DR POWER
Data Repository for Power system Open models With Evolving Resources

Olga A. Kuchar, Stephen T. Elbert, Mark J. Rice
PNNL

January 27, 2017

PNNL-SA-123465

Team

- PNNL
  - Mark Rice (Principal Investigator)
  - Steve Elbert (Co-principal Investigator)
  - Jenn Hodas (Tech-to-Market Lead)
  - Olga A. Kuchar (e-grids.org Web Portal Lead)
  - Laurentiu Marinovici (Data Curation Lead)

- NRECA
  - David Pinney (Open Modeling Framework Lead)
  - Justin Yang (Software Developer)

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

Global Project Overview

- Mission: design, develop and host a data repository and web portal to:
  - Provide open-access power grid datasets and the capability to review, annotate, verify, and search submitted datasets
  - Ensure sustainable model and dataset dissemination and evolution through user-defined dataset creation and validation
  - Integrate and extend NRECA's success with OMF to include transmission modeling
- Challenges
  - Evolving and proprietary models
  - No standard approach to models
    - Planning engineers use bus-branch models
    - Real-time operators use node-breaker models
    - Cutting-edge technology is not always defined in the models
  - DR POWER targeting support for:
    - OMF/GridLAB-D (high-resolution distribution models)
    - PTI and MATPOWER (planning models)

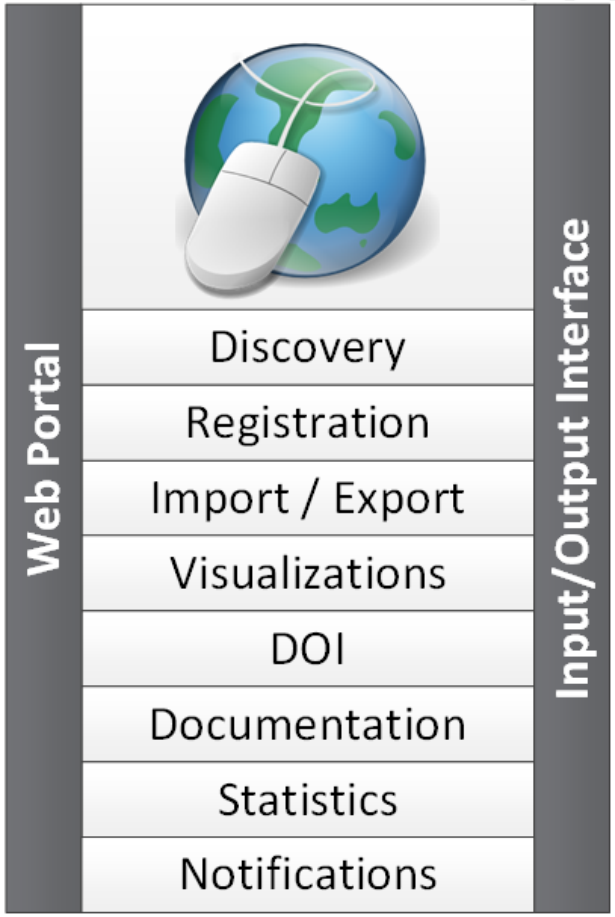
Goals

- Deliver the ability to collaboratively build, refine, review, and evolve high-fidelity power system models and accelerate grid optimization algorithm development
- Integrate community developed tools, esp. GRID DATA tools

Metric	State of the Art	Project Targets
Open Access	3 stale websites 10's of models Limited scenario sets	1 web portal 1,000s of models/data sets 1,000,000s of scenarios
Flexibility	Models are in limited formats No way to add new model details	Perform data transformations on-the-fly Ability to add new fields as needed and evolve the models, maintaining model history
Scalability	Total models are less than 1 Gigabyte	High-throughput scalable portal technology with petabyte storage
Sustainability	Static websites	PNNL is committed to building a dynamic community resource

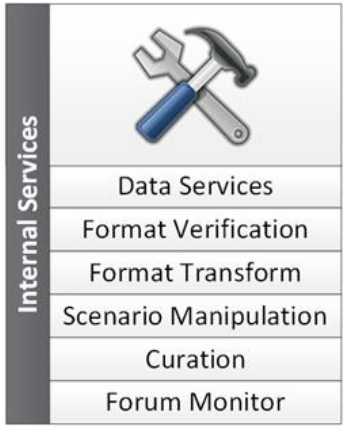
Web Portal

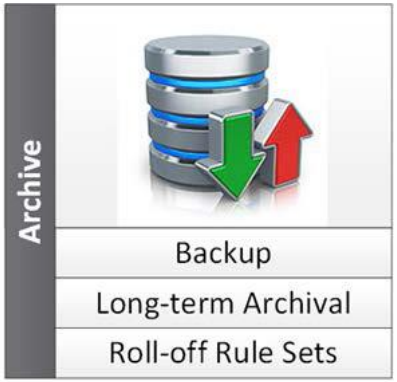
- Open-access web client for using data repository capabilities
- Execute tools for dataset generation, modification, citation, etc.
- Save results from such tools to repository and display the results to user
- Display various reports such as dataset version details, curation details, etc.
- Track upload,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



