CHANGING WHAT'S POSSIBLE

Sustainable Data Evolution Technology (SDET) for Power Grid Optimization

presented by

Ruisheng Diao, Ph.D., P.E.
Team Lead/Staff Research Engineer
Pacific Northwest National Laboratory

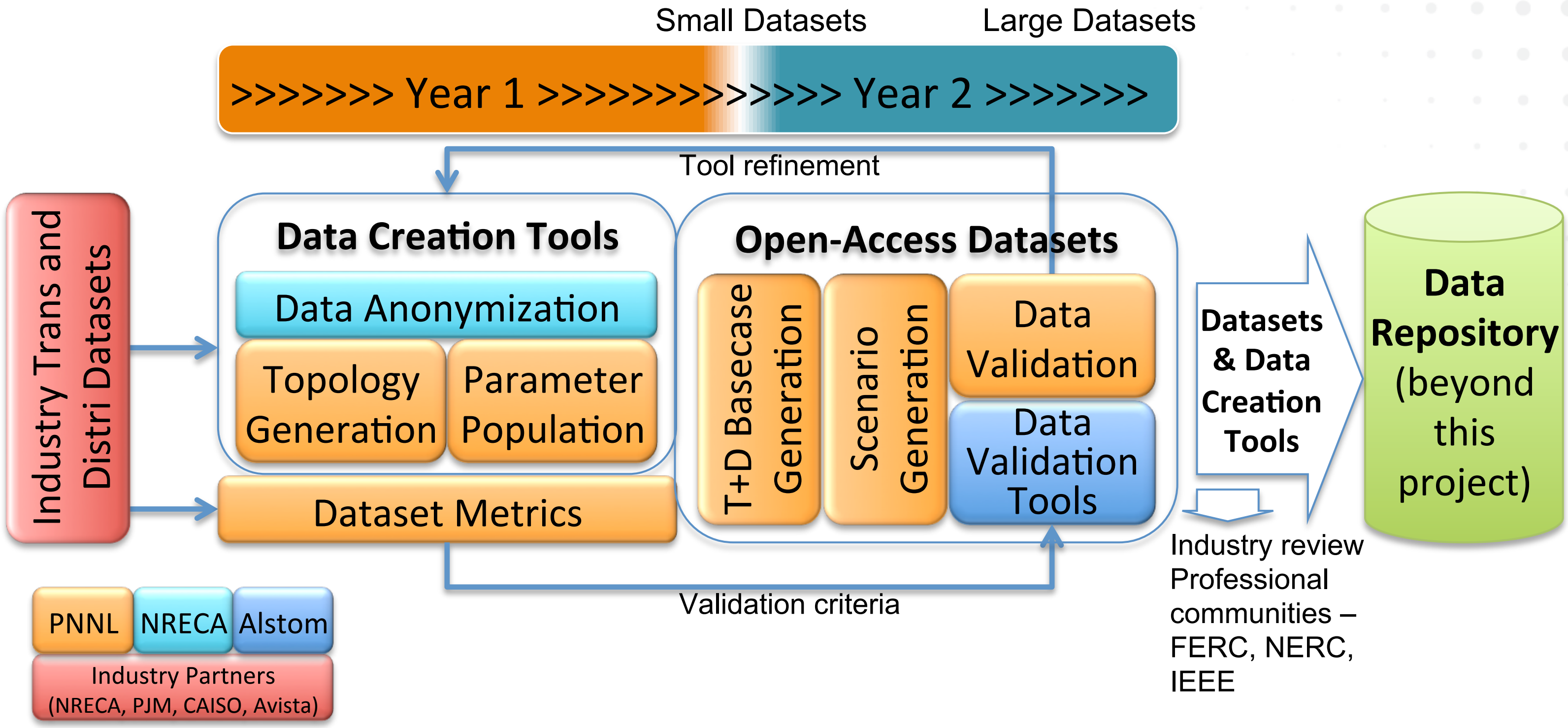
July 20th, 2017

Project Summary

- Objective: to deliver large-scale, realistic, **evolvable** datasets and data creation tools for optimization problems such as AC OPF and VVO
  - Derive data features/metrics for real T+D systems
  - Develop tools to generate large-scale, open-access, realistic synthetic datasets
  - Validate the created datasets using industry tools
  - Integrate with GRID DATA repositories
- A novel concept "***data evolution***", with long-lasting impact
  - Enable the datasets to evolve with the increasing grid complexity.
  - Accelerate development and adoption of grid optimization methods.
  - Improve the reliability, resiliency and efficiency of the power grid.
- Timeline: October 2016 - September 2018

Evolvable open-access large-scale datasets to accelerate the development of next-generation power grid optimization.

