902 Battelle Boulevard
P.O. Box 999, MSIN J5-05
Richland, WA 99352
(509) 375-2435
mark.rice@pnnl.gov
www.pnnl.gov

September 30, 2020

Pacific Northwest National Laboratory
Final Scientific/Technical Report
DR POWER
15 CJ000 09 04

DR POWER

Award:	DE-AR 15/CJ000/09/04
Sponsoring Agency	USDOE, Advanced Research Project Agency – Energy (ARPA-E)
Lead Recipient:	Pacific Northwest National Laboratory
Project Team Members	Mark Rice, Stephen Elbert, Olga Kuchar, David Pinney, Laurentiu Marinovici
Project Title:	Data Repository for Power system Open models With Evolving Resources (DR POWER)
Program Director:	Dr. Richard O'Neill
Principal Investigator:	Dr. Mark Rice
Contract Administrator:	Kendyl Foster
Date of Report:	09/30/2020
Reporting Period:	03/28/2016 – 09/30/2020

The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR 15/CJ000/09/04. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Please check the appropriate box:

✓ This Report contains no Protected Data.

☐ This Report contains Protected Data and the award allows data to be marked as protected. Refer to your Attachment 2 for guidance on how to appropriately mark Protected Data. The applicable notice is provided below:

PROTECTED RIGHTS NOTICE

These protected data were produced under agreement no._____with the U.S. Department of Energy and may not be published, disseminated, or disclosed to others outside the Government until 5 years after development of information under this agreement, unless express written authorization is obtained from the recipient. Upon expiration of the period of protection set forth in this Notice, the Government shall have unlimited rights in this data. This Notice shall be marked on any reproduction of this data, in whole or in part.

☐ This Report contains SBIR/STTR Data and the award allows data to be marked as SBIR data. Refer to your Attachment 2 for guidance on how to appropriately mark SBIR Data. The applicable notice is provided below:

SBIR/STTR RIGHTS NOTICE

These SBIR/STTR data are furnished with SBIR/STTR rights under [Award No. _____or a subaward under Award No._____]. For a period of [CHOOSE THE APPLICABLE QUOTED TEXT: for awards issued prior to May 2, 2019 "4 years or for awards issued on or after May 2, 2019 "20 years"], unless extended in accordance with FAR 27.409(h), after acceptance of all items to be delivered under this [Award or subaward], the Government will use these data for Government purposes only, and they shall not be disclosed outside the Government (including disclosure for procurement purposes) during such period without permission of the Contractor, except that, subject to the foregoing use and disclosure prohibitions, these data may be disclosed for use by support contractors. After the protection period, the Government has a paid-up license to use, and to authorize others to use on its behalf, these data for Government purposes, but is relieved of all disclosure prohibitions and assumes no liability for unauthorized use of these data by third parties. This notice shall be affixed to any reproductions of these data, in whole or in part.

Table of Contents

Table of Figures/Tables

Public Executive Summary

Pacific Northwest National Laboratory (PNNL) and the National Rural Electric Cooperative Association (NRECA) created a Data Repository for Power system Open models With Evolving Resources (DR POWER) to establish, curate, and evolve open-access power grid models and scenarios, and measurements and observations, collectively referred to as datasets. Existing open-access datasets are inadequate (too small and not representative of the complexity of the modern grid, and scattered all over the world wide web) for the investigation and benchmarking of the advanced optimization methods needed for next-generation grids. Open-access datasets are static and inadequate in other aspects: they do not provide format transformation tools; there is missing information needed for advanced applications; hard to cite; no model evolution history; difficult to update current collections; and little to no community involvement and feedback. Overall, the current landscape is hindering model evolution and benchmarking standards needed for transformational approaches to be developed, tested, and accepted by the power modeling community.

To address these inadequacies and revolutionize power modeling, DR POWER (https://egriddata.org) is an online, community-centered, modeling and scenario manipulation portal that supports large-scale realistic datasets and data tooling methods. The top right of the picture below illustrates the key components of the repository: the portal, the data management layer, internal services, and archive. DR POWER integrates three open-source products: the Open Modeling Framework (OMF); a digital object identifier framework via the Data ID Services provided by the Department of Energy Office of Science and Technical Information (OSTI); and DKAN, a web platform for publishing and sharing data. DR POWER enables the review, annotation, and verification of submitted datasets. This required: a data dictionary and ontology; benchmarking standards for comparing models; a curation process and supporting tools; new visualizations and model editors integrating transmission, distribution, and hybrid systems; and a research community that shares and develops models. There are several challenges: creating a citation model; integrating transmission and distribution models; data governance and metadata (provenance, set-up, assumptions, starting points, etc.); poorly defined curation techniques; and lack of community involvement.

Project Summary

Design, develop, and host a **data repository** and **web portal** to:

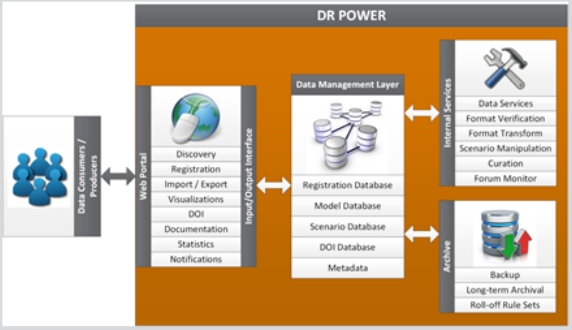
- Provide open-access power grid datasets and the capability to uniquely review, annotate, and verify submitted datasets
- Ensure sustainable model and dataset dissemination and evolution through user-defined dataset creation and validation

Proposed Project Impact

- Deliver the ability to collaboratively build, refine, review, and evolve high-fidelity power system models and accelerate grid optimization algorithm development

Proposed Targets

Metric	State of the Art	Proposed
Open Access	3 stale websites 10s of models Limited scenario sets	1 web portal 1,000s of models/data sets 1,000,000s of scenarios
Flexibility	Models are in a single format No way to add new model details	Perform transformations on the fly on the data Ability to add new fields as needed and evolve the models
Scalability	Total models under 1 Gigabytes	High-throughput scalable portal technology with petabyte storage
Sustainability	Stale websites with limited curations	PNNL will host the repository and build a community resource



Data repository capabilities:

- Download requested model and scenario data in any available format
- Upload and store datasets for different power grid models
- Import models of various formats, including currently available open models
- Save review of dataset and annotations performed by users
- Maintain dataset versioning after modification
- Save additional scenario information for time-series data generated

Web portal capabilities:

- Open-access web client for using data repository capabilities
- Execute tools for dataset generation, anonymization, expansion, etc.
- Save results from such tools to repository and display the results to user
- Display various reports such as dataset version details, curation details, etc.
- Track upload, download and access statistics

Acknowledgements

DR POWER website and the research behind it would not have been possible without the exceptional support of ARPA-E and the GRID DATA participants. Special thanks to NREL for providing our largest datasets. Finally, we would like to thank the energy research community whose interests and efforts in sharing data have made our repository a valuable resource.

