Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Welcome

Our webinar will start soon

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Topic 2. Cross-sector & Open Data Sharing and Risks

Eric Andersen, PNNL - Topic 2 Co-Moderator
Dr. Jim Follum, PNNL - Topic 2 Co-Moderator

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

POWER SECTOR TRANSMISSION & DISTRIBUTION DATA AND INFORMATION WEBINAR SERIES

TOPIC 1: T&D Information Sharing
Wednesday, October 11 | 10:00 a.m. to 12:30 p.m. PDT

TOPIC 2: Cross-sector & Open Data Sharing and Risks
Wednesday, October 18 | 10:00 a.m. to 12:00 p.m. PDT

TOPIC 3: Sensor Systems and Platforms
Wednesday, October 25 | 10:00 a.m. to 12:00 p.m. PDT

TOPIC 4: Sensor Data and Device Research
Wednesday, November 1 | 10:00 a.m. to 12:00 p.m. PDT

REGISTER TODAY!

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

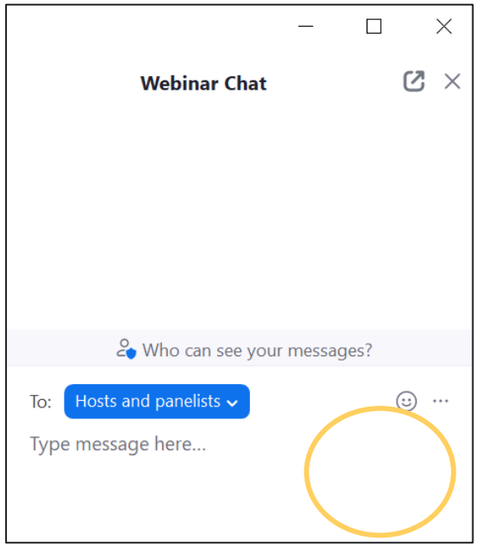
Los Alamos NATIONAL LABORATORY

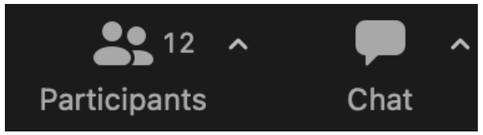
OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

Housekeeping items

- Recording the session (for internal purposes only)
- Slides will be made available on the event page
  https://www.pnnl.gov/events/power-sector-transmission-distribution-data-and-information-webinar-series
- Please type your questions in the chat box (two options)
  - Use “Host and panelists” option for posing questions only to presenters
  - Post questions and comments for all attendees to see
- Q&A
  - We will attempt to answer as many questions as we can, while considering the time.





Agenda

TIME (PDT)	TOPIC	PRESENTERS
10:00 – 10:05 a.m.	Welcome and Introduction	Eric Andersen, PNNL Sandra Jenkins, DOE Chris Irwin, DOE Roshanak Nateghi, DOE
10:05 – 10:25 a.m.	Utility Information Sharing: Challenges and Opportunities	Jim Ball, WAPA
10:25 – 10:45 a.m.	Utility Data Sharing Risks and Curating a National PMU Data Set	Eric Andersen, PNNL
10:45 – 10:55 a.m.	The Grid Event Signature Library (GESL)	Aaron Wilson, ORNL
10:55 – 11:05 a.m.	The Transmission Signature Library	Jim Follum, PNNL
11:05 – 11:15 a.m.	GridSweep® Instrument: Sharing the Data	Alex McEachern, LBNL
11:15 – 11:35 a.m.	Outage Data Initiative Nationwide (ODIN)	Supriya Chinthavali, ORNL
11:35 – 11:55 a.m.	Open Energy Data Initiative for Solar Data and Analytics	Kemal Celik, SETO
11:55 – 12:00 p.m.	Concluding Remarks	Eric Andersen, PNNL

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

DOE-Office of Electricity (OE) Introduction

Sandra Jenkins, DOE-OE

Chris Irwin, DOE-OE

Roshanak Nateghi, DOE-OE

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Utility Information Sharing

Challenges and Opportunities

James Ball
Cyber Program Advisor-WAPA

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

Western Area
Power Administration

Thesis Question

Why is it so bloody difficult to get utilities to share data with non-utility partners?

Western Area
Power Administration

Agenda

- History
- Laws and Regulations
- Cause and Effect
- Current Practice
- Challenges
- Possible Solutions
- Closing

Western Area Power Administration

History

- 1996 – E.O. 13010 established the Presidents Commission on Critical Infrastructure Protection
- 1998- PDD 63 sets initial strategy and collaboration framework
- 2002- FERC Order 630 establishes the CEII program
- 2013 FERC approves CIP-011
- 2015- Section 215 of the FAST Act directs FERC to issue updated regulations
- 2017- FERC Order 865 expands CEII regulations

Western Area Power Administration

Laws and Regulations

- FAST Act of 2015
  - Provided more comprehensive definition of CEII
  - Reaffirmed FERC CEII authorities
  - Directed further regulation (FERC Order 865)
  - Authorized FERC to impose sanctions for willful unauthorized disclosure
- FERC Order 865
  - Expanded Scope of CEII
  - Enhanced CEII Designation Process
  - CEII Sharing for Cyber Purposes
  - Protection Measures
  - Sanctions
- NERC CIP-011 (version)
  - BCSI concept
  - BCSI Access Controls
  - Audit and Enforcement

Western Area Power Administration

Cause and Effect

- Two Big Events
  - April 2013 Metcalf Incident
    - Subsequent discussion and reaction
  - May 2018 $2.7M fine levied against PG&E
    - CIP-011 violation
- Subsequent Actions For CIP-011
  - 2019-$100K fine
  - 2020-$150K fine
  - 2021-$200K fine

Western Area Power Administration

Current Practices

- BCSI and CEII governance policies
- Data Sharing Agreements
  - WIDSA (WECC)
- Data Protection Mechanisms (NERC Align)
- Anonymization (Neighborhood Keeper)
- CRADA/NDA’s etc.
- CEII Rule Development at DOE

Western Area Power Administration

Challenges

- Overclassification
- Overly Broad Definitions
- No Statutory Provision for R&D
- Corporate Risk Tolerance
- Technical Challenges
  - Encryption at rest/in transit
  - IAM Standards
  - Assurance Issues (common vetting processes, information storage standards)
- Trust Issues
- Nature of Proof in CIP/Auditing Standards

Western Area Power Administration

Possible Solutions

- Standard enforceable NDA/MOA that specifies information protection standards
- Contract Provisions
- “Get Out of Jail Free Card” for R&D Activities
- Responsibility on the holder of the information
- Technical Standards for Encryption, Access Management, Transmission
- Common Personnel/Facility certification processes and standards

Western Area Power Administration

Closing

- Information Sharing issues are not unique to the electric sector
  - Solutions exist
- We cannot build the grid of the future without transparency between components
- Look hard at what you think is sensitive and why
  - Memorialize these decisions
- Take some good faith risks.

Thank you

Jim Ball, WAPA
ball@wapa.gov

PNNL-SA-191429

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Utility Data Sharing Risks and Curating a National PMU Data Set

Eric Andersen
R&D Director
Electricity Infrastructure Operations Center
Pacific Northwest National Laboratory

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

Utility Data Sharing Risks and Curating a National PMU Data Set

Utility Data Sharing Risks and Economics

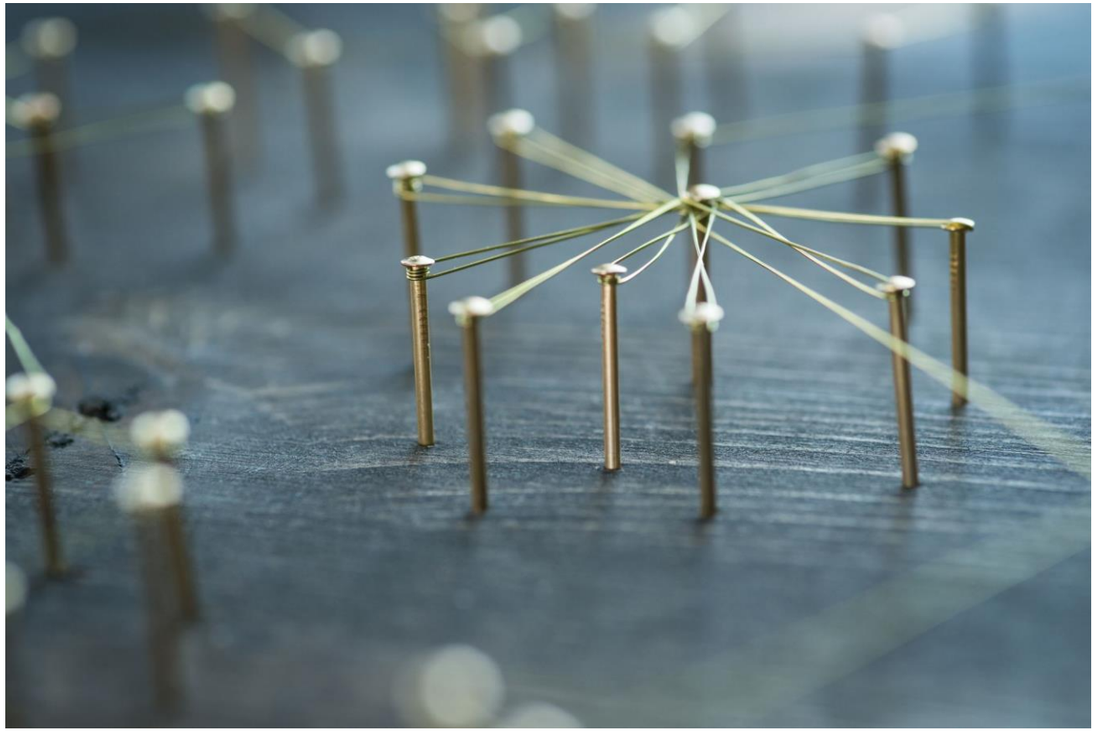
- Risk assessment framework
- Regulatory drivers
- Cost drivers
- Identify risks
- Offer mitigation strategies
- Outcome is a guide for data sharing for the PMAs

Curating a National PMU Data Set as a Use Case

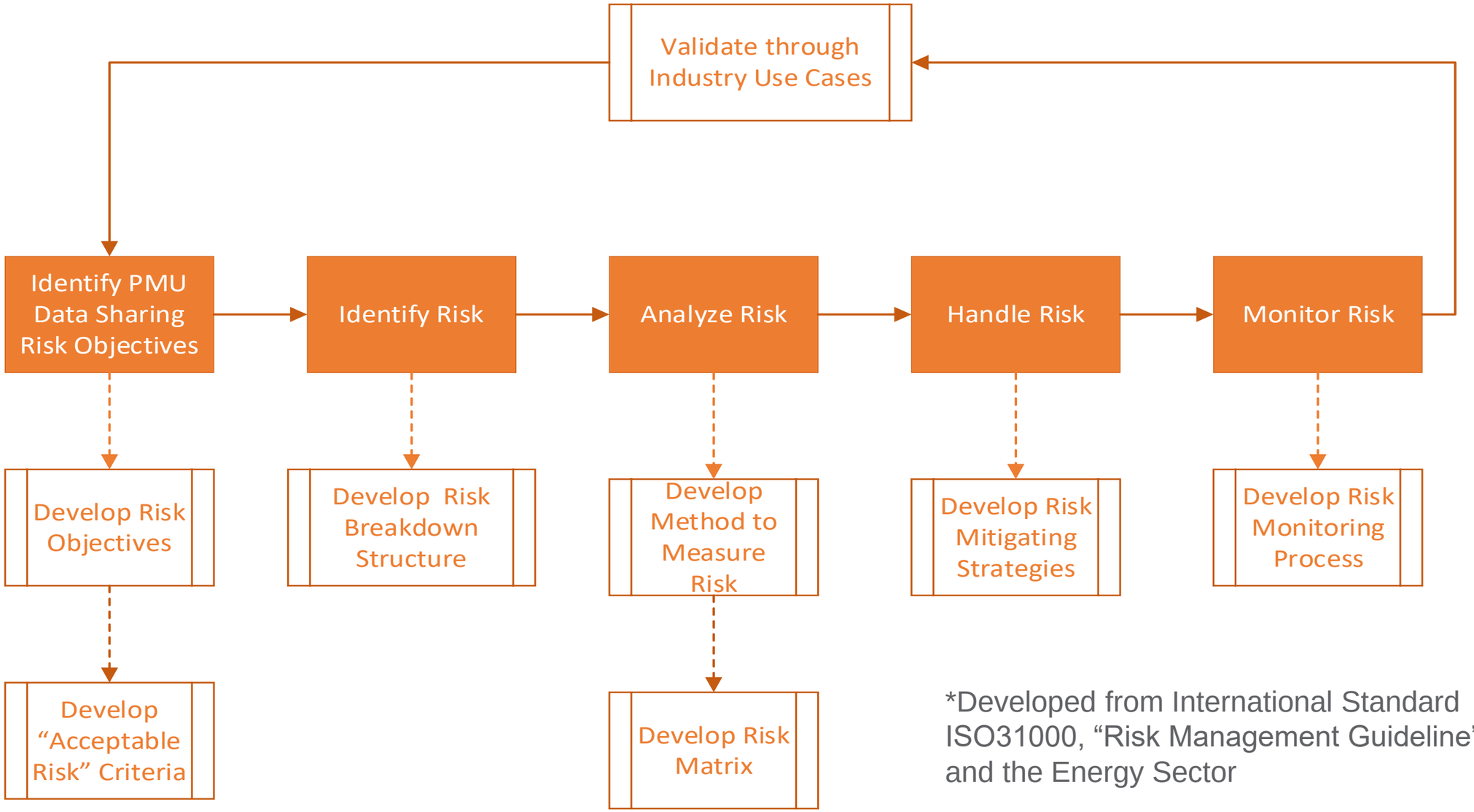
- Real-time operational data
- Data for research purposes

Data Sharing Stakeholders

- Between utility itself
  - Accessibility
- Between functional entities
  - TOPs to RC, BA
  - T&D boundaries becoming more important
- Between neighboring utilities
  - Regional Agreements (e.g., WIDSA, EIDSN)
- Third parties
  - Vendors – for system improvements, development
  - Research
    - ✓ Academia
    - ✓ Labs
  - Regulatory and Compliance
    - ✓ NERC, FERC



Utility Data Sharing Risk Framework*



*Developed from International Standard ISO31000, "Risk Management Guideline" and the Energy Sector

Utility Data Sharing Risks
Key Drivers

Regulatory and Legal Compliance

- What if I'm found to be in non-compliance NERC/FERC?
- Will we be held accountable if someone violates our NDA?

Economics and costs

- Technology advancements
- Reliability of grid operations and interconnectivity
- Data breach recovery costs
- Infrastructure costs
  - data pipelines
  - storage
- Historians and data retrieval
- Labor for data curation

Business Competitiveness

- Proprietary Intellectual Property
- Governance
- Reputation

Supply Chain

- Data quality
- Data formats
- Data security and misuse of data (e.g., topology, market trends, cloud services, etc.)
- Cybersecurity of contractors, vendors, and subcontractors

Some of the questions we're asking:

- Are these risks real or perceived?
- Are there examples of when they were realized, and what were the consequences?
- How do the risks change over time?

Regulatory Drivers

- **Utilities are obligated to share information following regulatory requirements**
  - NERC standards require information sharing between operational organizations to support reliable operations of the Bulk Electric System, and submit incident information to NERC and federal agencies
  - FERC requires utilities to submit market, financial, planning, and operational information
  - DOE requires utilities to submit operational and incident information
  - Other federal agencies, state, and local governments may require utilities to submit other kinds of information
- **Failure to share or submit this information can result in monetary fines or other penalties**
  - By federal statute, fines can exceed $1M per day per occurrence
  - NERC non-monetary penalties may include mandatory actions to prevent recurrence of the violation
- **Some information may be released by government agencies under FOIA, although other information may be protected from disclosure as CEII**

Mix S.R., and M.R. Weimar. 2023. Utility Data Risks and Economics: Assessing Government and Regulatory Impacts. PNNL-34193. Richland, WA: Pacific Northwest National Laboratory.

Economics | Exploring the costs for Data

- The costs of data sharing were explored with
  - 2 - Transmission Operators
  - 2 - Reliability Coordinators
- Data curation costs for data sharing were examined at PNNL

	Storage ($)					Total ($)	
Entity	Low	High	**Data Requests ($)**	**Historian ($)**	**WAN ($)**	Low	High
Utility 1	5,300	11,500	600	1,900	3,600	11,400	17,600
Utility 2	0	200	2,400			2,400	2,600
RC 1	14,600	31,600	5,300	1,900	3,600	25,400	42,400
RC 2*				1,900	3,600	25,400	42,400
PNNL				1,900			103,500

*-RC2 numbers were estimated from RC1 for storage and data requests

NOTE: these numbers are preliminary estimates and PNNL is still in the process of validating these numbers, and these numbers should not be construed to be final.

