

U.S. Department of Energy (DOE)
Energy Efficiency and Renewable Energy Office (EERE)
Solar Energy Technologies Office (SETO)
EERE T 540.111-02: Request for Information (RFI)
DOE RFI DE-FOA0002876

Request for Information on the cost and value of
acquiring, accessing, and sharing solar photovoltaic
(PV) system performance data.

SETO seeks responses from all interested
stakeholders in industry.

Cost Benefit of Data Interoperability and the Value of Performance Data

K. Dixon Wright
President
SRC Digital Insurance Services
www.SRC-Digital-Insurance-Services.com
Dixon@SRC-Digital-Insurance-Services.com
415-717-1092

November 18, 2022

Submission Overview

This RFI is focused on specific technical performance data related to solar asset management and the impact on system design and maintenance.

This submission offers a long-term objective to establish key general data elements for multi-industry data interoperability to enable digital ecosystems for infrastructure, including solar, by leveraging open data standards like XBRL, Orange Button Open API, FDX and others.

Participants in this submission may also submit responses to this RFI that address the technical aspects.

The Orange Button and OS2 initiatives have an extensive network of entities that are collaborating on several high-level strategies to improve the infrastructure ecosystem to deliver greater capital efficiency for owners while simultaneously improving profitability for all the various stakeholders engaged.

Creating a digital ecosystem that can leverage performance data beyond single system administration to include finance, insurance and surety that aligns public and private systems across market sectors requires both long term objectives for planning and short-term steps to generate revenue and momentum.

Executive Summary

Performance Data for Solar Photovoltaic Systems has as its primary purpose to inform the asset owner on how their system is functioning so that timely corrective actions can be taken if needed.

A secondary purpose and just as critical to the asset owner is the utilization of standardized data to access products and services related to finance, insurance and surety. Direct benefits can include better finance terms, better insurance coverage and access to surety credit for securing financial guarantees.

A third purpose is for benchmarking across a wide range of asset owners to develop a better understanding of best practices and to assist grid administrators with planning. While important there is no direct benefit to the asset owner, but a benefit to all stakeholders over time.

With adopted data standards the direct benefit will promote the transition to digital by developers and accelerate implementation of digital construction management systems as called for in both the Infrastructure Investment and Jobs Act (IIJA) and the Inflation Reduction Act (IRA).

Our responses to RFI questions are motivated by the vision of a model digital ecosystem from the perspective of all stakeholders, public and private, and based on the continued expansion of XBRL, Orange Button JSON, FDX and other open data standards to accelerate the transition to digital by expanding data interoperability.

Background Introduction

The structure and design of the digital ecosystem has been widely studied and the subject of various private and trade association initiatives along with multiple FOAs that have resulted in public/private collaborations, in particular: the initial DATA Act, the Orange Button efforts, Construction Progress Coalition, Financial Data Exchange (FDX), ConsensusDocs and the CURT-CII OS2 initiative at University of Texas Austin.