Accomplishments and Objectives

This award allowed Pacific Northwest National Laboratory to demonstrate several key objectives. The focus of the project was to build an open energy data repository that provides the ability to collaboratively build, refine, and review a range of high-fidelity power system models and datasets, while providing data reference and dissemination. As shown below, this website represents a quantum leap over today's technology of scattered, small-scale static models that do not represent the challenging environments being encountered by evolving power grids worldwide. The availability of unrestricted, advanced, citable models is expected to accelerate the pace of grid optimization algorithm development, leading to improved efficiency and reliability with reduced environmental impact.

Metric	State of the Art	Proposed
Open Access	3 stale websites 10s of models Limited scenario sets	1 web portal 1,000s of models/data sets 1,000,000s of scenarios
Flexibility	Models are in a single format No way to add new model details	Perform transformations on the fly on the data Ability to add new fields as needed and evolve the models
Scalability	Total models under 1 Gigabytes	High-throughput scalable portal technology with petabyte storage
Sustainability	Stale websites with limited curations	PNNL will host the repository and build a community resource

A number of tasks and milestones were laid out in Attachment 3, the Technical Milestones and Deliverables, at the beginning of the project. The actual performance against the stated milestones is summarized here:

Table 1. Key Milestones and Deliverables.

Tasks	Milestones and Deliverables
Task 1: Technology to Market - Broad outreach capability to build visibility for the project and related technology **1.1 Strategic engagement for adoption & impact** Identify stakeholders (users, software vendors, startups, ISOs, utilities, universities, FFRDCs, regulators, international organizations) and their interests. Develop plan to support current and future interests.	**M1.1.1** - Stakeholder analysis - Updated T2M plan delivered to ARPA-E and approved by the Program Director (PD). The T2M plan will: a. Map out broad landscape of possible repository use cases enabled by the team's proposed model repository. b. Identify potential early model repository adopters (by project end; at least 2) and other influential stakeholders (e.g. regulators, policy makers, system operators, software vendors, etc.; at least 3). Describe potential benefits of new model repository for each identified early adopter. c. Outline process for user adoption from start to finish, focusing on critical requirements and adoption barriers. **Actual Performance:** (04/12/2017) Stakeholder analysis completed and submitted for review to ARPA-E; presented findings during the project's quarterly review. **M1.1.2** - Targeted engagement - Update delivered to ARPA-E on targeted engagement activities to drive adoption of technology with potential early adopters. The update should include a summary of the team's efforts to: a. Engage with relevant model contributors, partners, repository users, and other stakeholders. b. Detail strategy for timing of future stakeholder engagement during the remaining project period. c. Provide tailored documentation and user support for operation of the repository.

	Actual Performance: (05/15/2017) GRID DATA teams are uploading data to the repository and we are working with early adopters on a case by case basis. Specifics will be presented during the next project quarterly review. **M1.1.2** - Targeted engagement – Update 2 **Actual Performance:** (01/12/2018) Milestone complete as noted on ARPA-E Memorandum for the Record dated March 22, 2018. **M1.1.2** - Targeted engagement – Update 3 **Actual Performance:** (09/18/2018) Presented overview of DR POWER and OMF integration work and GRID DATA technical goals at FERC's annual software conference. Updated progress on these activities will be presented at the Q3 review that will be held on September 18, 2018. PNNL has been doing continuous engagement with DR POWER through-out the first 2 years and has developed a plan for the next 2 years. **M1.1.2** - Targeted engagement – Update 4 **Actual Performance:** (07/12/2019) The PNNL team has ramped up community building in the last 6 months and the data repository is now complete. The DR POWER team is working to focus specifically on the modelers who have data that is of interest to the community. Adoption has been slow; the PNNL team has presented at several conferences. The PNNL team has briefed ARPA- E on this milestone in February and June 2019. **M1.1.2** - Targeted engagement – Update 5 **Actual Performance:** Not completed since this was after POP of Award Competition Date of 7/12/2020. No Comment currently 07/01/2020. **M1.2.1** User validation - Updated T2M plan delivered to ARPA-E and approved by the PD. The T2M plan should summarize the results of conducting interviews with stakeholders and potential users outside of core project advisors to validate that: a. Intended repository features are aligned with user needs and unanticipated community needs have been considered. b. Repository design and model management methodologies meet community requirements (e.g. model formats, editing features, model validation, etc.) and proposed user interface is sufficiently intuitive. c. Stakeholder concerns such as cyber security and data privacy have been properly addressed. **Actual Performance:** (03/08/2017) The plan was delivered to ARPA-E on 01/30/17 for ARPA-E review/comments and finalized during our quarterly review meeting held in March.