Take Aways

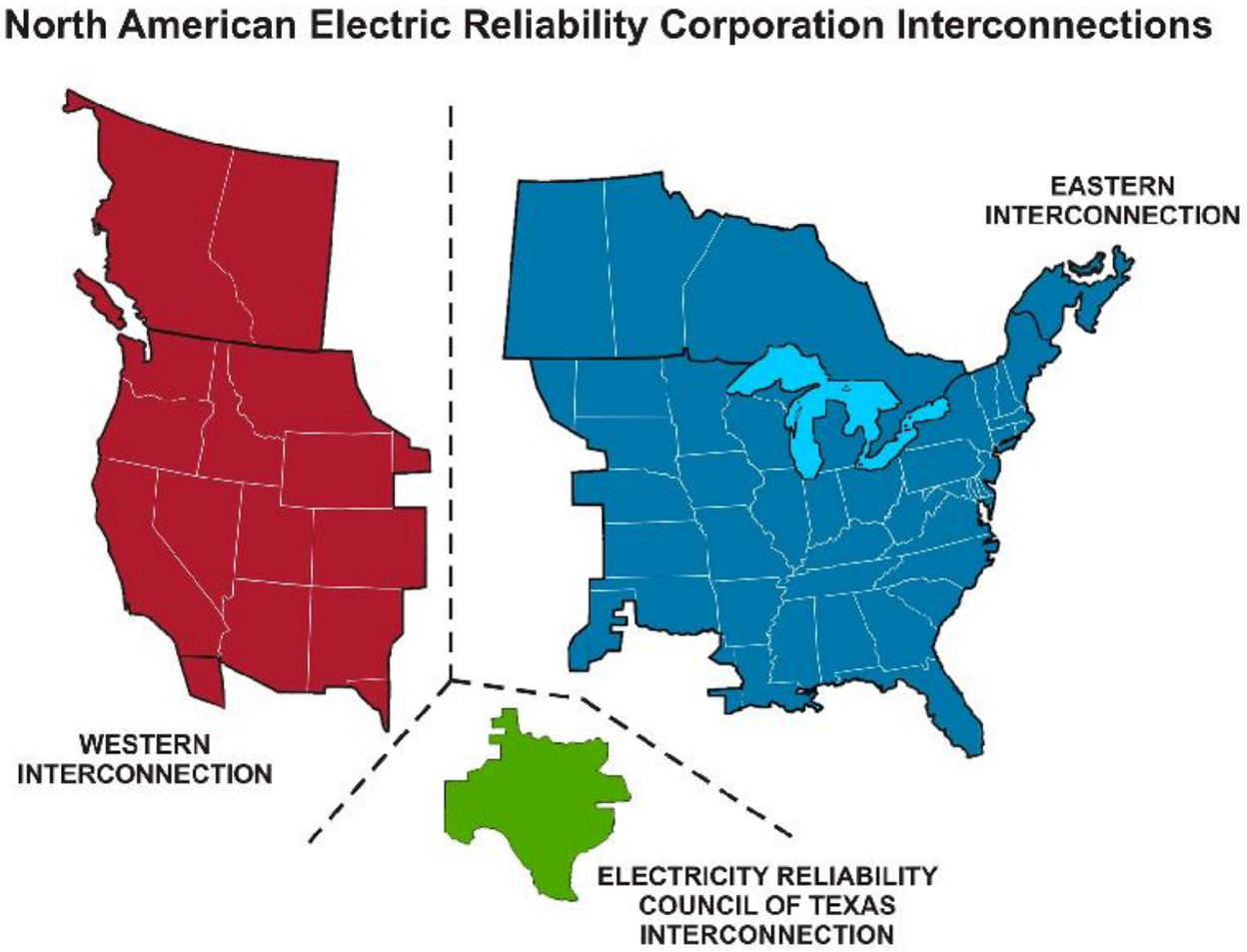
- Data curation represented the largest cost
- RC costs were driven by storage
- Utility 2 costs were smallest due to only small amount of storage on-prem

Risk Mitigation and Monitoring

- **Mitigation**
  - Develop data sharing governance and policies
  - Manage data being transferred
  - Implement appropriate access management controls
  - Create secure data transfer channels
  - Implement legal requirements (data transfer agreements and non-disclosure agreements)
  - Improve data quality
- **Monitoring**
  - Measure impact of sharing data
  - Periodic risk assessments
  - Reporting of data agreement violations

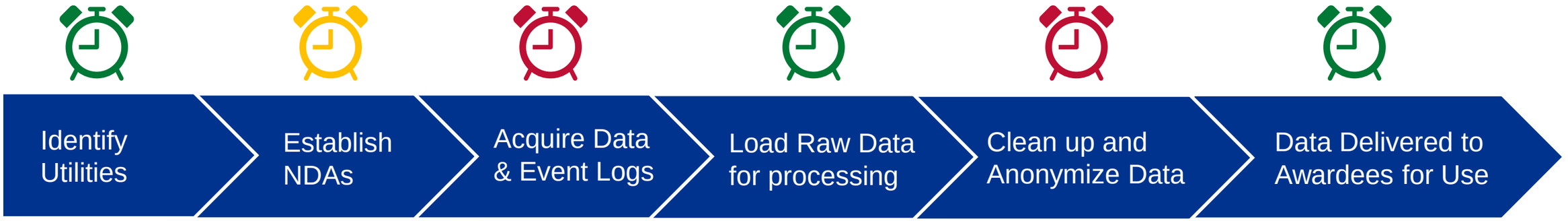
Curating a National PMU Data Set

- Department of Energy wanted data set to support AI/ML and big data analytics
- Look for ways to derive additional value from the vast amounts of sensor data already generated
- Curate real world data from each of the three US interconnections



Map source: https://www.energy.gov/oe/articles/north-american-electric-reliability-corporation-interconnections

Curation Process



Challenges:
NDAs with Data Providers

- Legal negotiations
- Providing assurance of data protection
- Providing details of the anonymization process before we had access to the data

Challenges:

- Pulling archive data was a heavy lift for some Data Providers
- Duration and age of data requested resulted in some providers having to obtain their data from third party archives
- Exported data came from a wide variety of archive tools (commercial, open-source, and custom) and were in different formats
- Event logs were all unique and lacked common taxonomy, and appear to be created manually

Challenges:

- Inconsistency of data between providers (e.g., different sequences and phases)
  Extracted data formats
  UTC timestamp formats
  significant digits
- Extensive manipulation was required to aggregate data
- Data quality varied by provider
- Duplicate data
- Anonymization required to remove topology metadata

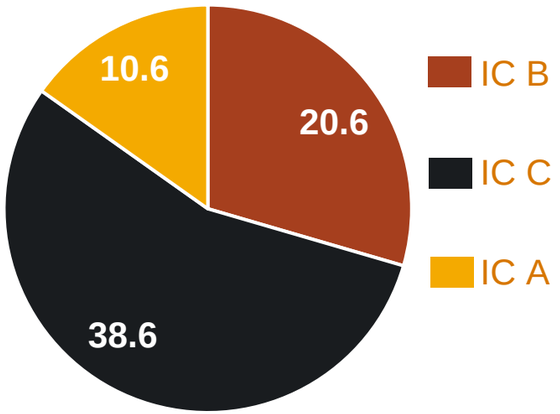
National PMU Data Set

Summary of Utility Synchrophasor Data Contributed				
Interconnection	**Dataset Range**	**Number of Data Providers (Utilities)**	**Total PMUs**	**Raw Data Size Received (TB)**
Eastern	2016-01-01 – 2017-12-31	5	250	38.6
ERCOT	2018-07-21 – 2019-08-24	5	221	10.6
Western	2016-01-01 – 2017-12-31	3*	43	19.0
	Total	13*	514	68.2

- There have been numerous requests for access to the anonymized data set
- NDAs with the Data Providers have been modified to enable additional use of the anonymized data set
- Additional synchrophasor data is being added, where possible
- The anonymized data set is available for others to use; however, the data users will have to agree to an NDA with PNNL

* NOTE: The western interconnection PMU data set is currently under expansion with several additional PMU data sets being curated into the National PMU data set

Storage Per Interconnection (TB)



Takeaways From the Curation Effort

- If you want data from a utility
  - Understand some of the utility data sharing risks and mitigation strategies
  - Improve your value proposition for utility participation
- Recommendations for utilities
  - Move toward standardizing the data retention and archive processes
  - Automate the creation of event logs and improve their consistency
  - Improvement is needed for inconsistent data across all utilities
    - ✓Data quality
    - ✓UTC formats and timestamps
    - ✓archive processes
    - ✓what data is archived

Thank you

Eric Andersen, PNNL
Eric.Andersen@pnnl.gov

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

The Grid Event Signature Library (GESL)

Aaron Wilson, Ph.D.
Research Electrical Engineer
Oak Ridge National Laboratory

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

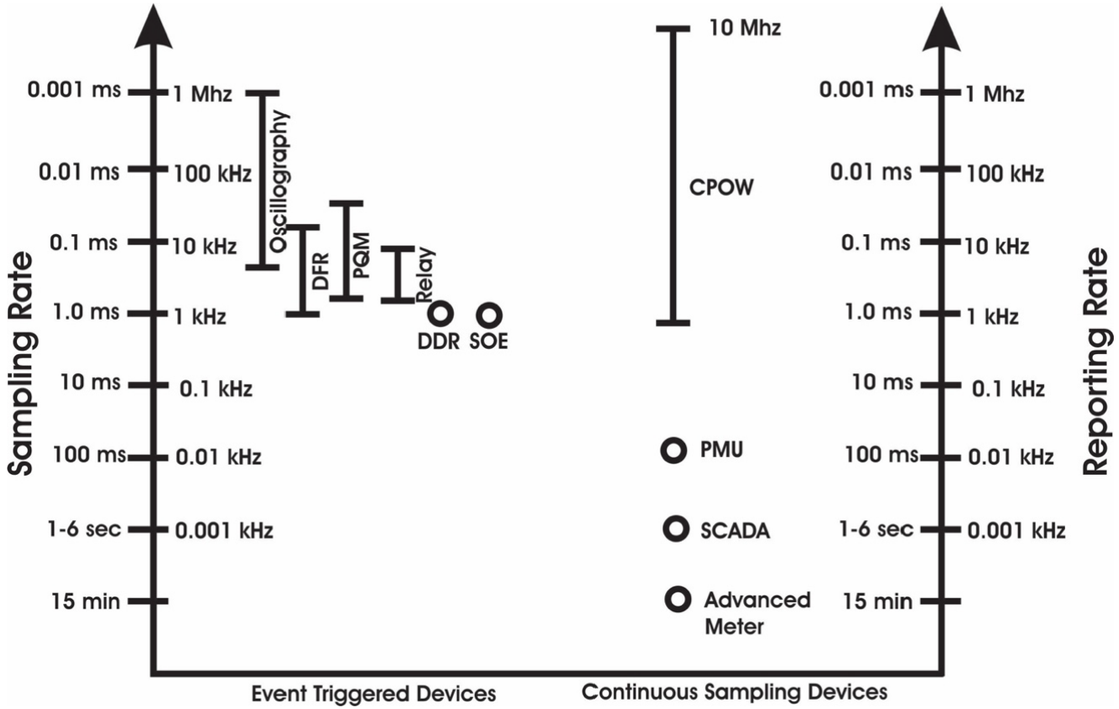
Pacific Northwest NATIONAL LABORATORY

Project Overview

- Funded by the DOE's Office of Electricity, the Grid Event Signature Library (GESL) was created to:
  - Facilitate, tag and fuse data feeds from multiple sources,
  - Implement a modular architecture for expandable design,
  - Anonymize event sources to enable open data sharing,
  - Provide an open-source, go-to resource for event identification and algorithm validation

Increased Grid Observability → Better Decision-Making

- Event triggered measurements
  - Relays
  - Digital fault recorders
  - Power quality meters
- Continuous measurements
  - SCADA
  - AMI (advanced metering infrastructure)
  - PMU (phasor measurement unit)
  - Point-on-wave (POW) measurements
- Event records
  - Outage and maintenance records
  - Device activation records



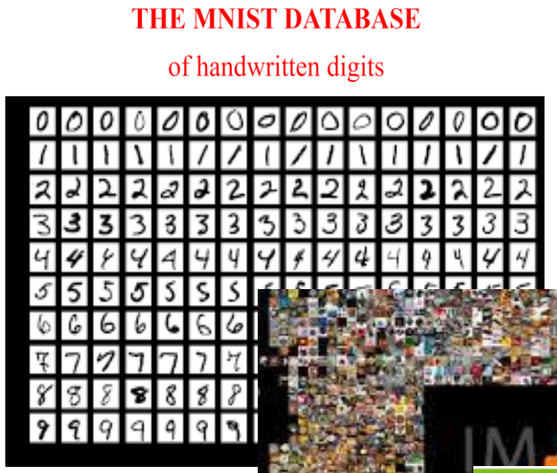
Grid Monitoring devices by resolution and data continuity*

*A. Silverstein and J. Follum, "High-resolution, time-synchronized grid monitoring devices," PNNL, Tech. Rep. PNNL-29770, Mar. 2020.

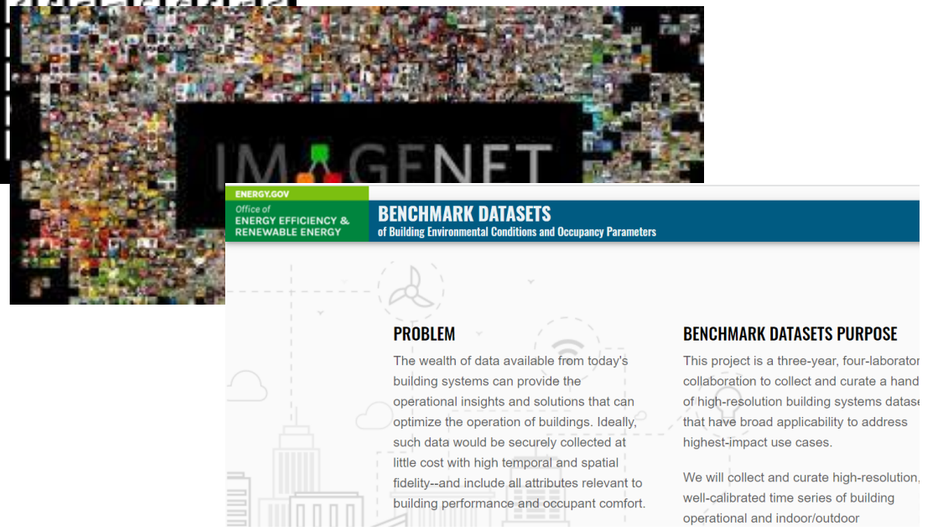
AI/ML-based Grid Health Monitoring

Public Datasets: Pivotal in other fields – why not the electric utility industry?

- Data labeling is critical to AI/ML
  - MNIST
  - ImageNet
  - BTO Building Benchmark Datasets
- Challenges exist for grid events
  - Data is decentralized and inaccessible
    - *Limits actionable data available for analytics*
  - Data is multimodal and unstandardized
    - *Prevents integration of different data sources*
  - Data is unprocessed and unvalidated
    - *Lacks critical metadata and proper labeling*



http://yann.lecun.com/exdb/mnist/



https://bbd.labworks.org/

https://syncedreview.com/

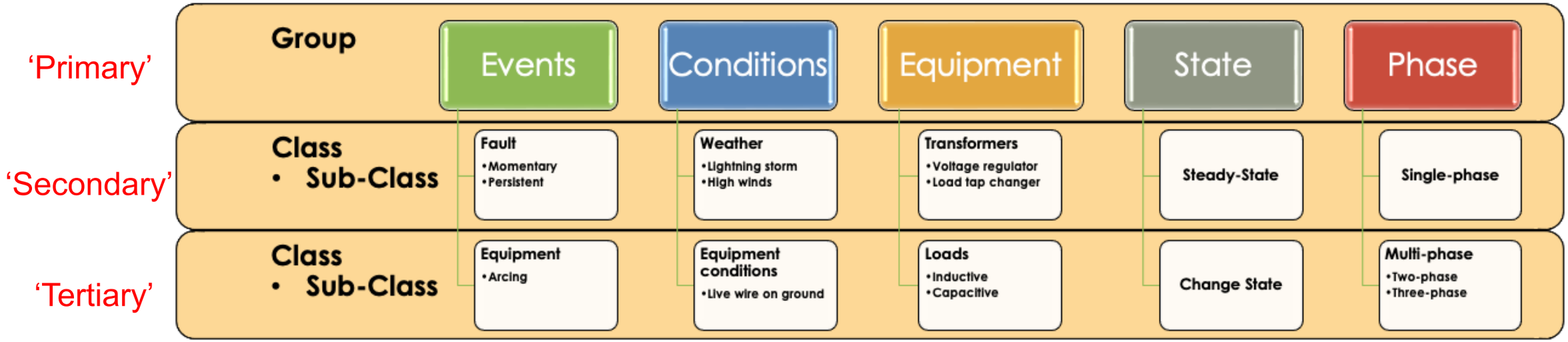
What is an “Event”?