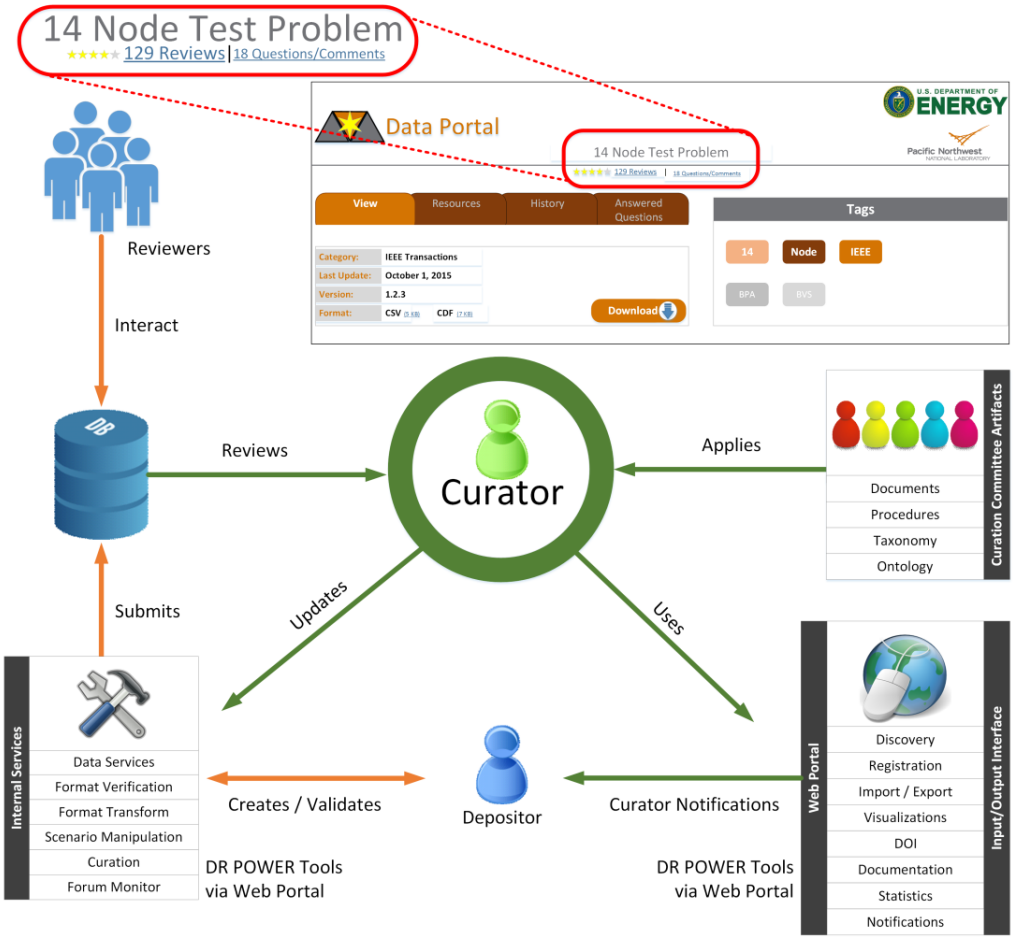
Data Repository (Back-End)

- Download requested model and scenario data in any available format
- Upload and store datasets for different power grid models; assign DOI
- 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





Active Curation



- Based on Digital Curation Center Lifecycle Model
- Curators will review uploaded models
- Curators will help guide model creation
- Community participation in reviews, questions, comments, etc.

