U.S. DEPARTMENT OF
ENERGY
Office of Electricity

# Transmission Research Program: FOA 1861 - Big Data Analysis of Synchrophasor Data

Sandra Jenkins

November 5, 2020

# Transmission Research Program at DOE

The original “Synchrophasor program ” goals have been met and the program exists in a new environment

- To a large extent, the reliability research themes and needs first outlined in the original Grid of the Future white papers have now been realized
- The research program is now a part of a the DOE Office of Electricity
- Significant, complementary research programs, including Advanced Grid Modeling, Energy Storage, Advanced Power Grid Components, Advanced Sensors, and Resilient Distribution Systems, have opened up new opportunities for collaboration.

Refocus the program through a series of commissioned white papers and symposium to push the boundaries of transmission research in the next 10-20 years

- Identify technical areas where DOE can take a leadership role in the transmission component of the Electric Delivery System.
- The white papers will be peer reviewed and discussed formally with a broad and diverse audience of stakeholders at a DOE-sponsored symposium in 2021.

# Data Development for Transmission

Address and **develop transmission system data** management and processing to ensure utilities are getting the full value from new sensors and existing data sources

- Developing partnerships for transmission system **data sharing** to facilitate wide area situational awareness, and emergency response. Also, develop research datasets and data platforms that reduce utility burden from data requests, and facilitate tool development with real data
- Examine methods of **data generation** and formatting at the data source, to find opportunities for improvements to better align those methods with the evolving end uses for the data
- Ensure the value of transmission system data can be realized by developing industry practices for **data interoperability** – so that disparate sources of data can be combined to form new insights for the transmission system, and to reduce data clearing and processing costs for utilities and researchers

# Tool development for Transmission systems

Develop tools that help system operators understand and **adapt to changes in supply and interdependencies** with the transmission grid

- Advance sensing and data analytics to be capable of fully capturing new system dynamics from the integration of renewable energy and inverter-based technologies
- Create new analytic methods to address natural gas and electricity interdependency

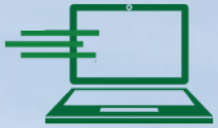
Modernization of transmission system tools through **human factors and cognitive science research** for transmission control

- Advance literature around control room decision processes (operators and support engineers) that tool developers can draw from
- Create advanced control room software that help ease operator burden and facilitate adaption to new operating challenges
- Develop and focus user testing and feedback from control room engineers through national lab partnerships

# Program Areas

Data Development for Transmission Systems

**High quality, accurate data for utilities, research community and policy makers**



Advanced Applications Research and Development

**Develop and demonstrate advanced grid applications and tools for transmission system visibility and controllability**

Human Factors, Visualization, and Tool Modernization for Grid Ops

**Visualize the complexity of Interconnected systems to provide more useful alarms, cues, and predictive information for grid operation and planning**

Transmission Measurement & Standards

**Develop new ways of measuring and characterizing the power system to inform data, model and tool development**

# Big Data Analysis of Synchrophasor Data

***"…explore the use of big data, artificial intelligence (AI), and machine learning technology and tools on phasor measurement unit (PMU) data to identify and improve existing knowledge, and to discover new insights and tools for better grid operation and management."***

## FOA 1861 Goals:

- Derive additional value from the vast amounts of sensor data already being generated
- Provide actionable information on the use of Machine Learning and Artificial Intelligence methods on large PMU data sets
- Enable faster grid analytics and modeling and better grid asset management though new tools

Press Release:
https://www.energy.gov/articles/department-energy-announces-20-million-artificial-intelligence-research
Fact Sheet:
https://www.energy.gov/sites/prod/files/2019/04/f61/Big%20Data%20Awards%20Fact%20Sheet%20FINAL_1.pdf

# National PMU data set

DOE and PNNL have worked to create a dataset that:

- Covers a significant number of PMUs and substations in each of the three US interconnections
- Covers multiple years and includes event logs
- Is real field data from a variety of sources that includes a variety of errors, inconsistencies, quality levels, and flaws
- Is anonymized to reduce the data's value in exposing potential information about power system operational vulnerabilities

Efforts are underway to further improve this dataset not only with additional data, but with more and higher quality labels as well.

# FOA 1861 Objectives

- Identify key events within each interconnection-specific dataset
- Identify unusual or anomalous events and patterns not in the event log.
- Catalogue the “signatures” (identifying patterns) including events that can be used to identify and diagnose events and grid conditions.
- Identify precursor conditions that warn about forthcoming events.
- Identify patterns that reveal the condition of power system equipment and insights into equipment modeling.
- Identify apparent ground-induced currents (GIC) relating to geomagnetic disturbances.
- Predict performance of power plants and other assets.
- Identify actual load event and patterns, potentially including detection of DERS.
- Identify factors that can improve wind integration and solar integration.
- Identify factors that reflect weather and seasonality without the use of site-specific weather data.
- Identify anomalies that may reflect cyber security issues and events rather than grid performance.