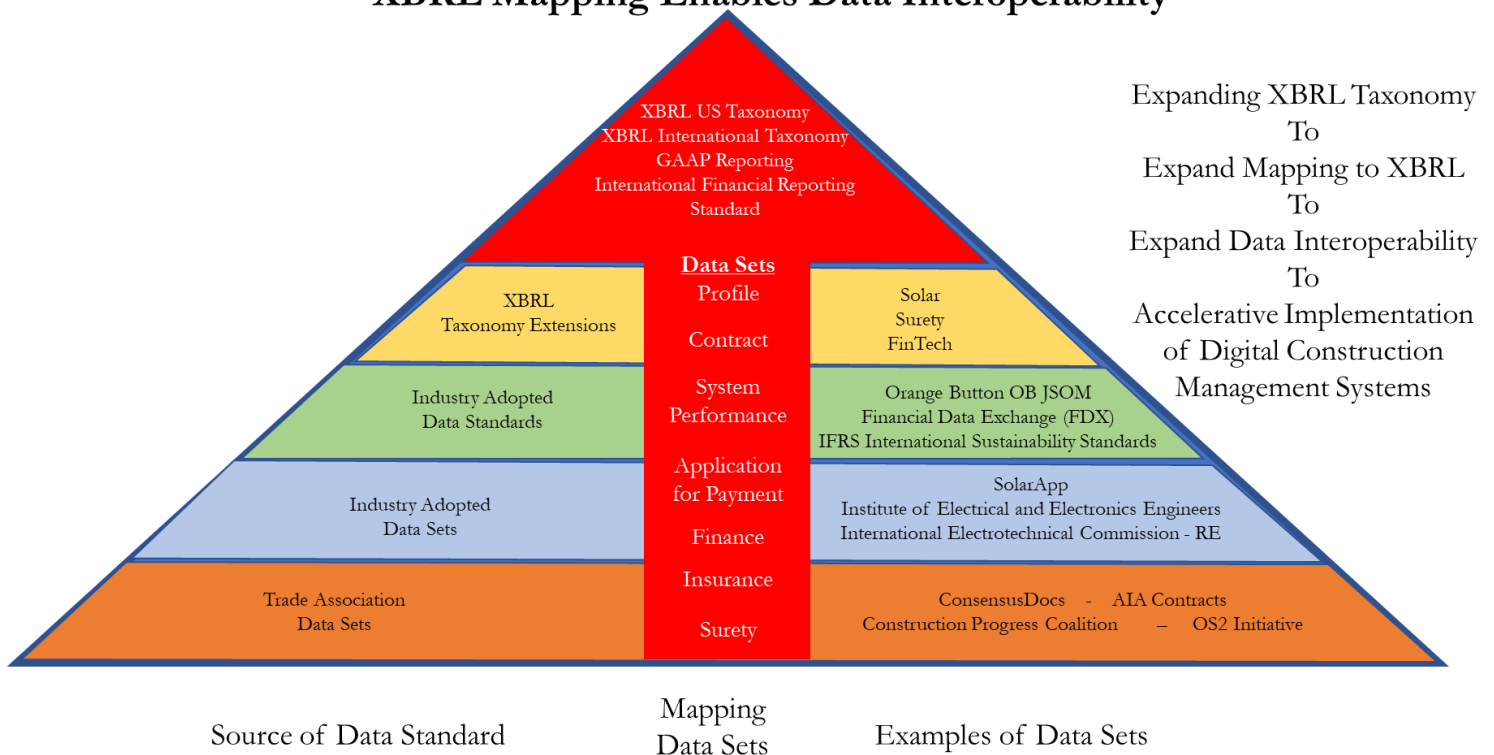
There are two separate uses for performance data relating to a solar facility –

1. Construction performance – Is contractor performance meeting requirements.
2. Ongoing operations – Is system performance meeting requirements.

The challenge is not technology, it is establishing consensus around data sets and data interoperability standards across market segments and public agencies and then implementing these standards.

Establishing key data sets and enabling mapping those data sets from any data source to XBRL will enable data interoperability across industry sectors, public and private.

Data Ecosystem – Hierarchy of Data XBRL Mapping Enables Data Interoperability



To accelerate implementation of data standards on May 24, 2022 Senators Mark Warner and Mike Crapo introduced S.4295 - [the Financial Data Transparency Act](#) to require all that federal agencies implement open data, including the requirement to “harmonize and reduce the private sector's regulatory compliance burden, while enhancing transparency and accountability, and for other purposes”. We refer to it as DATA Act 2.0.

In response to the DOE RFI regarding “Preventing Outages and Enhancing the Resilience of the Electric Grid” we set up a [“Digital Ecosystems for Energy Grid”](#) website for internal outreach and collaboration for our June 1st submission of recommendations and submitted a corresponding non-traditional grant request on September 25th based on those recommendations and in support of DAT Act 2.0.

Included in that response to the DOE are recommendations that include Performance Data for Solar Photovoltaic Systems.