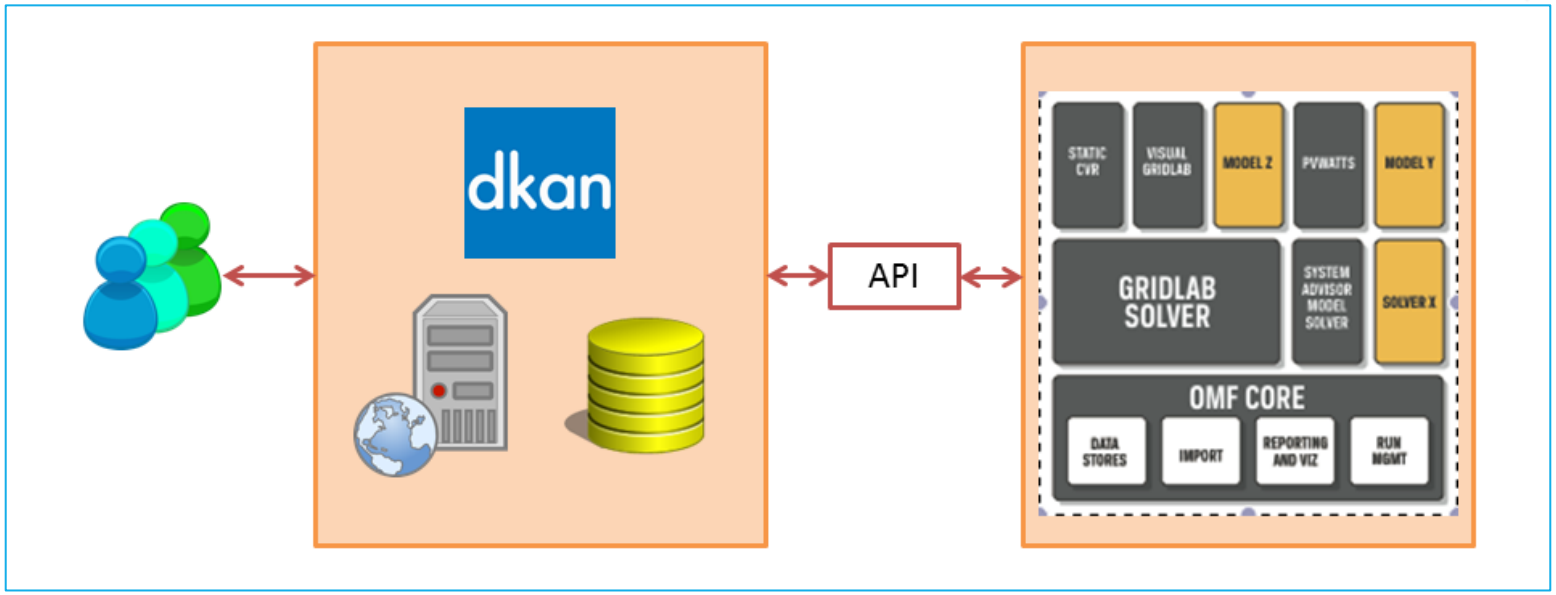
Short-term Goals

- Web Portal Version 1 deployed with following capabilities:
  - Register users
  - Upload/download models
  - Basic search
  - Populate with existing models
- DOI citations
- Tagging dictionary
  - Community engagement in curation process
- Format conversion and verification
  - PTI-MATPOWER
- Steering Committee engagement
- OMF
  - Updated Transmission User Interface
  - Integration into Web Portal

Long-term Goals

- Sustainability
  - PNNL/EIOC hosting
- Community Engagement/Adoption
  - Work with GRID DATA teams to make sure the repository meets their needs
  - Data repository for journals (e.g. IEEE)
- Functionality and maintainability
  - Ease-of-use
  - Citations

Overview and Docker Release



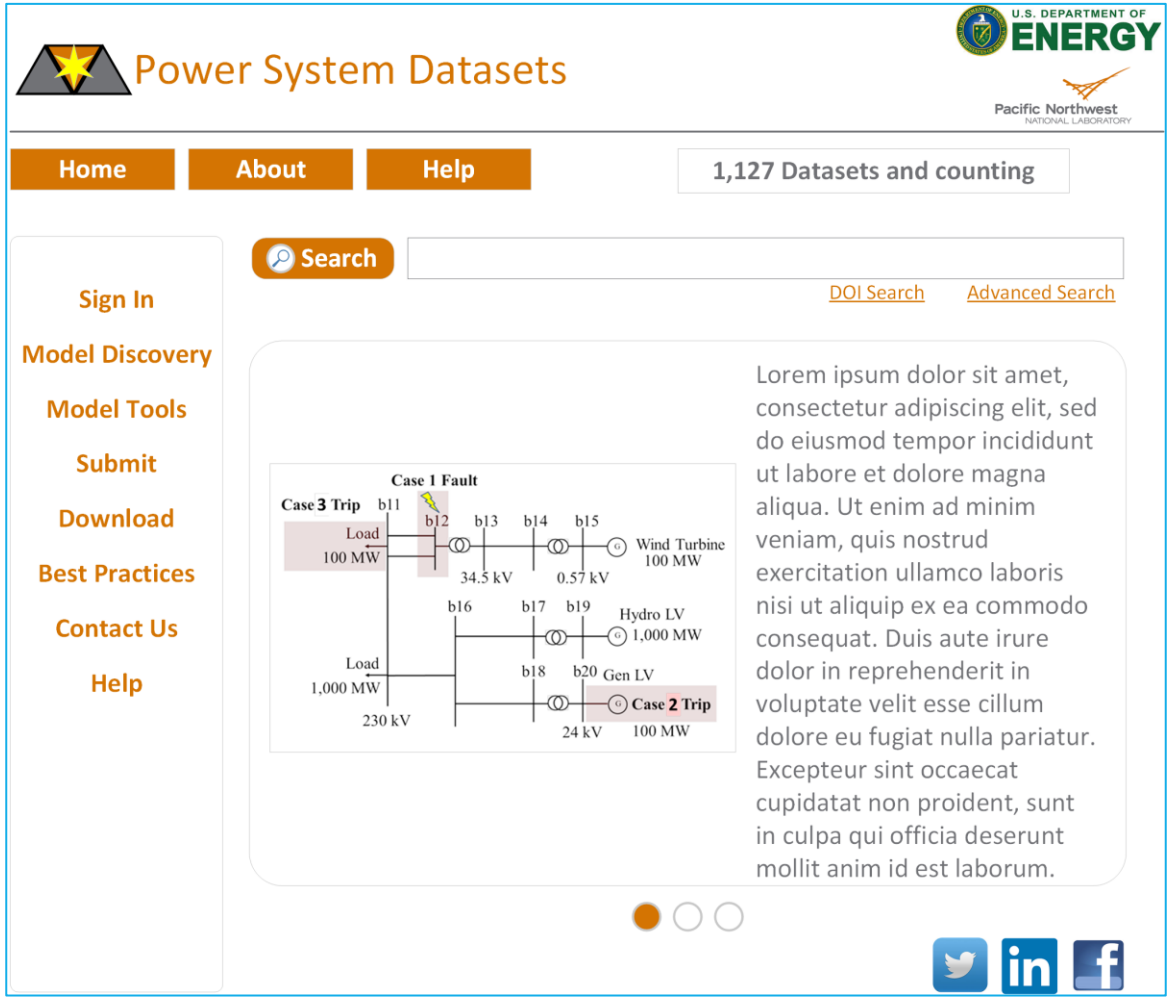
DKAN

- Drupal distribution to develop data portals
- Free, open-source open data platform with a full suite of cataloging, publishing, and visualization features that allows organizations to easily share data with the public
- Used to power many data portals, including:
  - Whitehouse (http://www.whitehouse.gov)
  - USDA (http://www.usda.gov)
  - California Data (http://data.ca.gov)
  - HHS (http://www.healthdata.gov)
- Meets open standards (e.g. Open Data, DCAT)
- Distributed by NüCivic, a subsidiary of GovDelivery

