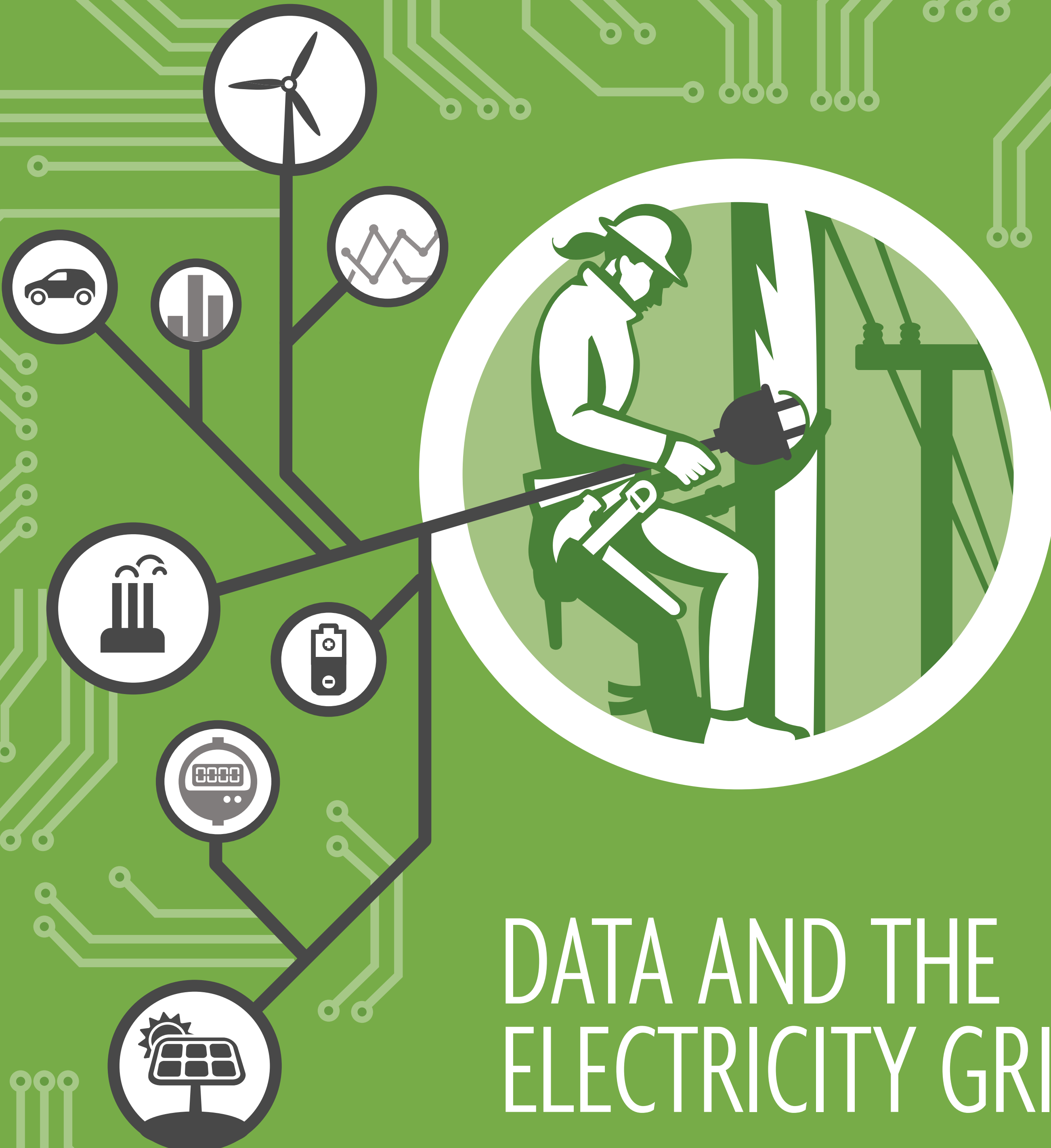


DATA AND THE ELECTRICITY GRID

A ROADMAP FOR USING SYSTEM DATA TO BUILD A PLUG & PLAY GRID

MORE THAN SMART

MORE THAN SMART

DATA AND THE ELECTRICITY GRID: A ROADMAP FOR USING SYSTEM DATA TO BUILD A PLUG & PLAY GRID

ARAM SHUMAVON, PAUL DE MARTINI, AND LAURA WANG

OVERVIEW

As customer adoption of Distributed Energy Resources ("DERs") increase, distribution system planning and operations will require additional data and information for grid operators and DER providers. Additionally, the opportunity to leverage DER into markets for distribution grid services and bulk power system services requires an exchange of certain information. This paper uses More Than Smart's Walk-Jog-Run Distribution Grid Planning architecture to frame a discussion of the data necessary to plan and operate a distributed grid as well as animate distributed markets as envisioned in California, New York, and other regulatory jurisdictions. In particular, this paper: (1) conducts a detailed inventory of the data needs of utilities, other market participants, and policymakers; (2) delineates the availability and usability of needed data and possible restrictions on this data; and (3) provides policy recommendations regarding grid data access strategies during the Walk-Jog-Run grid modernization stages.

INTRODUCTION

The technology and policy changes driving the increased adoption of DERs are also forcing a sea change in the role of data and technology in the electrical grid. Policymakers, utilities, and market participants must now consider the data necessary for the planning processes and operational practices of the electrical grid. Vanguard policy efforts like California's Distributed Resource Plan and New York's Reforming the Energy Vision are recognizing that a clear understanding of system and DER-related data is essential to both grid operators and DER providers. Robust data analysis and advanced informational capabilities present an opportunity to optimize the type and location of distribution grid investments as well as the operation of those resources. The exchange of system and DER data will also require navigating challenging questions about privacy, security, and market design.

Capturing the full value of the "Plug & Play" grid will require systems for exchanging large volumes of distributed information and the decision support analytics that can effectively utilize them. Increasing levels of DERs necessitates accessible and timely machine readable data to operationalize their integration into the Plug & Play Grid.

CONTENTS

ABOUT MORE THAN SMART

More Than Smart (MTS) develops solutions for building an electricity grid that provides opportunities for more parties to produce and use clean energy. A key focus is in providing assistance to states to follow the MTS Walk-Jog-Run® Framework for modernizing distribution grids through an engineering-based framework that acknowledges the unique energy policies of each state. More information can be found at www.morethansmart.org.

