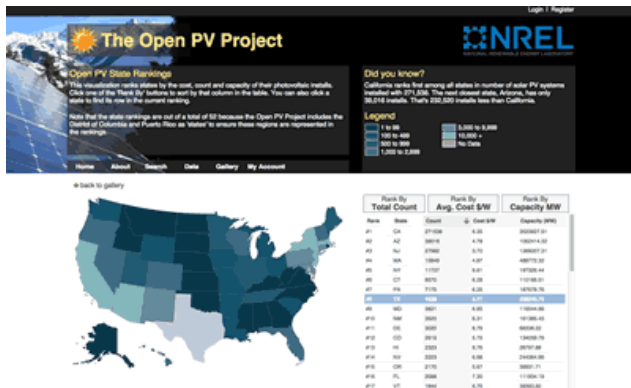


# The OpenPV Project: Crowdsourcing Solar Energy Data

## Case Study Overview



Photovoltaics convert solar energy into direct current. You can understand the photovoltaic market in the United States by estimating the installed capacity (or maximum possible amount) of photovoltaic energy. It's also important to understand how the prices of hardware and soft costs (including installation costs) related to photovoltaics change over time.

The *Open PV Project* is a joint effort by government, industry and the public to compile a comprehensive database of photovoltaic

installation data. Data for the project are voluntarily contributed from a variety of sources, including solar incentive programs, utilities, installers and the general public. The database is a Web-based resource that people can use to explore and understand current and past trends in the U.S. photovoltaics industry. Maintained by the contributors, the data are constantly changing, offering an evolving, up-to-date snapshot of the U.S. solar power market.

## Project Description

Data for the *OpenPV Project* come from two types of providers. The first type, which accounts for most of the data, comprises solar incentives programs, utilities and installers. They provide data in large volumes and have their own data validation processes.

The second type is the average photovoltaics consumer in the United States. Consumers usually furnish data from a single system and have no formal validation method.



## Challenges



Data standardization is a challenge for the *OpenPV Project*. Volunteers can give any data they have. Some data fields are required, such as system size, but users come up with many other data fields. One challenge was to create a data structure that could handle any submitted data but could also validate data for some general required parameters. The *OpenPV Project* has an automated process that handles data validation within a user interface. The system processes the data that volunteers upload and infers the columns that represent the required data, presenting them to the user in an intuitive interface. The user can then correct the data before submission.

Dealing with data duplication is an additional challenge. If a photovoltaics consumer enters the data for a residential system and then the installer or the incentives program also enters the

data, the system is counted twice. The *OpenPV Project* compares data submitted by users to find matches and thereby avoid duplication of data.

A final challenge is to make sure that people know about the *OpenPV Project* and have incentives to use it. The project meets the challenge in two ways. First, it includes useful and informative summary visualizations to draw people to the site, make sure that others reference it, and help proponents of solar energy keep their communities well represented in the database. Second, project staff get the word out in research papers and conference publications that are related to the project or that use its data.

## Benefits and Outcomes

Since 2011, the *OpenPV Project* has acquired data from over 400,000 installations of photovoltaic systems. The data are used to gauge the level of adoption of photovoltaics nationwide and to estimate the associated costs across the United States. The project also serves as the repository of these data, which can be downloaded with a high granularity of precision.

## Tips

The *OpenPV Project* case study illustrates the following steps in the Federal Citizen Science and Crowdsourcing Toolkit:

- **Scope Out Your Problem — Engage Your Stakeholders and Participants**  
Find a way to engage potential volunteers. If you offer some type of data download or visualization of summary statistics, people will find the site useful and point others to it.
- **Manage Your Data — Acquire Your Data**  
It's hard to get people to enter data into an empty system. By finding sources of data, even partial data, you can seed the system and draw users in to see the data you already have. If they like what they see, then they will be more likely to support your project.
- **Manage Your Data — Analyze Your Data**  
Ensure that people reference your project: Don't just summarize, inspire users. If your website shows your data in a way that is new or a very useful and intuitive, people will use your graphics to reference your project or the data it offers. This will draw attention to your project and build more support.

## Learn More

- Website: [Open PV Project](#)

## Contact Information

Name: Dan Getman

Email: [Dan.Getman@nrel.gov](mailto:Dan.Getman@nrel.gov)