Concluding Recommendations

1. Adopt the recommendations outlined in the June 1st submission to the DOE RFI for “Formula Grants to States and Indian Tribes for Preventing Outages and Enhancing the Resilience of the Electric Grid”
2. Support the non-traditional grant request of September 25th to fund:
 - a. \$2M for the continued expansion of XBRL
 - b. \$30M for the proposed Model Digital Ecosystem for Community Resiliency.
Proposal attached as an addendum.
3. Expand the Orange Button taxonomies (both XBRL and OpenAPI implementations) to enhance interoperability with existing IECRE and IEEE solar system performance data sets¹
4. Leverage SolarApp data sets for residential permit data. An Orange Button compliant interface for SolarApp is in development.
5. Leverage DATA Act 2.0 to align government and private data reporting utilizing the XBRL taxonomy and industry recognized data standards.
6. Engage with trade association initiatives like OS2 to establish data interoperability and new approaches to infrastructure construction based on capabilities enabled by standardized data.

Request for Information Categories and Questions

Category 1: Cost and Value of Data (from a data owner perspective)

Every stakeholder in a solar or clean energy project is a “data owner” with respects to the data. The perspective is unique for each stakeholder in a project, but the cost/benefits are the same.

The cost is relatively low if there are digital ecosystems enabled by open standards that can be implemented across industry segments.

The value or benefits can be extensive, depending on how much the data interoperability is leveraged to secure or offer products or services.

Example is the DOE SolarApp, a single data set for permitting solar projects that can also be used for finance, insurance, and surety. The cost of utilizing the data beyond just the permit is zero, but the benefits can include better financing, lower pricing for insurance or access to surety credit.

¹ [IECRE 61724](#) (UL)

IEC System for Certification to Standards Relating to Equipment for Use in Renewable Energy Applications

[IEEE 1526-2020](#) - Testing the Performance of Stand-Alone Photovoltaic Systems

The value of the data is in its interoperability with other stakeholders in a digital ecosystem to reduce administrative costs, improve risk management and lower the total project cost by eliminating waste and improving efficiency.

A reasonable cost and predictable outcomes are warranted by the value and enabled by XBRL.

The following are responses to the specific questions

1. What would be the estimated cost of curating a multi-year, high-quality dataset for a PV system, where that data set is captured at sampling rates of at least one reading every 15 minutes and contains readings from components such as the AC meter, the AC output of the inverter(s), the DC input(s) of the inverter(s), and the weather station?

Developing data without having it comply with an adopted data standard can be very expensive as the data would have little value outside of individual monitoring, so no offsetting benefits.

When the data is consistent with the Orange Button data standards then the data can be utilized by multiple stakeholders and service providers creating many options to leverage the data that offset the cost.

Many of the costs to add capability for performance data elements into the Orange Button taxonomy have already been incurred.

The one-time cost for platform developers to generate a report in a standardized Orange Button format is warranted by the value of the data individually for monitoring and system and compounded by the number of stakeholders that can also utilize the data for finance, insurance, and surety.

The ability for commercial system owners to monitor their systems is already provided by several independent service providers so the costs are known and predictable.

Many residential system owners are provided system performance data as part of their service, so no additional cost.

Primary examples of value are access to finance, insurance and surety for an immediate benefit to specific stakeholders.

Secondary examples are R&D for long-term benefits for all stakeholders.

2. What, if any, would be the added cost per MWDC associated with collecting, storing, and curating data for:

- a. tracker position
- b. string- or combiner-level DC information
- c. energy flows to and from energy storage (if applicable)

No Comment

3. Opportunity cost of data sharing:

- a. What is the perceived opportunity cost or burden for publicly sharing a data set containing historical values (with a lag of at least one year)?

The ROI for contributing to a standardized performance measurement process will be realized over time in better access to finance, insurance, and surety.

The more capital markets know about the risks they are assuming the better the risk management and lower pricing can be secured for financial products and services.

- b. What is the source of the opportunity cost?

Data that is already being generated for internal monitoring

- c. How does recency affect the opportunity cost? (i.e., is there a time lag of a certain number of years after which the opportunity cost drops to zero?)

Depends on the application.