Open Modeling Framework

- Free, open source software development effort led by NRECA's Cooperative Research Network (http://omf.coop)
- Provide model editing and verification capabilities for both transmission and distribution networks
- As of December 2015: 115 users from 40 utilities

Data Portal Wireframe Main Page



Data Portal Wireframe
Data Discovery

Data Portal Wireframe
OMF Example

Data Portal

U.S. DEPARTMENT OF ENERGY

Pacific Northwest NATIONAL LABORATORY

GridlabMulti 14 Bus Test Case

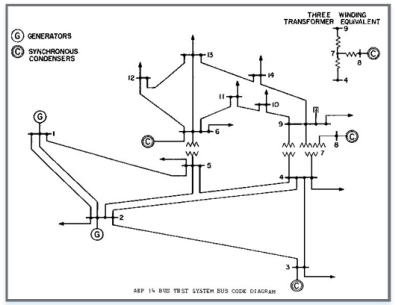
★★★★☆ 129 Reviews | 18 Questions/Comments

View | Resources | History | Answered Questions

Category:	IEEE Transactions
Last Update:	October 1, 2015
Version:	1.2.3
Format:	CSV (5 KB) CDF (7 KB)

Open Download

Preview



Tags

14 Node IEEE BPA BVS

Tell the world about the dataset!

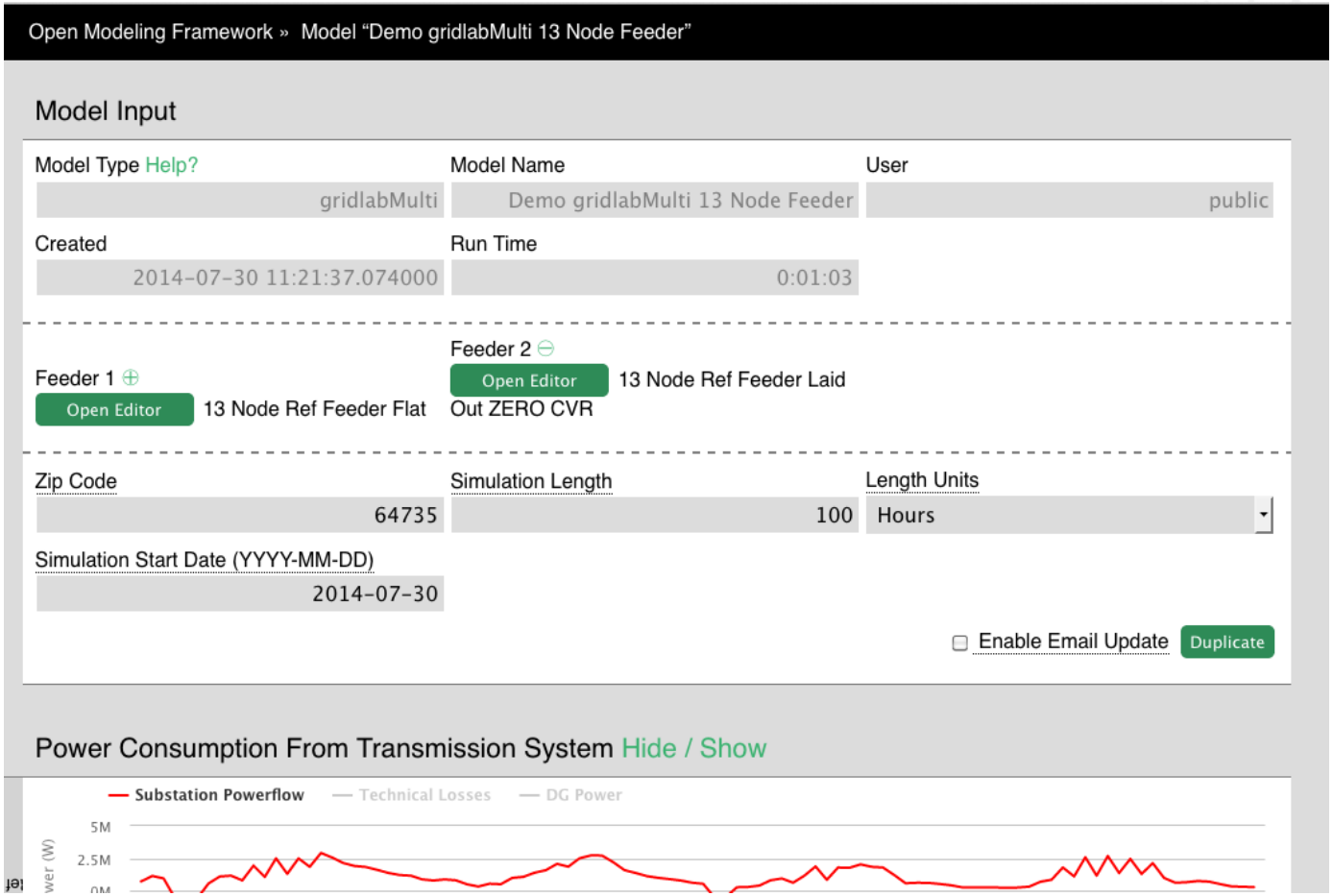
Developer Tools

Dataset URI:	.gov/pnwsgd/0E12F...	Copy URI
REST API:	.gov/pnwsgd/0G318...	Copy URI

