Overpowered

Connecting Energy to the Grid, Faster

The Team

UC Berkeley Masters of Information and Data Science

Paul Cooper

Hailee Schuele

Zhifei Dong

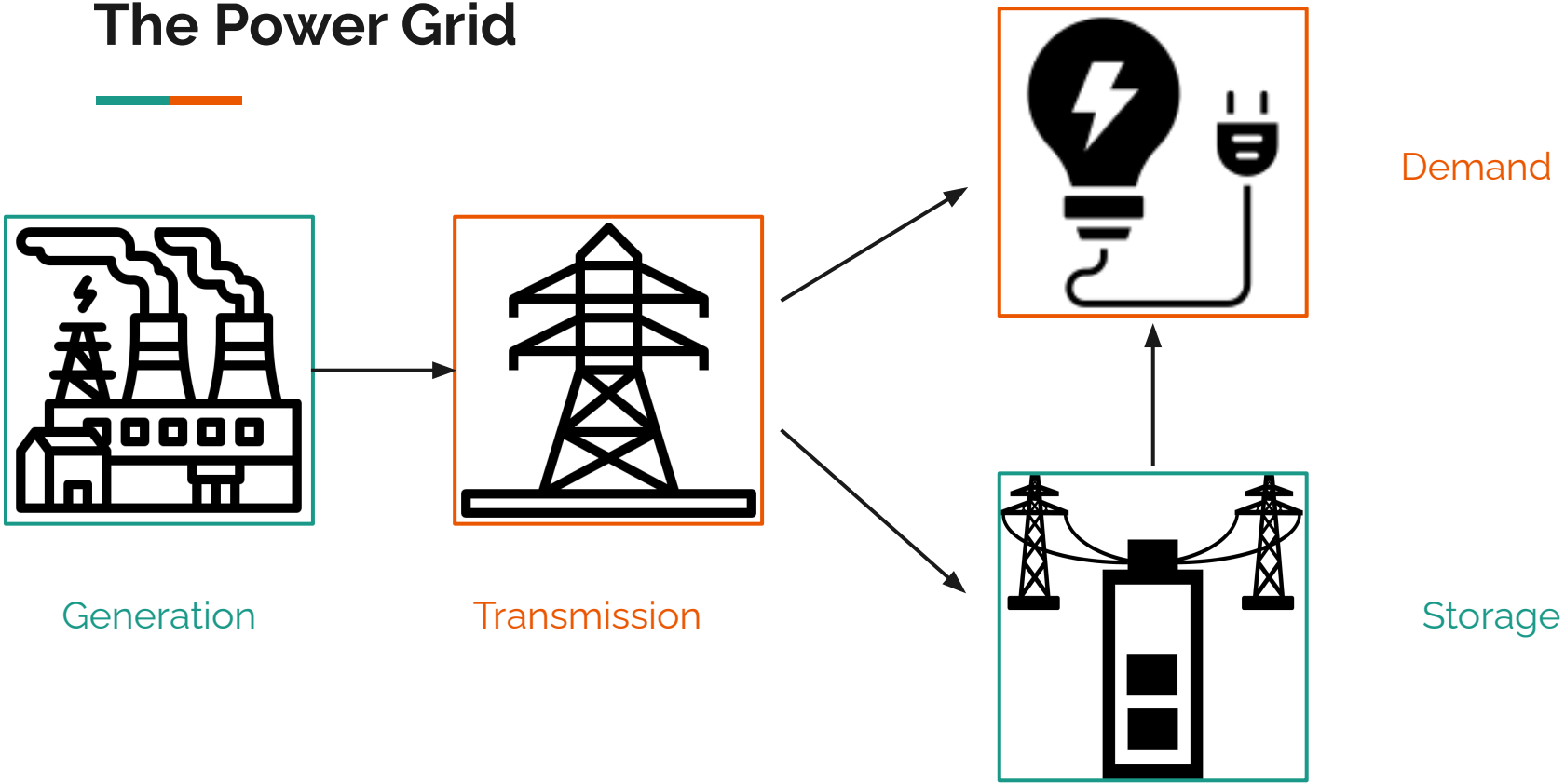
Adam Kreitzman

Our Mission

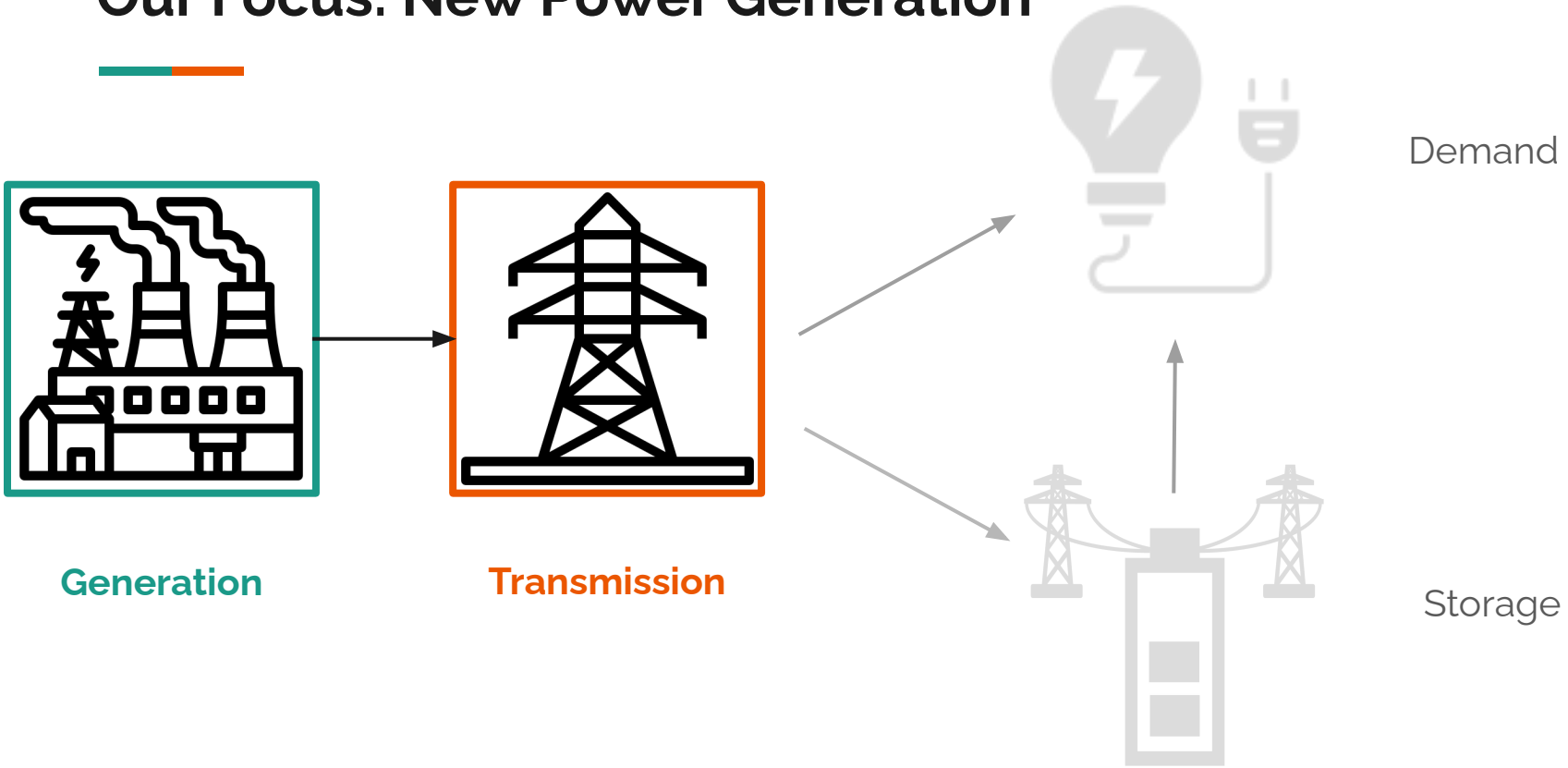
Provide a faster, more efficient way of connecting new energy sources to the power grid.

The Space

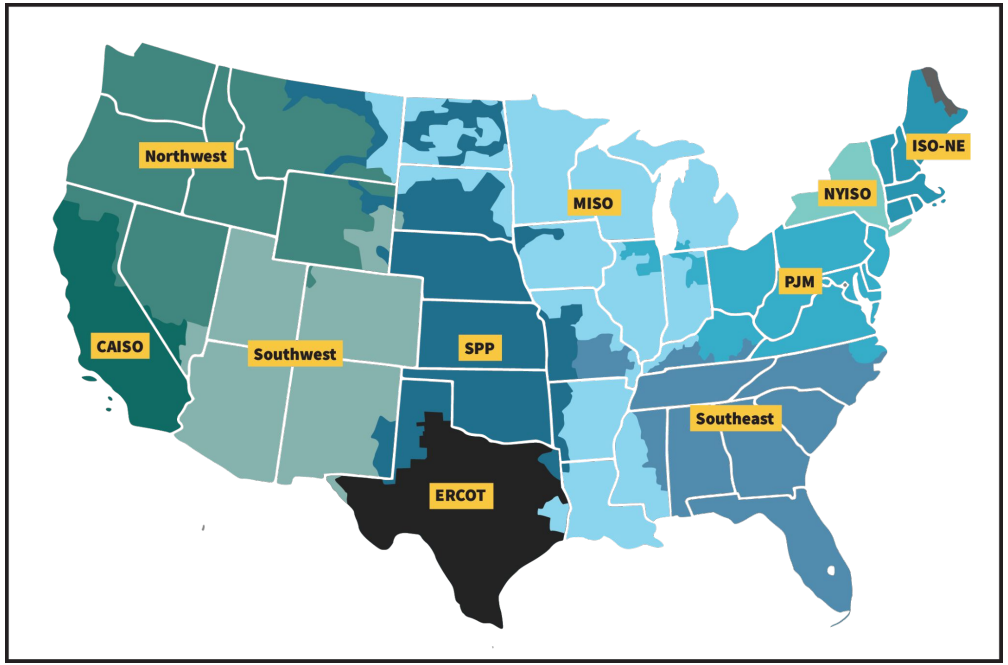
The Power Grid



Our Focus: New Power Generation



Regulatory Bodies



Geographic overview of RTOs and ISO. Source: FERC.gov

The Interconnection Queue

When a new power generation facility wants to come online, they submit an application to the **interconnection queue**

The Approval Process

- Each interconnection request is subject to **long, technical studies**
  - transmission capacity
  - market conditions
  - incident analysis
  - feasibility and impact
- One request may be **dependent** on other requests
- **Additional infrastructure investment** is often needed to accommodate a new power generation source

The Problem

Time, Money, and Dropout

1. 4+ year **wait times**
2. Unfeasible **infrastructure costs**
3. Wasted **time and resources**

