IEEE PES Big Data Panel
Sponsored by IEEE PES Big Data subcommittee

Finding the Right Grid Model for Your Research in the GRID DATA Repository Using Big Data Semantic Search

Terry Nielsen
EVP, GridBright
Chair of Tech Committee, BetterGrids.org

August, 2018

GRID DATA Program

Generating Realistic Information for the Development of Distribution And Transmission Algorithms

Duration	2016-2018
Projects	7
Investment	$11M

Goal:

Development of large-scale, realistic, validated, and open-access electric power system network models with the detail required for successful development and testing of new power system optimization and control algorithms.

BetterGrids Repository

- **A free library** of public grid model data
- **Supporting research** in grid optimization and reliability
- Enabling grid researchers to **collaborate and share data**
- Supported by a **community of volunteers** led by GridBright
- Funded by the **DOE ARPA-E** GRID DATA Program

BetterGrids Repository Status

Metric	Q2, 2017	Q3, 2017	Q4 2017	Q1 2018	Q2 2018
Model contributors (count)	5	5	6	8	9
Registerd accounts (count)	20	92	120	142	207
Registered curators (count)	3	3	3	3	3
Model collections (count)	12	12	13	13	13
Distinct models (count)	44	190	291	298	314
Model files (count)	55	630	817	833	881
Model files (Mb)	106	385	615	2906	2988
Downloads (count)	1,085	20,400	41,480	52,057	79,209

Capabilities

Contribute Models

Describe | Describe | Upload | Verify | License | Complete

Submit: Describe this Item

Please fill further information about this submission below.

Select the keyword(s) associated with this item. Hold down the "CTRL" or "Shift" key to select more than one keyword in the list.

Subject Keywords

- Power Grids
- Smart Grids
- Generators
- Distributed power generation
- Switches
- Power Quality

Select the data format from the list.

Data Format: MatPower

Enter Version of the selected Data Format.

Data Format Version: 1

Enter Model Class of this data.

Model Class:

Enter number of buses

Buses: 1000

Enter number of generators

Generators: 200

Enter number of loads

Loads: 500

Find Models

BETTERGRIDS.ORG Browse Search Help

Search

Search: All of BetterGrids

for: Go

Current filters: Data Format | Equals | MatPower | X

Start a new search

Add filters:

Use filters to refine the search results.

Title | Equals | Add

Results/Page 10 | Sort items by Relevance In order Descending Authors/record All Update

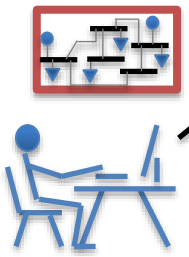
Results 1-6 of 6 (Search time: 0.001 seconds).

previous 1 next

Item hits:

Issue Date	Publisher	Title	Version	Data Format	Feeders	Loads	Buses	Generators
12-Mar-2013	*Edinburgh Test Case Archive*	39 bus New England test case with realistic cost data	*1*	*MatPower*	-	*46*	*39*	*10*

Model Curation



Model Submission

Model Submitted by Model Creator for purposes of sharing with research community

Submission Review

Submission reviewed for proper attributions, tags, description, etc.

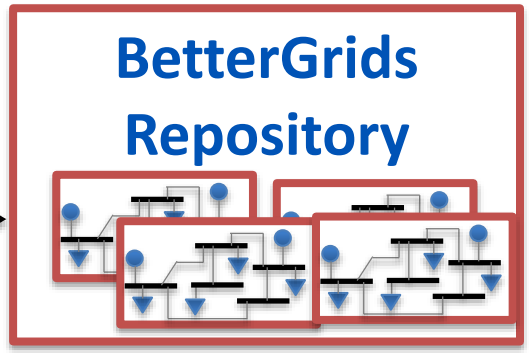
Model Review

Model reviewed for usability, data content, and annotation

Policy Review

Model reviewed for proper licensing, CEII data, and compliance with polices

Model published



Semantic Search Capabilities

Challenge

- Researchers often look for grid models with unique electrical conditions
- With a large number of models (100s-1000s) that are very large (10,000-1,000,000 nodes/buses) manual cataloging is impossible
- Unique model conditions can't be found with traditional file/database searches