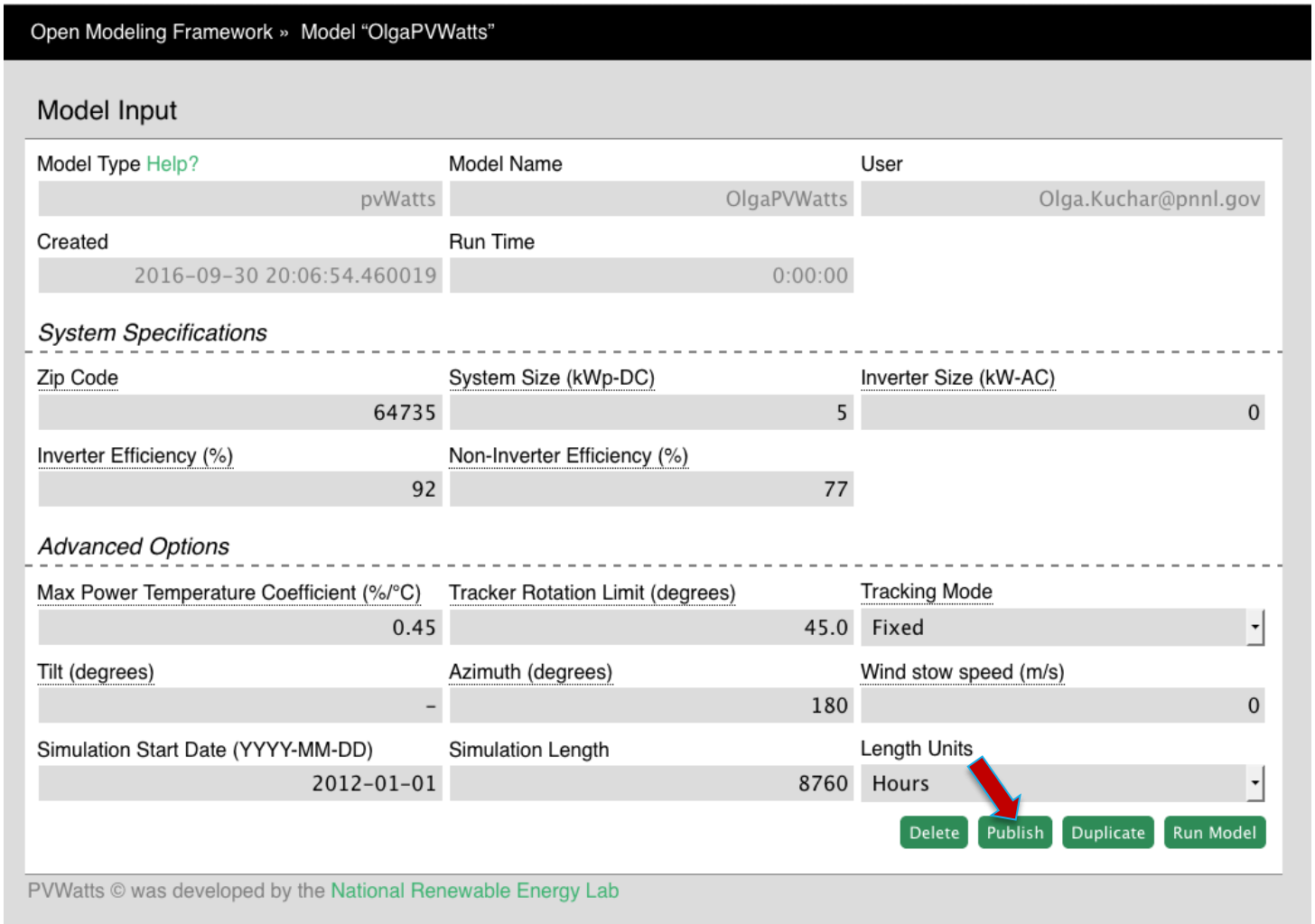
Do more with the data!

Suggest an idea | Request data

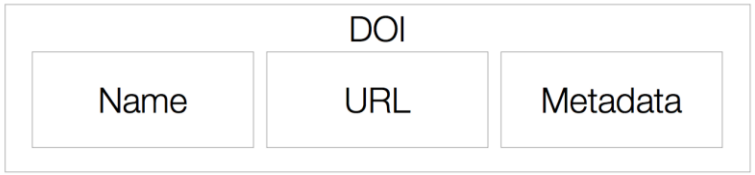
Data Portal Wireframe
OMF Example (2)



Data Portal Wireframe
OMF Example (3)



Digital Object Identifier (DOI)



- Uniquely identifies a digital data object (http://doi.org)
- Using DataCite's API to mint and update DOIs
- Name preferably includes an ORCiD ID (http://orcid.org)