1.2 User validation and feedback Get feedback from users, especially early users, on repository strengths, weaknesses, user friendly portions, least user friendly.	**M1.2.2** Preliminary results - Update delivered to ARPA-E on new model repository relative to community needs and hands-on feedback from early users: a. Obtain detailed hands-on feedback from early users. Validate that repository meets performance requirements, user interface is useful and intuitive, and repository design is reasonably future-proof. b. Quantify adoption of the repository (e.g. number of users, model uploads and downloads, forum activity, open-source contributors, industry inquiries, etc.) **Actual Performance:** (10/12/2017) This milestone was completed on October 12, 2017 and submitted via e-mail by Steve Elbert. **M1.2.2** Preliminary result – Update 2 **Actual Performance:** (1/29/2019) Since the launch of the DR POWER website in July 2017, we've seen almost 6,700 international users spanning across 123 countries. These users account for a total of more than 11,000 sessions and 4,600 data downloads. DR POWER currently has almost 129,800 curated datasets comprising more than 513,000 downloadable files for energy generation, transmission, distribution, and economics and markets.
1.3 Community building and organization Develop and implement a governance and sustainability plan.	**M1.3.1** Organizational planning - Update delivered to ARPA-E on the team's specific actions taken to develop organizational structure, processes, and resources for repository management. The team should summarize efforts to: a. Establish repository governance structure, including the drafting and publishing of bylaws, management hierarchy, and statement of purpose. b. Establish processes for model curation, including collection development, quality assurance, discovery and retrieval, archiving, and preservation for re-use over time. **Actual Performance:** (04/12/2017) Organization planning is complete and will be presented during the next project quarterly review. **M1.3.2** Community Building - Update delivered to ARPA-E on the team's specific actions taken to involve a broader cross-section of relevant stakeholders. The team should summarize efforts to: a. Solicit feedback from the community on the repository governance structures and organizational design. b. Solicit additional partners (ISOs, utilities, software vendors, etc.) to share additional models for inclusion in the repository. c. According to PD's discretion and guidance, conduct workshop(s) to engage collaborators from the technical community and educate potential users. **Actual Performance:** (06/28/2017) Steve Elbert attended the FERC Technical Conference.

	M1.3.2 Community Building - Update 2 **Actual Performance:** (1/12/2018) Update delivered to ARPA-E on the team's specific actions taken to involve a broader cross-section of relevant stakeholders. Milestone complete as noted on ARPA-E Memorandum for the Record dated March 22, 2018 **M1.3.2** Community Building - Update 3 **Actual Performance:** (09/18/2018) PNNL has ramped up community building in the last 2 quarters as we now have a complete data repository. We are working to focus specifically on the modelers who have data that is of interest to the community. This will be briefed out in the quarterly review. **M1.3.2** Community Building - Update 4 **Actual Performance:** (07/12/2019) The PNNL team has ramped up community building since the data repository is now complete. The DR POWER team is working to focus specifically on the modelers who have data that is of interest to the community. Adoption has been slow; the PNNL team has presented at several conferences (NCSU and ARPA-E Energy Innovation Summit). They will be attending the IEEE PES General Meeting and OSTI DOI conference. The PNNL team has briefed ARPA- E on this milestone in February and June 2019. **M1.3.2** Community Building - Update 5 **Actual Performance:** (Not completed after POP of Award, Reported 75% Complete; Completion Date of 7/12/2020) The NAERM project has created a community of electric grid modelers from 8+ national laboratories. The modelers are actively sharing models between laboratories and DR POWER is helping enable this in a safe and secure manner. **M1.3.3** Post ARPA-E transitions - Update delivered to ARPA-E on end-of-project transition activities. The update should detail the team's efforts to: a. Finalize user documentation (installation and maintenance, file formats, data processing tools, modeling applications, user interfaces, etc.). b. Secure funding for further repository development and/or maintenance. **Actual Performance:** (04/11/2019) We are completing the finalization of the documentation of the technology used in creation of the data repository. PNNL is still working on securing funding to further the repository after the end of this project.

	M1.3.3 Post ARPA-E transitions – Update 2 **Actual Performance:** (04/12/2019) PNNL has ramped up community building in the last 2 quarters, as we now have a complete data repository, but adoption has been slow. The PNNL DR POWER team has attended and presented at the following conferences: NCSU FREEDM April Symposium and ARPA-E TechConnect. The team will be attending the IEEE PES GM in Atlanta and the Data ID Service Workshop at OSTI. **M1.3.3** Post ARPA-E transitions – Update 3 **Actual Performance:** (10/12/2019) DR POWER is moving to a new server for maintenance and ease of upkeep after the project is completed. Documentation will go through a final update in CY2020, 6 months before project completion. **M1.3.3** Post ARPA-E transitions – Update 4 **Actual Performance:** (Not completed after POP of Award, Reported 90% Complete; Completion Date of 7/12/2020) The team will finalize the documentation after the last patch to https://egriddata.org that is being applied in this quarter. The sequential projects that will be funding this effort are NAERM ($335k for the next 6 months) and PNNL internal data hosting capabilities.

Task 2: Incorporating OMF Technologies - NRECA will enhance the OMF for use with DR POWER **2.1 OMF Extension Concept Validation** NRECA will enhance the OMF for use with DR POWER	**M2.1.1** GRID DATA Team Phone Calls - At least one phone call with every GRID DATA team conducted. The phone calls will seek feedback on the DR POWER team's proposed extensions to the OMF and verify there are no showstoppers in approach. The phone calls will also seek input regarding what users need to develop new sub-transmission classes and modules for OMF. A summary of feedback received will be presented to the ARPA-E Program Director. **Actual Performance:** (10/10/2016) All GRID DATA teams were interviewed to gather requirements for OMF technologies needed to support the program. The results were analyzed and used to inform specification process, and the analysis was delivered to ARPA-E 10/10/16. Full analysis reports continue to be available in the master document.
2.2 Transmission Modeling Support in the OMF Transmission model support will be added to the OMF by building out model data and formats. These transmission features will prepare the OMF code base for enhancement and deployment in DR POWER.	**M2.2.1** Transmission model data format - A transmission model data format is defined and ready for review. Format must be sufficient to determine if the data is consistent with the format, i.e., a schema definition is provided, and compatible with MATPOWER. **Actual Performance:** (10/10/2016) A transmission model data format (schema) was developed and documented in the "Data Schemas" section of the M2.4.1 Specifications document, and delivered to ARPA-E. **M2.2.2** Transmission model class - A new model class for transmission models in the OMF has been created and tested. **Actual Performance:** (01/12/2017) These features, which together offer the capability of importing MATPOWER transmission and generation models, visualizing them, editing them, and generating simulation results from them, were developed according to the M2.4.1 Specifications Document, committed to the master branch in the OMF source code repository, and deployed to the test environment on https://www.omf.coop/. The code is contained in the classes and HTML template interfaces listed in the "Module Architectures" section of the M2.4.1 Specifications Document and can be pulled from https://www.github.com/dpinney/omf/. **M2.2.3** Transmission user interface - OMF enhanced with an interface for users to perform in-situ validation of transmission models by running simple power-flow scenarios and visualizing the results. **Actual Performance:** (7/12/2017) Deliverables were submitted to ARPA-E.