Solution

- Construct a database of the models that "understands" the equipment & organization of the models so that it can be intelligently searched
- Achievable by translating & pre-processing the data, storing it in a NOSQL/graph database, and searching using **big data** techniques

Search Requirements Analysis

We identified four primary query *types* –

- Simple Model Attribute Queries – to find models of a specific type, format, author, title, description, or keyword
- Equipment Type Queries – to find models that contain desired numbers of specific equipment
- Model Hierarchy Queries – to find models that have criteria within a subset of the model based upon the model hierarchy
- Time Series Data Queries – to find time series data that meets a specific criteria

Definitions

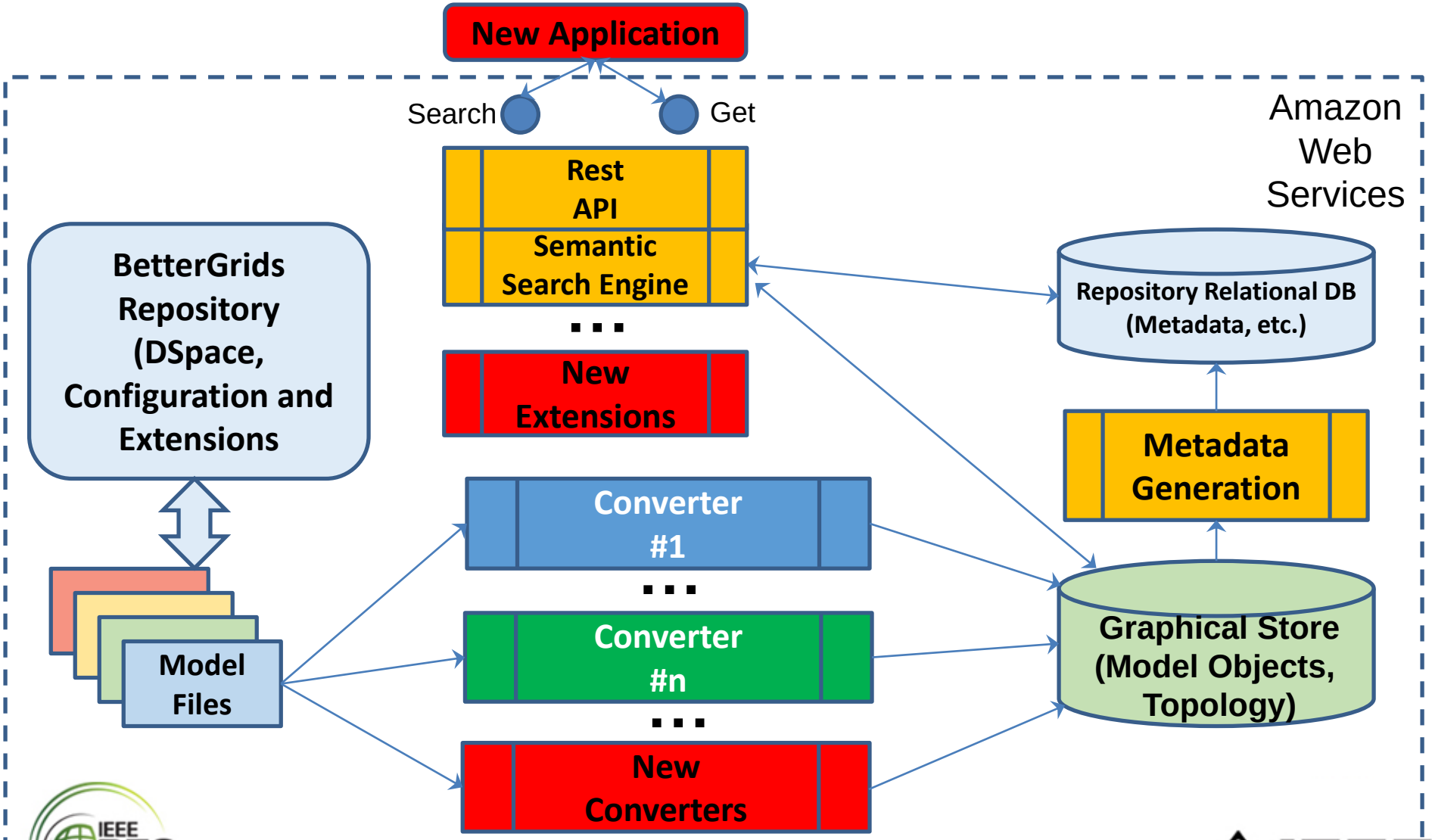
By Semantic search we mean that users can:

- Employ a relatively generic and natural vocabulary to find things of interest
- Without significant regard for the specific data attribute names and formats used by different network model storage formats

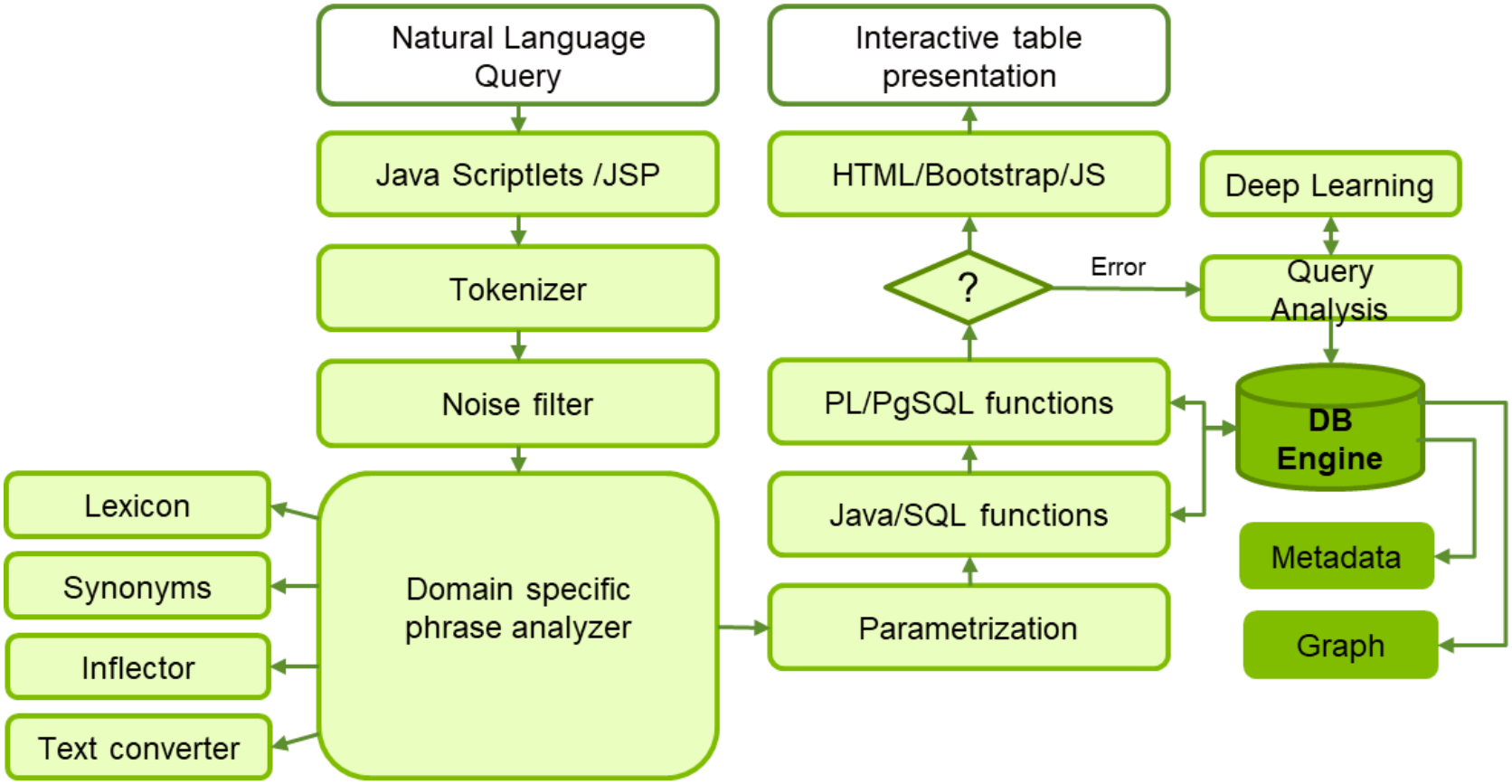
Definitions

- The use of a big data ‘graph database’ approach allows for efficient traversals of the network model, an activity where traditional databases often perform poorly.
  - This supports queries where the connectivity relationship between equipment matters.
  - Native graph processing is the most efficient means of processing graph like data because data elements physically points to each other
  - Non-native graph processing uses other processes: i.e Create, Read, Update or Delete (CRUD) operations