Web Portal API

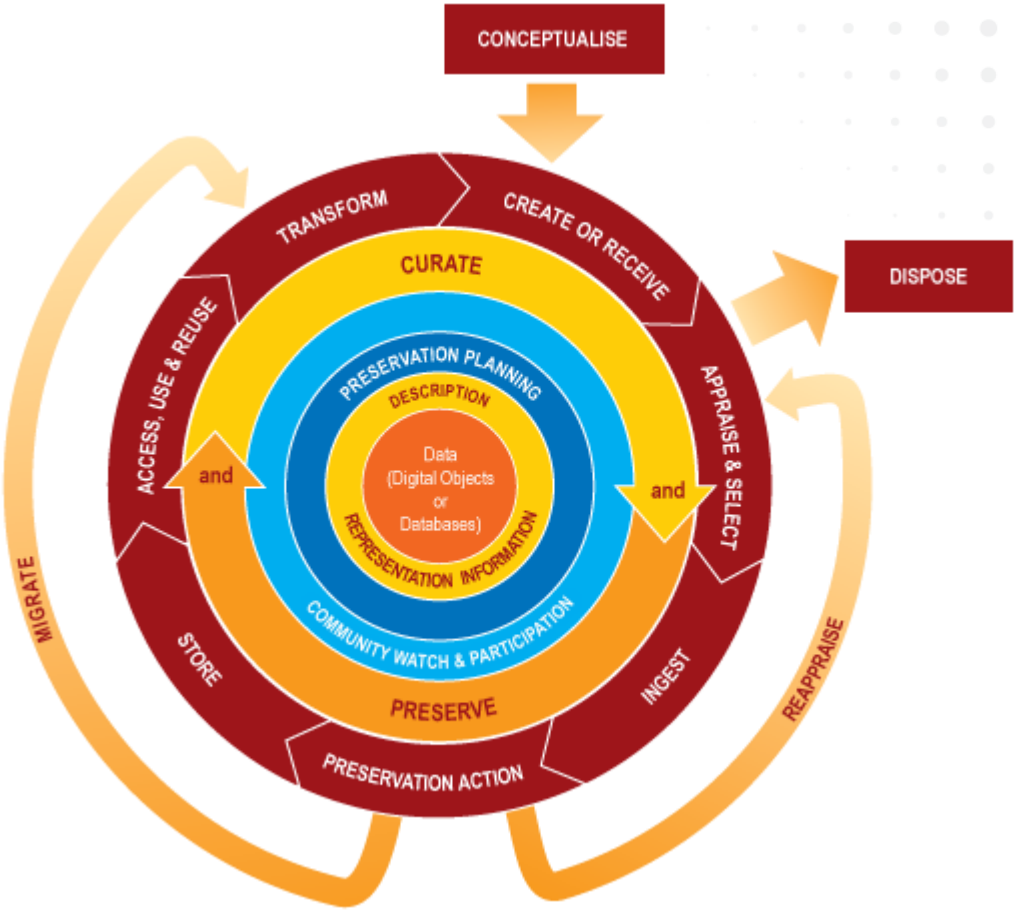
- Provide programmatic/automated access
- Read-only/download access (no writes or uploads via the API)
  - Upload is via user interface for CEII security
- Follows Open Data Schema

Web Portal Security

- FISMA compliant
- Content web proxy firewall
- Scanned, usually daily, for vulnerabilities
- Data in transit is protected via Open SSH and Open SSL
- Daily, weekly, and monthly backups
- Registration for data uploads and forum participation
- Terms and Conditions acceptance upon registration
- Additional Drupal's built-in security features
- RBAC user roles

Curation Working Group

- Appraising methodology based on inputs from experts in the group
- Engaging a community of users to bring in a larger set of models
- Transformation through advanced tagging and metadata addition to facilitate searching and classification