Tasks and Dependency



Datasets Requirements

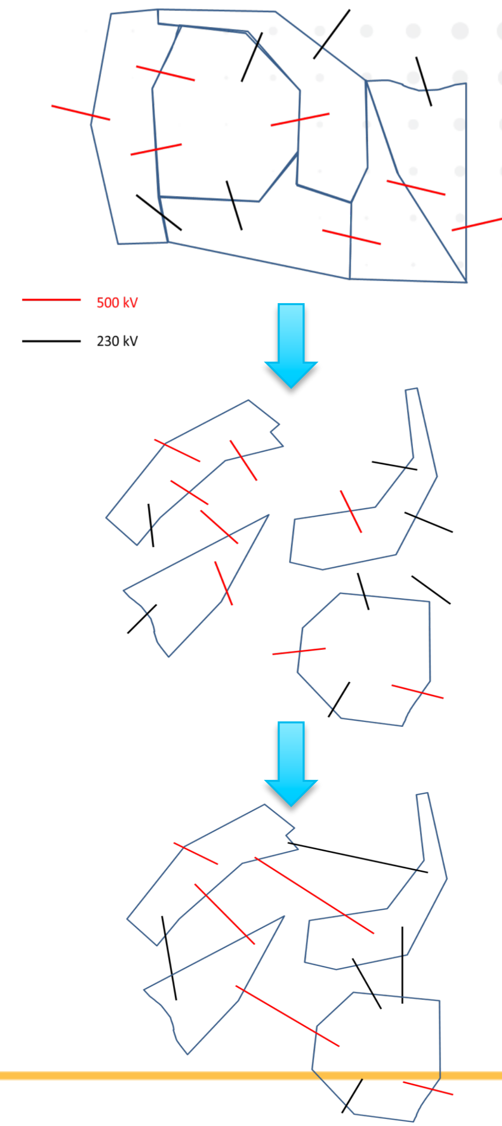
- Large-scale
- Realistic
- Open-access
- Sustainable (ARPA-E independent)
- Evolvable (datasets are not static)

