HCI 598  
M2: Problem Space  
Ashley Otero

## Users

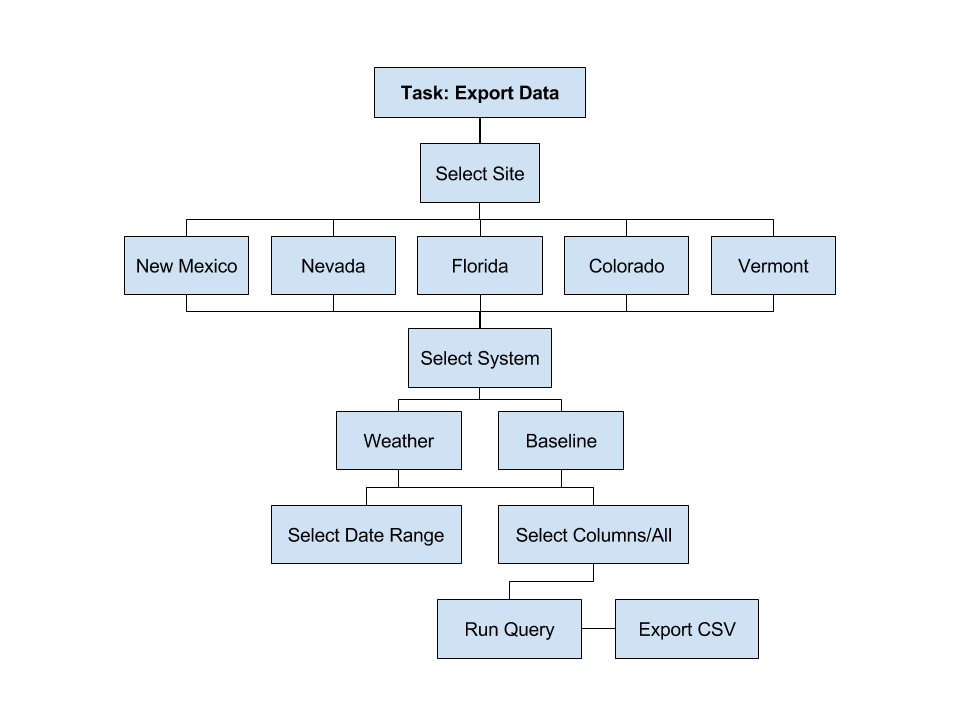
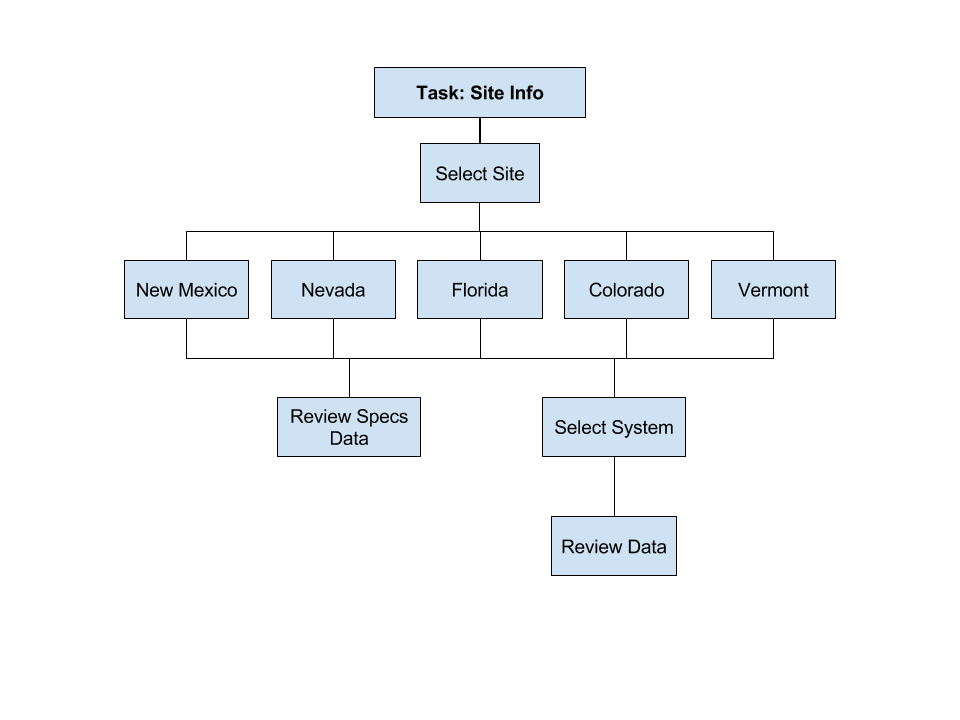
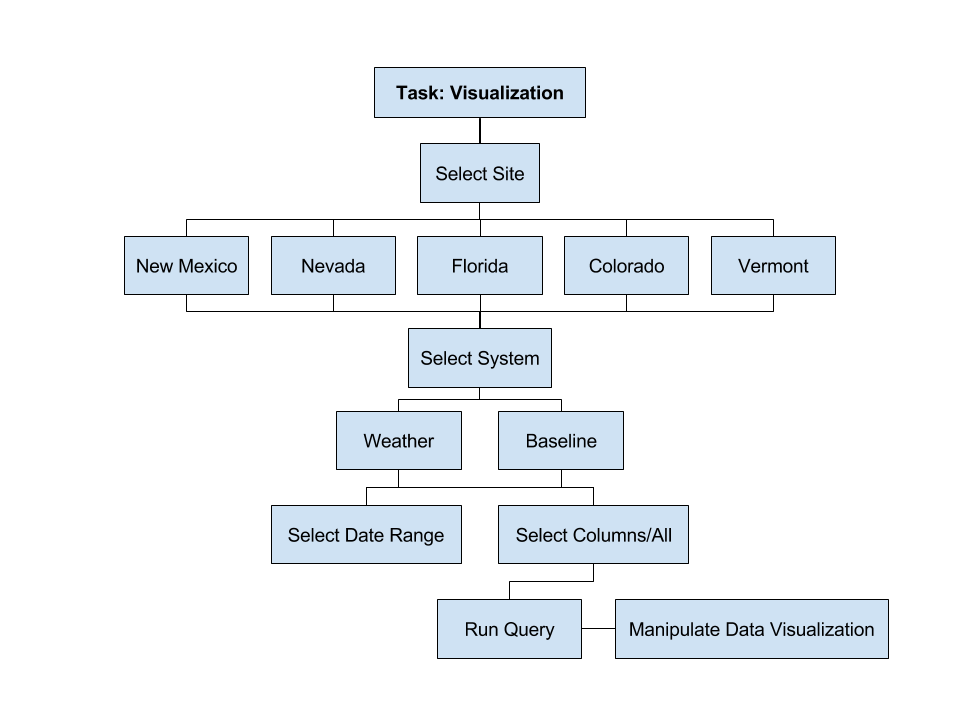
There are multiple stakeholders for my system, but the primary is the Department Of Energy, as they are the owner of the system, the final “customer,” and funding source for the development of the Regional Test Center (RTC) data project. Additionally, researchers at Sandia National Laboratories (SNL) in the Solar and Photovoltaic departments are also very interested in the data. Finally, the solar and photovoltaic industry partners such as module and inverter manufacturers, racking companies, and installers would also like to access and analyze the data. Some of this data may be proprietary, and therefore is unable to be shared publically.