Thank you to Energy Innovation for funding this report. We would like to acknowledge the contributions of Alexandria Smith for providing research support, and to acknowledge discussions from the MTS Working Group, the California Distribution Resources Plan Working Group (www.drpwg.org) facilitated by MTS, and the New York Reforming the Energy Vision process in developing the thinking and frameworks outlined in this report. The content of this report does not imply an endorsement by any individual or organization that participated in any MTS workshop or reflect the views, policies, or otherwise of any specific state.

AUTHORS

Aram Shumavon, Kevala Analytics
Paul De Martini, Newport Consulting Group
Laura Wang, More than Smart

This paper relies on More Than Smart's existing Walk-Jog-Run framework, outlined in Figure 1 below and described in other papers[1] to sequentially frame the evolution of the distribution system's data needs and capabilities. The Walk stage is representative of the current conditions on the grid in markets where existing grid limitations and the increasing adoption of distributed resources requires consideration of interconnection locations. The Jog stage involves greater need for situational awareness of the variability of certain DER on grid operations. The Jog stage also presents opportunities to leverage DER for grid services as non-wires alternatives to more traditional utility infrastructure upgrades through an operational market. The Run stage involves an expansion of the Jog stage's operational market to include distribution level energy markets involving bilateral forward transactions among DER providers and load serving entities across the distribution system.

SYSTEM DATA TYPES

The following system data types have been identified in California's More Than Smart working group and New York Distributed System Implementation Plan engagement group discussions among utilities, DER providers, and other stakeholders.

1 | **System planning data** refers to data and information used as inputs and assumptions in distribution system planning as well as the resulting output information. Planning functions include hosting capacity analysis, interconnection studies, and annual short and long term planning, as well as near-term operational planning.

System planning data also includes information that lowers the overall cost of DER integration including DER market development data as well as customer load data that supports efficient integration.

2 | **Regulatory compliance data** refers to information needed by regulators to make informed decisions about the performance of the distribution system, utility expenditure requests, and potential location and value of DER as a non-wires alternative or other system value. This includes data and information regarding assumptions and methodologies for determining locational and system value, as applicable. In later stages this may mean market surveillance data including procurement results and bid information.