Architecture



Semantic Search Engine

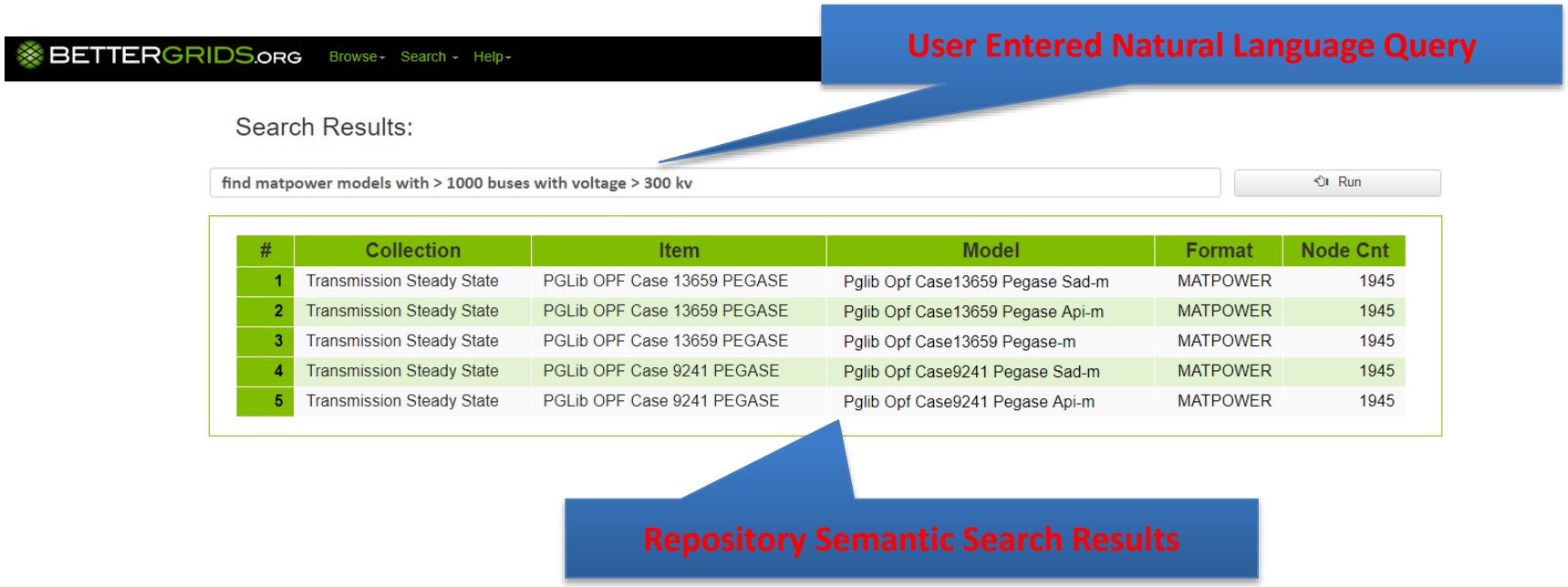


Semantic Search for Grid Models

- **Semantics** is about the relationship between phrases and their meanings
  - In contradictory to **syntax** or grammar
- For example, the phrase: ***"The water is triangle"*** is syntactically correct but has no semantic meaning
- Queries like:
  - *Show me all grid models with at least 10 generators;*
  - *Please return grid item files where the number of generators is more than nine;*