Deliverables

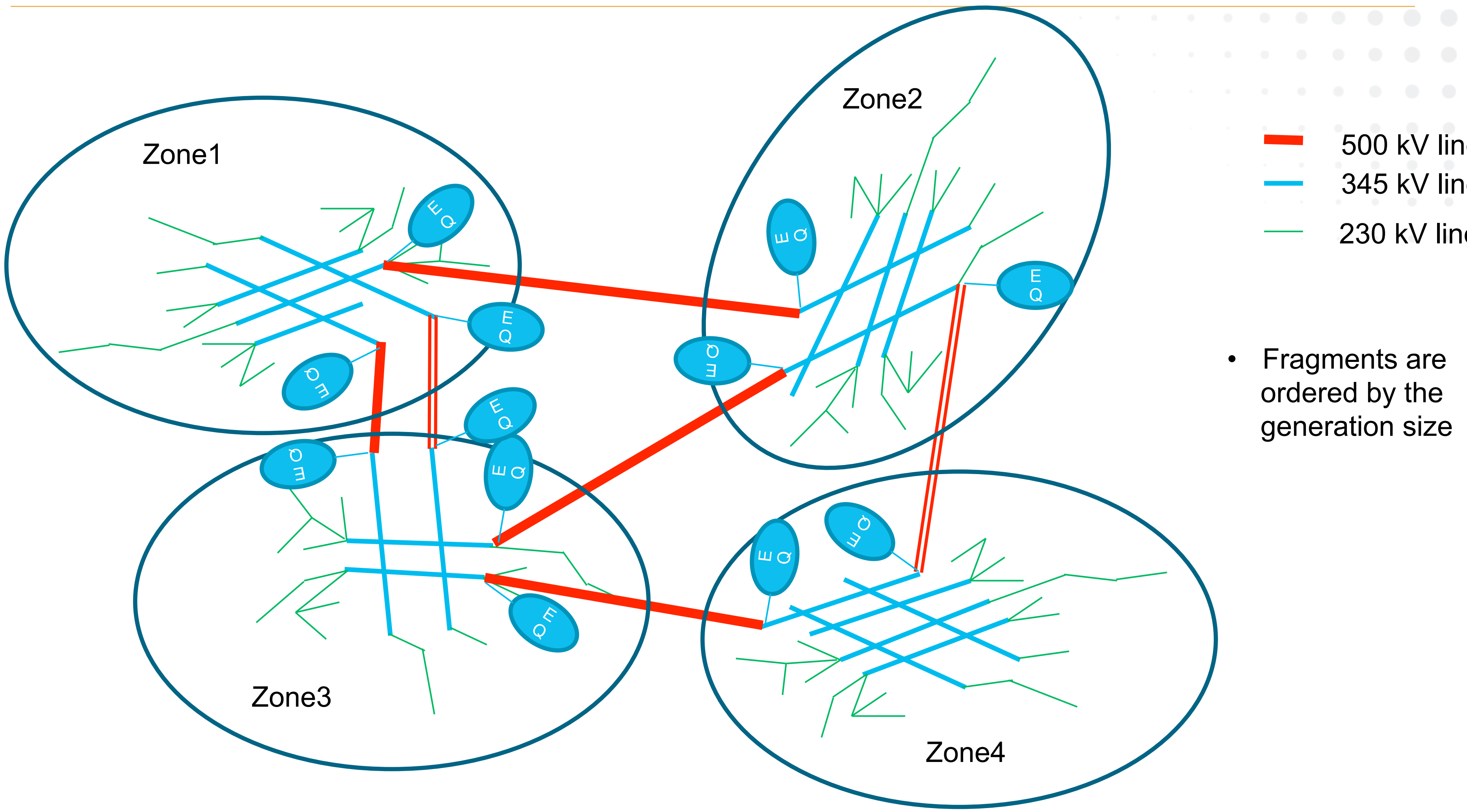
- Datasets
- Dataset creation tools

A Fragmentation Approach

- "Deterministic" approach on the system fragment level for the most of system parameters
  - Real-world systems will be used
  - Each system model will be fragmented into zones, preserving:
    - Generation, load level
    - Lines, transformers, controllers
  - Data anonymization approach will be used
  - The zones will be recombined to form the desired system model
  - Creating tie-lines between zones through a graph theory algorithms