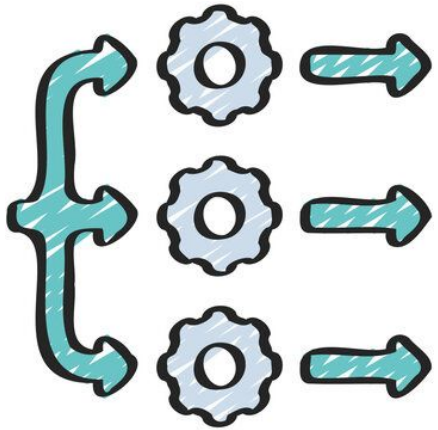
The Power Grid's Big Bottleneck

The queue process is outdated, inefficient, and expensive

The Solution

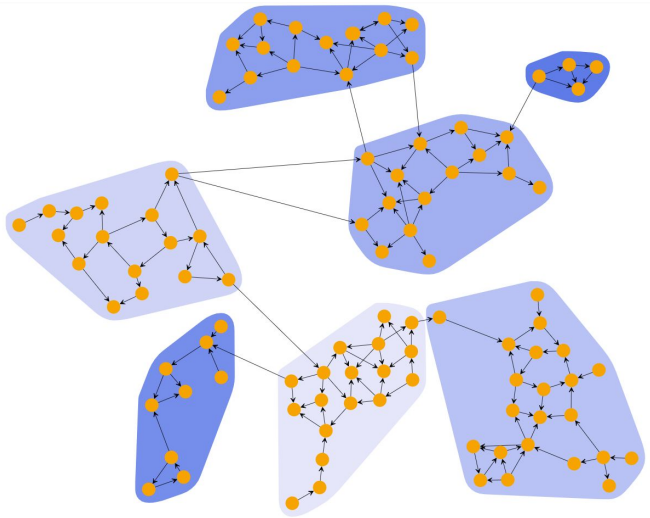
Cluster Processing

- New federal ruling allows for processing multiple applications together as a **cluster**
- Clusters allow:
  - Shared development investment
  - Reduced queue wait times
  - Fewer withdrawn applications



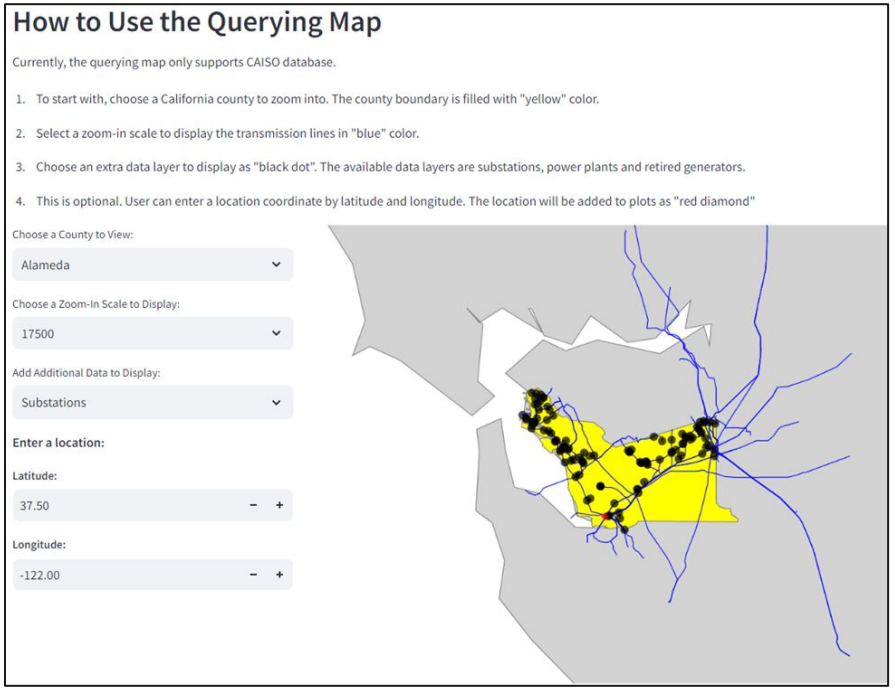
Improving Clusters

- If we can:
  - Identify efficient clusters
  - Understand how likely a given cluster is to be successful
- Then we can:
  - Speed up the queue
  - Reduce dropout
  - Save time and resources



Our Product

- Recommend **clusters of interconnection requests** to be studied together
- Provide **scoring mechanisms** to understand how strong the candidate batch is
- Provide tools for expert users to **customize their results**
- Provide **an integrated platform** for data querying, visualization and decision support



The Data

Piecing the Data Together

- CAISO interconnection queue data
- California substation GIS data
- California transmission line GIS data
- California power plant GIS data (including retired ones)
- CAISO future transmission projects
- *PowerWorld* - Power flow simulation software used by ISOs
- Simulated line capacity and load