- No standard definition of “event” in power system standards
- Per the “Glossary of Terms Used in NERC Reliability Standards,” the closest synonym is “Disturbance”:
  - An unplanned event that produces an abnormal system condition.
  - Any perturbation to the electric system.
  - The unexpected change in ACE that is caused by the sudden failure of generation or interruption of load.
    - “ACE” = Area Control Error, or “The instantaneous difference between a Balancing Authority’s net actual and scheduled interchange”

- However, an event may not necessarily indicate a disturbance (e.g., conventional line switching or source increase/decrease)

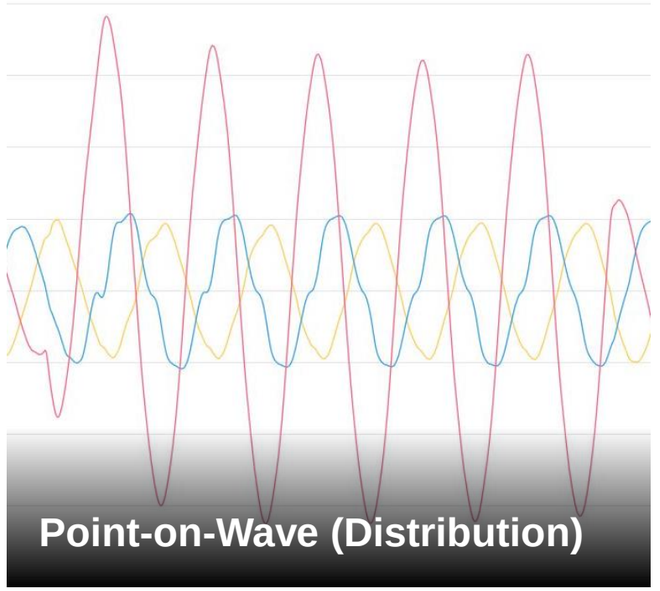
Event Labeling

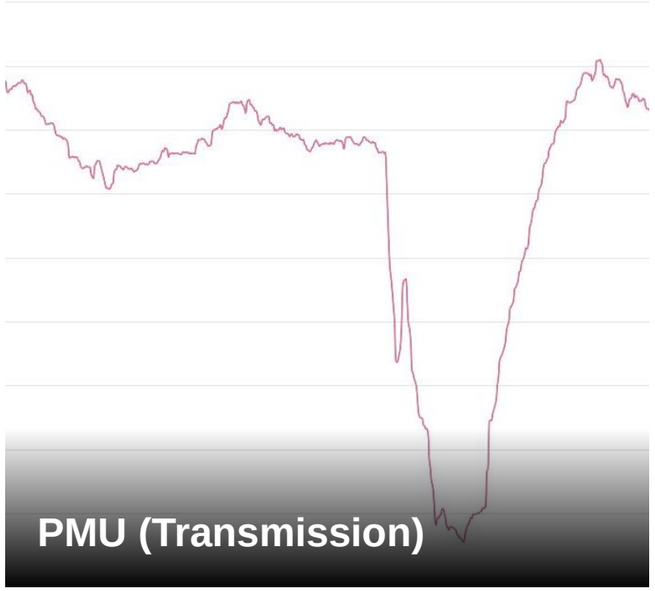
- Labeling schemes across data providers are not unique
  - Therefore, created a hierarchical “taxonomy” of events to help organize accordingly

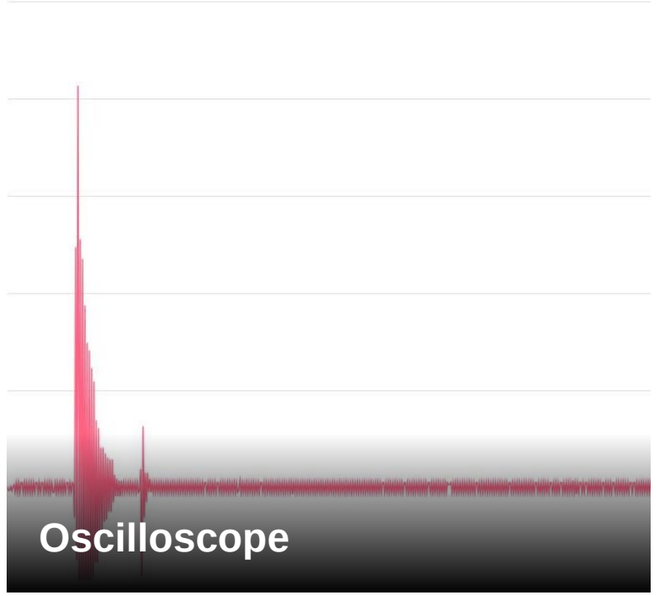


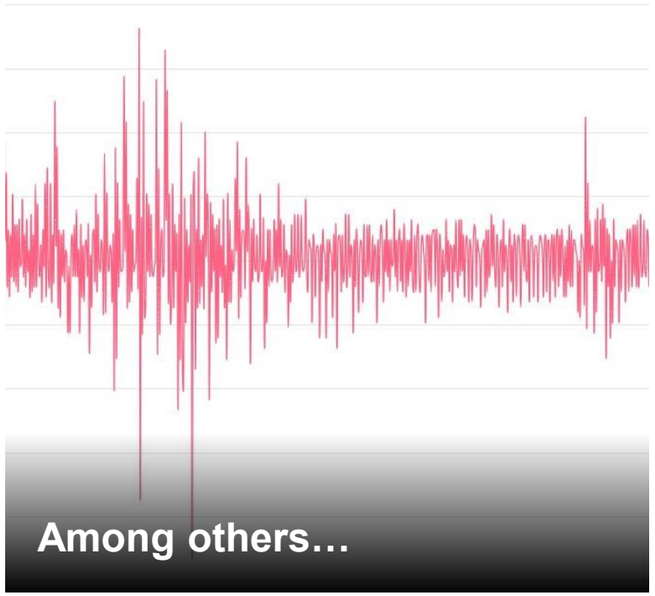
Data Types

- PoW events tend to be isolated disturbances from distribution systems
- PMU events come from transmission system locations
  - FOA 1861
  - FNET



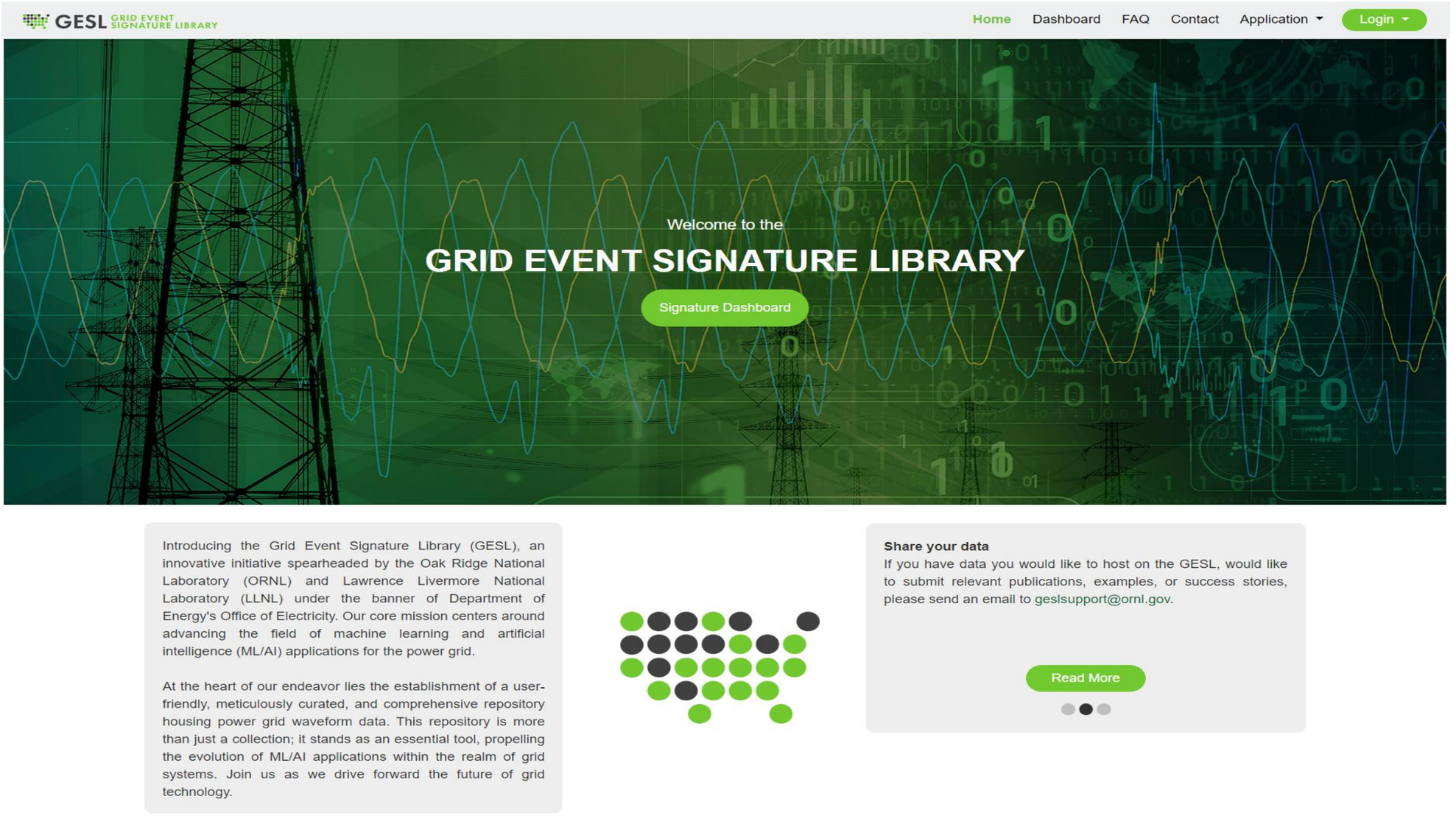




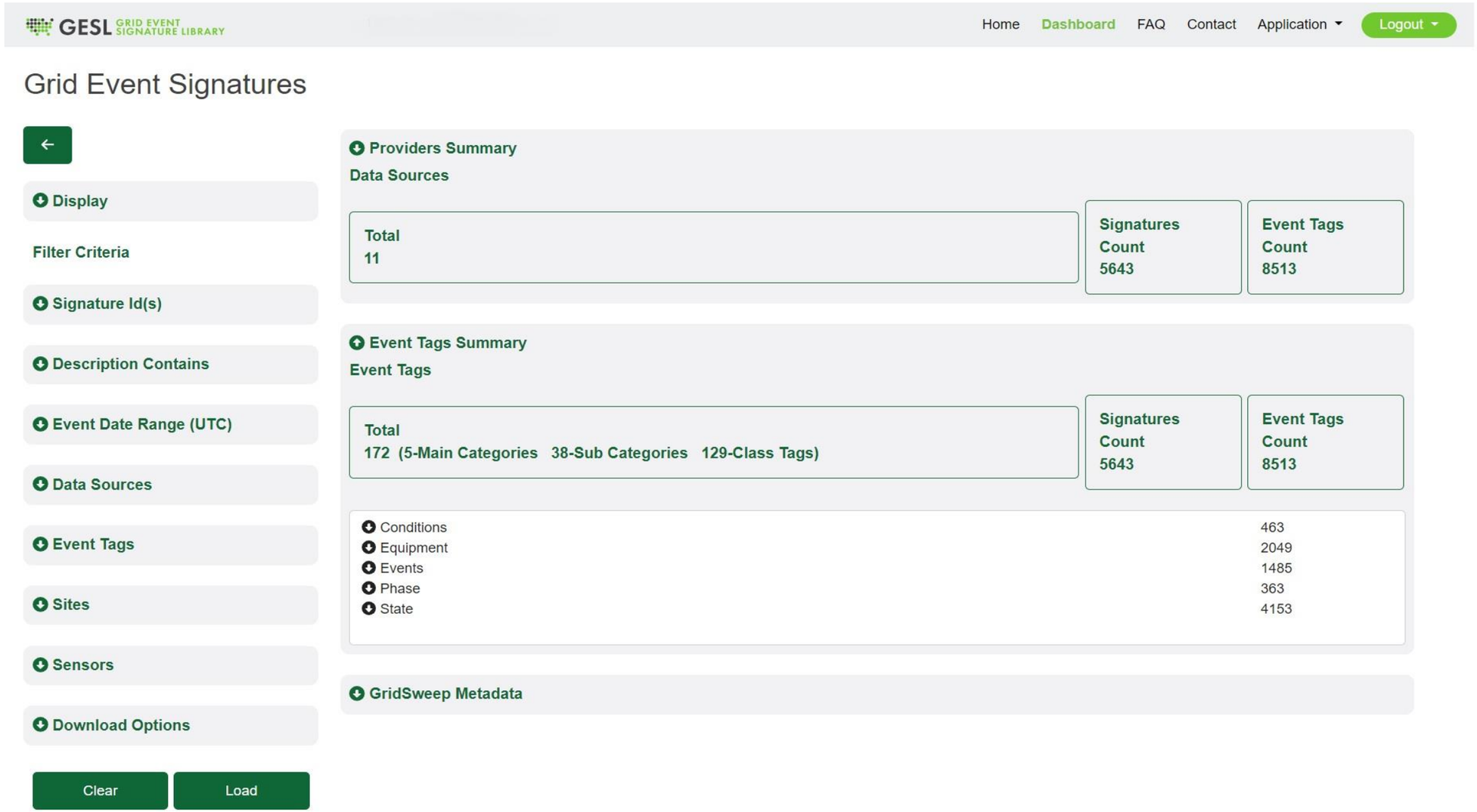


User Interface

https://gesl.ornl.gov

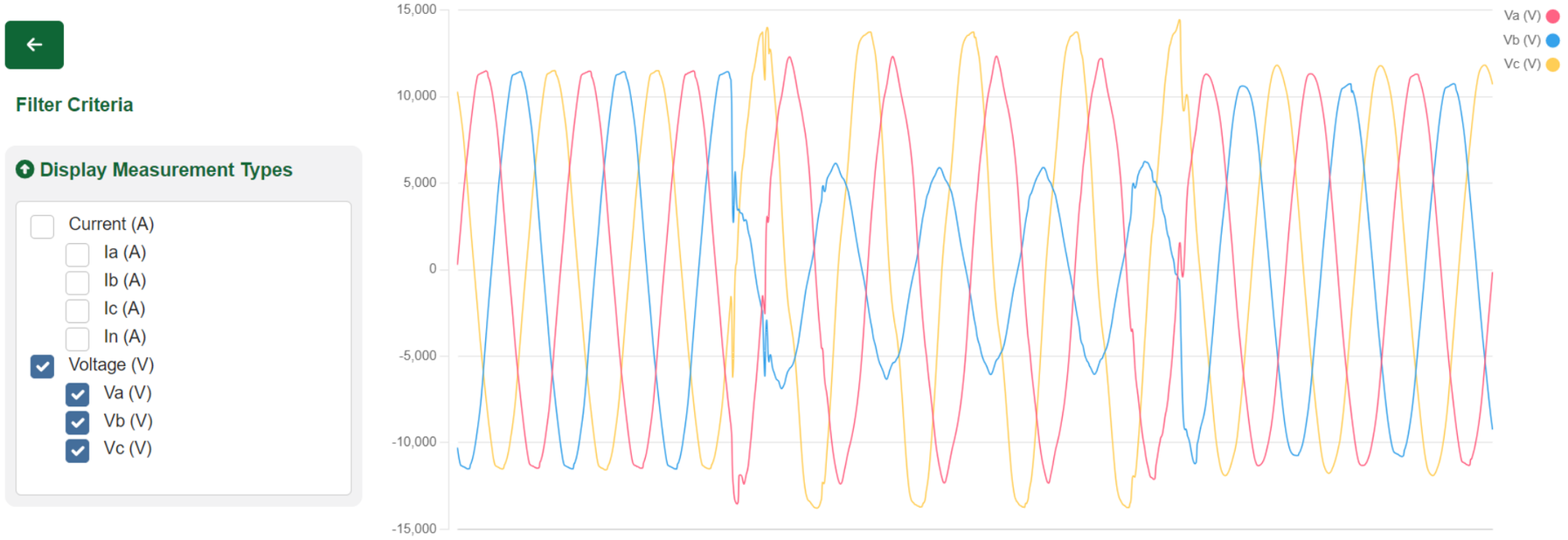


User Interface https://gesl.ornl.gov/dashboard

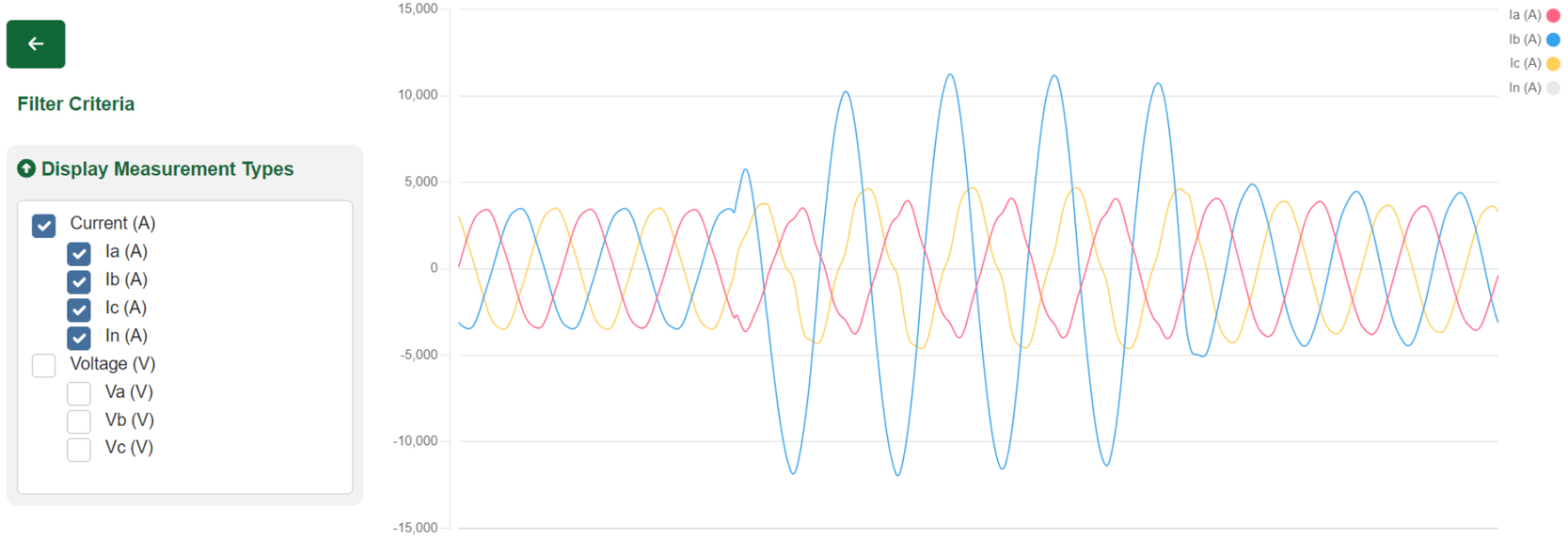


To load Signatures / data, enter filter criteria on left and click on the Load button