The types of users my system will be primarily targeting are researchers and industry representatives. I will be focusing my design and development time on researchers, as I have the most access to them, in terms of ongoing professional relationships. This will allow for multiple iterations and feedback cycles to be completed during the course of this project. These researchers are typically highly scientific and technical people, and most have a Ph.D. in their field. While not designers, developers, or programmers, they are intimately familiar with the workings and meaning of data collected by the photovoltaic systems and have demands on what type of visualizations will be most persuasive and valuable for the RTCs. All have basic computing skills and immeasurable knowledge in the varying data types across the five different RTC sites. I will be conducting one on one interviews as well as using internal and external internet capabilities to gain greater insight into their background and priorities

## Tasks

There are three main tasks that all users who access the RTC data system will want to accomplish:

1. Export data based on a specific set of user-defined criteria.
2. Find site specific information such as technical specifications and location.
3. Manipulate a visualization of data based on a specific set of user-defined criteria.

These tasks are displayed in the task analysis diagrams below:  


* **Data Types - Baseline:** 
  + TIMESTAMP
  + RECORD
  + Sys1Vdc\_Avg
  + Sys1Str1Idc\_Avg
  + Sys1Str2Idc\_Avg
  + Sys1Vac\_Avg
  + Sys1Iac\_Avg
  + Sys1Wac\_Avg
  + Sys1VARac\_Avg
  + Sys1PowerFactor\_Avg
  + Sys1Frequency\_Avg
  + ModTemp1\_Avg
  + ModTemp2\_Avg
  + ModTemp3\_Avg
  + ModTemp4\_Avg
  + ModTemp5\_Avg
  + ModTemp6\_Avg
  + ModTemp7\_Avg
  + ModTemp8\_Avg
  + Sys2Vdc\_Avg
  + Sys2Str1Idc\_Avg
  + Sys2Str2Idc\_Avg
  + Sys2Vac\_Avg
  + Sys2Iac\_Avg
  + Sys2Wac\_Avg
  + Sys2VARac\_Avg
  + Sys2PowerFactor\_Avg
  + Sys2Frequency\_Avg
  + ModTemp9\_Avg
  + ModTemp10\_Avg
  + ModTemp11\_Avg
  + ModTemp12\_Avg
  + ModTemp13\_Avg
  + ModTemp14\_Avg
  + ModTemp15\_Avg
  + ModTemp16\_Avg
  + RefCell1Irrad\_Avg
  + RefCell1Temp\_Avg
  + RefCell1R\_Avg
  + RefCell2Irrad\_Avg
  + RefCell2Temp\_Avg
  + RefCell2R\_Avg
  + LocalAmbientTemp\_Avg
  + POAIrrad1\_Avg
  + ICP7019\_1CJCTemp\_Avg
  + ICP7019\_2CJCTemp\_Avg
  + ICP7019\_3CJCTemp\_Avg
  + ICP7019\_4CJCTemp\_Avg
  + ICP7019\_5CJCTemp\_Avg
  + ICP7019\_6CJCTemp\_Avg
  + ICP7019\_7CJCTemp\_Avg
* **Data Types - Weather:**
  + TIMESTAMP
  + RECORD
  + Global\_Wm2\_Avg
  + Direct\_Wm2\_Avg
  + Diffuse\_Wm2\_Avg
  + Pressure\_mBar\_Avg
  + WS\_ms\_Mean
  + Wdir\_Mean
  + Wdir\_Std
  + WS\_ms\_Std
  + WS\_ms\_3sec\_Max
  + WD\_deg\_SMM
  + Temp\_C\_Avg
  + RH\_pct\_Avg
  + Panel\_Temp\_C\_Avg
  + Battery\_V\_Avg
  + ProcessTime\_S\_Max
  + Global\_mV\_Avg
  + Rain\_mm\_Tot
  + Rain\_mm\_Daily
  + Temp\_CMP22\_C\_Avg