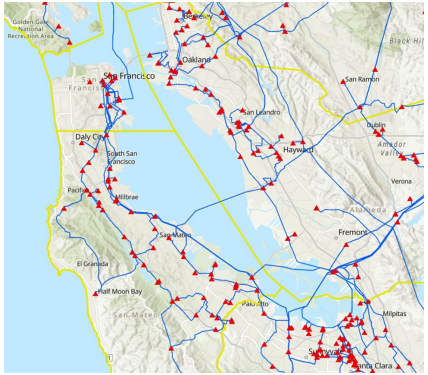
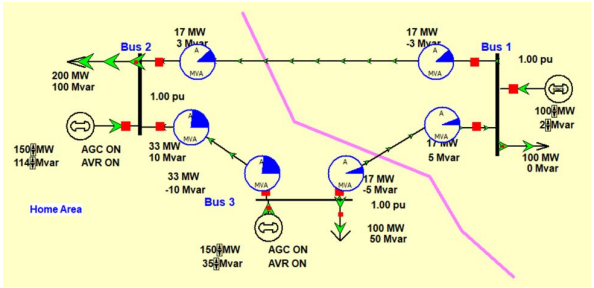
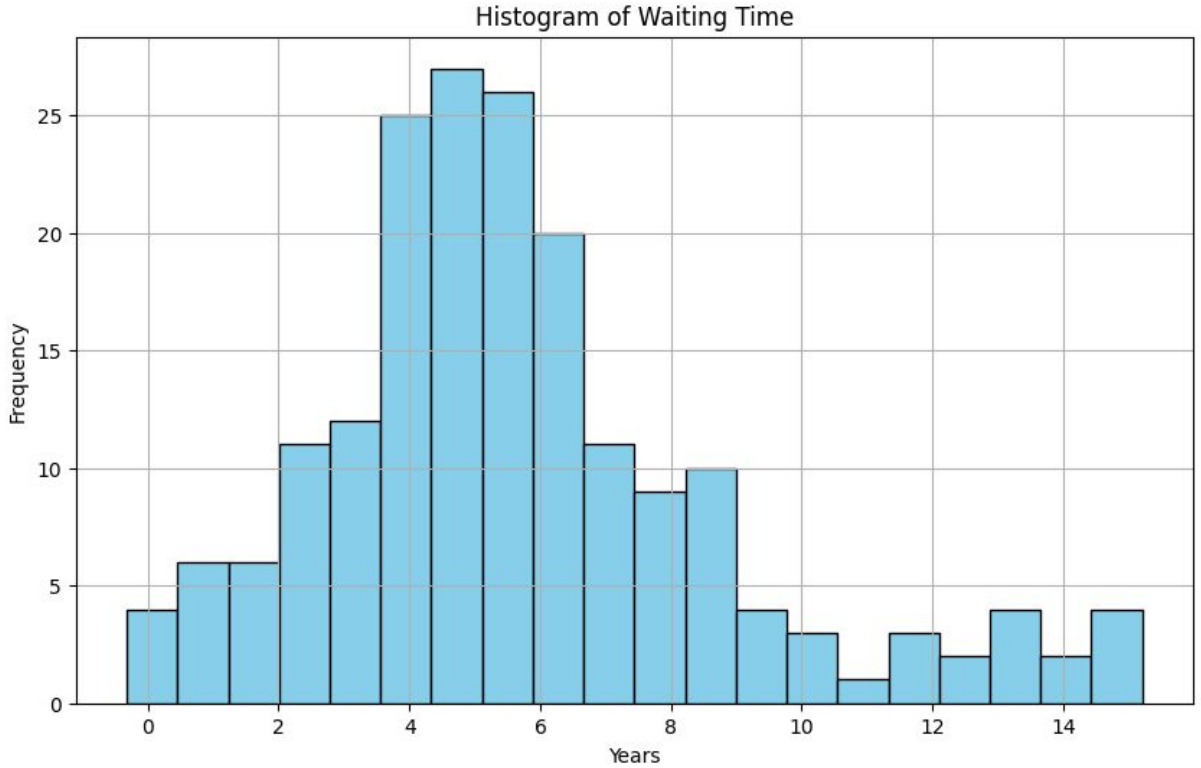


Table 8.2-1: New Reliability Projects Found to be needed

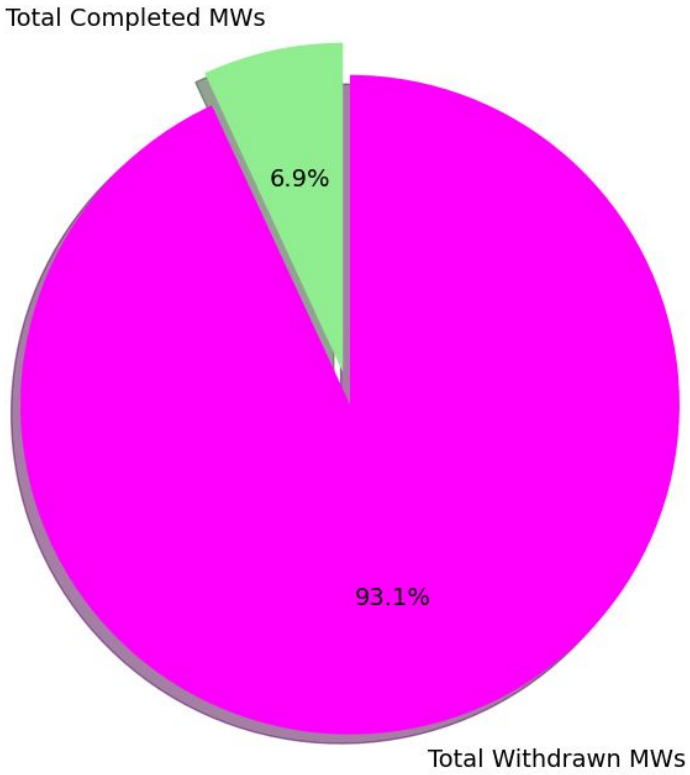
No.	Project Name	Service Area	Expected In-Service Date	Project Cost (in millions of dollars)
1	Garberville area reinforcement project	Humboldt	3032	204
2	Tulucay-Napa #2 60 kV line Reconductoring project	NCNB	2028	4.6
3	Santa Rosa 115 kV lines Reconductoring project	NCNB	2028	74
4	Tesla 115 kV Bus Reconfiguration Project	CVLY	2030	55
5	Banta 60 kV Bus Voltage Conversion	CVLY	2024	17.5
6	Metcalf 230/115 kV Transformers Circuit Breaker Addition	GBA	2026	15
7	South Bay Area Limiting Elements Upgrade	GBA	2027	11
8	Redwood City Area 115 kV System Reinforcement	GBA	2030	110.8
9	Lone Tree – Cayetano – Newark Corridor Series Compensation	GBA	2027	25
10	Pittsburg 115 kV Bus Reactor project	GBA	2032	26
11	Equipment Upgrade at CCSF Owned Warnerville 230 kV Substation	Fresno	2024	1.6
12	Los Banos 70 kV Area Reinforcement Project	Fresno	2029	60
13	Los Banos 230 kV Circuit Breaker Replacement	Fresno	2032	66
14	Panoche 115 kV Circuit Breaker Replacement and 230 kV Bus Upgrade project	Fresno	2032	184
15	North East Kern 115 kV Line Reconductoring Project	Kern	2032	296
16	Mesa 230/115 kV spare transformer	CCLP	2032	50.48



