed on the left.  If not, no code exists yet. |
| Siemens Electric Data: 17 bldgs | Carleton Meters | Daily | *01/04/2009* | *11/26/2011* | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addElectricity();  Data filename:  “CAMPUS ELECTRIC.txt”  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | If the new data are in the same format as “CA MPUS ELECTRIC.txt”, use the same code listed on the left.  If not, no code exists yet. |
| Siemens kwdemand (main elec) | Carleton Meters | 5 min. | *01/22/2006* | *11/12/2011* | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addkwDemand ();  Data filename: compiled as “20xx.txt”  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addkwDemand ();  Data filename: compiled as “20xx.txt”. If the format does not change, the same code can be used again. |
| Siemens Weekly Useage (main elec) | Carleton Meters | Daily | *01/04/2009* | *11/26/2011* | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addkwWeekly ();  Data filename: compiled as “20xx.txt”  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addkwWeekly ();  Data filename: compiled as “20xx.txt”. If the format does not change, the same code can be used again. |
| Xcel Bills: Campus Electricity | Xcel Meters | Monthly | *Date* | *Date* | *Python / PHP*  *Filename*  *File Location* | *Python / PHP*  *Filename*  *File Location* |
| Xcel Bills: Campus Natural Gas - Firm | Xcel Meters | Monthly | *Date* | *Date* | *Python / PHP*  *Filename*  *File Location* | *Python / PHP*  *Filename*  *File Location* |
| Xcel Bills: Campus Natural Gas – Interr. | Xcel Meters | Monthly | *Date* | *Date* | *Python / PHP*  *Filename*  *File Location* | *Python / PHP*  *Filename*  *File Location* |
| Xcel Bills: Off-Campus Houses - Electrical | Xcel Meters | Monthly | *Date* | *Date* | *Python / PHP*  *Filename*  *File Location* | *Python / PHP*  *Filename*  *File Location* |
| Xcel Bills: Off-Campus Houses – Natural Gas | Xcel Meters | Monthly | *Date* | *Date* | *Python / PHP*  *Filename*  *File Location* | *Python / PHP*  *Filename*  *File Location* |
| Solar PV  (Cassat-Mem) | TBD | Daily or Hourly  (TBD) | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE |
| Solar Thermal  (Cassat-Mem) | TBD | Daily or Hourly  (TBD) | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE |
| Wind Turbine #1 Data | Vestas SCADA System | Hourly | 04/20/2011 | 12/19/2011 | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addWindTurbine ();  Data filename: compiled as “201xHourlyWindTurbineData.txt”  Note: Many helper functions are inside senddata.php. This file should be in the same directory as campusData.php. | Filename: campusData.php  File location: Moodle: **Siemens Point data and Building data**  Specific function: addWindTurbine ();  Data filename: compiled as “201xHourlyWindTurbineData.txt”.  If the format does not change, the same code can be used again. |
| Wind Turbine #2 Data | GE SCADA System | 10 min. | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE |
| Monthly Diesel Generator Testing Logs (hand written) | Operator Logs | Monthly | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE | FUTURE PHASE |

The data is stored in a [ server type ] at [ server I.P. address, other location data ] in the Carleton Data Warehouse. Database administrators as of the date of this plan include [list all names]. To become a database administrator one must [outline the process, i.e. contact so-and-so, get approval from so-and-so, etc…]

* 1. **Project #2: User Interface – Data Query and Export**

Theprimary purpose of this project is to create a user interface that allows users other than database administrators to query the database and export specific data sets to \*.xls or \*.csv file format based on a set of user-defined criteria. The user interface is currently located at [D/J – provide web address here]. Specifically:

* + 1. Users must be able to use a drop-down menu to select the data type, units, data source (if multiple available, i.e. electrical data from Xcel invoices vs. Carleton meters), start date, end date, time increment (hourly, daily, monthly, annually) and divisor (none, per square foot, per person)
    2. Users must then see a checkbox list of all buildings for which the selected parameters are available. They must be able to select one or multiple checkboxes to pull data for multiple buildings and/or the main campus.
    3. If a user selects a data increment less than the smallest available increment, an error message must appear directing the user to choose a larger increment.
    4. If a user selects a start date earlier than the earliest available start date, an error must appear or the resulting data set must read “No Data” for dates prior to the available data start date.
    5. If a user selects a time increment that is larger than the smallest available, data should roll up from the smallest available increment. The calculation should be performed internal to the query.
  1. **Project #3: User Interface – Data Input**