	M2.2.4 Integration into DR POWER - OMF Integrated into DR POWER. (Code has been written that allows for successful integration into DR POWER; all interfaces in OMF v2.4 have been coded and tested). All requirements in M2.4.2 implemented. **Actual Performance:** (07/12/2018) OMF is integrated into DR POWER. Providing support to PNNL engineers to improve the user experience of the integration (interface tweaks and automation of conversion steps).
2.3 Network Visualization and Editor Enhancement The team will upgrade OMF's web-based distribution feeder editor to include transmission (this is focused on transmission models only; no enhancements for distributions work). Edits needed: conductor radius, load models, and anything that is modeled. The team will verify that changes to transmission models within the OMF visualization environment are retained after model export.	**M2.3.1** Initial transmission editor - OMF enhanced with ability to view and edit transmission networks; additional component classes for transmission assets; OMF demonstrated working with at least one 25,000 bus model. **Actual Performance:** (1/12/2017) These features, which together offer the capability of importing MATPOWER transmission and generation models, visualizing them, editing them, and generating simulation results from them, were developed according to the specifications in M2.4.1, committed to the master branch in the OMF source code repository, and deployed to the test environment on https://www.omf.coop/. The code is contained in the classes and HTML template interfaces listed in the "Module Architectures" section of the specifications M2.4.1 and can be pulled from https://www.github.com/dpinney/omf/. **M2.3.2** Final transmission editor - The scale and performance of the network editor improved. Network editor demonstrated working with larger transmission and distribution networks; OMF demonstrated working with at least one 100,000 bus model. **Actual Performance:** (04/12/2018) All features listed in the specification have been implemented and all release candidate bugs have been fixed. Extensive scalability testing was completed and met the SOPO goals. Reported in March - May 2018 Quarterly Report. **M2.3.3** Integration into DR POWER - Enhanced (transmission-capable) OMF network visualization and editor tools integrated into DR POWER. **Actual Performance:** (07/12/2018) Enhanced OMF integration was completed into DR POWER.
2.4 Architectural Support for DR POWER Subject matter expert support throughout performance period;	**M2.4.1** Initial specifications and initial launch model - Initial draft of software architecture plan for integrating OMF into DR POWER delivered to and approved by the ARPA-E Program Director. The plan will Include all class diagrams and data model schema details that are needed to enable a software engineer to implement the initial system in code by the end of Q2. (NOTE:

define architectural support for DR POWER and GRID DATA program.	There is no expectation that any code will exist at this point.) The document will also tag all classes as belonging in either DR POWER or OMF Codebase. **Actual Performance:** (10/10/2016) The initial architecture plan for extending and integrating OMF technologies into DRPOWER was developed. Feedback from PNNL was used to enhance the design. Full specifications for the OMF enhancements to support GRID DATA were developed in collaboration with PNNL and delivered to ARPA-E on 10/10/16. Feedback from ARPA-E was received and incorporated in the specification document. **M2.4.2** Complete specifications and refinements **-** Final software architecture plan for OMF and DR POWER integration delivered to and approved by the ARPA-E Program Director. **Actual Performance:** (07/12/2017) Deliverables were submitted to ARPA-E and approved.
Task 3: DR. POWER System - Build, test, deploy, and maintain repository functions	**M3.1.0** Data Management Plan - **Actual Performance:** (09/28/2016) The Intellectual Property and Data Management Plan was submitted on September 28, 2016 and uploaded to ePIC; plan was approved by ARPA-E on 12/7/2016.
3.1 Documentation The team will use best practices for software development and include design APIs.	**M3.1.1** Requirement Specification Document - Q1: Initial requirement specification document for the overall system architecture and interfaces delivered to and approved by the ARPA-E Program Director. This document is a strawman to be reviewed by others whose buy in is critical to the success of GRID DATA. **Actual Performance:** (10/12/2016) The document was submitted to ARPA-E on 10/12/16 and approved by the Program Director on 12/7/16. **M3.1.1** Requirement specification document - Q4: Team will deliver to the ARPA-E Program Director an update to the requirement specification document based on feedback from other GRID DATA teams and other stakeholders. The document will include a list of mandatory requirements and verification checks to be performed in Q8. These requirements will form the basis of the Q8 Go/No-Go decision point. ARPA-E Program Director's approval of the document is required to meet this milestone. **Actual Performance:** (7/12/2017) The updated Requirement Specification Document was submitted to ARPA-E.

	M3.1.2 Design and implementation document - The DR POWER design and implementation document delivered to and approved by the ARPA-E Program Director. This document must include details on system interfaces, APIs for additional tools contributed by community, and definitions for the roles of those responsible for project development. This is a strawman document for further development with the community. **Actual Performance:** (10/12/2016) The document was submitted to ARPA-E on 10/12/16 and approved by the Program Director on 12/7/16. **M3.1.3** Testing document - Deliver documentation for test procedures. This document will describe new features being added to the next production version and will specify validation to ensure the features are working properly. When possible publish the results from the features. **Actual Performance:** (01/12/2017) The document was delivered to ARPA-E on 01/12/17. The document holds the current test plan for the capabilities in version 1.0 of the DR POWER web portal. **M3.1.3** Testing document 2 **Actual Performance:** (07/12/2017) The updated document was submitted to ARPA-E. **M3.1.3** Testing document 3 **Actual Performance:** (01/12/2018) The updated document was submitted to ARPA-E. **M3.1.3** Testing document 4 **Actual Performance:** (7/12/2018) Testing Document updated to provide a test plan for the initial integration of OMF into DR POWER. This test procedure checks to see if a user can upload an OMT extension file and open the file in the dockerized version of OMF running on the DR POWER server.
3.2 Web Portal Deliver the single-point access to the data repository	**M3.2.1** Web Portal Version 1 - Initial version of the web portal and repository deployed with at least the following capabilities: registering a user, creating an account, uploading/downloading models, basic search, DOI, and populated with UW models. **Actual Performance:** (01/12/2017) Registering a user; creating an account; uploading/downloading models; basic search (tag searching is implemented); and populated with initial UW models. Submission of a DOI is being implemented. The PHP code for DOI submission request to OSTI has been united tested. We integrated the web portal with the PHP code and was