# FOA 1861 Awards

| Performer | Team Members | |
|---|---|---|
| Iowa State | Electric Power Group (EPG), Google Brain, IBM | Robust Learning of Dynamic Interactions for Enhancing Power System Resilience |
| SEL | Oregon State University | Machine Learning Guided Operational Intelligence from Synchrophasors |
| University of California Riverside | EPG, Michigan Technological University | Discovery of Signatures, Anomalies, and Precursors in Synchrophasor Data with Matrix Profile and Deep Recurrent Neural Networks |
| University of Nevada, Reno | Arizona State University (ASU), IBM, Virginia Tech | Robust Event Diagnostics Platform: Integrating Tensor Analytics and Machine Learning Into Real-time Grid Monitoring |
| GE | GE Grid Solutions | PMU-Based Data Analytics using Digital Twin and PhasorAnalytics Software |
| Siemens | Southern Methodist University, Temple University | MindSynchro |
| Ping Things | NA | Combinatorial Evaluation of Physical Feature Engineering and Deep Temporal Modeling for Synchrophasor Data at Scale |
| Texas A&M | Temple University, Quanta Technology | Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART) |

# FOA 1861 Project Reporting

Two phases of reporting

- Summarize patters, insights, event signatures, data quality, analytical tools
- How electric industry can leverage results, and future research recommendations.

| REQUIRED TOPICS | CONTENT REQUIRED FOR TRAINING DATASET REPORT | CONTENT REQUIRED FOR TRAINING AND TEST DATASETS REPORT |
|---|---|---|
| **Key Findings** | Describe in summary what was found within the training dataset including **patterns, insights, and a catalog of event signatures**. These should include many of the 11 items listed in Section I of Funding Opportunity Announcement 1861. | Describe in summary what was found within the full dataset, including patterns, insights, and a catalog of event signatures. These should include many of the 11 items listed in Section I of Funding Opportunity Announcement 1861. |
| **Causal and Correlated Factors** | Describe the **factors within the training data that the analysis shows are correlated or causal to the insights and patterns identified, including event precursors**. These should include many of the 11 items listed in Section I of Funding Opportunity Announcement 1861. | Describe the factors within the full dataset that the analysis shows are correlated or causal to the insights and patterns identified, including event precursors. Provide additional detail on catalog of event signatures and many of the 11 items listed in Section I of Funding Opportunity Announcement 1861. |
| **Data Quality** | Review the **quality of the training dataset**, describe any unique quality problems in each electrical interconnection's data, and explain any problems resulting from bad data. | Review the quality of the full dataset and describe any unique quality problems in each interconnection's data. Describe any patterns in the bad data that could reveal why and how the bad data was created. Describe how Recipient's tools handled data quality problems, and explain any problems resulting from bad data. |
| **Analytical Tools** | Describe the **analytical methods and tools used to identify the above insights**, including whether and which of those tools use proprietary rather than generic analytical methods. | Describe the analytical methods and tools used to identify the above insights, including whether and which of those tools use proprietary rather than generic analytical methods. Explain how much of the analytical findings could be achieved with similar tools, outside of those of the Recipient, and how much may be due to proprietary capabilities of the Recipient's software and firm. |
| **Applying These Insights** | **Need not be addressed in the Training Dataset Report.** | **Explain how** PMU owners **and users** and electric industry software **developers can use these insights and similar tools to develop better real-time and automated tools** for power system and asset management and other purposes. Explain **what additional data would be desirable to develop better insights and findings** for power system operation, asset management, cybersecurity monitoring, or related high value purposes. Discuss any potential cybersecurity implications from applying these insights. |
| **Rationale for Further Work** | **Need not be addressed in the Training Dataset Report.** | **What research topics and scope would you recommend DOE fund** in the area of big data, artificial intelligence, and machine learning and why? **Is there specific work that would be a valuable follow-up to this FOA and its findings?** |

# FOA 1861 Meta-Analysis

PNNL is conducting a meta-analysis of the portfolio findings that will:

- Facilitate understanding across the projects to aid dissemination and adoption of the research findings
- ML/AI experts at PNNL will review key events and evaluate the impact of anonymization
- Projects began in September 2019 and were originally set to be completed in April 2021; completion of some projects may be delayed due to impacts from Covid-19.

PNNL will summarize results from the eight awardees in a report that:

- Highlights key findings
- Compiles event signatures
- Highlights unusual/impactful events and precursors
- Discusses how to operationalize findings
- Compares and contrasts methods

# Questions

Thank you.