Researchers will be using their work computers to access the RTC data application. This will either be at their office environment with sit/stand desks and ergonomic chairs, or at their telecommute office, which may have any number of setups. Likely, this data, though it will be able to be pulled and manipulated on the fly, will not be accessed in a rushed manner, as it will require the experience and careful analysis of the researcher.

## Analysis

Currently the system does not actually exist in any meaningful way. The data is compiled by each system and sent to a centralized repository where it stores daily data in the form of .dat files. These .dat files are not human-readable and often contain previous day or days of information, leading to an overlap if compared side by side. This data is not easily converted to a usable form such as a CSV, rendering it functionally useless for comparison over time and comparison between systems. My system will ingest current daily data into a database to allow for easier data cleanup tasks, such as resolving daily data redundancy.

The system will also allow users a way to export data as a CSV, which will be helpful in allowing them to manipulate the data outside of the system in whatever way they need to. Since it is impossible to know what specific tools a user may need in the future, this feature allows an open-ended supply of data that is not restricted by the current system solution. The flat .dat files make it impossible to understand and compare the data on the fly, which is why my system will also create on-screen visualizations that can be manipulated to drill down based on user criteria. Being able to visually represent a user’s query to the database of information will allow for researchers to quickly analyze large and small variances, incomplete data, hardware or system failures, and provide a quick snapshot of the overall performance of a system.

The DOE mandates this data be publically available for consumption by industry, government officials, other research institutions, and the general public (DOE, 2014). While the stakeholders may not be interested in usability as a primary goal of the system, it is a nice to have requirement. Meaningful application and visualization of the data allows for greater dissemination of the photovoltaic performance and reliability, which in turn can either: 1) potentially engage with external industry partners to allow for a more self-sustaining business model for the RTCs, or 2) validate the necessity and value in the research taking place at the RTCs for the DOE.

The data itself requires a certain amount of baseline knowledge to understand the meaning and function. This introductory knowledge is assumed, as a primer on photovoltaic systems is outside the scope for the system. The strength of the system is the sheer amount of data available for dissemination and analysis, and by hosting the application externally, literally anyone in the world with an internet connection can perform analysis. As governmental systems have traditionally been locked down behind firewalls and in classified environments, the ability for the public at large to access this data, even if it is not generally useful in everyday life, is a step forward in overall transparency of the final purpose and use of taxpayer dollars (Okamoto, 2017).

I will not create a fully customized system, but will be utilizing various open source libraries to successfully implement the data comparison and visualization functionality. Also, due to the amount of potential data elements, I will focus on a desktop-first application, rather than my typical focus of mobile first. While I will make sure that responsive elements are used, it will not be my primary requirement.

## Measures

The success of my system will be measured in multiple ways to assure the most objective critique. First, I will evaluate the success of the system based on user perception by using the System Usability Scale to register the user’s perceived usability of the system. Since these satisfaction scores are typically correlated with higher success rates (Kortum, 2014), I will then contrast those results with a Likert scale that asks questions regarding the perceived ease of use. While engaging directly with users, I will be measuring the effectiveness of my design by recording user behavior and actions during task scenarios, which will result in task success rate data. After the direct user feedback is collected, I will also measure the overall engagement of the system based on data-driven analytics including the amount of files exported, amount of unique visitors to the application, amount of repeat visitors, individual site traffic, and number of visualizations per system.

## 

## 

## References

1. United States, Department of Energy, Office of Science and Technology Policy. (2014). *Public Access Plan*.
2. Okamoto, K. (2017). Introducing Open Government Data. *Reference Librarian*, *58*(2), 111-123. doi:10.1080/02763877.2016.1199005
3. Kortum, P., & Peres, S. C. (2014). The Relationship Between System Effectiveness and Subjective Usability Scores Using the System Usability Scale. *International Journal Of Human-Computer Interaction*, *30*(7), 575-584. doi:10.1080/10447318.2014.904177