	demonstrated at the GRID DATA meeting held in January. In addition to this work, a subset of the Curation Working Group has delivered an initial tagging dictionary for the static UW test cases. This tagging dictionary is implemented in this version of the web portal. **M3.2.2** Web Portal Version 2 - Version 2 of web portal deployed with expanded user interface for file transformations during upload/download and the addition of forums, portal statistics, and user manual/documentation. **Actual Performance:** (07/27/2017) Version 2 is complete and in production at https://egriddata.org. **M3.2.3** Web Portal Version 3 - Version 3 of the web portal deployed with expanded format verification and conversion capabilities, initial curation, and forum monitoring. (The list of file formats that are supported by the repository will also be updated.) **Actual Performance:** (01/12/2018) Version 3 of the web portal deployed with expanded format verification and conversion capabilities, initial curation, and forum monitoring. The web portal accepts all file formats for upload; it is format agnostic. The required features outlined in milestone M3.2.3 were released as part of Web Portal Version 2. Web Portal Version 2 was published on January 28, 2018 and demonstrated to ARPA-E on 02/20/2018. **M3.2.4** Web Portal Version 4 - Q8 (Go/No Go): Version 4 of the web portal deployed with support for additional models and tools, bug fixes (service updates). The repository must now meet or exceed all feature and performance requirements specified in M3.1.1 at end of Q4. **Actual Performance:** (07/12/2018) The current production site includes all features noted in the WAS for this milestone; see https://griddata.org **M3.2.4** Web Portal Version 4 Update - Q12: Service update 1 (bug fixes, community requests) (v. 4.1) **Actual Performance:** (07/12/2019) The PNNL DR POWER team is rolling software updates as needed. No issues have been identified at this time. **M3.2.4** Web Portal Version 4 Update - Q16: Service update 2 (bug fixes, community requests) (v. 4.2) **Actual Performance:** (05/31/2020) The last feature to the DR POWER web portal was the integration of Los Alamos National Lab (LANL) Python-based Grid Research for Good (GRG) converter. The converter resides in a Docker

3.3 Data Management Data storage; data architecture documentation, including database schemata, business rules; and development, test, and production sites.	image and uses REST API for communication between it and the website. Additionally, our docker containers got "upgraded" for automatic restarts due to environmental issues (i.e. power outages) and do self-checks for liveness (i.e. confirm converter is up and ready to receive a file for conversion). **M3.3.1** Initial data architecture -Initial data architecture document for flat file storage, user account information, and basic model metadata, including vision and use cases completed and delivered to ARPA-E. **Actual Performance:** (12/07/2016) The document was submitted to ARPA-E on 10/12/16 and approved by the Program Director on 12/7/16. **M3.3.2** Flat file storage and database schema - Create flat file storage and database schema to save models, user accounts, model versions, and model statuses. This will include a document showing the requirements met and identify any requirements not met. **Actual Performance:** (1/12/2017) The current version of the web portal already implements a flat file storage and database schema to save models, user accounts, model versions, and model statuses (along with a lot of metadata). The document is a summary of these features and includes a table showing the requirements met and identify any requirements not met; additionally, this information is now included in the M3.1.2 Design Document. **M3.3.3** Unstructured data storage - System for forum data storage and other unstructured data implemented. **Actual Performance:** (7/12/2017) This document has been rolled into M3.1.2 Design Document at APRA-E's request, which was submitted to ARPA-E on July 12, 2017. **M3.3.4** Define common model schema - Updated data architecture document delivered to ARPA-E. The updated file must include a common model schema definition that will allow for easy transformations between different model formats. The document will also specify which features being considered by other GRID DATA repository team will be supported or not supported in Q4 and Q8. **Actual Performance:** (Cancelled during Q3 FY2017 07/12/2017) Based on discussions with ARPA-E, this milestone no longer applies and has been cancelled from our project plan.

	M3.3.5 Deploy common data storage for current models. **Actual Performance:** (06/15/2017) Our model and data storage is in two locations: NREL's data is on our institutional computing file storage and all other data is located on our web server. **M3.3.6** Final data architecture - Updated data architecture document for new model formats delivered to ARPA-E. The team will compare supported models to previously documented requirements from M3.3.4 (Q4). The team must formulate a plan that supports storage of all models being produced by other GRID DATA teams. (In some cases, this can require the other teams to develop model format conversion tools.) **Actual Performance:** (01/12/2018) Updated the data architecture document for new model formats delivered to ARPA-E on January 16, 2018. **M3.3.7** Update common data storage - Complete expanded/updated common data storage deployed. The data storage must meet or exceed all requirements for the overall DR POWER system in M3.1.1. All GRID DATA model formats must now be supported (either directly or with converters supplied by other teams.) **Actual Performance:** (01/30/2020) Decision to leave the data models in two locations due to system performance and maintenance. The large NREL solar and wind data will continue to be stored on our specialized data array for large data sets and our smaller system models are stored locally on our webserver.

3.4 Format Conversion Tool Format conversion tool developed (using GOSS) and tested. The tool will not only do conversion but state what fidelity is lost or assumptions used in each type of conversion.	**M3.4.1** Format conversion tool requirements and testing plan - Updated documentation for tools in GOSS and a comprehensive testing plan delivered to and approved by ARPA-E PD. **Actual Performance:** (03/28/2017) This milestone was revised to focus on developing automation capacities for the other aspects of the curation process, i.e., format verification and model curation. A discussion with Pat McGrath was scheduled in April to discuss the changes and issues related to Task 3.4. In conclusion, GOSS was removed from our deliverables and replaced with a REST API with Docker containers to integrate any converter provided by a user into DR POWER. **M3.4.2** PTI and IEC TC57 CIM model conversion - PTI and IEC TC57 CIM model conversion capabilities implemented and tested. Successful demonstration of bidirectional conversion of models in both formats. **Actual Performance:** (06/12/2017) Conversion between PTI 23 and MATPOWER 6.0: Created a Conversion Request workflow so users can request file format conversions. This was rolled out in Web Portal Version 1.5; this milestone was reopened on September 28, 2017. ARPA-E expanded the scope of the milestone to include an example of automated conversion between CDF and PTI (raw only) to MATPOWER. This feature was developed in Web Portal Version 2.0 and demonstrated to ARPA-E on November 21, 2017. This was released to the website in January 2018. **This milestone shows complete because it was complete when PNNL implemented the request feature; but then this milestone was expanded, as noted above. **M3.4.3** IEEE Open Data Model, IEEE Common Data Format, PSS/E, and UCTE model conversion - IEEE Open Data Model, IEEE Common Data Format, PSS/E, and UCTE model conversion capabilities implemented and tested. Successful demonstration of bidirectional conversion of models in all four formats. **Actual Performance:** (07/12/2017) Conversion is no longer a milestone for DR POWER based on discussions with ARPA-E. ARPA-E would like to see at least one conversion tool in the next release (Web Portal 3.0) due in Q6. This feature was demonstrated to ARPA-E on 11/21/2017. **M3.4.4** Additional formats identification - The team will document other model conversion tools that are needed to support the community. A plan for supporting these formats in the second version of the format conversion tool must be described. (In some cases, format conversion tools may be required to be provided by other GRID DATA teams.)

