

Ensuring Bioenergy Data Can Be
Accessible, Usable, and Useful

Debbie Brodt-Giles
July 21, 2020

Why is data access and usability important?

Data should adhere to the FAIR principles:

- Findable
- Accessible
- Interoperable
- Reusable

It is important that federally funded data be shared openly with the public

Data that can be shared with the public will spur innovation, reduce duplicative work, and rapidly advance industries and technologies



Findable Data

Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

- Metadata should be used
- Metadata should have globally unique and persistent identifier or document object identifier (DOI)
- Metadata and identifiers should be register or indexed in a searchable resource



Accessible Data

Once the user finds the required data, people need to understand how to access the data.

- Metadata, and the raw data, should be retrievable by their identifier using a standard communications protocol
- The protocol should be open, free, and accessible.
- Proper authentication allows for access and use
- Data and metadata need to be available for long-term use and citations (even if the data is out of date or has been replaced by a newer dataset).



Interoperable Data

The data needs to be easily integrated with other data and/or applications or workflows for analysis, storage, and processing.

- Data should use formats that are easily shared using broadly acceptable languages and formats.
- Metadata should use vocabularies that follow FAIR principles.
- When possible, utilize data content models to ensure consistent vocabularies and taxonomies.



Reusable Data

Data should be optimized for reuse and reutilization of the data. Data should be well-defined, provide provenance about the data sources, and allow for others to build on the data.

- Metadata has deep descriptions
- Data has a clear license assigned so people know how the data can be used
- Data has a clear provenance model
- Data uses proper domain-relevant community standards



Examples of Data Work



Linked energy information on hundreds of topics crowdsourced from industry and government agencies.

Share Data



Tracking the Sun

Abstract

Berkeley Lab's Tracking the Sun report series is dedicated to summarizing installed prices and other trends among grid-connected, distributed solar photovoltaic (PV) systems in the United States. The present report, the 11th edition in the series, focuses on systems installed through year-end 2017, with preliminary trends for the first half of 2018. As in years past, the primary emphasis is on describing changes in installed prices over time and variation in pricing across projects based on location, project ownership, system design, and other attributes. New to this year, however, is an expanded discussion of other project characteristics in the large underlying data sample. Future editions will include more of such material, beyond the report's traditional focus on installed pricing. The trends described in this report derive primarily from project-level data reported to state agencies and utilities that administer PV incentive programs, solar renewable energy credit (SREC) registration systems, or interconnection processes. In total, data were collected and cleaned for more than 1.3 million individual PV systems, representing 81% of U.S. residential and non-residential PV systems installed through 2017. The analysis of installed pricing trends is based on a subset of roughly 770,000 systems with available installed price data.



Data from October, 2019
Submitted Apr 2, 2020

Contact

Lawrence Berkeley National Laboratory (LBNL)

[Galen Barbose](#)
510.495.2593

Status

Awaiting curation

License

Authors

Galen Barbose
Lawrence Berkeley
National Laboratory

Naim Darghouth
Lawrence Berkeley
National Laboratory

Keywords

energy, power, solar, pv installations, grid, oedi

Share

Lawrence Berkeley National Laboratory

Copy Citation

Export Citation to RIS

1 Resource



Tracking the Sun Dataset

AWS S3 Bucket

View



Related Datasets

Datasets associated with the same DOE project

Submission Name

Resources

Submitted

Status

No submissions found.

Additional Info

Last Updated: 3 months ago



Marine and Hydrokinetic Data Repository U.S. DEPARTMENT OF ENERGY

search: MHKDR data Search

- Wave Energy
- Wave Energy Prize
- River Energy
- Current Energy

Featured Data

180 submissions

Wave Energy Prize

1 resource

ALFA: SWIFT Buoy Data Under Extreme Wave Conditions at th...

5 resources

Underwater M: Seabotix v1.BV

U.S. Department of Energy GDR Geothermal Data Repository

search: GDR data Search

- Direct Use
- EGS
- Forge
- Hydrothermal

Featured Data

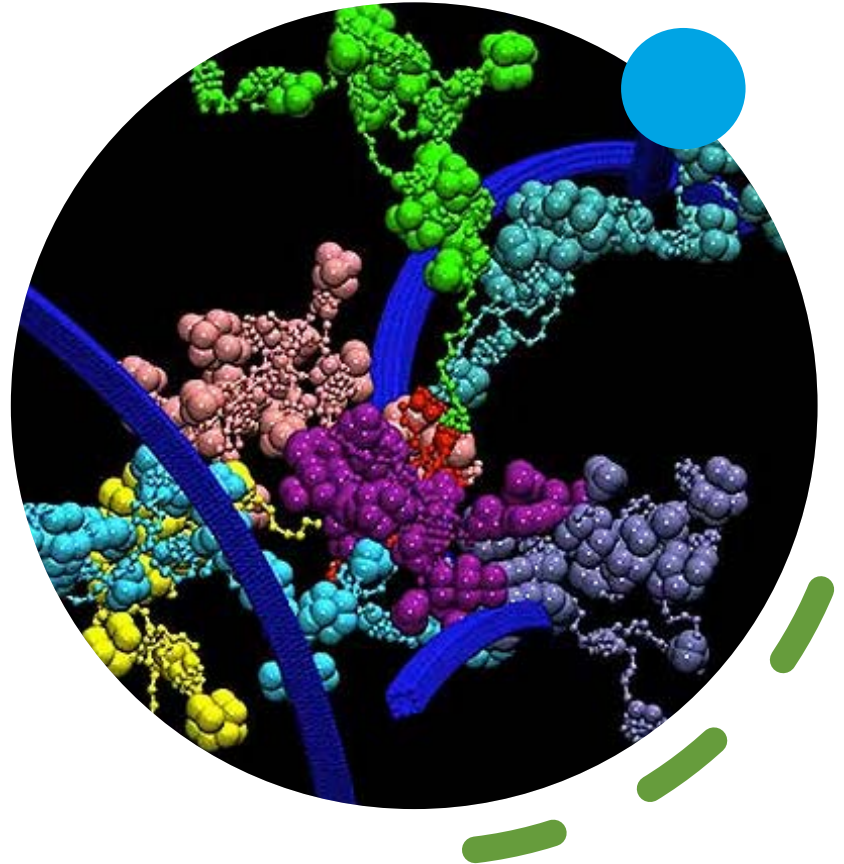
81 submissions

15 submissions

1 resource

How can this be applied to Bioenergy Data?

- Let's make bioenergy data more FAIR
- We can try to find interesting data that has not been readily accessible and reusable and make it available for public use
- We can use a carrot or a stick to secure data
 - Pay data owners for their high-value data to
 - Require all federally funded data to be collected and stored for reuse



Let's find the data!



One important step is to find the data that exists but is not currently accessible.

I would love to brainstorm about:

- What data would be highly valuable to industry?
- What laboratory data exists, where does it live now, and how can we make it more accessible?
- What other key data owners have valuable data that could be shared?
- How can we avoid duplicative work, advance industry, and accelerate innovation?

Let's Work Together

I'm always happy to talk data and create a path forward that works for the Bioenergy industry!

- Debbie.brodt.giles@nrel.gov
- 303-476-0600



This Photo by Unknown Author is licensed under [CC BY-ND](#)

Thank you!

www.nrel.gov

Debbie.Brodt.Giles@nrel.gov