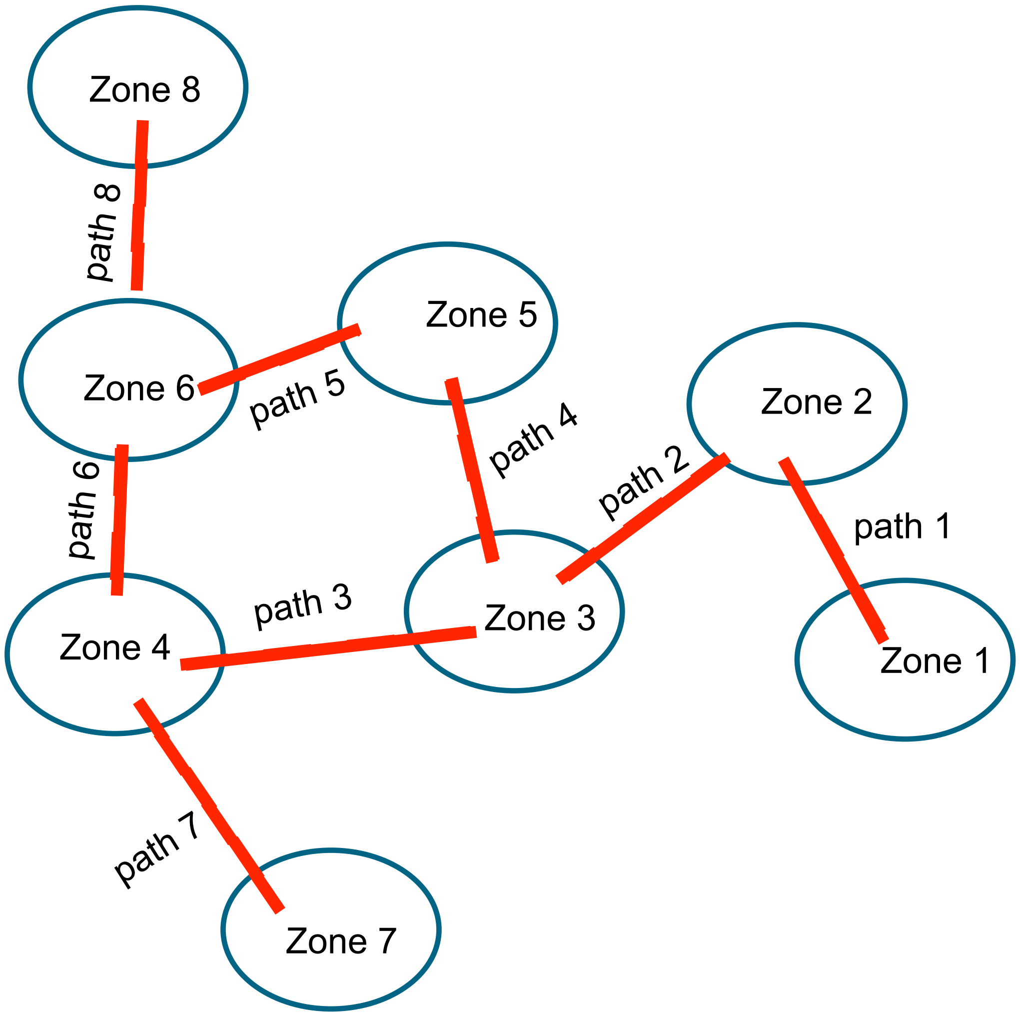


Prepare fragments from real-world models



- Fragments are ordered by the generation size

An iterative process to build the “kernel”



Inputs

- Desired size of the model
- A number of fragments with connectors

Outputs:

- Synthetic skeleton/ topology (picture in the middle)
- Minimizing voltage difference between zones
- Minimizing line crossings
- Satisfying graph metrics
- Paths are ordered such that smallest zones are connected first.

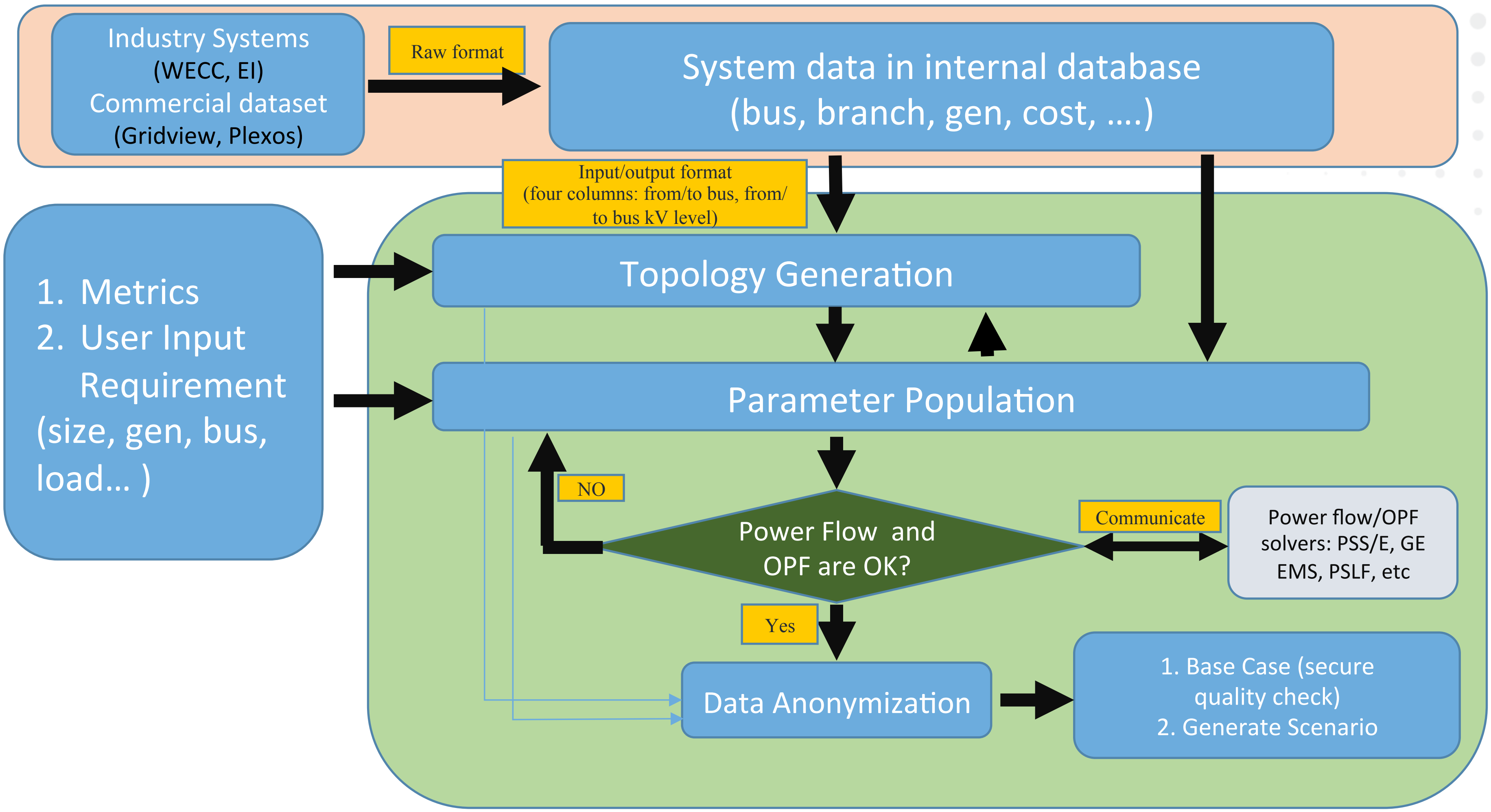
Creating Key Grid Information

- A "probabilistic" approach for
  - Production cost/market bid data
  - Variable resources
  - Random factors added to the system load
- Distribution System Model Creation for VVO
  - Real-world feeder models and data will be collected
  - Applying a data anonymization approach