Theprimary purpose of this project is to create a user interface that allows a limited user group (only those with administrative rights) to enter data into the database and edit incorrect or missing data entries. Specifically:

* + 1. For single entries, users should be able to call up a form that allows them to enter data date / time, source, time increment, and value(s).
    2. For bulk entry, users should be able to download an Excel template with field names and sequence that match the database schema. The template should clearly indicate which fields are mandatory. Users will copy / paste new data into the template (leaving the column headers intact) and click a button that will run a script to import the data into the database.
    3. Where possible, data entry into the template should be limited by a drop down menu (i.e. source, fuel type) or a format constraint (i.e. where dates or numbers are required).
    4. If a user attempts to upload a data entry template with missing data in mandatory fields, an error message must appear stating that there is missing data. If possible, the error message should state which row or column has the missing data.
    5. If a user attempts to upload a data entry template with the wrong data format in a particular field (i.e. text where a date or number are expected), an error message must appear stating that there is data in a particular row or column that does not match the required format.
  1. **Project #4: User Interface – Graphing and Reporting**

Theprimary purpose of this project is to develop a stock grouping of standard graphs and reports and/or a link to a data visualization and graphing software program.

* 1. **Project #5: User Interface – Web Integration**

Theprimary purpose of this project is to develop a web accessible “dashboard” to display “real-time” energy data and graphs, allow broader distribution of Carleton energy data and provide a “Data Repository” page on the Carleton sustainability website.

1. **Project Schedule:** 
   1. **Spring Break Objectives (March 19 – 23 2012):**
      1. Complete the progress matrix in section IV.a above – DA, JL
      2. Practice database demonstration – DA, JL, PL, ML
      3. Database demonstration for Rich Goerwitz – DA, JL, PL, ML
   2. **Spring Term Objectives (March 26 – June 1 2012):**
      1. Complete Project #1, Project #2, Project #3 – DA , JL
      2. Beta test Project #1, Project #2, Project #3 – DA, JL, PL, ML, RN, MT
      3. Present results to secondary stakeholders
      4. Develop graphics / reporting mock-up – ML, RN, MT, PL
   3. **Summer / Fall Term Objectives (June – November 1 2012):**
      1. Beta test Project #2 and Project #3 with tertiary stakeholders
      2. Develop detailed work plan for Project #4 and Project #5
      3. Research options for long-term database
      4. Develop detailed work plan for long-term database
      5. Implement new work flow for users of Project #3 data entry interface
2. **Document Control:** 
   1. Primary file repository:

The campus network will be the primary repository for documentation related to this project. All documents related to this project will reside in the following parent folder (contact Martha Larson at [mlarson@carleton.edu](mailto:mlarson@carleton.edu) to request access to this folder):

I:\Committees\Sustainability\Energy Information Database

* 1. Relevant sub-folders within the Energy Information Database folder are as follows:
     1. Raw data files (uncompiled):

I:\Committees\Sustainability\Energy Information Database\**Energy Data**

* + 1. Documentation related to the “Stop Gap” database:

I:\Committees\Sustainability\Energy Information Database\**MySQL Database**

* + 1. Documentation and files related to the future permanent database:

I:\Committees\Sustainability\Energy Information Database\**Future Database**

* + 1. Documentation related to research or demos of third-party software systems that may be suitable candidates for the future permanent energy database:

I:\Committees\Sustainability\Energy Information Database\**Database Research**

* 1. File Naming Conventions:
     1. All files will be named as follows: YYMMDD\_document title
     2. All file revisions will be made by the original author (unless specifically delegated by that person) retain the same original YYMMDD but add revision number to the title as follows: YYMMDD\_document title\_rev#
  2. File Edits + Revisions:
     1. Anyone with access to the file system can upload new documents.
     2. Edits to existing word files should be made using “track changes” and submitted to the original author for review and entry into the master copy stored on the campus network.
     3. Edits to existing Excel files should be highlighted and submitted to the original author for review and entry into the master copy stored on the campus network.
     4. Edits to database files [Paula, Jie, Daniel – please define process here]