Waveform Signature ID: 2



Waveform Signature ID: 2



Thank you

Contact: Aaron Wilson: wilsonaj@ornl.gov, Jhi-Young Joo: joo3@llnl.gov

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

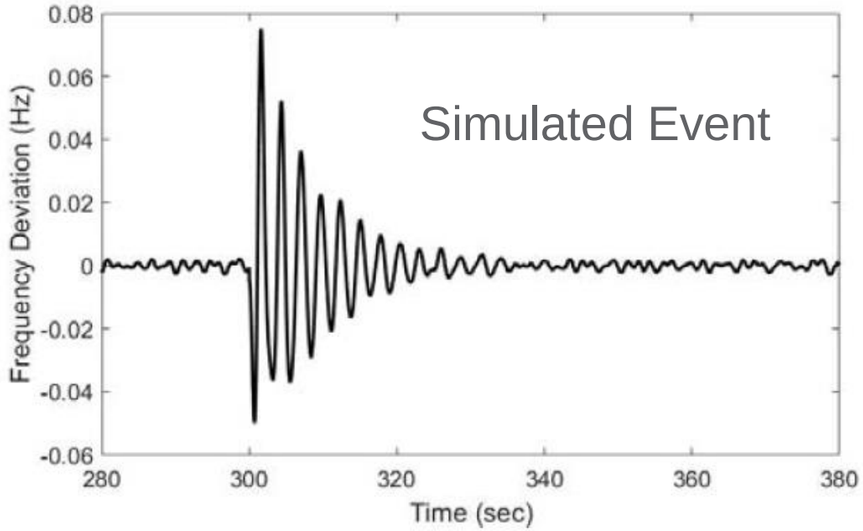
The Transmission Signature Library

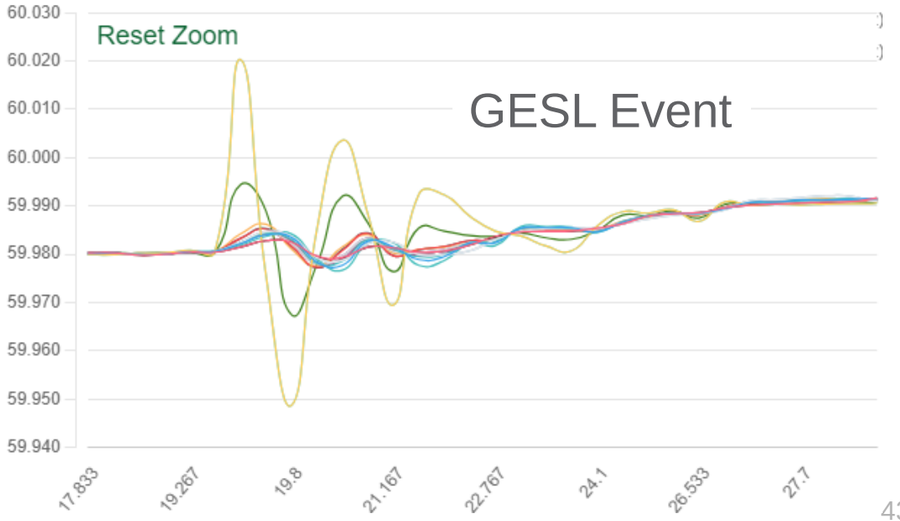
Jim Follum, PNNL

PNNL-SA-191306

Motivation for the Transmission Signature Library (TSL)

- DOE's Big Data Synchrophasor Analysis[1] program highlighted the value of providing researchers with field-measured data
  - DOE requested that portions of the data gathered for the program be made publicly available
- Common datasets improve the reproducibility of research and comparison of methods
- Simulated data cannot capture the diversity and complexity of real measurements
- The Grid Event Signature Library (GESL)[2] provided an excellent way to make real-world synchrophasor measurements publicly available





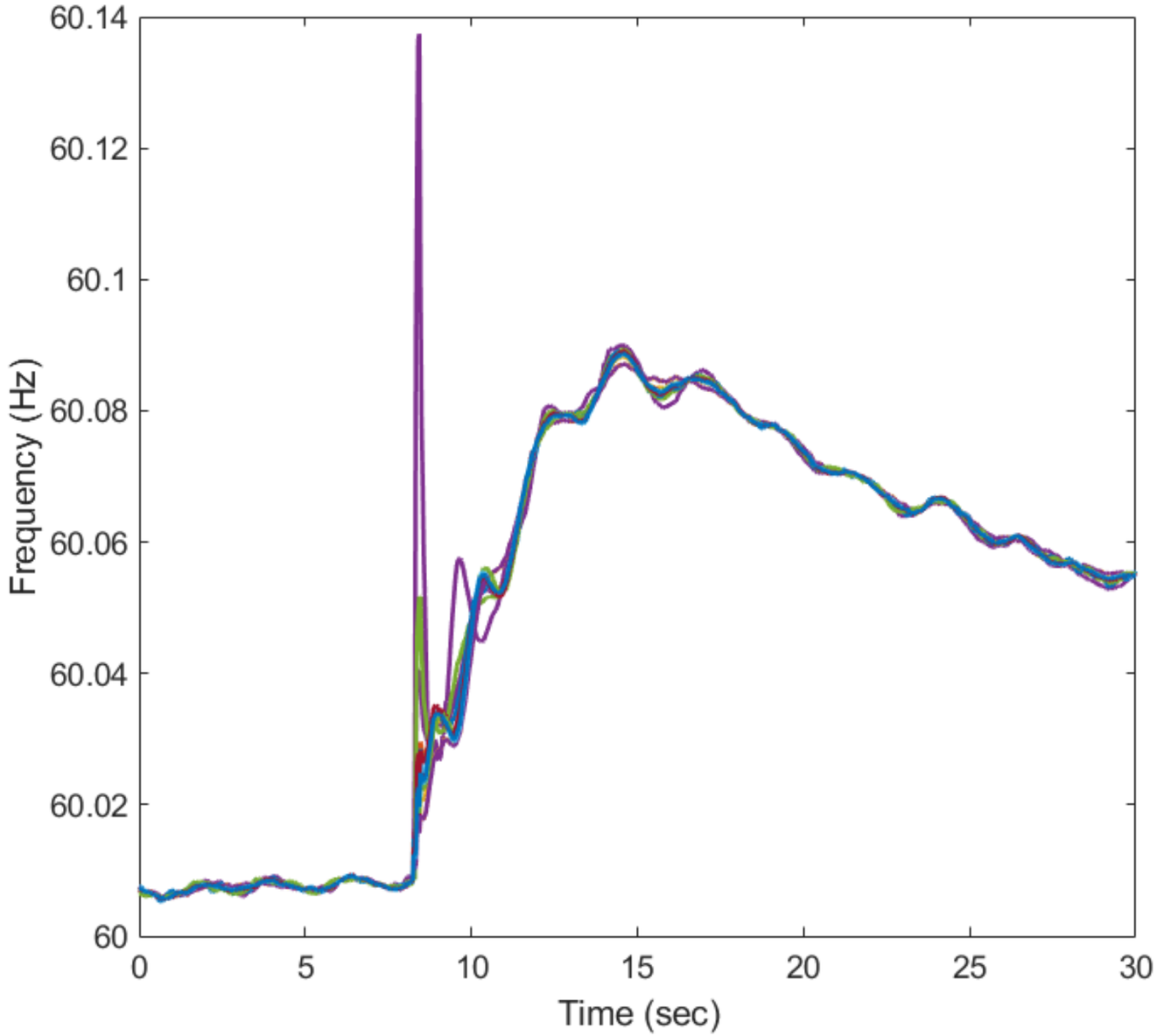
[1] https://www.energy.gov/oe/big-data-synchrophasor-analysis
[2] https://gesl.ornl.gov/

Curation

- 1694 event records were selected from the two-year dataset created for the Big Data Synchrophasor Analysis program
- Event identification
  - Event logs provided by utilities
  - Events detected by researchers participating in the Big Data Synchrophasor Analysis program
  - Sustained oscillations detected by PNNL
- Each entry was labeled using the GESL syntax
- Stored under Provider 9 (Eastern Interconnection) and Provider 10 (Western Interconnection)
- Data quality
  - Signals with severe data quality problems were removed
  - Signals with small amounts of missing data were retained

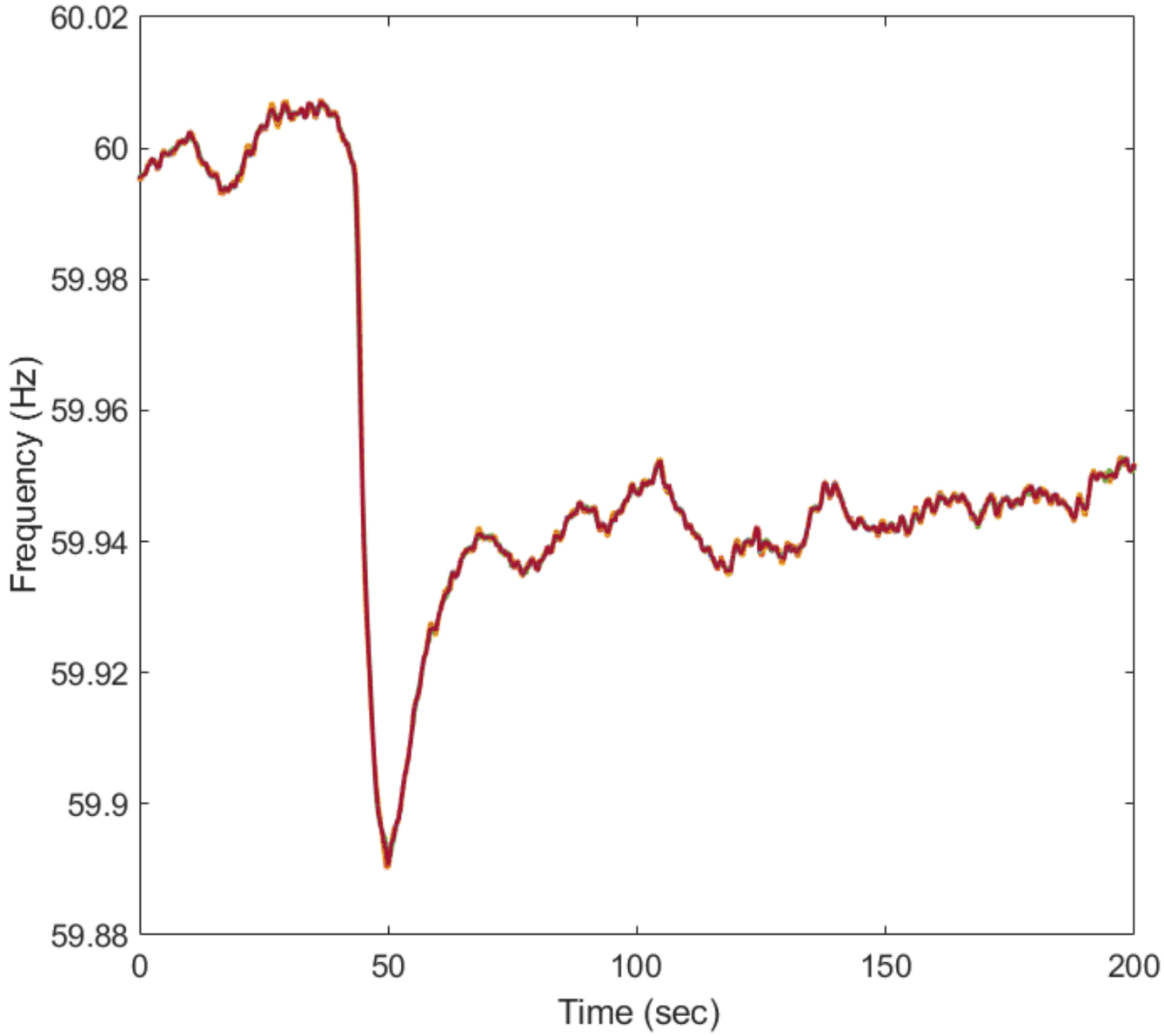
Event Types

- **Frequency Deviations**
- Generator Trips
- Oscillations
- Faults
- Ambient
  - 1 minute
  - 30 minute



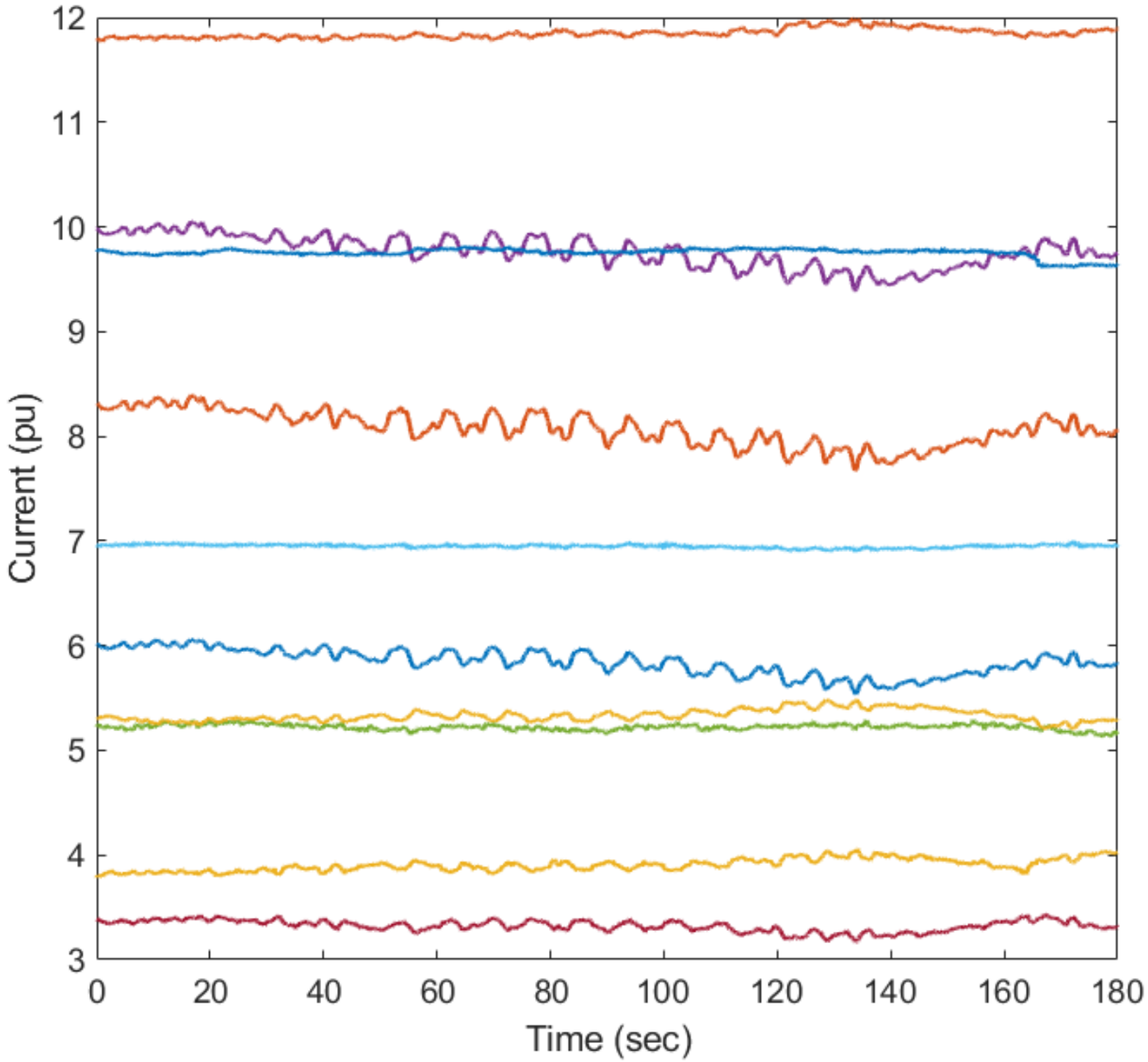
Event Types

- Frequency Deviations
- **Generator Trips**
- Oscillations
- Faults
- Ambient
  - 1 minute
  - 30 minute