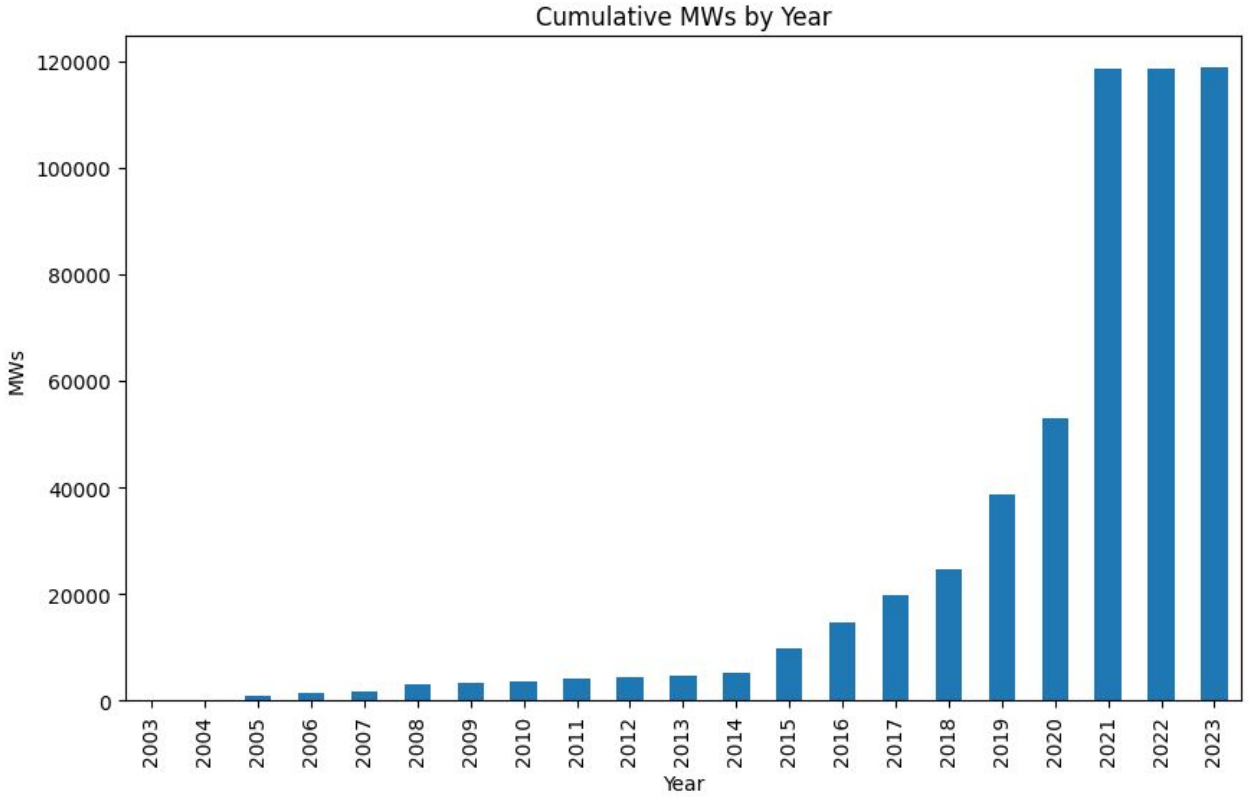
Key Data Takeaways - Completed Queues



Key Data Takeaways - Completed vs. Withdrawn



Key Data Takeaways - Current Queue



The Technical Approach

Concept Abstraction & Operationalization

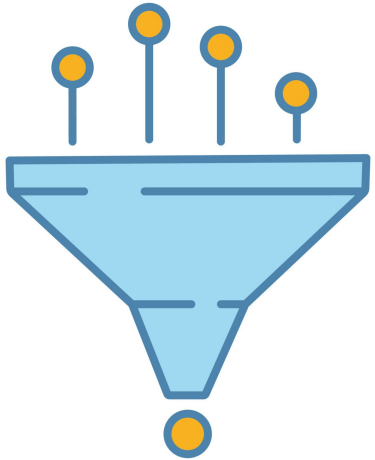
Concept	Variables
Location	• Geolocation of Interconnection Point • Proximity to Existing Infrastructure (retired plants) • Proximity to Planned Infrastructure
Infrastructure	• Included Storage • Transmission Availability • Generator Type(s) • Amount of Energy Output
Process	• Permit Status • Expected On-line Date • Queue Position • Utility company

Output Variables

Measurement	Meaning
Likelihood Scores	• Likelihood of success based on features learned from historical data
Cluster Strength	• The strength of the recommended cluster as a whole using aggregated similarity calculations
Total MegaWatts	• The total MWs provided to the grid by the cluster minus the current available transmission capacity

Data Pipeline

- Databricks & Blob Storage
- Feature Selection/Engineering
  - SME feedback
  - Inference from past interconnection requests
- Preprocessing for Machine Learning:
  - VectorAssembler + Scalers
  - Generating a numeric representation