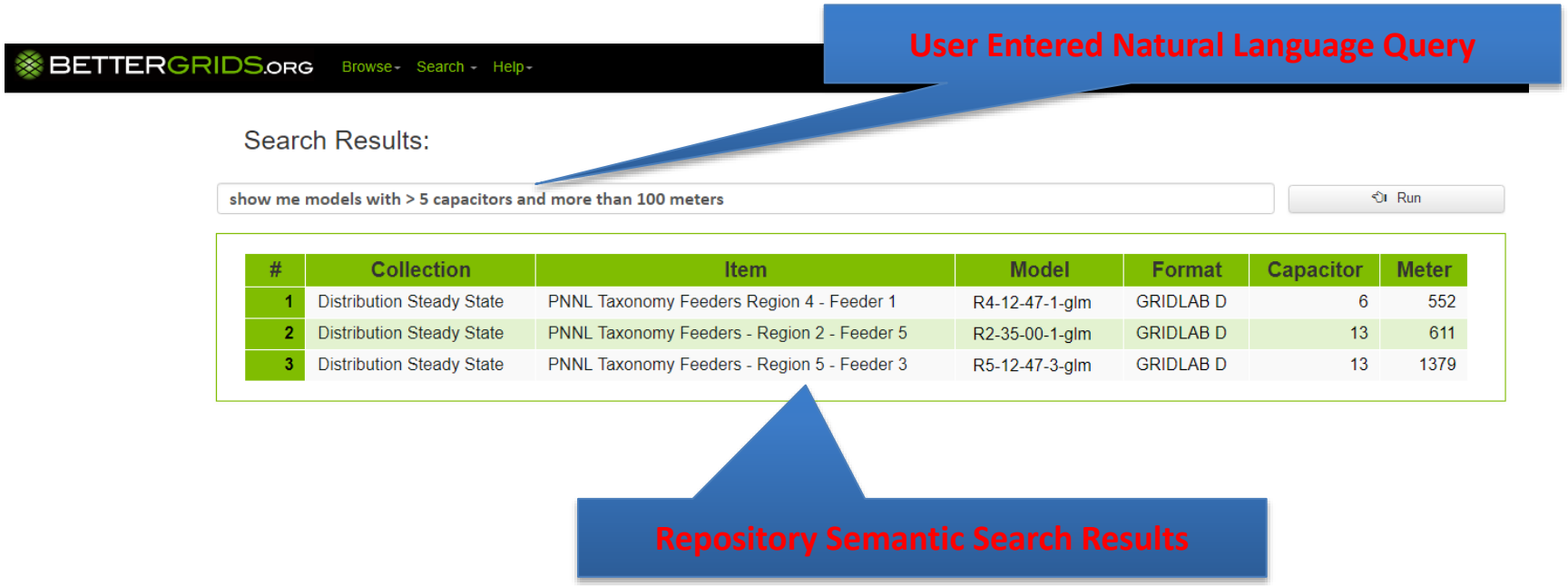
  look different, but semantically are absolutely identical (in a scope of Power Grids terminology).

Semantic Search Example #1



#	Collection	Item	Model	Format	Node Cnt
1	Transmission Steady State	PGLib OPF Case 13659 PEGASE	Pglib Opf Case13659 Pegase Sad-m	MATPOWER	1945
2	Transmission Steady State	PGLib OPF Case 13659 PEGASE	Pglib Opf Case13659 Pegase Api-m	MATPOWER	1945
3	Transmission Steady State	PGLib OPF Case 13659 PEGASE	Pglib Opf Case13659 Pegase-m	MATPOWER	1945
4	Transmission Steady State	PGLib OPF Case 9241 PEGASE	Pglib Opf Case9241 Pegase Sad-m	MATPOWER	1945
5	Transmission Steady State	PGLib OPF Case 9241 PEGASE	Pglib Opf Case9241 Pegase Api-m	MATPOWER	1945

Semantic Search Example #2



#	Collection	Item	Model	Format	Capacitor	Meter
1	Distribution Steady State	PNNL Taxonomy Feeders Region 4 - Feeder 1	R4-12-47-1-glm	GRIDLAB D	6	552
2	Distribution Steady State	PNNL Taxonomy Feeders - Region 2 - Feeder 5	R2-35-00-1-glm	GRIDLAB D	13	611
3	Distribution Steady State	PNNL Taxonomy Feeders - Region 5 - Feeder 3	R5-12-47-3-glm	GRIDLAB D	13	1379

Thank You

Terry Nielsen
Executive Vice President, GridBright, http://gridbright.com
Technical Committee Chair, BetterGrids, http://bettergrids.org
Email: terry.nielsen@gridbright.com
Phone: (612) 978-1381

Join the Community at: www.bettergrids.org
More info at info@bettergrids.org