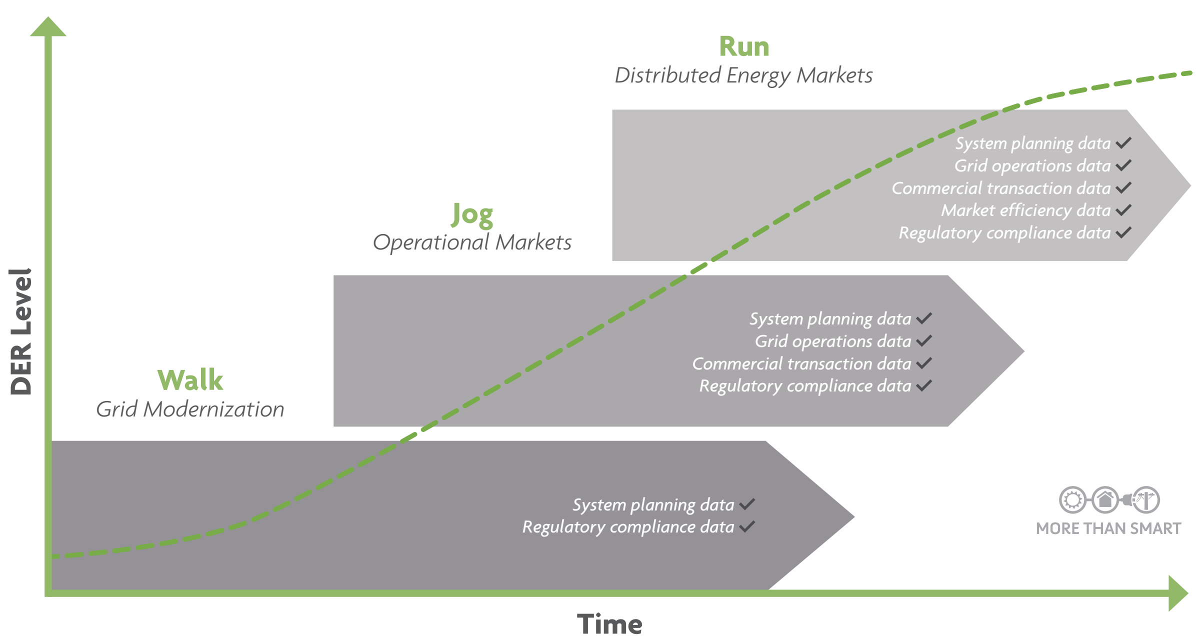
3 | **Grid operations data** refers to real-time (day-ahead through day of operation) information about the state of the distribution system and availability and operation of interconnected DER. This includes utility data and information from grid sensing and measurement devices and operational systems as well as generated by DER providers' and customer's systems and devices.

4 | **Commercial transaction data** references the data and data formats necessary to drive efficiency into procurement activities while also laying the foundation for compensating grid activities that may be procured in the future.

5 | **Market efficiency data** refers to information that supports market oversight and evaluation of market effectiveness to achieve desired objectives regarding customer choice, system performance, reliability, and environment.

FIGURE 1.

DATA TYPES ACCORDING TO THE WALK-JOG-RUN FRAMEWORK



1 http://morethansmart.org/wp-content/uploads/2016/09/plug-and-play-report_online_v2.pdf

GRID DATA EVOLUTION

The five system data categories described above will apply in varying degrees to system grid data needs as a more distributed electric system evolves. The Walk-Jog-Run framework, outlined in Figure 1, helps stakeholders understand the data needs that are essential functions for managing the evolution of the grid. The Walk stage requires system planning and regulatory compliance data. The Jog stage incorporates grid and DER operational data and commercial transaction data. The Run stage requires market efficiency data. Each of these requirements comes from discrete business needs of stakeholders as we have observed through MTS, New York REV, and other stakeholder settings across the country.

In the Walk stage, data needs are focused on planning for increased uptake and setting the right framework for future market functions. Regulatory and planning activities related to enabling higher penetration distributed generation resources are the drivers for data needs. The Walk stage is crucial for establishing the frameworks upon which markets will function, quantitatively supporting the answers to the questions of why and where the market should be capturing the full value of distributed energy resources. System planning data and regulatory compliance data are key to this process as are clear descriptions of avoided cost methodologies and inputs. The Walk stage requires surfacing of localized or localizable data in order to begin the process of quantifying the system's inherent potential of minimizing costs.