If purchasing insurance or surety where the coverage is the performance of the solar system, then the timing must be close to real time.

If it is for R&D, actuarial purposes, or improvements to risk management, then the timeline can be broader.

- d. Would collecting less precise information help reduce that cost? (e.g., location recorded as gross coordinates with just two decimal places instead of an address or 4 decimal place precision, normalization of power and irradiance values instead of raw values)

Depends on the cost, but the more complete the data the better.

4. Operations and maintenance information:

- a. Aside from cost concerns, what would stand in the way of sharing operations and maintenance information that is concurrent with the time series datasets?

Lacking any standardization of O&M reporting information, shared data is at substantial risk of misinterpretation. O&M data records are not in any standard format, so any sharing requires bespoke interpretation.

Confidential information could be restricted to just those stakeholders that warrant the information.

- b. What would be the estimated cost of curating that dataset?

No Comment

Category 2: Access, Availability, and Value of Data (from a data user perspective)

1. If you are performing or plan to perform R&D related to the performance of PV systems, do you have sufficient data for analysis and validation? If not, what is the most critical missing information? (e.g., length of collection time, variety of system types, sizes, and locations, number of observed variables, completeness and/or fidelity of data)

It is not just about the data; it is the ability to import the data in a format that has the information.

Capital markets are constantly undertaking R&D to determine risk and pricing. Standardized consistent data enables that R&D and performance data sets like [IECRE](#) and [IEEE](#) provide the level of detail.

The missing component is having the various data compliant with the Orange Button taxonomy so they could be utilized more effectively by more stakeholders.

2. What are the minimum and optimum sets of PV system performance variables collected from a standard non-residential PV system (anything bigger than 50 kWDC) that are necessary to perform and validate your R&D project (e.g., modeling of output of DC field, inverters, or entire systems, modeling of impact of soil and snow, identification, classification, and prediction of performance anomalies)?

A system providing the IECRE or IEEE data sets should have all the data necessary to provide R&D.

If not, those data elements needed should be added to the industry adopted data sets and incorporated into Orange Button.

3. What is the value of collecting high-quality time series data from residential systems at higher sampling rates (at least one reading every 15 minutes), given the lack of ground sensor-based irradiance measurements from such systems?

That would depend on its indented use for the data.

4. What are unique parameters regarding the performance of residential systems that cannot be estimated from models or lab-scale prototypes?

No response

5. Accessibility of large datasets:

- a. How much value would be added by an interactive interface for accessing large datasets if that data is already otherwise accessible and accompanied by high-quality metadata (e.g., through an Application Programming Interface—API—or directly from a data repository)?

No response

- b. Do you know of any examples of interactive interfaces for accessing time series data that you consider exemplary?

No response

6. DOE and other federal, state, and local government organizations collect and publicly share some environmental data (e.g., high-quality irradiance data available from the Surface Radiation Budget (SURFRAD) and SOLar RADiation (SOLRAD) networks). Is there a need for additional environmental data that could be collected by government agencies that is not currently captured at a sufficient granularity, rate, and quality, or at all?

No response

Category 3: Value-add Ancillary Datasets (from a system developer/owner perspective)

1. Unavailable data

- a. What data is not currently collected by the operators of your assets that could provide additional value?

Data elements that need to be included include unique identifiers (UUID's) like the [SolarApp site ID](#), The [Legal Entity Identifier \(LEI\)](#), Surety Bond Validation Number (BVN) and the federal [SAMS](#) number.

- b. What advantage could this data provide for the operation and long-term value of the assets?

Connecting various stakeholders to an individual project or portfolio.

- c. What are the barriers to collecting or accessing those data sets?

Data interoperability

2. How does an aerial inspection of the asset impact its optimal operation?

No response

- a. In your experience, what (if any) is the desired frequency of such an inspection?

No response

3. Extreme weather impact

- a. Do you collect data about damage inflicted by extreme weather events, such as hurricanes, derechos, large hail, or floods?

The insurance industry does collect this type of data

If so, what would be the optimal way to leverage the data to support more resilient PV systems?

Defer to insurance colleagues