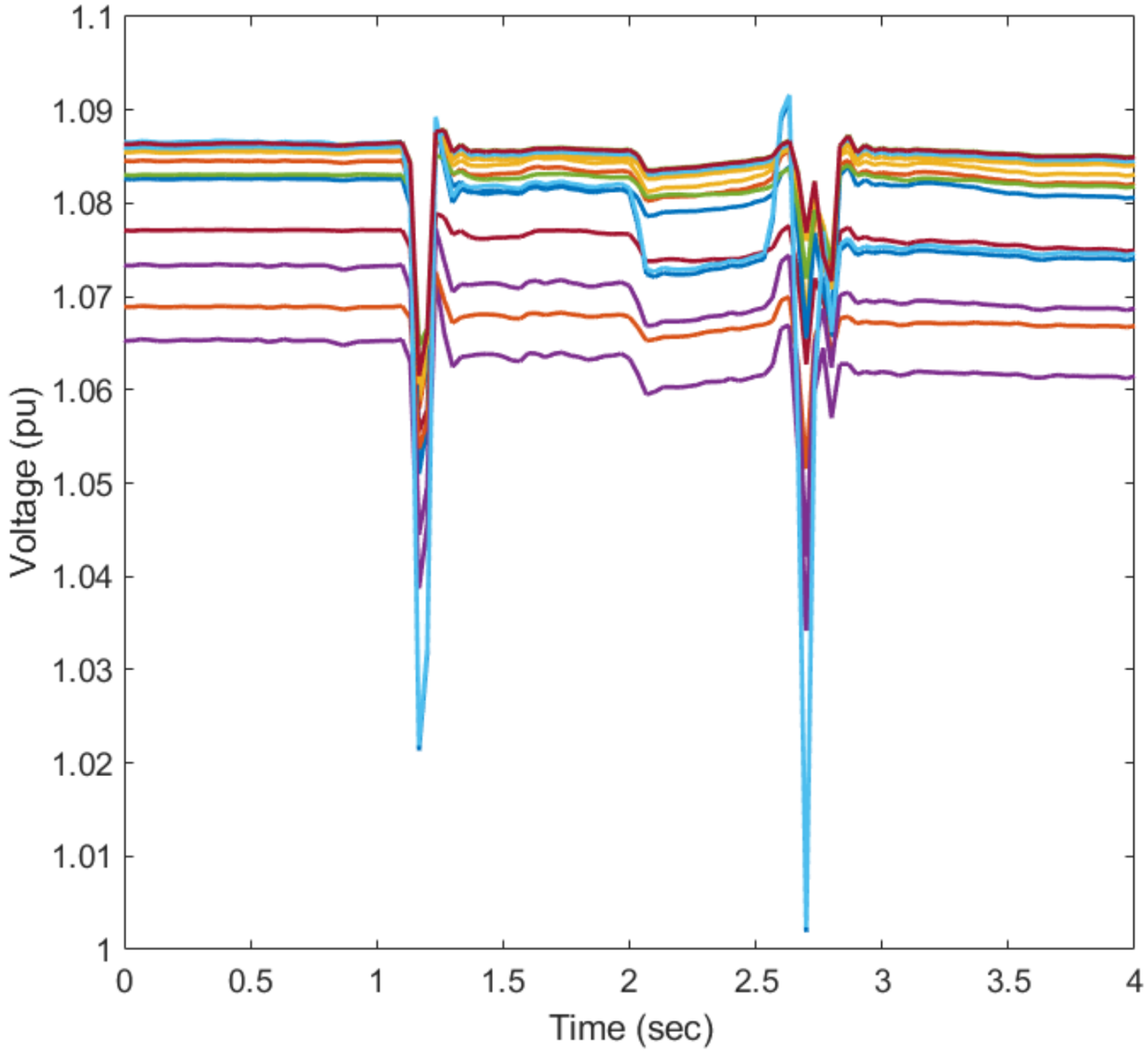
Event Types

- Frequency Deviations
- Generator Trips
- **Oscillations**
- Faults
- Ambient
  - 1 minute
  - 30 minute



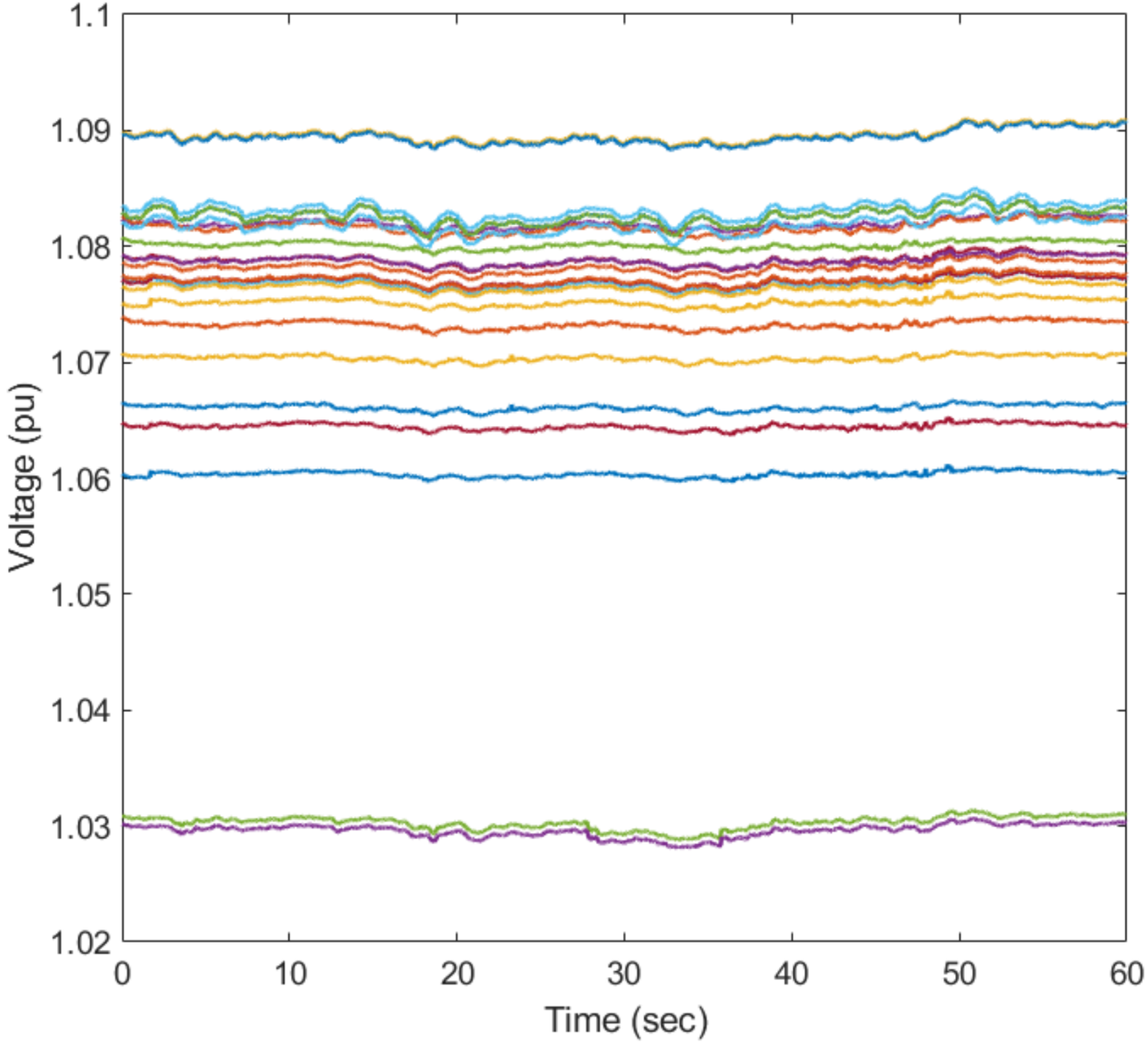
Event Types

- Frequency Deviations
- Generator Trips
- Oscillations
- **Faults**
- Ambient
  - 1 minute
  - 30 minute



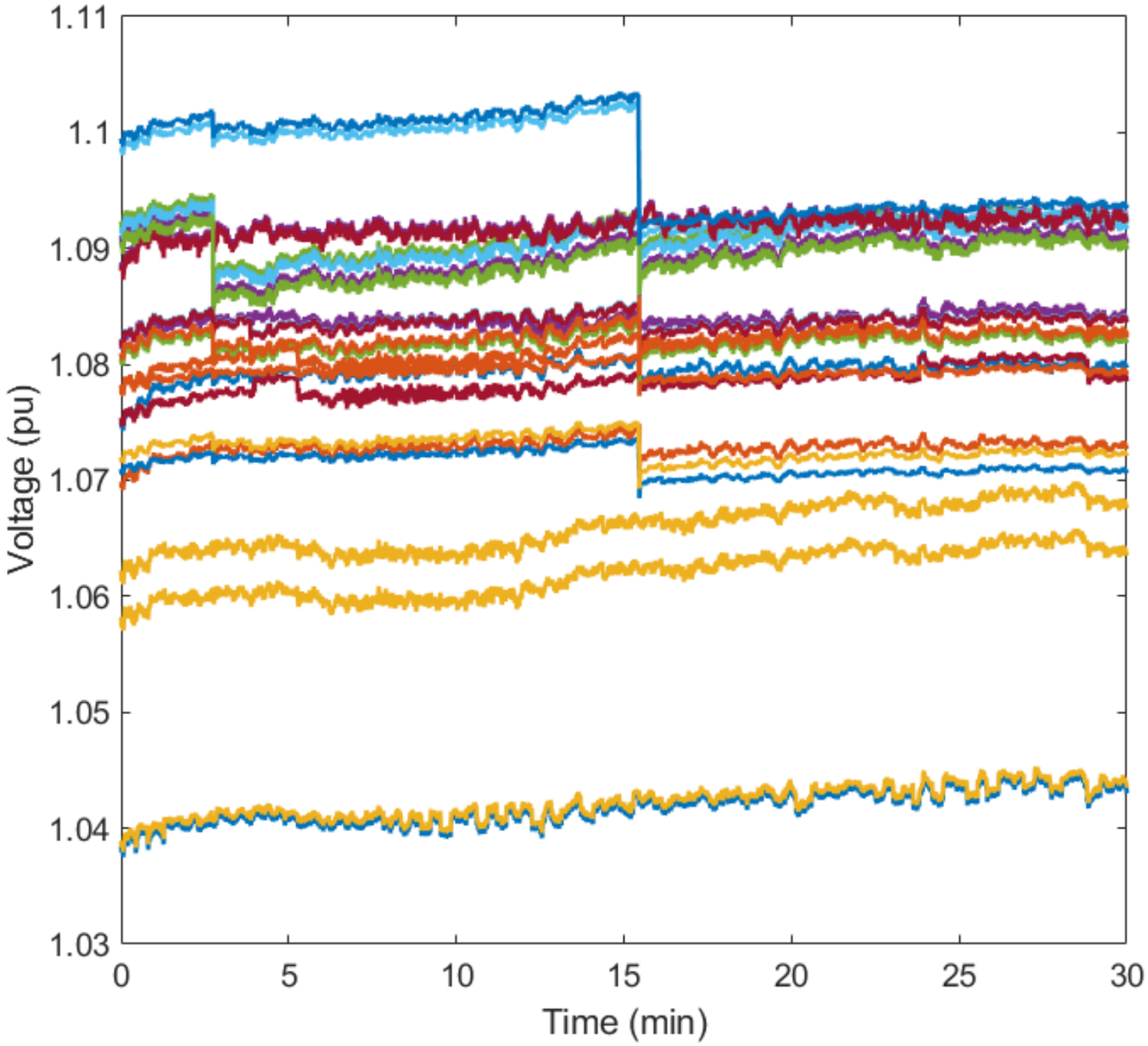
Event Types

- Frequency Deviations
- Generator Trips
- Oscillations
- Faults
- Ambient
  - **1 minute**
  - 30 minute



Event Types

- Frequency Deviations
- Generator Trips
- Oscillations
- Faults
- Ambient
  - 1 minute
  - **30 minute**

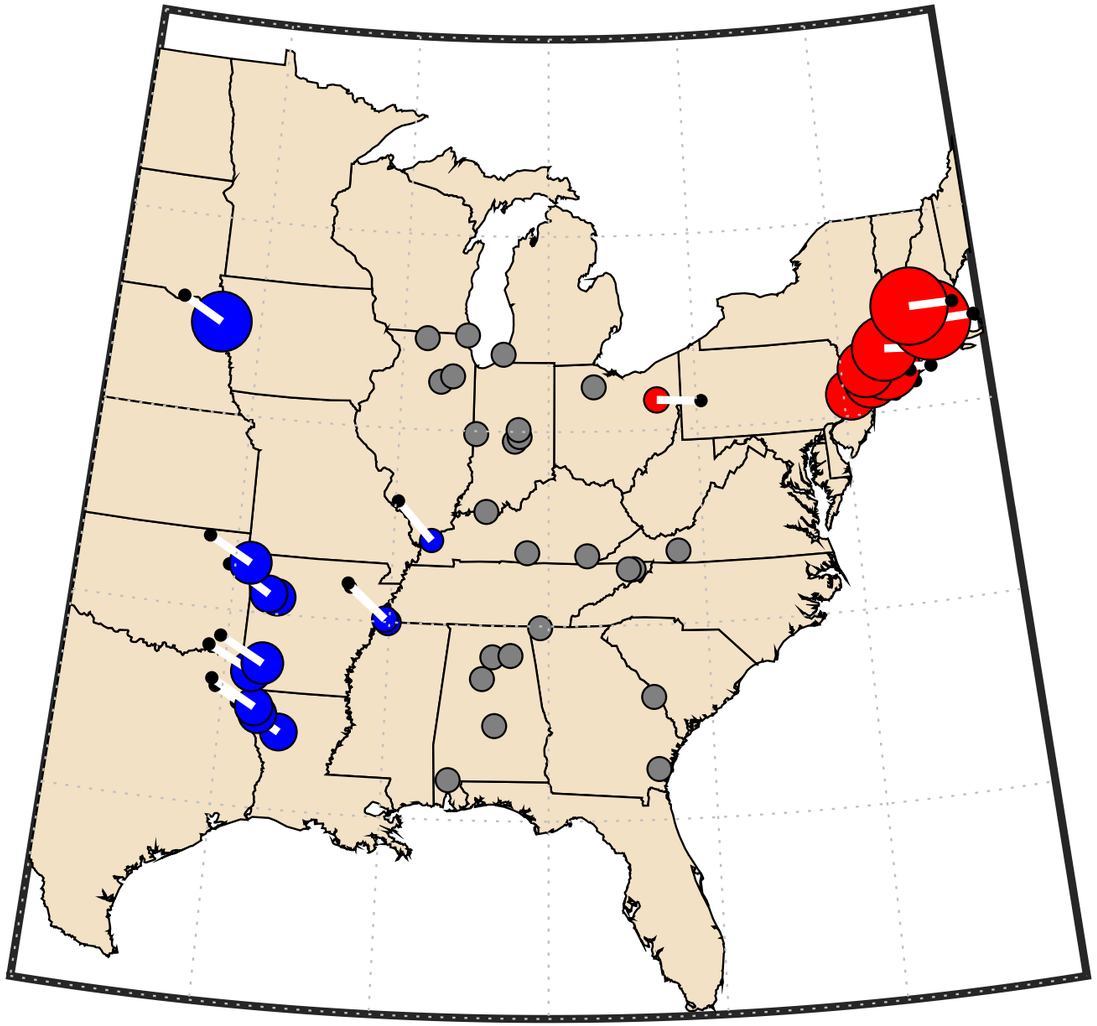


Obfuscation Procedure

- The TSL's anonymization procedure was informed by:
  - A study reviewing the effectiveness of the anonymization process implemented for the Big Data Synchrophasor Analysis program
  - Previously published Test Case Library of Power System Sustained Oscillations - https://web.eecs.utk.edu/~kaisun/Oscillation/actualcases.html
  - Most importantly, meetings with several data providers
- Anonymization steps:
  - No topology information
  - Single current measurement per PMU
  - Substation names removed
  - PMU identifiers randomized for each event
  - All data at 30 frames per second
  - UTC timestamps removed; month and year retained
  - Voltage and current magnitudes converted to per unit

Obfuscation Limitations

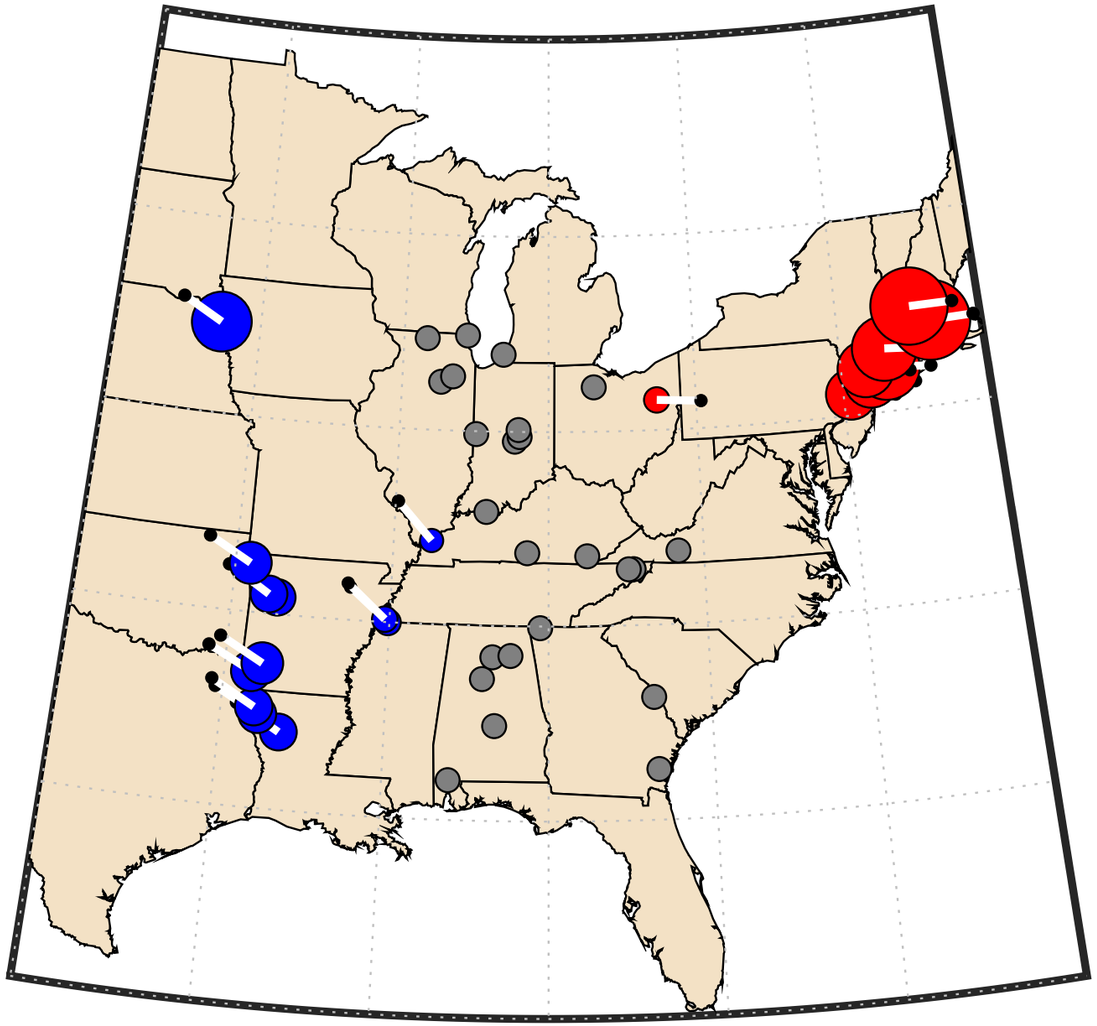
- Can PMU data be perfectly anonymized and remain useful? No.
  - Each interconnection has distinct characteristics
  - A system's dynamics cannot be removed without destroying the data



Shape of the Eastern Interconnection's Northeast-Midwest Mode of Oscillation

Obfuscation Limitations

- Can PMU data be perfectly anonymized and remain useful? No.
  - Each interconnection has distinct characteristics
  - A system's dynamics cannot be removed without destroying the data
- Is it possible to obfuscate data to the point where data owners feel comfortable making it publicly available? Yes.