The Jog stage is characterized by expanded data needs to animate operational markets for grid services provided by DER. Establishing operational market mechanisms to capture local benefits of DER involves sharing economic and engineering data and information related to specific location value. As these markets mature, the accessibility of data and information related to DER availability and performance under a tariff, program, or contract requirements must increase. These operational markets also require grid operational data to coordinate physical transaction schedules between distribution operators, wholesale market operators, and DER providers. Finally, as higher levels of DER develop, greater situational awareness is required, involving dynamic grid connectivity models and real time power flow characteristics.

In the Run stage, data is needed for market coordination and constraint management of energy transactions across the distribution system and into the bulk power system. In order to provide market oversight and assess market efficacy, new tools for processing grid data become necessary. The Run stage involves additional energy market development beyond the operational markets: bilateral energy transactions supported by DER supply resources will be expected to provide visibility of their status, as well as behavior, to multiple entities.

GRID DATA NEEDS

Turning any data into actionable information requires an understanding of the data's availability, accessibility, and usability. Delineating system data in these contexts is crucial to understanding how the role of data will change over time.

- *Data availability* means knowing the origin of the data, such as who collected it and the manner in which it was collected.
- *Data accessibility* means understanding who has access to it and, when access may be limited, how to provide transparent alternatives, such as aggregating or anonymizing data.
- *Data usability* describes the format of the data as it relates to computational processes as well as the lag, durability, and frequency of those data. Data usability is a metric to delineate the fact that data included in static maps is less user-friendly than a downloadable, machine-readable data set, which is less user-friendly than data made available via an Application Programming Interface ("API").

Clearly addressing issues related to grid data needs early in regulatory processes is an essential step to achieving desired policy objectives as identified in California, Minnesota, New York, and other jurisdictions moving toward greater integration of DER.

Importantly, data are required from both utilities and DER providers to support planning, operations, and markets. The More Than Smart framework also recognizes there is a crucial role for modeled, de-identified, anonymized, and aggregated data where certain market participants believe they cannot make necessary information available. Privacy and security concerns should also be addressed in the context of availability. While this paper addresses the current discussions on system data, the sources, cost, and value of data remains an area of rapid development.

As these markets mature it is also important to recognize that there is also the potential for developing value-added information beyond that which is required. The processes supporting these markets should recognize that some data may have value even if they are not essential for the operation of the market. This information may be subject to different availability, accessibility, and usability requirements.

In Figure 2 we address common data types, their sources, availability, and usage. This list is in no way exhaustive and elements will necessarily change over time as products and markets evolve.

FIGURE 2.

DATA TYPE AVAILABILITY, USABILITY, AND APPLICATION MATRIX

DATA TYPE	SOURCE: UTILITY	SOURCE: DER PROVIDER	SOURCE: PUC	FREQUENCY & ACCURACY	AVAILABLE?	PUBLIC ACCESS?	MACHINE READABLE?	PURPOSE FOR DATA: SYSTEM PLANNING	REGULATORY COMPLIANCE	COMMERCIAL USE	GRID OPERATIONS	MARKET EFFICIENCY
1. Distribution Capital investment	✔			Annual planning process Forecasts (ex: 5-10yr forecast period)	●	○	✖	✔	✔		✔	
2. Circuit capacity (nominal)	✔		✔	Annual 15 minute or hourly interval	●	○	●	✔	✔		✔	✔
3. Circuit connectivity models	✔			Real-time on circuit configuration; connectivity of customers and DER is not automated and accuracy varies by utility	○	○	✖	✔			✔	
4. Customer data (individual)	✔			Monthly billing 15 minute interval for C&I, hourly for residential is commonly available to customers and their designees	●	✖	●	✔	✔	✔	✔	✔
5. Customer data (aggregate)	✔			As requested, typically using 15/15 method	●	○	●	✔	✔	✔	✔	✔
6. DER capacity (existing and queued)	✔	✔	✔	Monthly for interconnection queues, but existing capacity is not well known as DER providers have the most accurate information that is not compiled	●	○	○	✔		✔	✔	✔
7. DER services performance	✔	✔	✔	This data is not yet available, but expected to become available as DER grid services are provided	○	✖	✖	✔	✔	✔	✔	✔
8. Distributed generation adoption forecasts	✔	✔	✔	Annual planning process	●	●	✖	✔	✔		✔	✔
9. Grid conditions (historic)	✔			Daily, information is provided in rate cases and annual reliability reports	●	○	✖	✔	✔	✔	✔	
10. Grid conditions (real time)	✔			Real time at substation level, not yet available consistently at more granular levels in distribution	○	✖	✖	✔	✔	✔	✔	✔
11. Hosting capacity	✔			Monthly updates based on changes reflecting DER interconnections	●	●	●	✔	✔	✔		
12. Hourly DER gross profiles	✔	✔		As recorded by DER provider, utilities do not consistently have this information	●	✖	○	✔		✔	✔	
13. Interconnection cost data	✔			Utility interconnection tariff (Rule 21), costs are included in rate cases	●	●	✖	✔	✔	✔		✔
14. Load growth forecast	✔			Annual planning process Forecasts	●	●	●	✔	✔			
15. Locational net benefits or value	✔			Utility developed values available to PUC	●	○	○	✔	✔	✔	✔	✔
16. Market potential or saturation studies		✔			●	✖	✖	✔				✔
17. Planned resiliency and reliability projects	✔			Annual planning process	●	●	●	✔			✔	
18. Project attributes	✔	✔		Annual planning process	●	●	●	✔			✔	
19. Reliability statistics	✔		✔	Annual planning process	●	●	○	✔	✔		✔	
20. Utility rates	✔		✔	General ratemaking	●	●	○		✔			✔
21. Voltage & power quality	✔				●	●	○	✔	✔		✔	✔