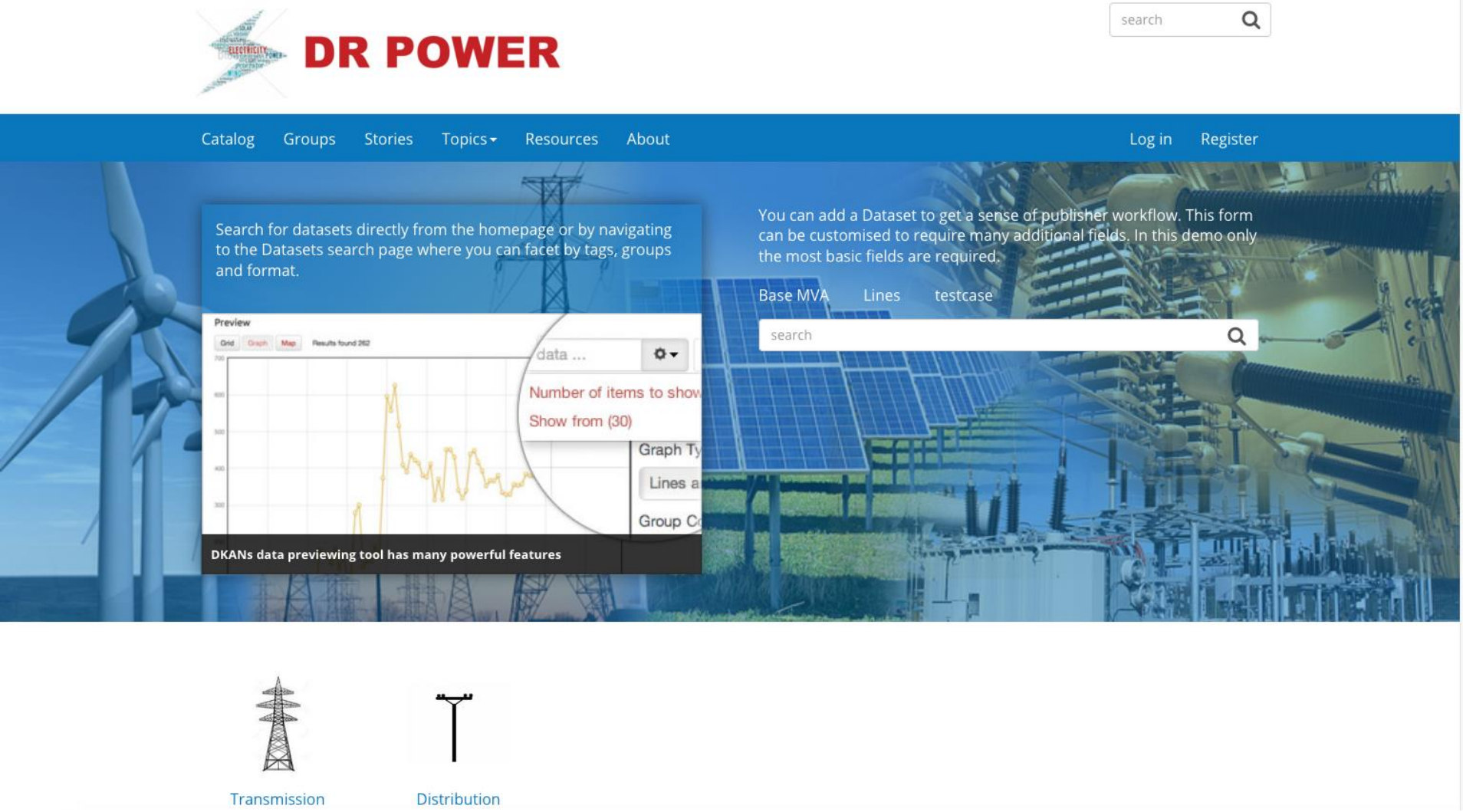


DCC Curation Lifecycle model
http://www.dcc.ac.uk/resources/curation-lifecycle-model

CEII

- A data provider will certify that the uploaded file does not contain any CEII or Proprietary Information
- Human curators will review the uploaded file(s) in the context of possible CEII and other data quality issues
- Data found to be CEII or Proprietary will be immediately removed from public access and any backups
- DOI will remain with a status update

Web Portal Version 1



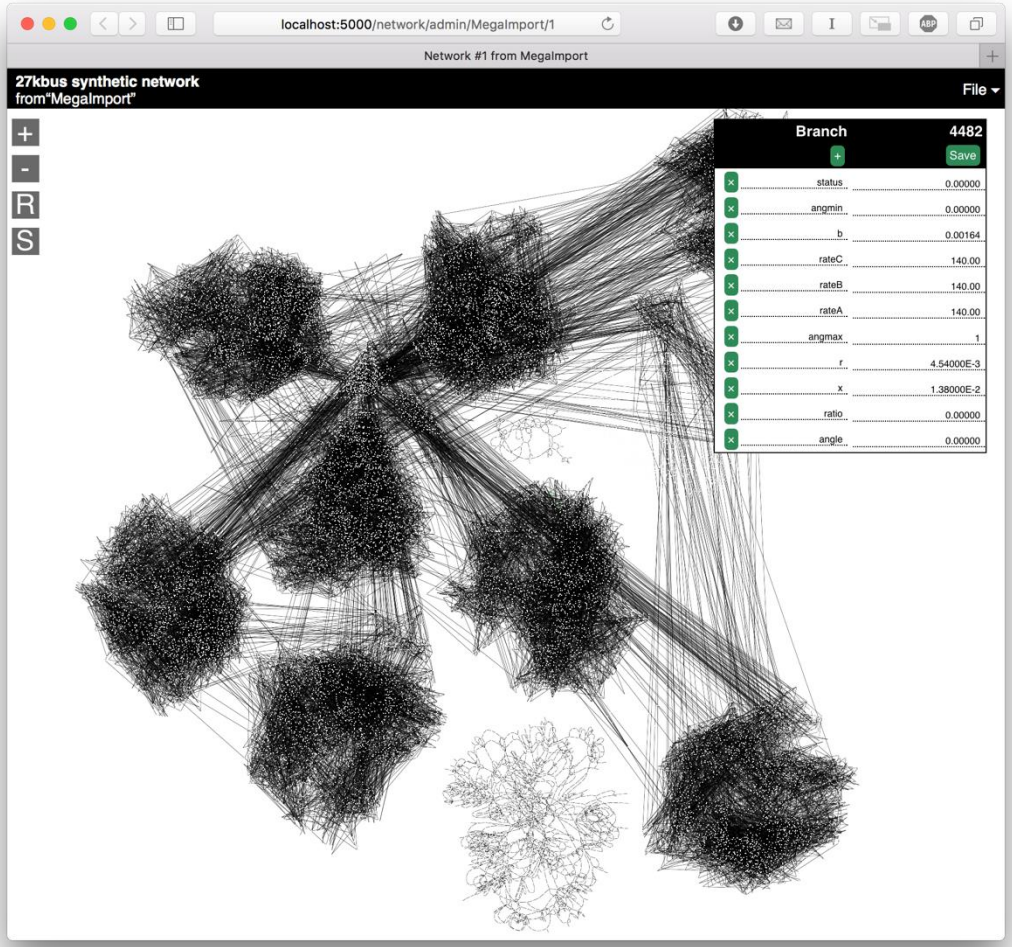
Functionality

- Register users
- Upload/download models
- Basic search
- Populated with Richard Christie's models
- Tagging dictionary
- CEII certification