Shape of the Eastern Interconnection's Northeast-Midwest Mode of Oscillation

Is Obfuscated Data Useful?

- Anonymized data has limited value for analyzing the underlying **system**
- It has significant value for evaluating **algorithms**
  - Event detection and classification
  - Automated frequency response analysis
  - Oscillation detection and analysis
  - Managing missing data
- In our ongoing work, we are highlighting how the TSL can support a variety of use cases

IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

An Open-Source Library of Phasor Measurement Unit Data Capturing Real Bulk Power Systems Behavior

SHUCHISMITA BISWAS (Member, IEEE), JIM FOLLUM (Member, IEEE), PAVEL ETINGOV (Member, IEEE), XIAOYUAN FAN (Senior Member, IEEE), and TIANZHIXI YIN

Pacific Northwest National Laboratory (PNNL), Richland, Washington, USA (e-mail: shuchismita.biswas@pnnl.gov)

Corresponding author: Shuchismita Biswas (e-mail: shuchismita.biswas@pnnl.gov).

The Pacific Northwest National Laboratory is operated by Battelle for the U.S. DOE under Contract DE AC05 76RL01830.

ABSTRACT This paper describes an open-source library of transmission-level synchrophasor measurements, curated with the aim of accelerating data-driven research and development in the power systems domain. This dataset contains measurements describing both disturbances and ambient conditions, spans two years in time, and is sourced from electric utilities across the United States. Comprised of 1694 unique events, this is the largest open-source repository of real transmission-level phasor measurement unit (PMU) data to date, and will be invaluable for benchmarking new algorithms, testing tools and approaches developed by vendors and researchers, and developing educational tools for university students and control room operators. This paper additionally highlights several potential applications of the dataset that may be useful to the research community.

INDEX TERMS Phasor measurement unit, dataset curation, benchmark dataset, Synchrophasor dataset

I. INTRODUCTION

PHASOR measurement units (PMUs) record time-synchronized high-resolution measurements of power systems quantities, providing a granular insight into power grid dynamic behavior [1]. Ever since commercial PMUs became available and were installed by electric utilities around the world in the last few decades, PMU data has provided wide-area situational awareness to system operators and enabled a plethora of applications such as event detection and localization [2], [3], natural oscillation monitoring [4], [5], forced oscillation detection [6], [7], equipment failure prediction [8], [9] and dynamic model validation [10]–[12]. However, given the sensitive nature of the critical infrastructure that PMUs monitor, electric utilities are reluctant to share PMU data with researchers without contractual safeguards in place. This is a barrier to resource-constrained researchers who may lack institutional relationships with utilities, and thus end up using simulated/synthetic measurements for their research. This prevents researchers from addressing unique challenges present in real power systems measurements and creates considerable obstacles in improving the technology readiness levels of research-grade algorithms. Further, the lack of publicly available labeled datasets has limited the development of machine-learning based applications in the power system domain.

Recognizing this challenge while seeking to accelerate data-driven research and development in power systems, the US Department of Energy (DOE) and Pacific Northwest National Laboratory (PNNL) aggregated and anonymized two years of transmission-level PMU data along with event logs from electrical utilities across the United States [13]. Eight DOE research grant awardees (henceforth referred to as awardees) were provided access to this National PMU Dataset (NPDS) under non-disclosure agreements (NDA). The awardees developed several approaches aimed at data cleaning, event detection, and disturbance classification that were used to identify additional events not documented in the original event logs [14]. Building on these efforts, a library of representative events has been extracted from the NPDS, validated by domain experts, and made publicly available through the Grid Event Signature Library (GESL) framework maintained by the Oak Ridge National Laboratory (ORNL) and Lawrence Livermore National Laboratory (LLNL) [15].

This paper describes this dataset of event signatures (referred to as the Transmission Signature Library or TSL in the rest of this work) and how it has been curated, labeled, and validated. Some of the many potential applications of the TSL dataset are illustrated and future avenues of ex-

VOLUME 4, 2016

1

Conclusion

- The Transmission Signature Library has several benefits
  - It will help researchers develop algorithms that can handle the real-world diversity and complexity of field-measured data
  - It provides researchers with a common dataset to facilitate comparison of methods
  - It will make research more reproducible by making the underlying data available
- Public release required obfuscation
  - The anonymization procedure was informed by prior work and discussion with data providers
  - No practical anonymization is perfect, but it can be acceptable to data owners
  - Despite anonymization, the dataset remains valuable in supporting research
- See our open-access journal paper to learn more about the TSL:
  - S. Biswas, J. Follum, P. Etingov, X. Fan and T. Yin, "An Open-Source Library of Phasor Measurement Unit Data Capturing Real Bulk Power Systems Behavior," in IEEE Access, vol. 11, pp. 108852-108863, 2023, doi: 10.1109/ACCESS.2023.3321317.

Thank you

James Follum, PNNL
James.Follum@pnnl.gov

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Topic 2 - Cross-sector and Open Data Sharing and Risks

GridSweep® instrument: Sharing the data

Alex McEachern
Affiliate, Lawrence Berkeley National Lab
ALEX@MCEACHERN.COM

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

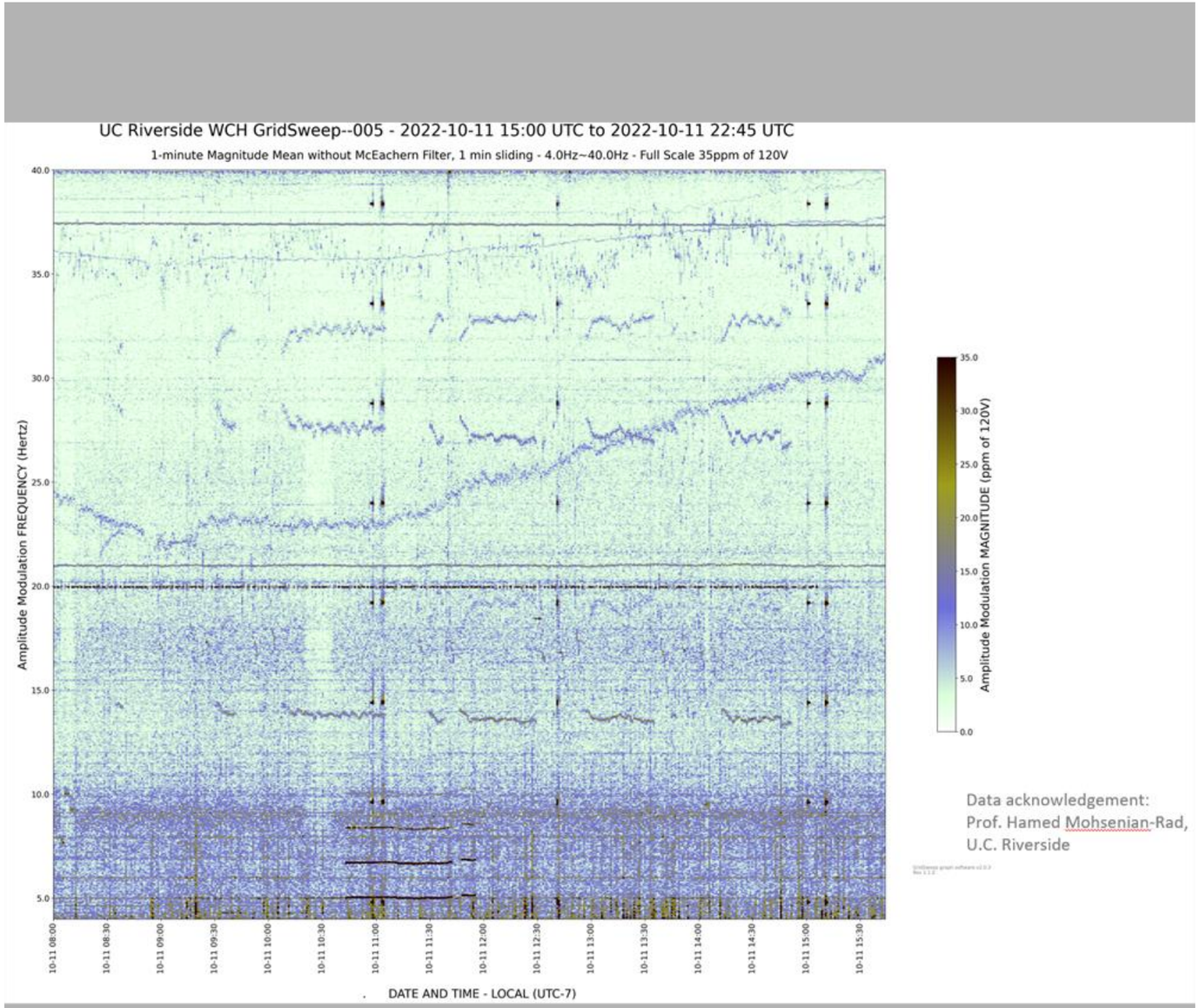
GridSweep Instrument

- A new instrument for probing distribution grid stability (GMLC/SETO/OE)
- Parts-per-billion voltage resolution
- Applications
  - Reduce risks of increased solar/battery deployment
  - Detect **rehearsals** of substation cyberattacks
  - Novel information about distribution grid behavior
  - **Fleet available at LBNL**

GridSweep – 4-year DOE-funded development & deployment

GridSweep Instrument Data

- Creates open-source data for **research**, optimized for voltage oscillations below 60 Hz
- GPS-sync Continuous-Point-On-Wave voltage measurements, at **parts-per-billion** resolution
- **All measurements:**
  - **Made at 120-volt outlets**
  - **Voltage only**
  - **Stored locally**



GridSweep data – 8-hour, 5Hz - 40Hz ,heat map at U.C. Riverside

GridSweep Instrument: Initial Deployments

McEACHERN LABORATORIES INC

ALAMEDA MUNICIPAL POWER
A Department of the City of Alameda

BERKELEY LAB

IDAHO POWER®
An IDACORP Company

Hawaiian Electric

UNIVERSITY OF CALIFORNIA
UC RIVERSIDE

Dominion Energy®

GridSweep deployment – Hawaiian Electric Co

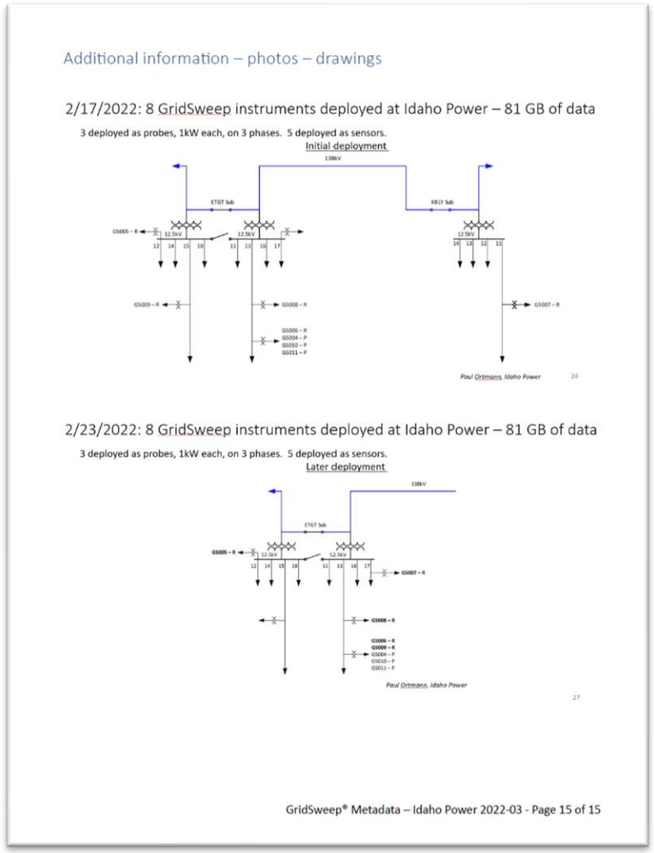
GridSweep Instrument: Design Challenges

GridSweep design goal	Solution	Challenges
Measure distribution grid characteristics, while avoiding extensive utility approvals.	All measurements (and probing) only on 120-volt outlets.	Develop instrument capable of parts-per-billion voltage measurements, plus probing device, plus advanced signal extraction algorithms.
Avoid extensive utility IT approvals.	Absolutely no internet/intranet connection. Hard drive.	Works well for research. (But not acceptable for control.)
Plug into 120-volt outlets at utility office buildings, homes, substations, etc.	GridSweep instrument is UL-listed, FCC-certified, CE-marked	Aside from design challenge, UL-listing requires UL-certified manufacturing location.
Open-source release all raw data, processed data, and software.	**ORNL's Grid Signature Library.**	**Typical "signature" in ORNL library is 0.1 megabyte. Typical GridSweep measurement is 100's of gigabytes.**

GridSweep Instrument: Data Sharing Results

- GridSweep analysis software is now released, open-source.
- Approximately two terabytes of GridSweep data is pending release on ORNL “Grid Signature Library”
  - Metadata
  - Raw GPS-CPOW voltage recordings – Excel files
  - Processed voltage recordings – Excel files

Examples:
GridSweep experiment metadata



Additional information – photos – drawings

2/17/2022: 8 GridSweep instruments deployed at Idaho Power – 81 GB of data

3 deployed as probes, 1kW each, on 3 phases. 5 deployed as sensors.

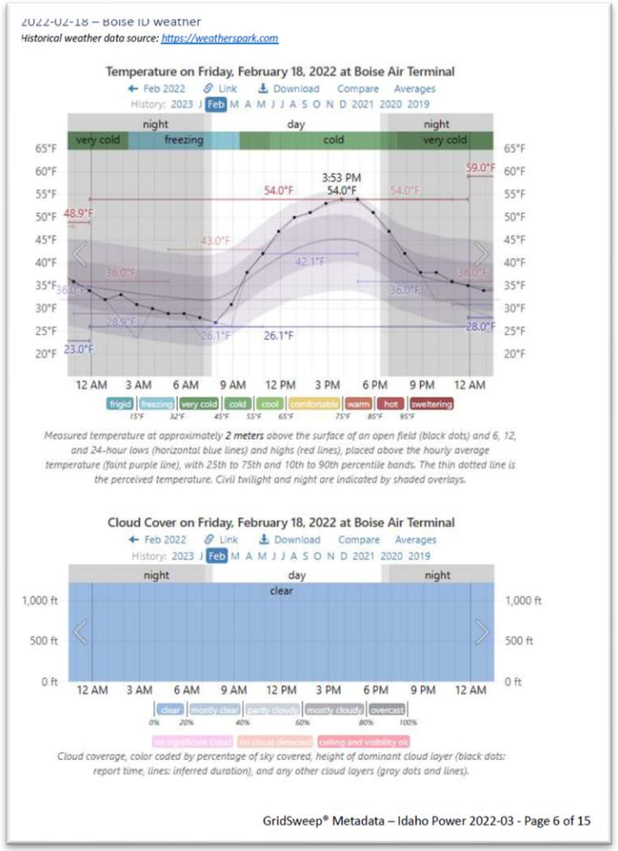
Initial deployment

2/23/2022: 8 GridSweep instruments deployed at Idaho Power – 81 GB of data

3 deployed as probes, 1kW each, on 3 phases. 5 deployed as sensors.

Later deployment

GridSweep® Metadata – Idaho Power 2022-03 - Page 15 of 15



2022-02-18 – Boise ID weather

Historical weather data source: https://weatherspark.com

Temperature on Friday, February 18, 2022 at Boise Air Terminal

Measured temperature at approximately **2 meters** above the surface of an open field (black dots) and 6, 12, and 24-hour lows (horizontal blue lines) and highs (red lines), placed above the hourly average temperature (faint purple line), with 25th to 75th and 10th to 90th percentile bands. The thin dotted line is the perceived temperature. Civil twilight and night are indicated by shaded overlays.

Cloud Cover on Friday, February 18, 2022 at Boise Air Terminal

Cloud coverage, color coded by percentage of sky covered, height of dominant cloud layer (black dots: report time, lines: inferred duration), and any other cloud layers (gray dots and lines).

GridSweep® Metadata – Idaho Power 2022-03 - Page 6 of 15

GridSweep Instrument: Data Sharing – Lessons Learned

1. GridSweep measurements reveal previously unknown characteristics of distribution feeders and may revise the way we think about IBR deployment limits.
2. By restricting GridSweep measurements to 120-volt outlets, we discovered that utilities had little to no concern about open-source releasing the data (even though, given GridSweep's precision, characteristics of the distribution feeders and substations might be revealed).
3. Avoiding the utility IT requirements, i.e. completely eliminating any instrument internet/intranet connection, made it a LOT easier to install instruments.
4. The two big technical challenges
    1. Making parts-per-billion voltage measurements in residential/office environments
    2. UL listing, FCC certification, CE marking, and especially UL-certified manufacturing
5. It was difficult to get guidance from DOE regarding what we perceived as possibly sensitive uses of this publicly-released data.
6. It was challenging, as newbies, to work through the open-source release requirements for the software that we had written.