● AVAILABLE ○ LIMITED OR IN-PROGRESS ✖ NOT AVAILABLE

MORE THAN SMART

THE VALUE OF SYSTEM DATA

The value of data, like the value of distributed resources, may be geographically variable. States, including New York and California, have explicitly included locational value of distributed resources by the inclusion of wholesale locational marginal prices and geographically discrete avoided distribution capital upgrades. As a result of the geographic variability of the value of distributed resources, some areas may explicitly benefit more from the availability of certain kinds of data, and the costs and benefits of data provisioning should be analyzed in light of that variability. Also, the net benefits of data must also recognize that the availability and usability of data will be dependent on improvements to current system sensing, measurement, connectivity, and locational information. These data and related analytics, as well as the telecommunications infrastructure supporting them, may require increased utility capital investment, but also may be provided by other market participants. Understanding the full value of data requires due diligence into the costs of providing that data, including the costs of alternatives to that data.

There is also the potential for value-added data to be created, as is being discussed in New York. Such data may be beyond that which is required for compliance and is generally thought to have value for a limited set of entities. The allocation of costs for providing this information should be addressed separately from data that benefit the broader market. For example, certain utility-created information may support an individual DER provider's market development activities and therefore may be considered value-added. In New York this type of value-added information is being considered a potential new revenue stream for utilities. Likewise, DER providers may develop information beyond that required for interconnection, operations, and market participation that has value to others. The data sets or information that may be considered necessary or standard system data versus which ones are value added will evolve over time.

RELATED NATIONAL DEVELOPMENTS

Several related national efforts are underway that are enabling a more distributed energy future. These efforts are provided as context to a broader range of data and information efforts that have been identified to enable federal and state environmental and energy goals.

- **Green Button℠ and Orange Button℠:** *standardizing production and consumption data*

  The standardization of consumer production and consumption data formats is essential to driving scalable software solutions to grid needs. Consumer consumption data standardization has been implemented in several markets, including California, as a result of the Green Button effort[2], which provides improved access and usability of consumption data for consumers and their designates.

  Consumer distributed generation production data standardization is also underway as a result of the Orange Button efforts.[3] Creating unified data standards will enable the solar industry to reduce market inefficiencies and lower costs for customers. Orange Button supports the creation and adoption of industry-led open data standards to facilitate the growth and expansion of distributed solar. Orange Button strives to reduce soft costs by streamlining the collection, security, management, exchange, and monetization of solar datasets across the value chain of solar. It is expected that the standards will be adopted by stakeholders representing 60% of the U.S. solar market.

- **Machine Readable Tariffs:** *from static PDFs to downloadable datasets*

  The Open Energy Information OpenEI Utility Rate Database (URDB) provides rate structure information from more than 3,000 Energy Information Administration (EIA) recognized utilities.[4] Currently, most utilities report their rates in PDF format, and rates must be manually converted into a machine-readable format for use by market participants. National Renewable Energy Laboratory (NREL) has been developing machine-readable formats that all utilities can use to improve the URDB.