☐ This file contains no Critical Electric Infrastructure Information (CEII) *

By checking this box, I hereby acknowledge that the data in this file contains no Critical Electric Infrastructure Information (CEII) as defined in 18 C.F.R. § 388.113(c)(1); as well as "Critical Electric Infrastructure Information" and "Defense Critical Electric Infrastructure" as defined in Sec. 215A(a)(3) of the Federal Power Act (16 U.S.C 824 et seq.) as amended by Division F, Section 61003 of the Fixing America's Surface Transportation Act or the "FAST Act", 2015, Pub. L., No. 114-94 and as further defined, designated or identified in any regulations or orders promulgated thereunder.

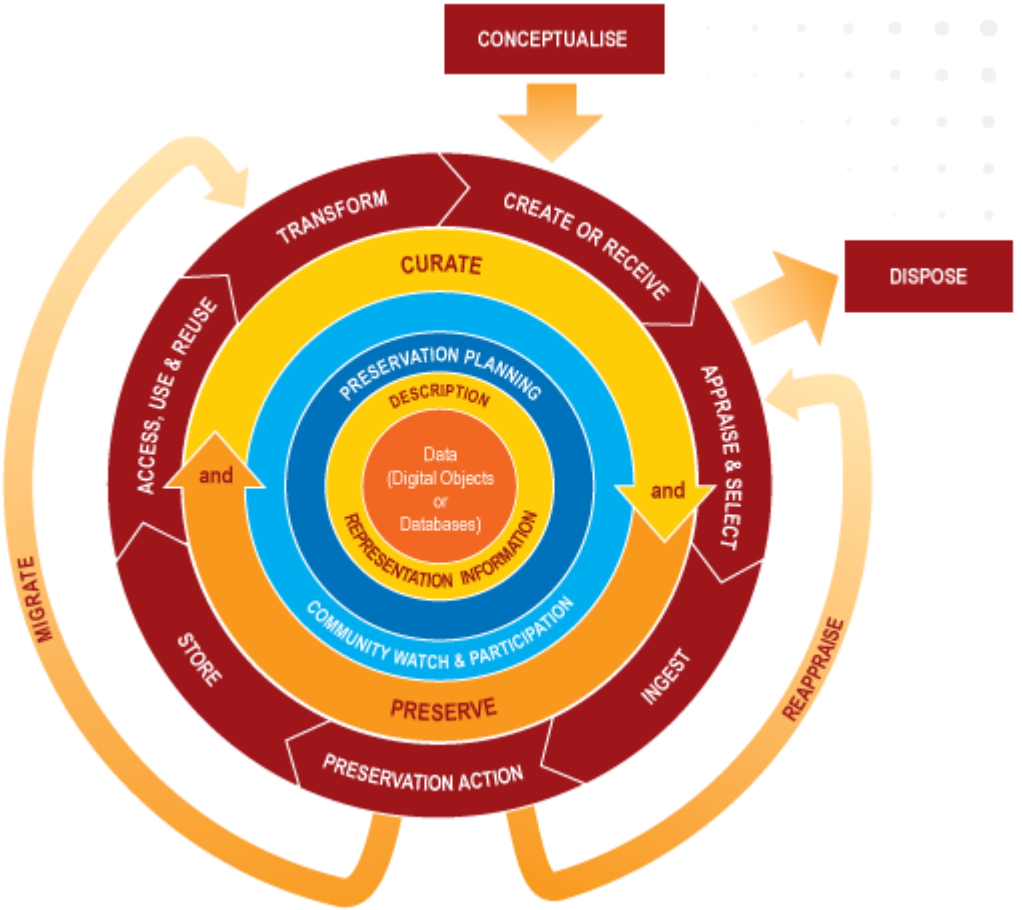
OMF Improvements



https://drive.google.com/open?id=0B0hLAR4WQ4X4RlFfcFNISnA4UUk

Curation Working Group

- Conceptualize
  - Power System data dictionary
  - Taxonomy and tagging
- Create or receive
  - Selection of different models in particular formats



DCC Curation Lifecycle model
http://www.dcc.ac.uk/resources/curation-lifecycle-model

Changes in Approach

- IEC Technical Committee 57: Common Information Model
  - Not used by GRID DATA teams
  - Expensive to maintain
  - Commonly used to represent node-breaker models
  - Bus-branch models preferred for steady-state (e.g. power flow, OPF)
- OMF uses MATPOWER as base transmission class

Challenges Ahead

- Community engagement
- Taxonomy expansion
- Curation Guidance
- Format conversions and verification

Focus for the Year Ahead

- Live Web Portal (awaiting approvals)
- Community engagement
- Evolution of curation process
- Updated Transmission User Interface (OMF)
- Improvements to the portal:
  - PTI-MATPOWER conversions
  - OMF access
  - GRID DATA tools

Sustainability

- Institutional funding
- Continual community development
- Synergistic opportunities:
  - End-use load energy consumption (building data)
  - DARPA RADICS program
  - Hosting solar and wind data (e.g. NREL)

CHANGING WHAT'S POSSIBLE

Thank you!