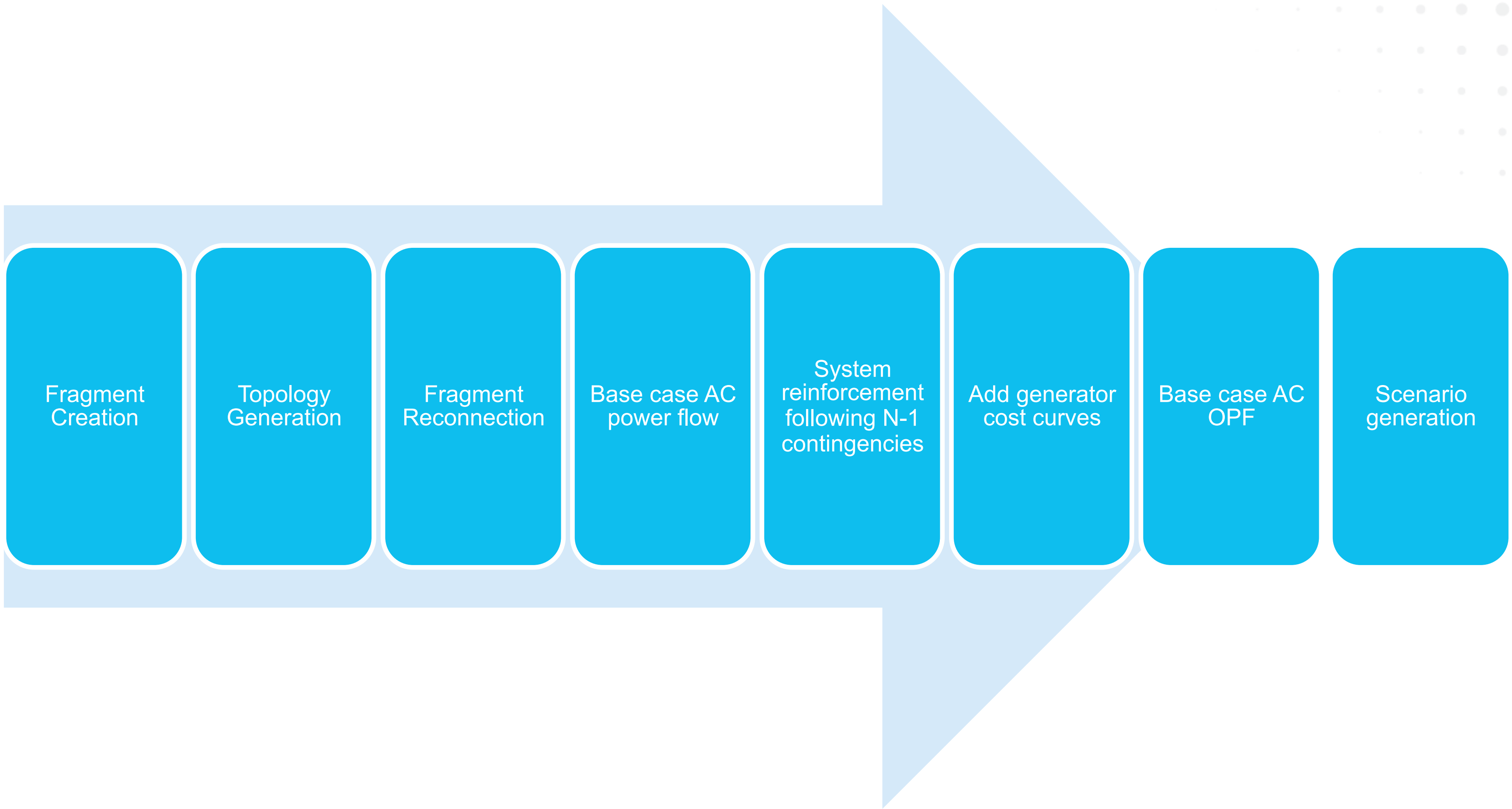
Metrics for measuring realism

- **Graph-theoretic metrics**
  - Degree distribution
    - The degree of a node is the number of connections it has to other nodes
    - Degree distribution is the probability distribution of these degrees over the whole network
  - Average shortest path length
    - A path between two vertices (or nodes) such that the sum of the weights (number of edges in a path) of its constituent edges is minimized
    - Average number of branches between 2 buses
  - Diameter
    - The longest shortest path between any pair of vertices
    - The max number of branches between 2 buses (is a function of system size)