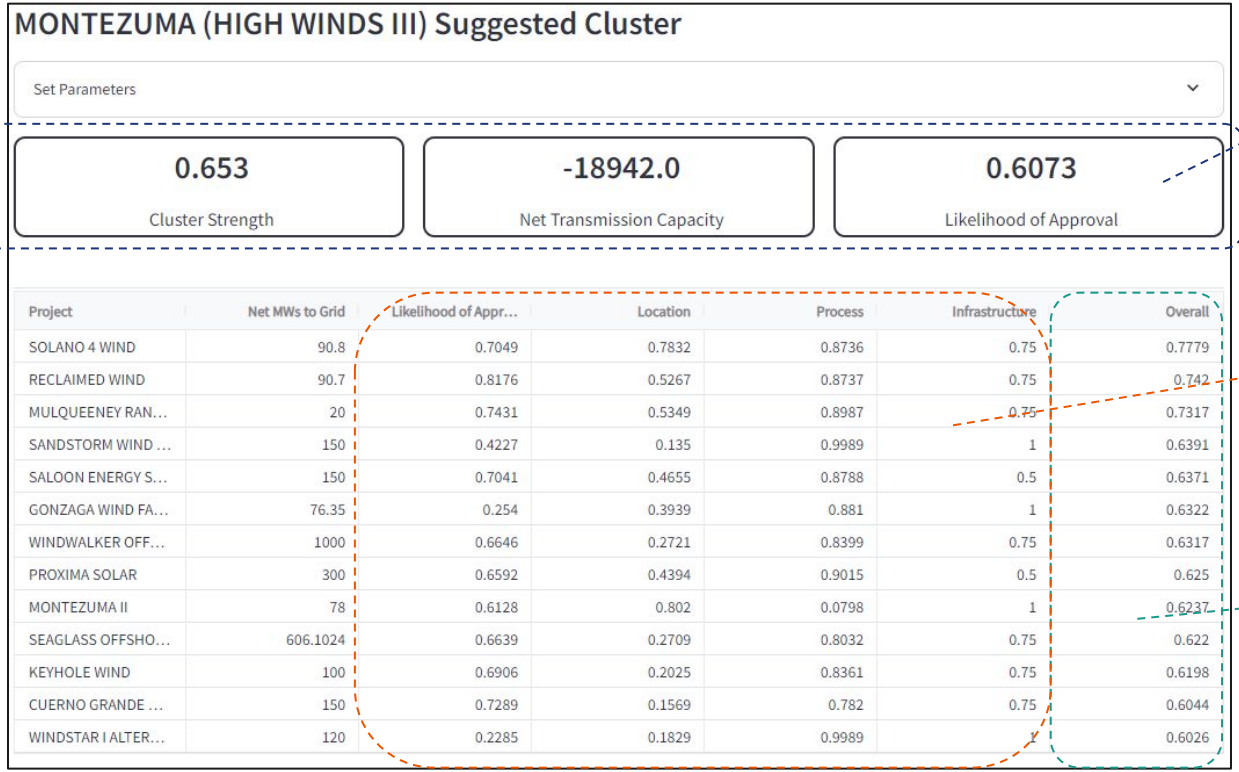
Likelihood of Approval

- A method of judging the strength of an application based on supervised learning past data and future grid infrastructure development.
- Combines coefficients from supervised learning with proximity to retired power plants and future transmission projects to gauge relative strength of a grid applicant.
- Allows a decision maker to consider the probability of application success in cluster-building process.

Custom Clustering Algorithm

- Crafting a similarity algorithm that involves group subsets + process subtleties – not a simple or intuitive process to create.
- Variable-centric similarity inspired by Cosine Similarity.
- Custom similarity algorithms built to handle all variables.
- Only considering projects downstream for each project (i.e in a FIFO approach, what projects could be added to this one to form a cluster, maintains fair process)

Sample Results



MONTEZUMA (HIGH WINDS III) Suggested Cluster

Set Parameters

0.653	-18942.0	0.6073
Cluster Strength	Net Transmission Capacity	Likelihood of Approval

Project	Net MWs to Grid	Likelihood of Appr...	Location	Process	Infrastructure	Overall
SOLANO 4 WIND	90.8	0.7049	0.7832	0.8736	0.75	0.7779
RECLAIMED WIND	90.7	0.8176	0.5267	0.8737	0.75	0.742
MULQUEENEY RAN...	20	0.7431	0.5349	0.8987	0.75	0.7317
SANDSTORM WIND ...	150	0.4227	0.135	0.9989	1	0.6391
SALOON ENERGY S...	150	0.7041	0.4655	0.8788	0.5	0.6371
GONZAGA WIND FA...	76.35	0.254	0.3939	0.881	1	0.6322
WINDWALKER OFF...	1000	0.6646	0.2721	0.8399	0.75	0.6317
PROXIMA SOLAR	300	0.6592	0.4394	0.9015	0.5	0.625
MONTEZUMA II	78	0.6128	0.802	0.0798	1	0.6237
SEAGLASS OFFSHO...	606.1024	0.6639	0.2709	0.8032	0.75	0.622
KEYHOLE WIND	100	0.6906	0.2025	0.8361	0.75	0.6198
CUERNO GRANDE ...	150	0.7289	0.1569	0.782	0.75	0.6044
WINDSTAR I ALTER...	120	0.2285	0.1829	0.9989	1	0.6026