	Actual Performance: (11/15/2017) PNNL is working with NREL and has received some of their data and are working with them on cataloging. PNNL staff are targeting the tool from the University of Michigan as the first external conversion tool to be hosted by DR POWER. Since this milestone depends on GRID DATA teams and portal user feedback, our plan is to continually evaluate the needs of our community. **M3.4.5** Second release of format conversion tool. - Second release of conversion tool. Verification that all required model formats are now supported in the format conversion tool. Any loss of fidelity noted in documentation. **Actual Performance:** (04/12/2018) PNNL has not released a second version of the conversion tool since we currently are not experiencing any problems with it; and only 1 user has requested a conversion. Not all formats are supported in the current tool; only the agreed upon RAW and CDF. Additionally, as part of the format conversion tool integration to help with data preservation, the LANL Python-based Grid Research for Good (GRG) JSON format converter was integrated within the portal. The converter was Dockerized in a Python 3 container and wrapped within a REST API to allow for transfer of file to be converted inside the container and streaming out the converted result through HTTP requests. Upon requests, curators can perform conversion of certain version of PTI/PSSE RAW models to the new GRG type.
3.5 Format Verification Tool Model format verification tools developed using GOSS.	**M3.5.1** Format verification tool requirements - Requirements document for format verification V1 delivered to and approved by ARPA-E PD. **Actual Performance:** (3/28/2017) The requirements document was submitted and approved on March 28, 2017. **M3.5.2** Format verification tool design - Format verification tool design document delivered to and approved by ARPA-E PD. **Actual Performance:** (02/27/2017) This was incorporated in M3.4.1 and submitted to ARPA-E. **M3.5.3** Initial release of format verification tool - Initial verification tool meeting requirements specified in M3.5.1 released. **Actual Performance:** (Cancelled 03/12/2017) Based upon discussion with ARPA-E, this milestone no longer applies. **M3.5.4** Release of verification tools API - Verification tools API document released.

	Actual Performance: (Cancelled 03/12/2017) Based upon discussion with ARPA-E, this milestone no longer applies. **M3.5.5** Second release of format verification tool. Second release of verification tool. **Actual Performance:** (Cancelled 03/12/2017) Based upon discussion with ARPA-E, this milestone no longer applies.
3.6 Model Curation Tool Development Development of curation tools for annotating, reviewing, and editing the models and scenario datasets.	**M3.6.1** Data dictionary and ontology definitions - Implementation of data dictionary and initial ontology definition completed. **Actual Performance:** (07/12/2018) An initial data dictionary was documented for power grid models in collaboration with Dr. Weaver of Trusted Cyber Infrastructure for the Power Grid (TCIPG) and Dr. Blackwell, The Louis G. Forgione University Professor of Classics at Furman University. The initial dictionary was reviewed by the Curation Working Group in February 2017. A subset of the dictionary is implemented as a taxonomy in Web Portal Version 1. **M3.6.2** Initial annotation and review tool development - Implementation and initial testing of annotation and review tool completed. **Actual Performance:** (07/11/2017) An initial tool is complete and incorporated in Version 2.0. **M3.6.3** Initial release of annotation and review tool - Initial release of annotation and review tool (align with web portal); The team will confirm that this tool meets or exceeds all requirements specified by the team's strategic plan for curation (as specified in M4.2.2). **Actual Performance:** (1/31/2018) The annotation and review tool was demonstrated to ARPA-E in June 2017. The tool meets all requirements specified by the team's strategic plan for curation (as specified in M4.2.2). This tool was released on the public website on January 31, 2018. **M3.6.4** Curation editing tool - Release model editing tool defined in M1.2.1. **Actual Performance:** (07/12/2018) OMF integration allows curators the ability to edit some of the models (OMT). We are continually improving the curation workflow based on feedback from our curators.

3.7 Forum Monitor User forums can be created around a topic, model, or interest. This task will provide tools to aid in monitoring forum.	**M3.7.1** Forum monitoring requirements - Forum monitoring requirements document delivered to and approved by the ARPA-E PD. **Actual Performance:** (01/12/2017) The document was submitted on January 12 and is integrated into our M3.1.1 Requirements Document as of 4/15/2017. **M3.7.2** Forum monitoring design **Actual Performance:** (01/12/2017) Forum monitoring design document delivered to and approved by the ARPA-E PD and is found in M3.1.2 Design Document. **M3.7.3** Forum monitoring tool **Actual Performance:** (03/08/2017) Forum monitoring tool released to the website.
3.8 Archive Develop and document policies and rules for system backup	**M3.8.1** Long-term data storage policy and rules - Documentation for long-term data storage policy and roll-off rules will be delivered to and approved by the ARPA-E PD. **Actual Performance:** (07/12/2018) Long-term data storage policy and rules are in the M.3.1.2 Design document. This document was sent to DOE for approval on July 12, 2018.
4.1 Governance Design Establish term lengths for members and recruitments policy.	**M4.1.1** Steering Committee Established. First meeting (Vision and Use Cases). DR POWER Steering Committee established with at least the following confirmed members: 3 academic, 2 industry, 1 Non-PNNL FFRDC, 1 Government. First meeting of the committee conducted, with a focus on overall system vision and use case definition. **Actual Performance:** (12/12/2016) After having submitted the Steering Committee Charter on October 2016 and having it approved in December 2016, its members were briefed with the scope and the long-term goals of the project, while presented with the current stage of the development. Members of the committee have accepted the roles as strategic advisors, insight and recommendation providers, such that the project continues to make progress and grows its impact and visibility. **M4.1.2** Steering committee feedback on initial repository. **Actual Performance:** (02/23/2017) The web portal https://egriddata.org version 1 was presented to the steering committee whose feedback on the initial repository was positive and constructive. Suggestions and possibly future-added features were made, such as hosting time-series data (e.g.,