  - Average clustering coefficient
    - Ratio of actual edges between its neighbors to all possible edges
    - Clustering coefficient tells how well the graph nodes are connected with each other
- **Power grid parameter metrics**
  - Based on real-world characteristics, we will use the following typical statistical measures: 1) Mean value; 2) Standard deviation (STD); 3) Min value, and 4) Max value

SDET Tool Architecture

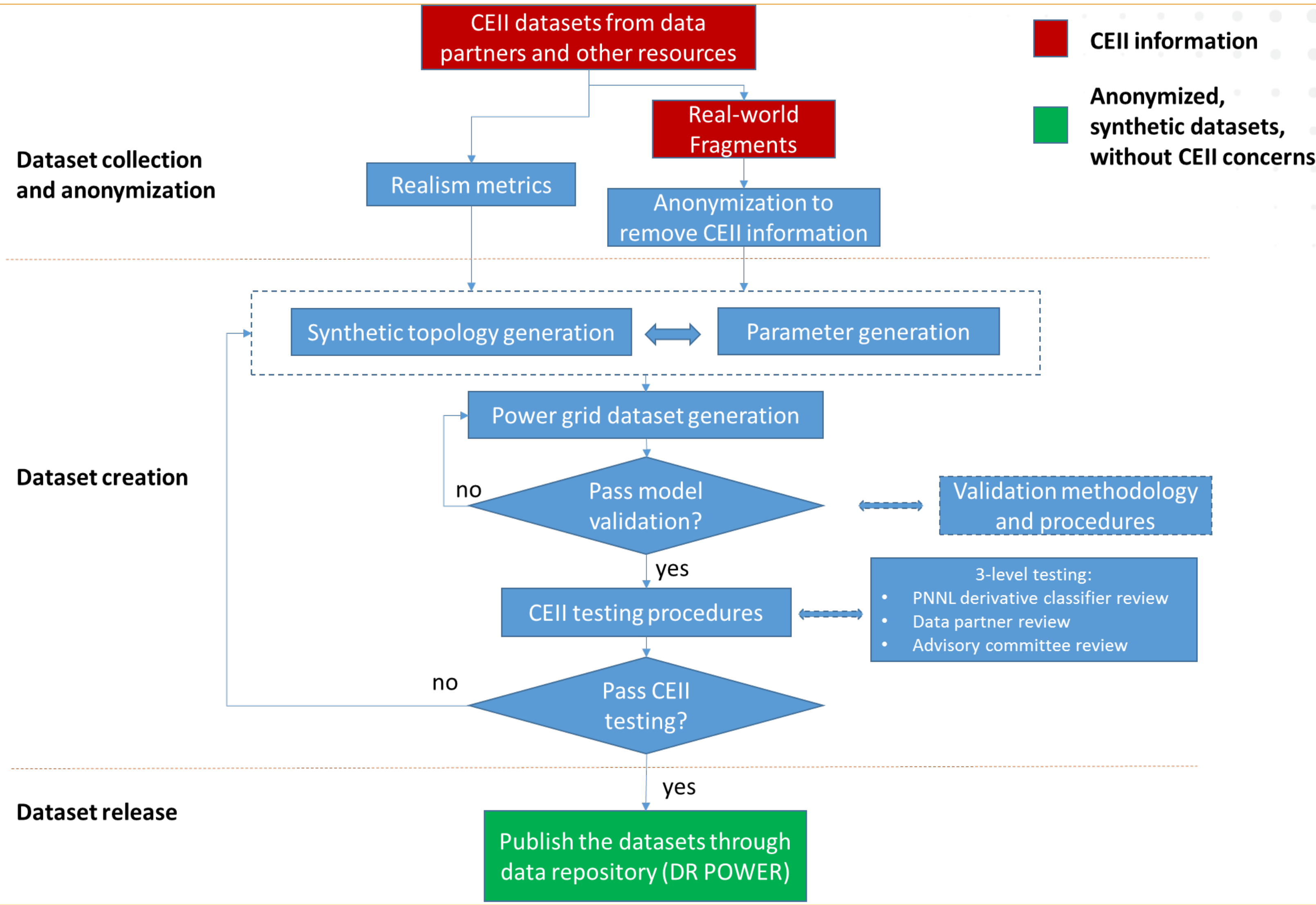


Key Steps



Fragment Creation
Topology Generation
Fragment Reconnection
Base case AC power flow
System reinforcement following N-1 contingencies
Add generator cost curves
Base case AC OPF
Scenario generation