  Additionally, there are ongoing proceedings to increase EIA's processing efficiency through Form EIA-861 and Form EIA-861S (Short Form) which collect annual information on the retail sale, distribution, transmission, and generation of electric energy in the U.S. Proposed changes to the forms include a new respondent type entitled "Behind the Meter", adding new questions asking for the capacity of small-scale storage associated with net metered (NEM) and non-net metered distribution capacity and for the virtual NEM capacity and customer counts of NEM programs. Comments overall indicate that DER data needs to be properly accounted for in the Annual Electric Power Industry Report.[5]

- **Public Building Datasets:** *better insight into building performance*

  Several efforts are underway to drive greater visibility into building performance from an energy intensity perspective. Shared, anonymized datasets like Lawrence Berkeley National Lab's developed Building Performance Database help quantify how individual buildings are likely to be able to provide grid benefits based on energy efficiency investments or onsite distributed generation. Other efforts, such as Stanford University's VISDOM project performing consumption pattern analysis work are helping categorize common consumption patterns with increased temporal and geographic granularity. These data sets, and the algorithms that generated them, are expected to support important modeling and forecasting efforts in the industry.

2 http://www.greenbuttondata.org/

3 http://www.orangebuttondata.org/

4 http://en.openei.org/wiki/Utility_Rate_Database

5 http://www.eia.gov/survey/frn/electricity/electricity2017_05192016.pdf

RECOMMENDATIONS

Information will drive the transformation of the electricity industry in this century, just as it has for other business sectors. Pervasive, low cost computing combined with increasingly more powerful analytics will play a crucial role in a more distributed energy future. The speed with which the external technological advancements and business transformations are occurring will present challenges to the electricity industry unless the availability, accessibility, and usability of data is addressed in a direct and comprehensive manner. Below are recommendations for electric industry and policy leaders to consider regarding the creation, use, and access of system data through a step-wise approach that can evolve with the growth of DER in each state.

1 | **Walk stage:** The walk stage should focus on three main areas, including: (1) establishing a thorough assessment of data collection needs answering the questions: What type of data is needed and for what purpose? Who provides what types of data? Who pays for what types of data? What type of accessibility is required for different data?; (2) providing more transparency into the availability of required types of data by all relevant parties; and (3) initiating consistent system and DER provider reporting information within a jurisdiction.

Key steps include:

a. *Comprehensive data surveys:* Stakeholders should be engaged to develop a list of their respective data needs and availability of data they collect or produce. Discussions should focus on the types of data referenced in Figure 2, including data accessibility based on privacy, security, and market sensitivity concerns and the development of publicly available alternatives as proxies for restricted data. The data referenced in Figure 2 should help frame the spectrum of data.

b. *Learn by doing:* Operational market pilots play a crucial role in preparing for subsequent stages. Soliciting the data that will inform where pilots should occur, what their objectives should be, and the data used to realize them should be a core element to the Walk stage.

c. *Create transparent and replicable data utilization methodologies:* The process of developing operational markets requires all stakeholders to be able to analyze the implications of those markets. Some data will have restrictions, but where such restrictions occur, methodology should be developed which allows others to replicate the same calculations with public data.

d. *Utilize standardized data formats and protocols within a jurisdiction:* It is essential to ensure early adoption of common standards to reduce seams issues where possible. Data portals, permissions, and platform-oriented data standards including interoperability protocols are critical issues to be considering during the Walk stage.

2 | **Jog stage:** Data requirements in the Run stage should be driven by operational coordination of DER participation in wholesale markets and as needed for distribution operational market mechanisms.

a. *Transparency and accessibility of reporting and settlement tools:* These tools should build on the data standards established during the Walk stage. There should be particular emphasis on visibility into prices for grid services and tools for comparing regulatory mechanism, such as rates and programs, to the efficacy of utility procurement mechanisms at meeting specified needs.

b. Data types that should be addressed in this stage are grid operational data, grid services related commercial transaction data, and regulatory compliance data. Permissions should be fully addressed in this stage and the role of non-utility-provided data should be well understood. Data-driven non-wires alternative analysis will be common practice in the latter years of the Jog stage and policymakers and market participants should expect to consider needed data accessibility.

3 | **Run stage:** Data requirements in the Run stage should be driven by the evolution of distributed energy markets and the need to physically coordinate bilateral transactions between multiple parties across a distribution system and into wholesale markets. Over time, spot markets may emerge, as will the need to dynamically manage constraints and balance loads and resources for a local distribution area. If this develops, it will require additional information exchanged between market participants and the distribution market operator.

MORE THAN SMART

www.morethansmart.org | info@morethansmart.org
426 17th Street, Suite 700, Oakland, CA 94612