	aggregate and non-aggregate load), with certain degree of categorization (e.g., low, medium, or high). **M4.1.3** Third steering committee meeting - Detailed plan on how to continue to engage steering committee, documented and approved by steering committee and ARPA-E PD. **Actual Performance:** (06/12/2017) During the third steering committee meeting on June 12, 2017, the web portal version 1.5 was presented. Aside from its functionality, the portal was introduced from the curator perspective to point out the importance and necessity of a continuous curation process involving both the members of the curation working group as well as the users. The committee was satisfied with the current stage of the repository front-end development, content, testing examples, curation strategy and workflow. Suggestions were made and questions raised in order to steer future actions, such as: capture more account information upon registration, e.g., name, contact information, and affiliation and provide links to published and/or cited papers.
4.2 Curation Guidance Establish policy and procedures for proper curation of the repository.	**M4.2.1** Curation Working Group - Collaborative curation working group established. The working group is expected to consist of the five PIs from the GRID DATA model teams or their representatives plus at least one academic data curation researcher/library scientist. **Actual Performance:** (03/01/2017) After having established the curation working group, its members were instructed on their role in maintaining, preserving, and adding value to the repository by inspecting the uploaded data and curate it with feedback from the uploaders. Constructive discussions developed around automated versus human-in-the-loop curation, and possible tools to help curators in their process. A first draft of the document containing specifications for the curation model was also reviewed. It concluded upon some of the main topics the document should include, that is: metadata; file formatting; file versioning; automation tools; and DOI workflow. **M4.2.2** Curation strategic plan - Strategic plan for curation defined and documented by the curation working group. This document will follow requirements specified by the curation professional. The document will specify minimum technical features required in each repository system component to support curation. **Actual Performance:** (06/20/2017) The 1st version of the Curation Strategic Plan was finalized and submitted to the team members for review on June 20,

	2017. It highlighted the type of data users are expected to find in the DR POWER repository, that is:
	- electrical power system models and simulations:
	  o transmission and generation models,
	  o distribution models,
	  o MATPOWER, Power System Toolbox, GE's PSLF, etc. models;
	- observations and measurements:
	  o PMU, wind, solar, weather data, etc.;
	- accompanying material:
	  o graphics and tables,
	  o instrumentation,
	  o data dictionaries,
	  o references and bibliography, etc.
	The document also provided guidelines for the curators to ensure the curation process is unified and unbiased. The curation model is based on the Digital Curation Centre (DCC) curation lifecycle model, and gives details on how to perform:
	- creating/receiving data, that is uploading, revisioning, modification, tagging, etc.,
	- appraising through
	  o validation of what is allowable and unallowable content,
	  o verification by doing an in-depth check for quality assurance, metadata, tags, readability, and performance,
	  o certification, and/or
	  o transformation and type conversion by request.
	
	The Curation Strategic Plan document also describes the types of metadata DR POWER requires, such as descriptive, structural, contextual, technical, and administrative. These types are detailed to highlight their use and responsibility.
	
	M4.2.3 Guidelines for model organization - Guidelines for model organization - - tags and taxonomy of data searching created and documented.
	Actual Performance: (01/12/2018) The Curation Strategic Plan document was updated to include specific guidelines for model organization to make data identification easier and more efficient. Thus, the following categories have been proposed and adopted:
	- generation (supply),
	- transmission (transfer),
	- distribution (demand/use/consume),
	- economics and markets,

	- weather. Each model in these categories was then scoped for different applications, each performed for different study types, and requiring certain software environments. The document presents curators with the accepted taxonomy, that is the IEEE taxonomy at https://www.ieee.org/content/dam/ieee-org/ieee/web/org/pubs/taxonomy_v101.pdf, and also guidelines on how to discern and accept specific folksonomy. Curation work also focused on developing and integrating a conversion model that would help during the preservation stage. To begin with, converters from transmission and generation models in Common Data Format (CDF) and/or PSSE (raw format) to MATLAB format used by Power System toolbox were studied, documented and integrated withing the portal to be used by curators at the user's request. **M4.2.4** Implement model organization - Model organization guidelines reviewed and updated (as necessary). Significant changes are documented and delivered to ARPA-E PD. **Actual Performance:** (06/29/2018) DR POWER model organization has been reviewed without major changes, then implemented on egriddata.org. Since no major changes have been done, no updates to the written document have been made. For the last quarter, the curation team has focused on actual data upload, review and curation, and DOI process improvement. Curators have also been involved in community outreach, in order to recruit new users to upload data to DR POWER. **M4.2.4** Implement model organization - **Actual Performance:** (07/12/2019) As DR POWER model organization proved to be efficient as initially implemented, no changes have been made to it; rather its features have been tested through constant upload, curation, and download activities. Curators have also been involved with testing new features and updates of the portal given technology evolvement and providing feedback to the developers. **M4.2.4** Implement model organization. **Actual Performance:** (04/14/2020) As the new Grid Research for Good (GRG) model for optimal power flow problems with unprecedented fidelity has emerged, allowing for conversion to and from this model was a major upgrade for the portal. Implementing and integrating the LANL Advanced Network

	Science Initiative (ANSI) converter followed the guidelines documented by the Curation Strategic Plan.

Project Activities

Pacific Northwest National Laboratory (PNNL) and the National Rural Electric Cooperative Association (NRECA) created a Data Repository for Power system Open models With Evolving Resources (DR POWER) to establish, curate, and evolve open-access power grid models and scenarios, collectively referred to as datasets; this repository is located at https://egriddata.org.

Some of the key features of the website are:

- Visualizations and model validation via the Open Modeling Framework.
- Digital Object Identifiers (DOIs) for dataset reference.
- Data and metadata curation.
- File format converters.
- API for integrating third-party software with the repository, as shown via the NREL solar and wind application tool.

Since going live on July 1, 2017, the repository contains over 272,500 curated datasets, comprising over 1,000,000 curated files. Some website statistics for DR POWER:

- 61 registered users.
- 20,524 unique visitors that created 30,134 sessions with a total of 108,288 pageviews.
- 12,353 downloaded files.

It is interesting to note that the small models described in MATLAB files are still the most popular with our energy community, as depicted in Table 1; however, users are finding the metadata information on these pages to be of interest, as noted in their page duration, found in Table 2.