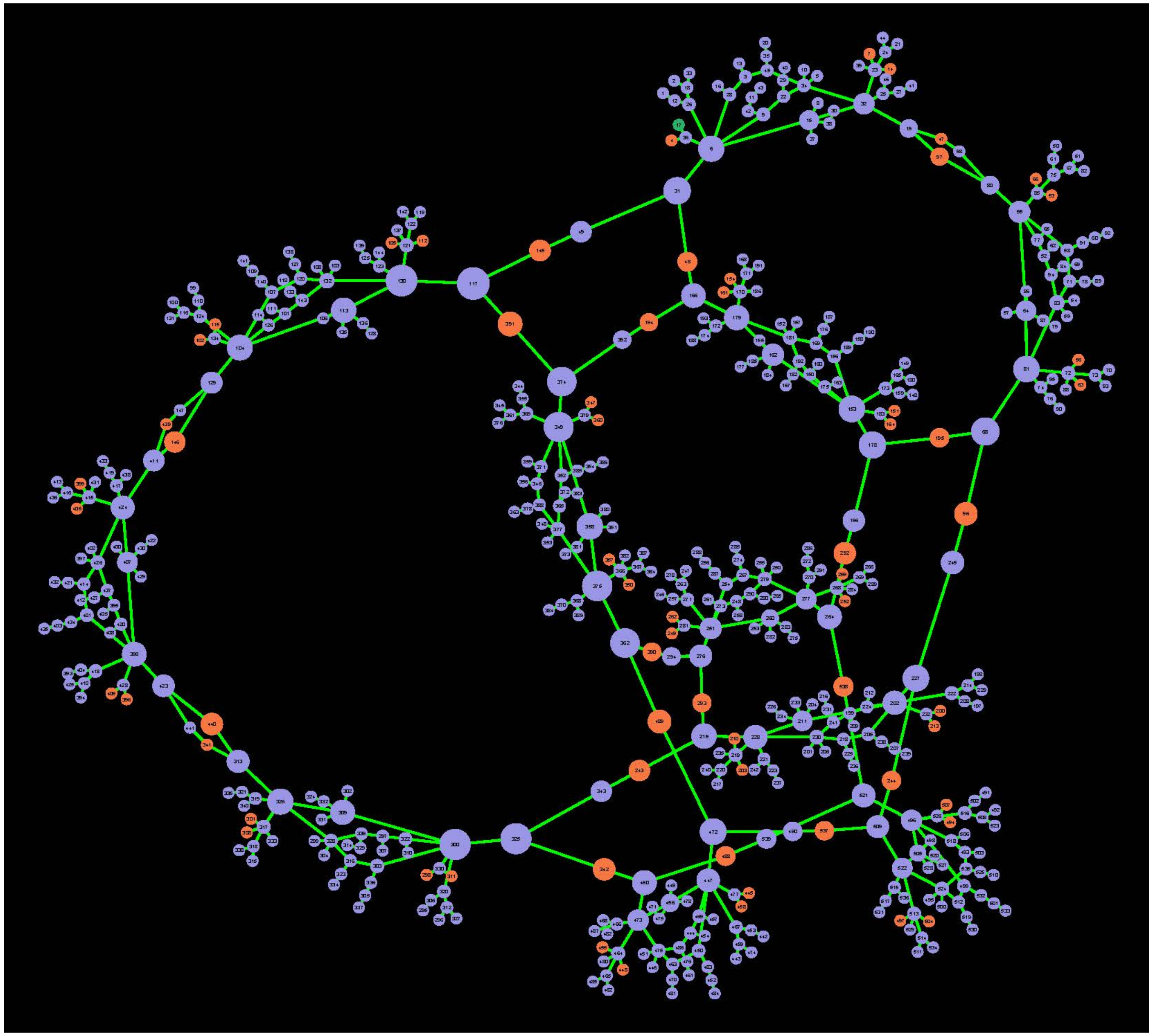
Curation Process



Achievements So Far

- Key modules ready
  - SDET framework in C++
  - PTI file parser, v33
  - Fragment creation code in python
  - Topology creation code
  - Fragment reconnection
  - Validation module through PSSE
  - Creation of the generator cost curves
- A few power system models with ~500 buses
  - Good convergence
  - Meeting metrics requirements
- A few power system models with ~3000 buses
  - Good convergence
  - Meeting metrics requirements