Thank you

Alex McEachern
Alex@mceachern.com

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Outage Data Initiative Nationwide (ODIN)

Supriya Chinthavali
Group Leader, Geospatial Sciences and Human Security Division (ORNL)

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

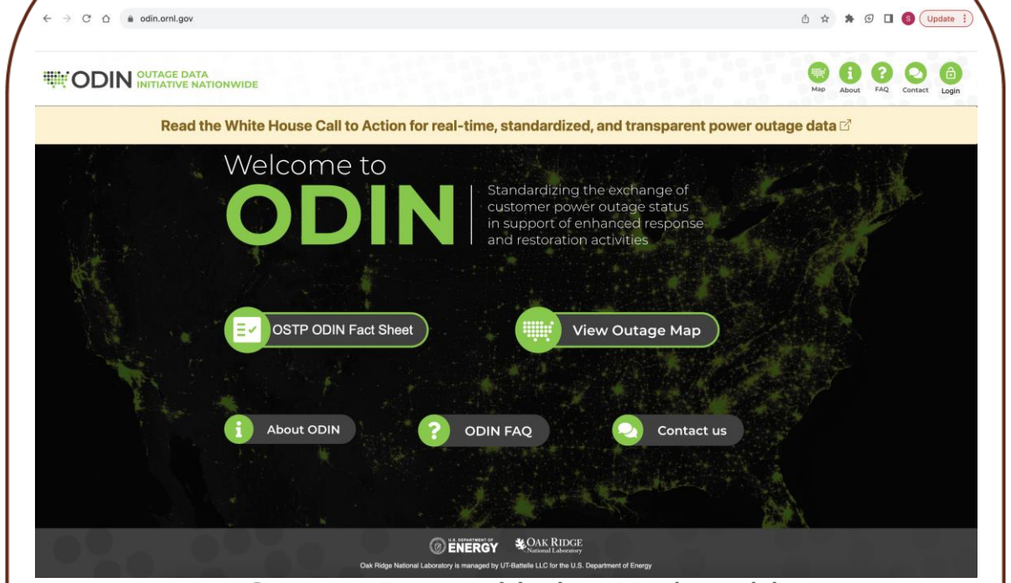
Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

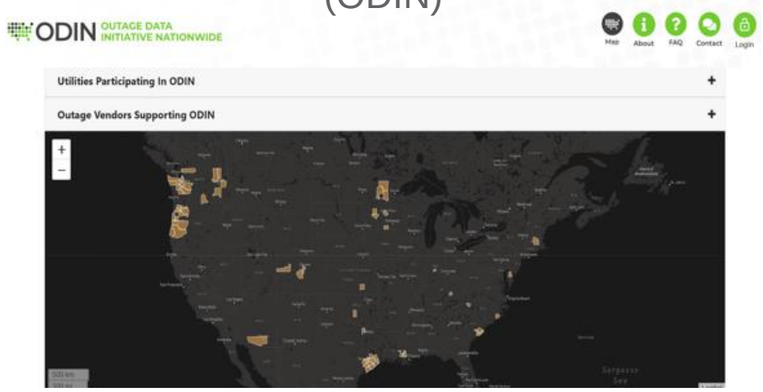
Pacific Northwest NATIONAL LABORATORY

ODIN - Outage Data Initiative Nationwide

Data & Interoperability



Outage Data Initiative Nationwide (ODIN)

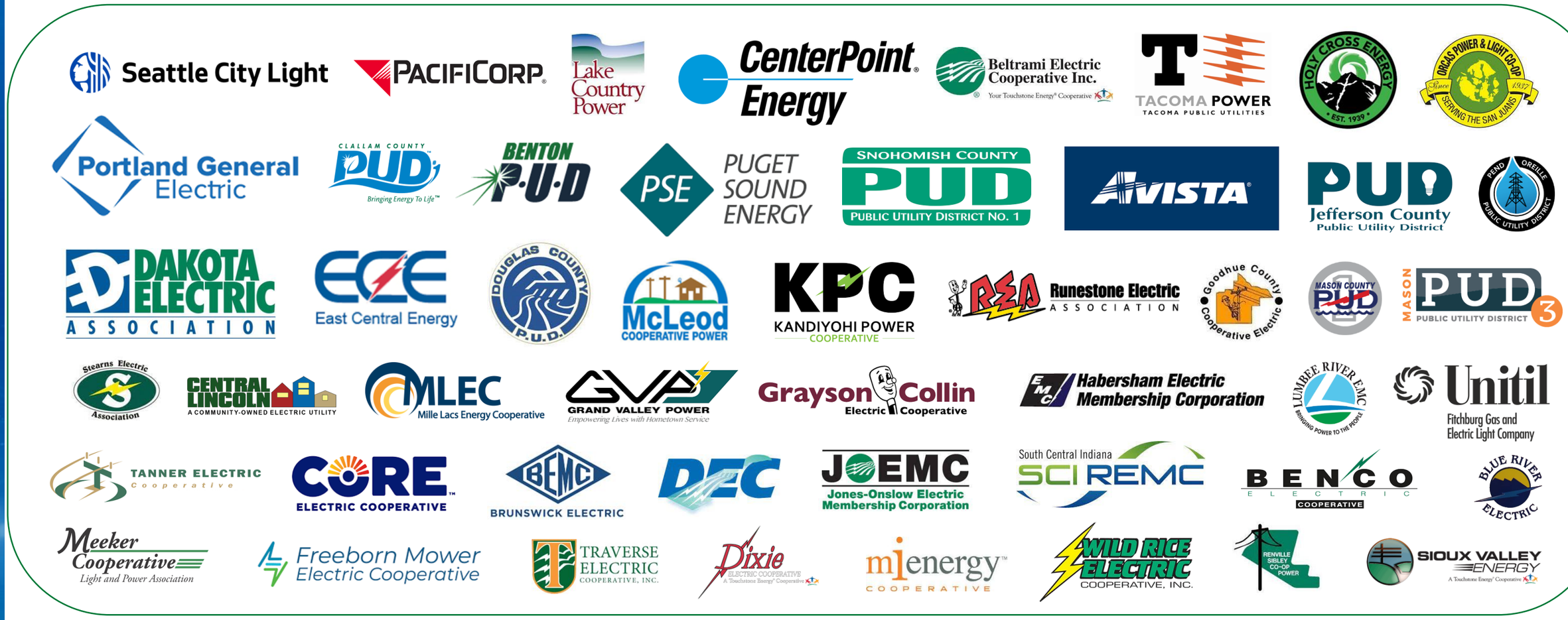


https://odin.ornl.gov

Problem: Outage data from utilities is valuable to customers, neighboring utilities, and regional emergency management partners, but data is too often fragmented, unavailable, and/or lacking commonalities

Solution: ODIN is a network of leading electric service providers who are committed to providing comprehensive interoperable power outage data that enables utilities and others to exchange data freely with designated stakeholders at all levels — helping restoration, reliability, risk mitigation, emergency response, and more

The ODIN Network



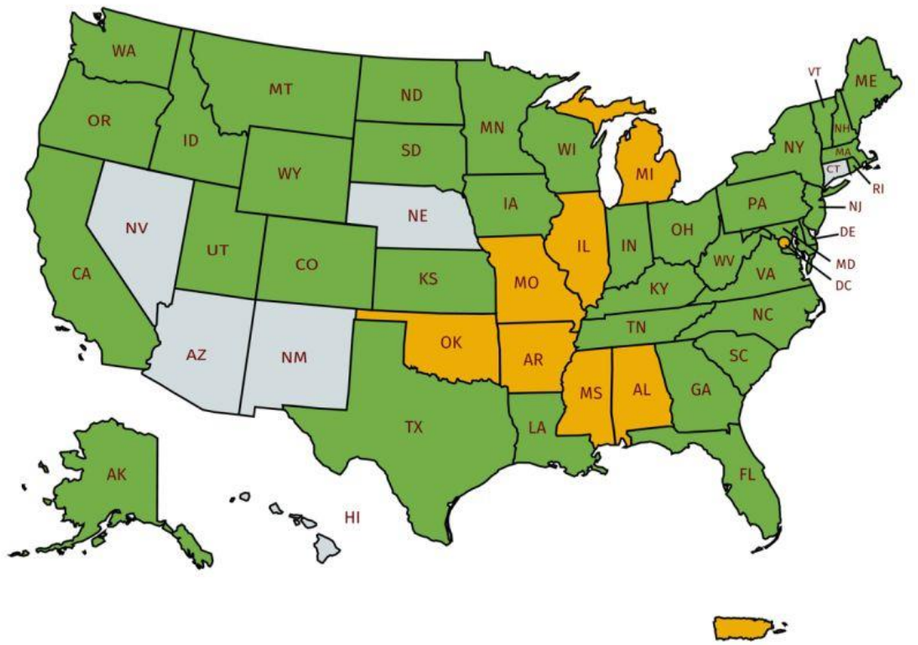
White House Call to Action

- ODIN effort expanded real-time collection of data to **nearly 44 states and Puerto Rico**
- Committed utilities were highlighted at the **White House Electrification Summit** on December 14, 2022

ODIN Network

125
Utilities Participating

14
Supporting Vendors



Coverage in 48 states and PR!

Benefits of Data Sharing

Unlock opportunities for **federal funding**

Control and **authoritatively** share your data using ODIN

Spend less time on the phone in an emergency, allowing you to concentrate on **crucial** restoration efforts

Save lives in underserved communities and for the electricity-dependent

Commit to leading on transparent sharing and resilience

On the Same Page When It Matters

ODIN standardization allows you to effortlessly share authoritative, real-time data with:

NEIGHBORING UTILITIES

EMERGENCY MANAGEMENT

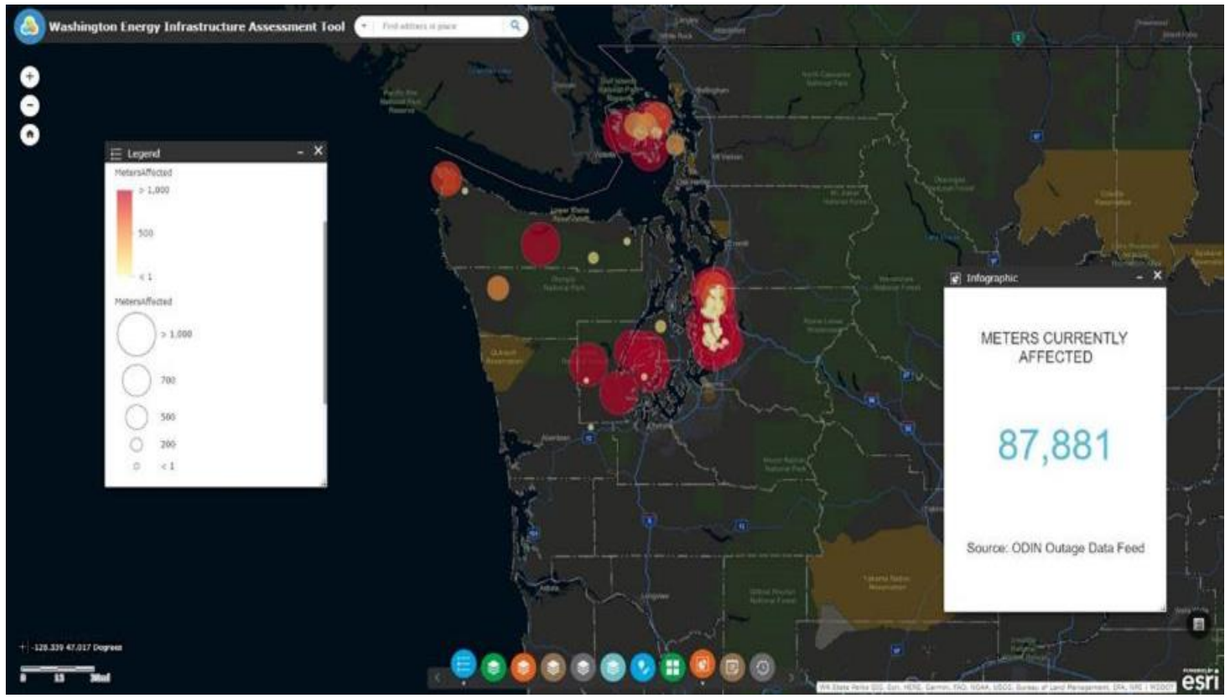
GENERAL PUBLIC

Federal Funding

Funding Source	Funding Amount	Purpose
Grid Resilience Innovation Partnership program (GRIP) (BIL)	**$10.5 Billion** to Utilities and Related Operators, participation, and ODIN participation is a Program Policy Factor in Topic Area 1, Grid Resilience Grants	Support innovative approaches to transmission, storage, and distribution infrastructure to enhance grid resilience and reliability.
Inflation Reduction Act	**$40 billion** in loan authority	Transmission expansion projects and emerging technologies, including HVDC deployment and GETS manufacturing and deployment.

Use Case: Washington Department of Commerce

WASHINGTON ENERGY
INFRASTRUCTURE
ASSESSMENT TOOL



Collaborative efforts operational since 2021

- WA can monitor outages from storms and other events in a single “pane of glass,” including wildfire and utility service territories
- This supports resilience planning and coordinates mitigation planning between utilities and local emergency management

Use Case: Minnesota Rural Electric Association

A statewide association representing 50 not-for-profit coops

85%	⅓ Pop.	138k Miles
of Minnesota's Landmass	of Minnesotans Served	of Electric Line

Standardized outages provide visibility to multiple stakeholders

"Creating a state-wide map of outage data will benefit stakeholders in Minnesota before, during, and after high-impact grid and weather events."

~ Darrick Moe, President & CEO, Minnesota Rural Electric Association

Source: https://www.mrea.org/

Participation Made Easy

Integrate through your outage data vendor

AERINET SOLUTIONS	datacapable	esri	FUTURA SYSTEMS	MILSOFT Utility Solutions
nisc	ORACLE	OSI An AspenTech Business	SIENA	STR energy services

Contact odin@ornl.gov today to see if your vendors support the ODIN standard

Integration Options

Standards

- Common Information Model (CIM) IEC 61968-3
- MultiSpeak v4.1 (and greater)

Methods

- Vendor Supported Integration
- ESB Integration
- Utility Development via API

As fast as a 30-minute integration setup
one-on-one support available at no additional cost

Secure Exchange of Data

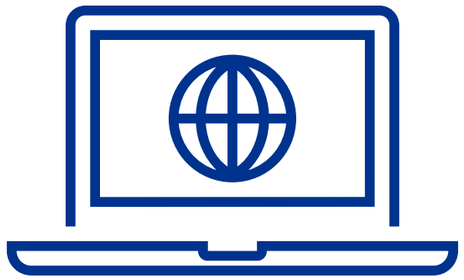
The Outage Data Initiative Nationwide is part of the the U.S. Department of Energy and Oak Ridge National Laboratory.

Oak Ridge National Laboratory is managed by UT-Battelle LLC for the U.S. Department of Energy

Commit to Lead Resilience Today

1. Fill out and send the Participation Letter to ODIN

2. The ODIN team will schedule a 30-minute kickoff meeting to discuss implementation



Visit https://odin.ornl.gov

Thank you

Supriya Chinthivali
Chinthivalis@ornl.gov

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Open Energy Data Initiative for Solar Data and Analytics

Kemal Çelik, Ph.D.
DOE/SETO/SI Technology Advisor

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

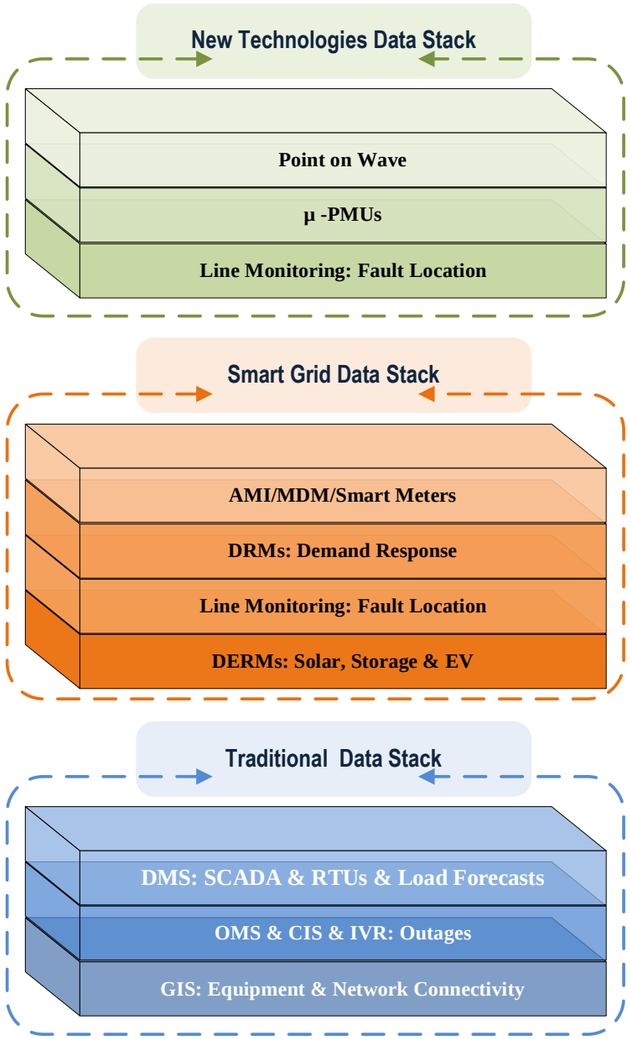
Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