Table 2. Top 5 downloaded files for DR POWER between July 2017 – September 2020.

Rank	Downloaded Dataset	Downloads
1	33-bus radial distribution system https://egriddata.org/sites/default/files/case33bw.m	2,657
2	IEEE RTS 24-bus case https://egriddata.org/sites/default/files/case24_ieee_rts.m	767
3	IEEE 30-bus system https://egriddata.org/sites/default/files/case_ieee30_0.m	766
4	IEEE 14-Bus Power Flow Test Case https://egriddata.org/sites/default/files/ieee14cdf.txt	762
5	6-bus, 3-generator transmission network model https://egriddata.org/sites/default/files/case6ww.m	534

Table 3. Top 5 visited pages for DR POWER between July 2017 – September 2020.

Rank	Dataset	Pageviews	Avg. Time on Page
1	33-bus radial distribution system https://egriddata.org/dataset/33-bus-radial-distribution-system	5,227	00:04:13
2	IEEE RTS 24-bus case https://egriddata.org/dataset/ieee-rts-24-bus-case	2,110	00:03:35
3	IEEE 14 Bus Power Flow Test Case https://egriddata.org/dataset/ieee-14-bus-power-flow-test-case	1,787	00:02:22

Rank	Dataset	Pageviews	Avg. Time on Page
4	IEEE 30-bus system https://egriddata.org/dataset/ieee-30-bus-system	1,695	00:03:27
5	IEEE 14-bus case https://egriddata.org/dataset/ieee-14-bus-case	1,330	00:02:45

Project Outputs

A. Journal Articles

B. Papers/Presentations

Rice M.J., O.A. Kuchar, S.T. Elbert, L.D. Marinovici, and D. Pinney. (2019) **How New Open Data Efforts are Elevating Grid Modeling.** Presented by M.J. Rice, D. Pinney at TechAdvantage, Orlando, Florida.

Kuchar O.A. (2018) **Grid Optimization (GO) Competition: Solution Submission/Using the Platform.** Presented by Olga A Kuchar at FERC Technical Conference, Washington, District of Columbia.

Pinney, D., Rice, M., Elbert, S., Kuchar, O., and Marinovici, L. (2018) **Increasing Real-Time and Day-Ahead Market Efficiency through Improved Software.** Presented by David Pinney at FERC Technical Conference, Washington, District of Columbia.

Kuchar O.A., S.T. Elbert, M.J. Rice, and M.C. Marinovici. (2017) **DR POWER: System Design Specification.** Pacific Northwest National Laboratory, Richland, WA.

Rice M.J., O.A. Kuchar, S.T. Elbert, and L.D. Marinovici. (2017) **NASPI Talk about Data Repository for Power system Open models With Evolving Resources (DR POWER).** Presented by Mark J Rice at NASPI March 2017 WG meeting, Gaithersburg, Maryland.

Elbert S.T., M.J. Rice, O.A. Kuchar, and L.D. Marinovici. (2017) **DR POWER: A DATA REPOSITORY FOR POWER SYSTEM OPEN MODELS WITH EVOLVING RESOURCES.** Presented by Stephen Elbert at FERC

Kuchar O.A., S.T. Elbert, and M.J. Rice. (2017) **DR POWER** ARPA-E Meeting. Presented by Olga A Kuchar, Stephen T Elbert at GRID DATA Annual Technical Review Meeting, San Diego, California.

Marinovici, L., Kuchar O.A., S.T. Elbert, and M.J. Rice. (2017) **DR POWER – Curation Strategic Plan.** PNNL-26158. Richland, WA: Pacific Northwest National Laboratory.

Kuchar O.A., S.T. Elbert, M.J. Rice, and M.C. Marinovici. (2017) **DR POWER: System Requirements Specification.** Pacific Northwest National Laboratory, Richland, WA.

Kuchar O.A., S.T. Elbert, and M.J. Rice. (2017) **DR POWER - Forum Design.** PNNL-26158. Richland, WA: Pacific Northwest National Laboratory.

Kuchar O.A., S.T. Elbert, and M.J. Rice. (2017) **DR POWER - Forum Requirements.** PNNL-26159. Richland, WA: Pacific Northwest National Laboratory.

Rice M.J., Z. Huang, S.T. Elbert, O.A. Kuchar, and M.C. Marinovici. (2016) **Open Ecosystem for Creation and Distribution of Realistic Data Sets.** Presented by Mark J Rice at 2016 IEEE PES Innovative Smart Grid Technologies, Minneapolis, Minnesota.

C. Status Reports

Quarterly presentations provided to ARPA-E.

D. Media Reports

E. Invention Disclosures

F. Patent Applications/Issued Patents

G. Licensed Technologies

H. Networks/Collaborations Fostered

Steering Committee:

Type of Organization	Member	Institution
Academic	Daniel Kirschen	U. Washington
Academic	Warren Powell	Princeton University
Academic	Andy Sun	Georgia Tech
Government	Richard O'Neill	FERC
Industry	Yonghong Chen	MISO
Industry	Mani Vadari	Modern Grid Solution
Non-PNNL FFRDC	Jean-Paul Watson	Sandia National Laboratory

Curation Working Group:

Institution	Member	Expertise
NREL	Bryan Palmintier	GRID DATA Project
U. Michigan	Pascal Van Hentenryck	GRID DATA Project
UIUC	Gabriel Weaver	Model evolution
Furman University	Christopher Blackwell	Multi-versioned data sets with attributes
PNNL	Ruisheng Diao	GRID DATA Project
PNNL	Justin Day	Research Librarian

Other:

Carleton Coffrin from Los Alamos National Laboratory (LANL) Advanced Network Science Initiative (ANSI) to ensure Grid Research for Good (GRG) format converter is fully functional and correctly integrated as a tool in DR POWER

Members of the Data Service ID from Office of Scientific and Technical Information (OSTI) for Digital Object Identifiers to aid in technical integration and message passing between DR POWER and OSTI for DOI registration with DataCite.

Michael Rossol from National Renewable Energy Laboratory (NREL) for wind and solar data.

I. Websites Featuring Project Work Results

https://egriddata.org

J. Other Products (e.g. Databases, Physical Collections, Audio/Video, Software, Models, Educational Aids or Curricula, Equipment or Instruments)

K. Awards, Prizes, and Recognition

Follow-On Funding

None at the time of this writing.