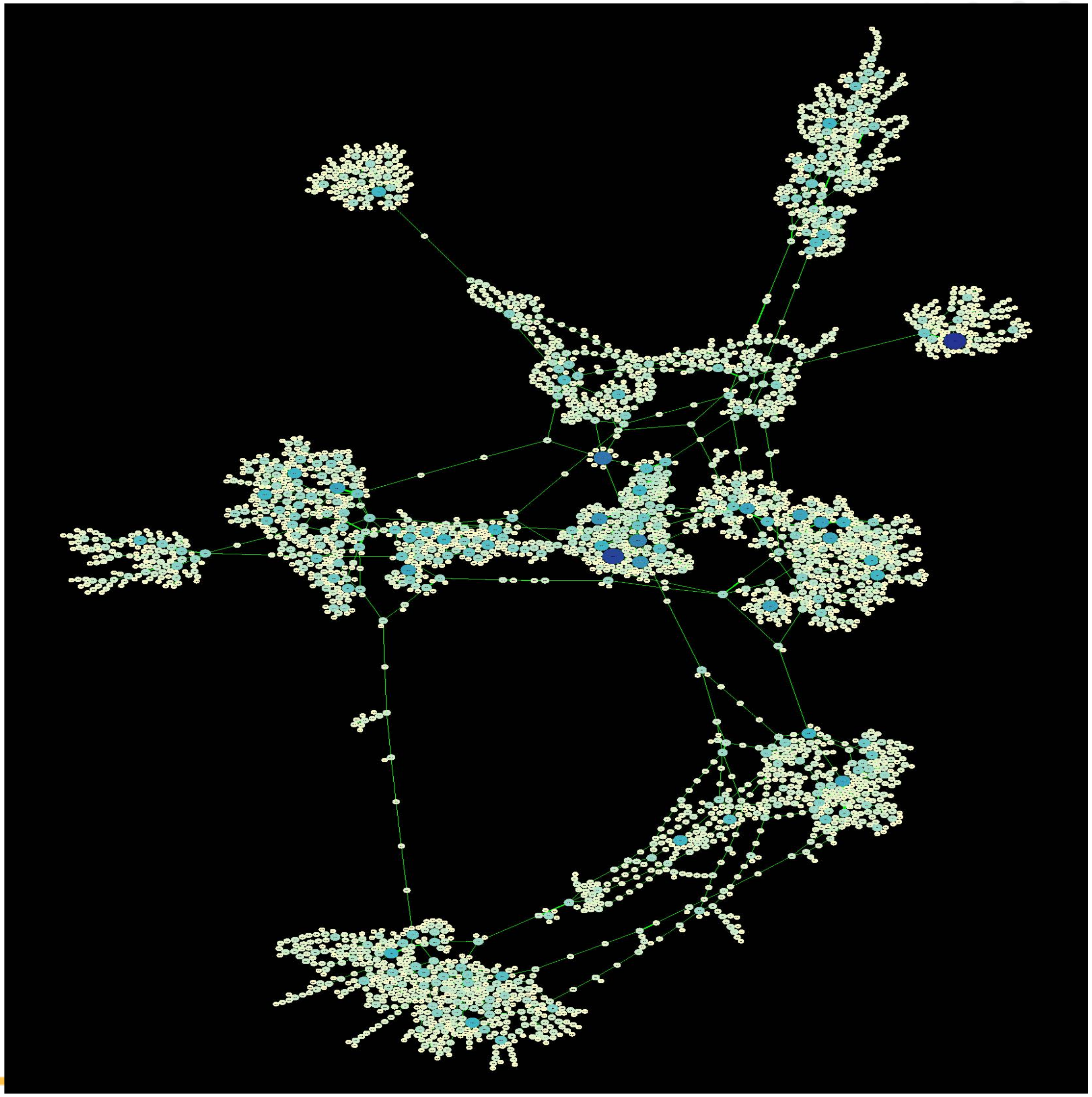
500 bus model generated



- 10 real-world fragments
- 45 tie lines created
- 528 buses
- 66 generators
- 6.3 GW of load

3000 bus model generated

- 21 real-world fragments
- 116 tie lines created
- ~3000 buses
- ~500 generators
- 36.5 GW of load



Conclusions

- Making datasets evolving is important to keep up with grid development and enable technology advancement
- Delivering datasets is important, but delivering data creation tools can enable data ***evolution***
  - Topology generation tool
  - Parameter population tool
  - Data anonymization tool
- Datasets and data creation tools are to be shared through GRID DATA repositories and professional communities

Questions?