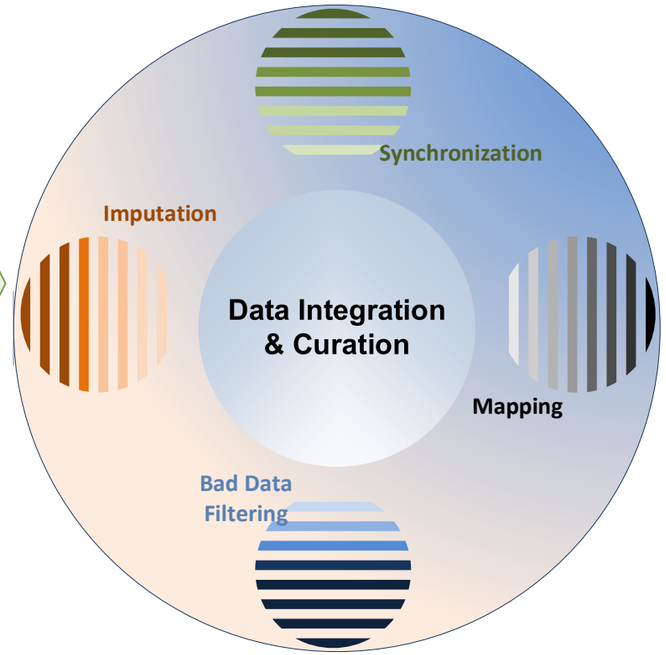
Pacific Northwest NATIONAL LABORATORY

OEDI SI Overview

Data from Multiple OT/IT Technologies



Data Compilation & Scenario Setup



Solar Integration Analytics

DMS

- SCADA & Monitoring
- Short-term Load Forecasting
- Load Allocation & Voltage Calculation
- Fault Location
- Service Restoration & Switching

- Short-term Load Forecasting
- Load Allocation & Voltage Calculation
- Fault Location
- Service Restoration & Switching
- Optimal Power Flow
- Unit Commitment

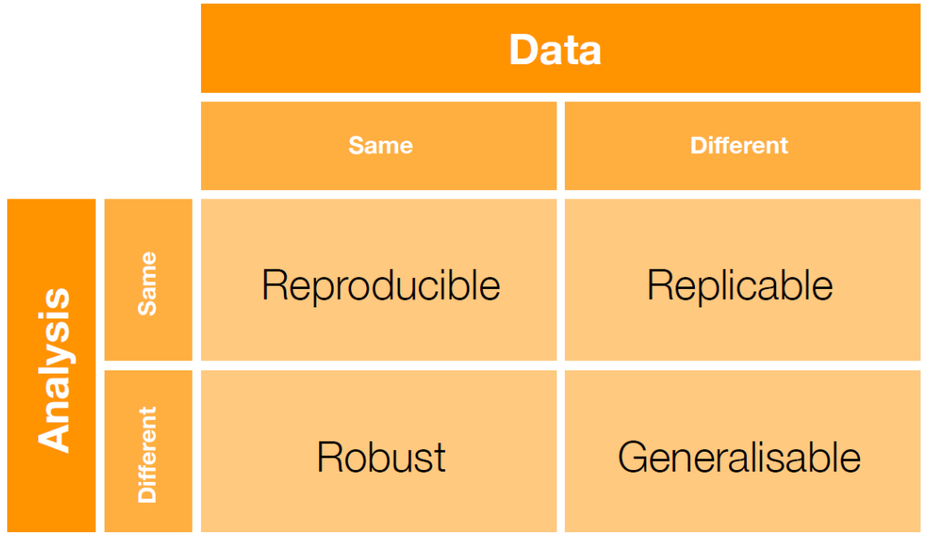
Planning

- Load Growth Analysis
- Renewables Impact
- Electric Vehicles
- Demand Response Allocation

SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

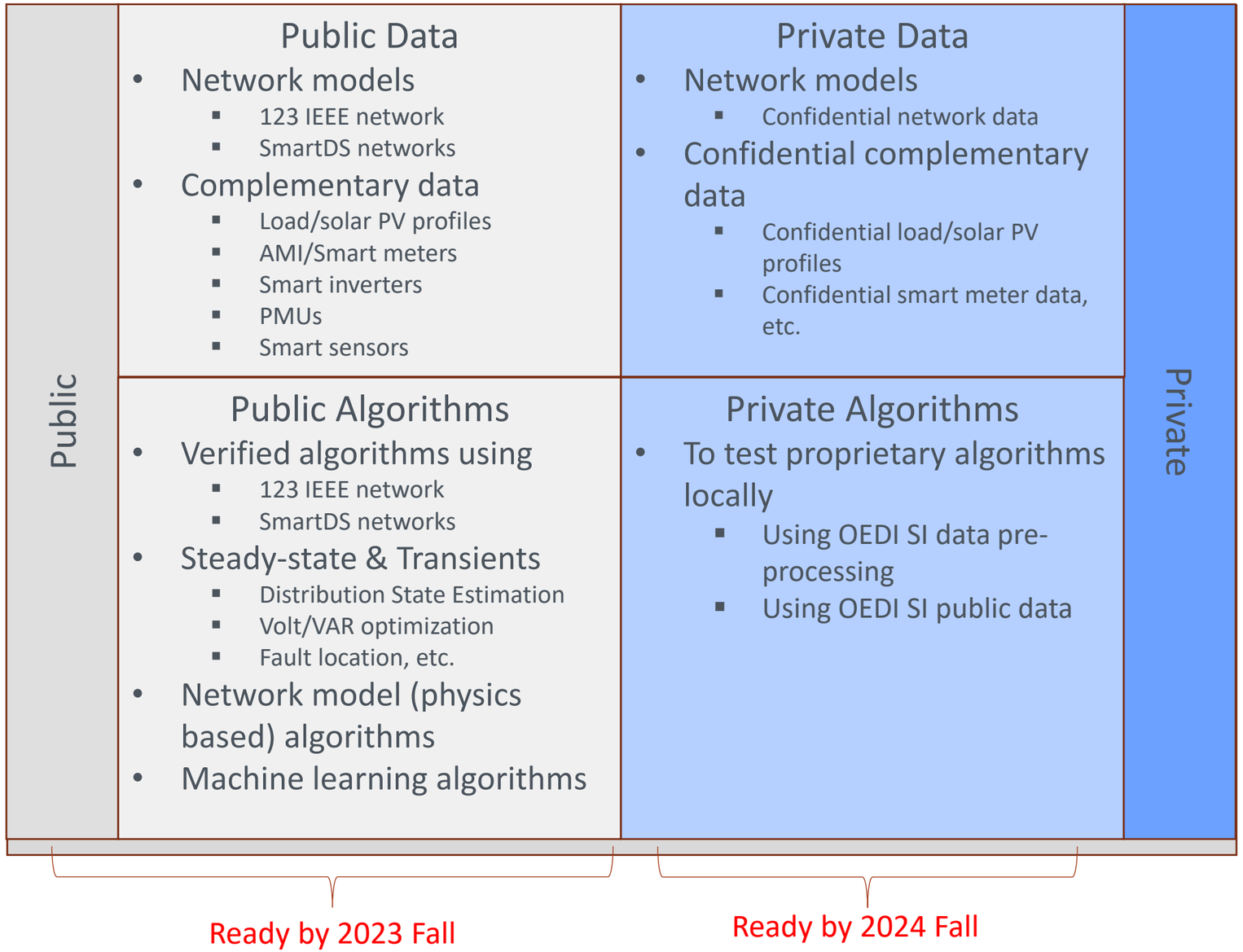
Provide Easy Access to Data and Algorithms for Solar Integration Simulations

- Adaptation of power systems analytics for distribution networks with high distributed solar generation participation
- Robust physics/network model-based algorithms
- New machine learning algorithms based on large data sets
- Steady-state and transients' analysis
- Data interfaces (CIM, OpenDSS, Gridlab-D)



Whitaker (2018) https://doi.org/10.6084/m9.figshare.7140050.v2

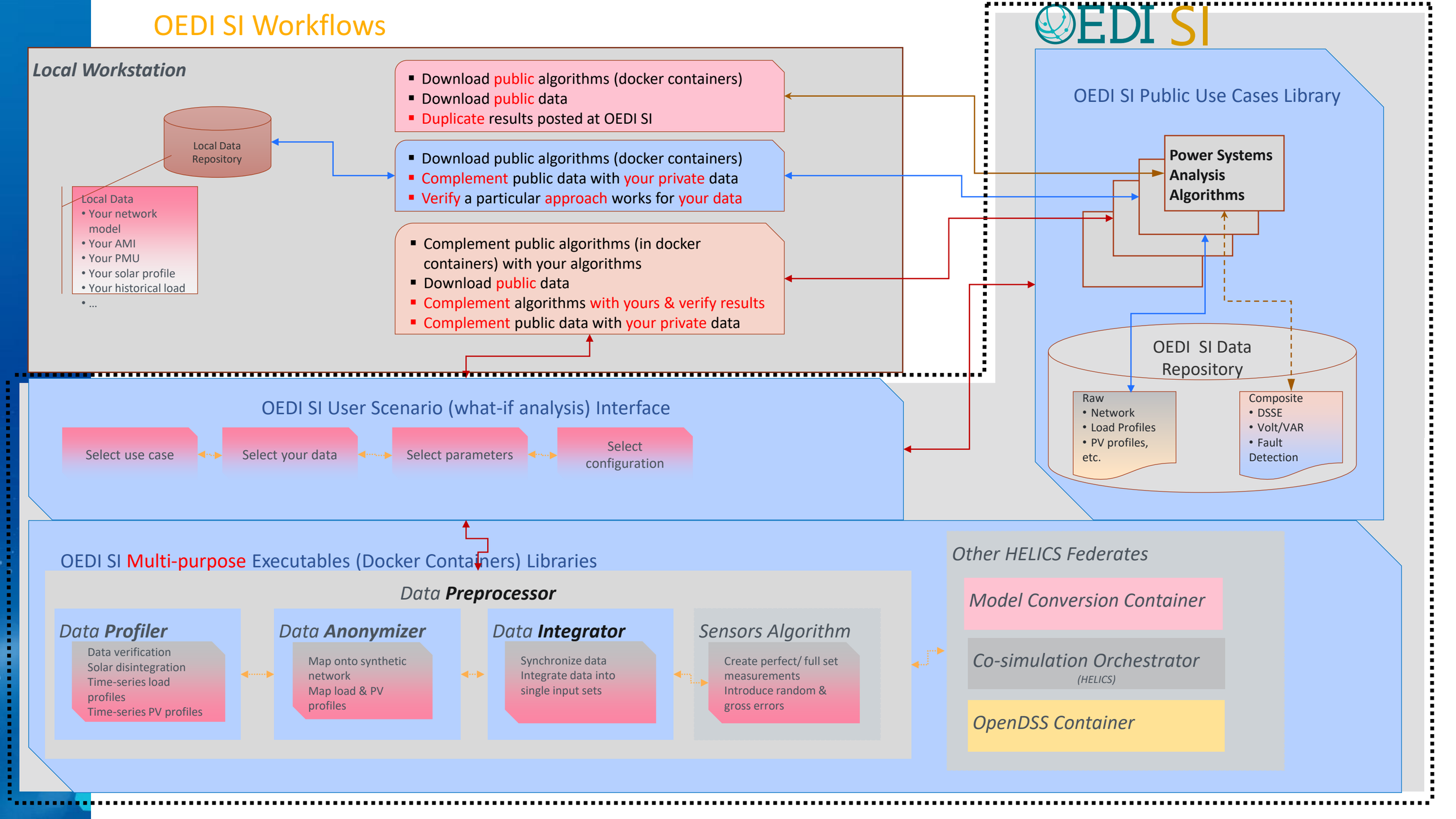
OEDI SI – OEDI FY22 Lab Call, Core Topic



- SETO Core Lab Call Program
  - National Labs collaboration
  - ANL, NREL, ORNL, PNNL
  - 2022 Oct. to 2024 Sept.
  - Approaching end of 2nd Budget P
  - User interface is rudimentary
  - But private data/private algorithm proof-of-concept implemented & tested
  - Actively promoting to public
  - Users group kick-off
  - In BP3, more data & UI & algorithms/data from other SETO/SI programs/projects

SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

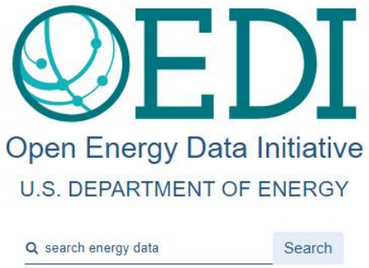
OEDI SI Workflows



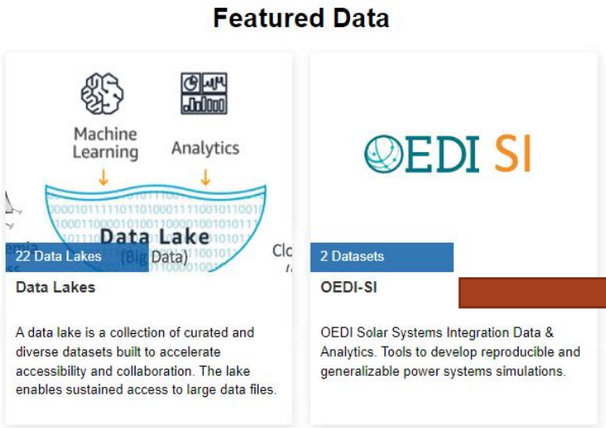
Workflows

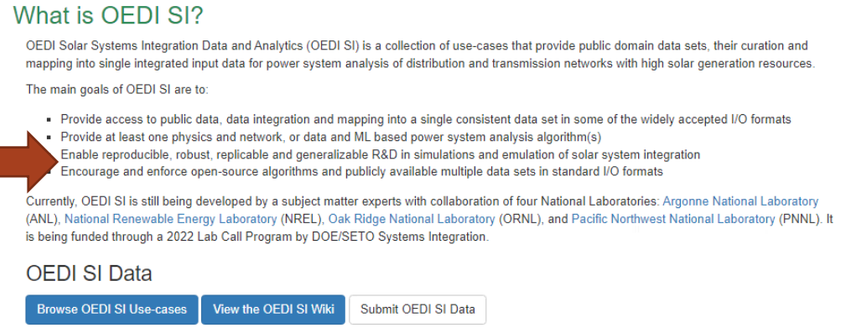
1. Use OEDI SI for replicating models and algorithms
   - Simplest use-case.
   - Use the public data & public algorithms to see if you can replicate the posted results. You do not need to run the data pre-processing. The composite (integrated and cured) data is ready to be used.
   - Download the posted data and docker containers, run the executable.
   - Compare your results to the posted results using (publicly available) different metrics. They should be very similar.
2. Use OEDI SI to check your algorithms that you do not want to share
   - More complicated use-case.
   - Download the public data. You do not need to run the data pre-processing. The composite (integrated & cured) data is ready to be used. All you need is to use a data standards OEDI SI supports in your input algorithm (API)
   - Download the docker containers and swap the modularized a public algorithm(s) with your own.
   - Compare your results to the posted results using different publicly available metrics. They should be quite similar.
   - Later on, when/if you are sure about your algorithm, you can share it on OEDI SI.
3. Use OEDI SI to check your algorithms and your data that you do not want to share
   - Most complicated use-case.
   - Download the public data if you want to use certain (raw data) components. Run the data pre-processing.
     - If you would like, you can run the anonymizer and publish your data on OEDI SI later on.
   - Download the docker containers and swap the modularized a public algorithm(s) with your own.
   - Compare your results to the posted results using (publicly available) different metrics. They should be similar.
   - You can post your algorithm as a robust & scalable approach.

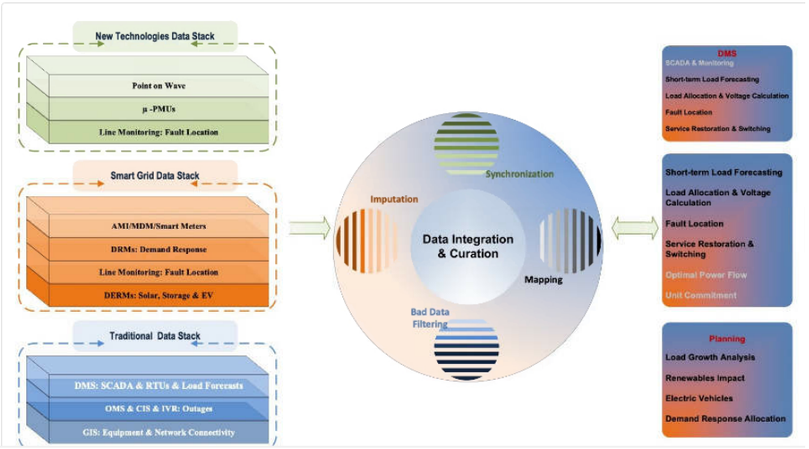
OEDI SI Web Portal



- Leverage existing OEDI architecture
- OEDI is from previous Lab Call
- Typically, single source data







https://data.openei.org/
https://openei.org/wiki/OEDI-SI/Overview

Thank you

Power Sector Transmission & Distribution Data and Information

WEBINAR SERIES

Concluding Remarks

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY

POWER SECTOR TRANSMISSION & DISTRIBUTION DATA AND INFORMATION WEBINAR SERIES

TOPIC 1: T&D Information Sharing
Wednesday, October 11 | 10:00 a.m. to 12:30 p.m. PDT

TOPIC 2: Cross-sector & Open Data Sharing and Risks
Wednesday, October 18 | 10:00 a.m. to 12:00 p.m. PDT

TOPIC 3: Sensor Systems and Platforms
Wednesday, October 25 | 10:00 a.m. to 12:00 p.m. PDT

TOPIC 4: Sensor Data and Device Research
Wednesday, November 1 | 10:00 a.m. to 12:00 p.m. PDT

REGISTER TODAY!

U.S. DEPARTMENT OF ENERGY
OFFICE OF ELECTRICITY

Los Alamos NATIONAL LABORATORY

OAK RIDGE National Laboratory

Pacific Northwest NATIONAL LABORATORY