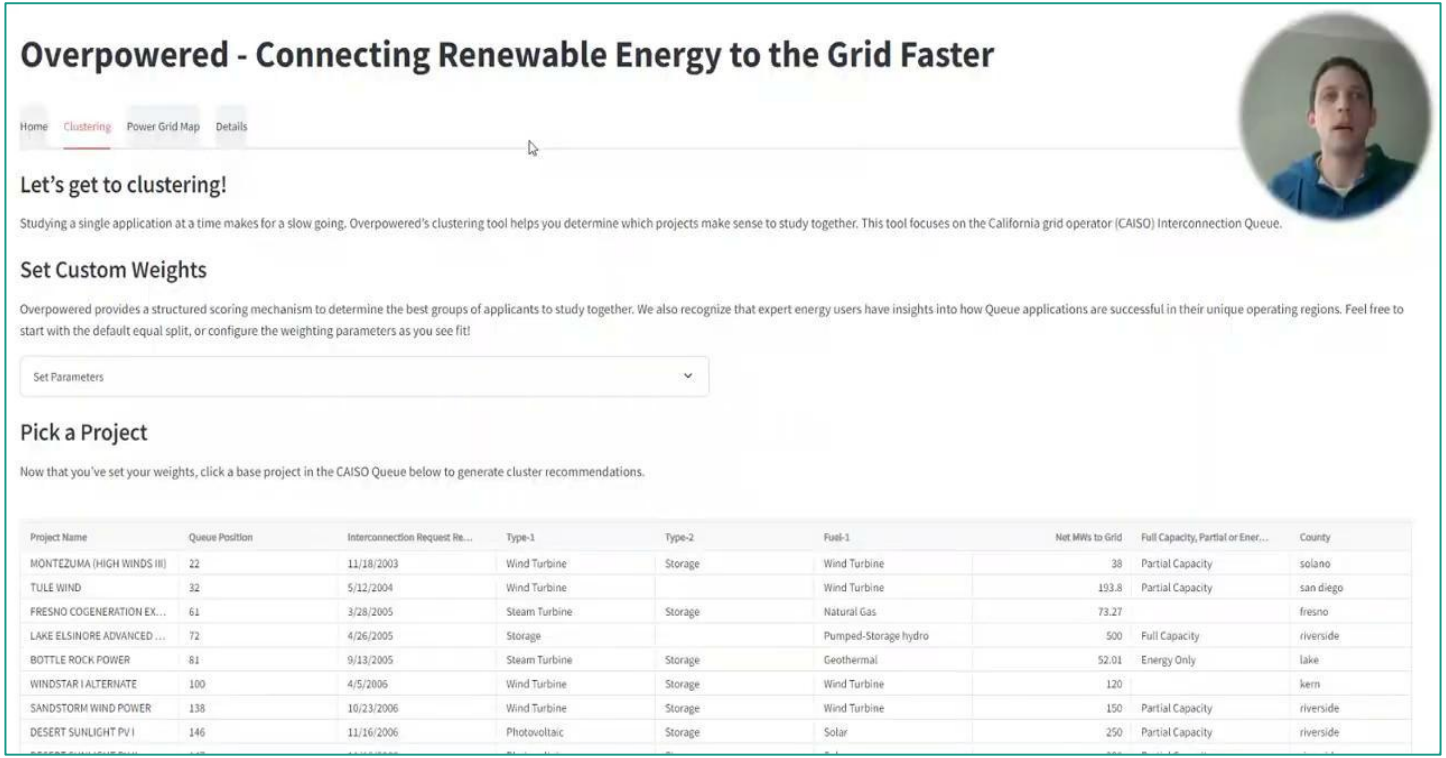
Cluster Metrics

Component Scores

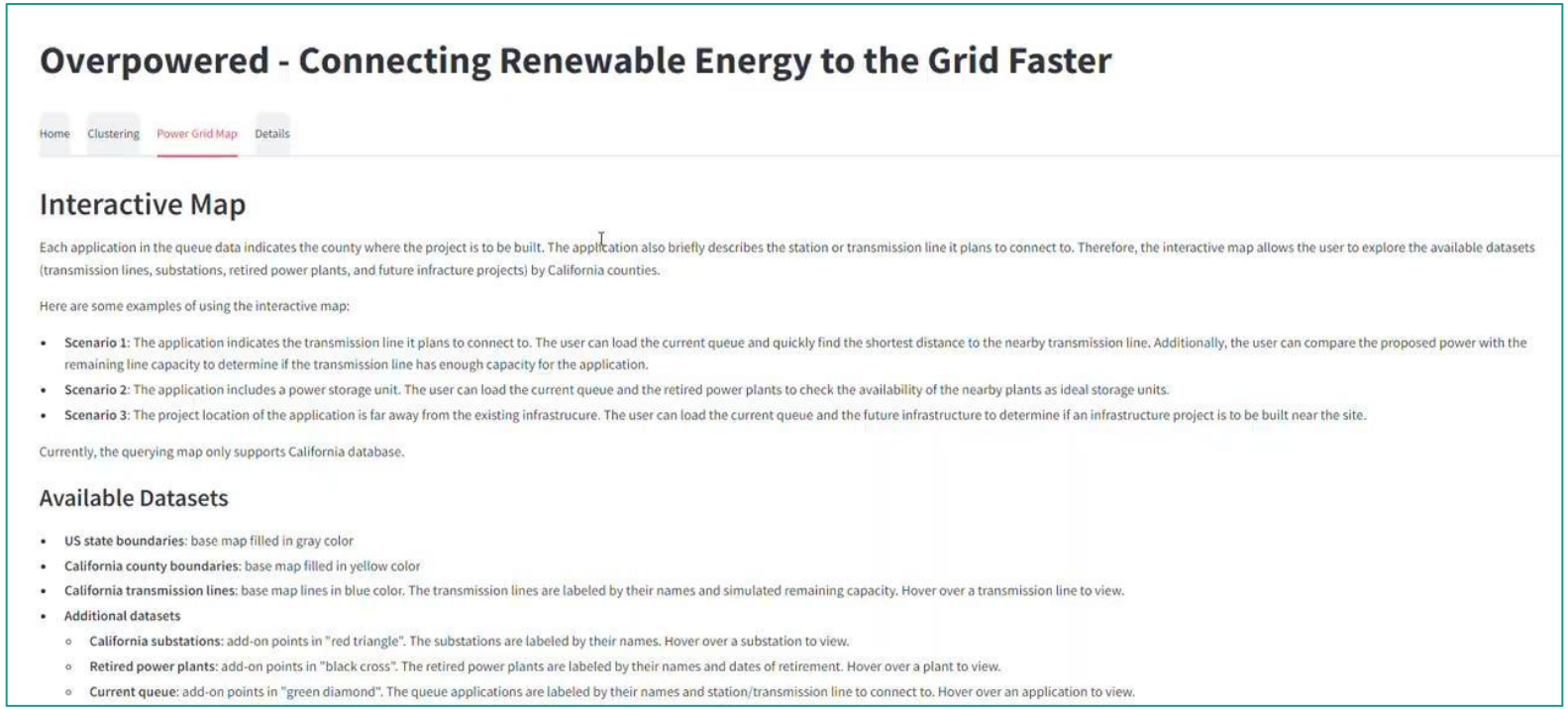
Aggregate Score

The Demo

Demo - Clustering Tool



Demo - Interactive Map



Key Takeaways

TULE WIND Suggested Cluster

Set Parameters

Assign relative weights

Location: 4

Infrastructure: 2

Process: 1

Likelihood of Approval: 3

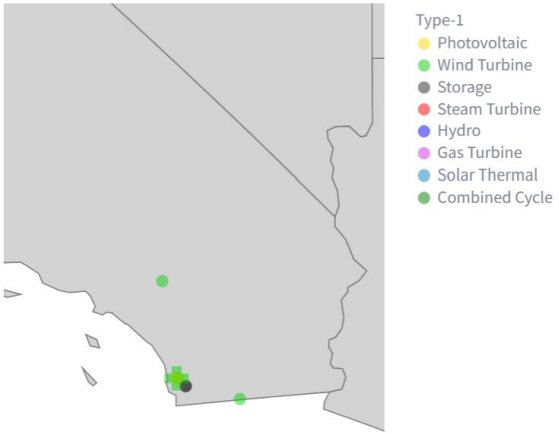
Rerun

0.6676	-18709.1	0.7159
Cluster Strength	Net Transmission Capacity	Likelihood of Approval

Project	Net MWs to Grid	Likelihood of Approval	Location	Pro
MOUNT LAGUNA WIND 2	400	0.7134	0.8203	0.8
RUGGED SOLAR FARM	71.88	0.718	0.8368	0.3
STARLIGHT SOLAR	20	0.6611	0.837	0.8

Choose a Zoom-In Scale to Display:

3000

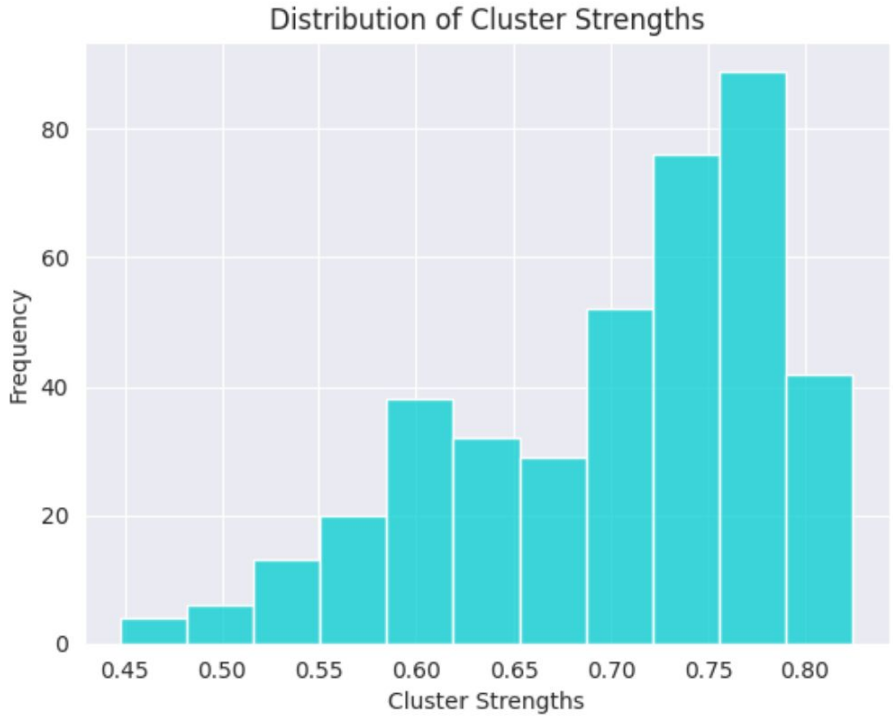


The Evaluation

Measuring Impact

- Our ultimate goal is to **improve queue efficiency**
- We want to measure:
  - Cluster strength
    - How compatible are the projects within a cluster?
    - How likely is the cluster to be approved?
  - Improved wait times
    - Are we speeding up the queue?

The Tool's "Goodness"



SME Feedback

“The old way has its limitations and everyone complains about it”

“When we’re trying to advise clients, the want to know where the best place to connect would be… and we really don’t know”

“Having historical data boiled down is the value of your tool. You don’t have to rerun the study every time (you assess a new applicant)”

The Overpowered Value

Breaking Down a Complex Problem

- Technical Challenges
  - Piecing together various data sources of different types
  - Calculating unknown variables such as transmission line capacity and usage load
- Human-Centric Challenges
  - Abstracting technical and human-centric process considerations
  - How to allow tool flexibility in an inexact problem-space
  - Machine learning expertise in an antiquated system

Roadmap

- Extend the application to developers
- Incorporate additional data points that grid operators use
- More SME feedback
- Additional data outside of California
- Incorporate outcomes from initial cluster studies
- More supervised ML-focused approaches to grid expansion built on top

The Mission

Impact

Provide a faster, more efficient way of connecting new energy sources to the power grid.

Acknowledgements

- Damian Berger (SME at Atwell Group)
- Annie Polakowski (SME at Atwell Group)
- Dr. Roy McCann (Principal Investigator at GRAPES)
- Thomas Dempsey (SME at MYNO Carbon)
- CAISO Customer Support
- UC Berkeley MIDS Capstone class and professors

> *"Striking a balance between a slow and certain approach and considering all-comers fairly is really the heart of the solution."*
>
> - Damian Berger (SME at Atwell Group)