Project Ideas

# Background

A National Infrastructure for AI on the Grid (NI4AI) is an ARPA-E funded project that provides open access to electric grid measurement data. The goal of the project is to eliminate barriers to accessing and analyzing data that could help us improve and modernize the grid. The project is led by a startup called [PingThings](http://pingthings.io), and is built on an open access version of their platform which offers a data visualization interface and API access via Python, Julia and Go.

We are looking for student teams to explore un-solved analytical challenges that the industry is facing. Successful projects will be featured on the ni4ai blog, and will have opportunities to connect with industry stakeholders about their work. Interested students should let us know by filling out [this google form](https://forms.gle/W1HYgXXNjBMT7ttc8).

Have questions? Looking for ideas? Send an email to [laurelndunn@gmail.com](mailto:laurelndunn@gmail.com) with [NI4AI] in the subject line.

# Sample Projects

### Real-Time Analysis of National Grid Data

We are installing sensors throughout the country to stream real-time voltage data from measurement sites across the country. The data can be used to study how the frequency of the grid changes across different sites, or to study phase angle differences indicative of macro-scale power flows. Look for sensors in the **ni4ai** and **texas\_pmus** data collections.

[This video](https://www.youtube.com/watch?v=bdBB4byrZ6U) shows how a frequency deviation caused by a generator trip in Florida made waves throughout the entire Eastern United States. For more ideas about how to analyze macro-scale energy trends, read through the The [Texas Synchrophasor Network reports](https://web.ecs.baylor.edu/faculty/grady/texas_synchrophasor_network.html). You’ll find several years of reports, including information about what happened on the grid during the recent ice storm in Texas, and during the Superbowl.

### Fault Classification

High-frequency sensors on the grid automatically record voltage and current measurements during electrical faults. Faults can be caused by a number of issues -- such as lightning, equipment malfunction, or external contact with animals or trees. If left unaddressed, some of these issues (particularly equipment malfunction and tree contact) can pose a wildfire risk.

Today, to determine the cause of a fault requires sending a field engineer out to look. This is expensive and time consuming, and isn’t guaranteed to locate the problem -- if there is one. The **PQdata/epri** collection has voltage and current measurements for hundreds of faults, and metadata about what caused the fault.

Can you develop a supervised learning algorithm for determining the cause of a fault from voltage and current measurements?

### Phasor Calculation

Phasors are a compressed representation of waveform data on the grid. You can learn more about phasors on our blog posts [What’s in a phasor](https://blog.ni4ai.org/post/2021-01-08-whats-in-a-phasor/) and What’s the angle [Part 1](https://blog.ni4ai.org/post/2020-07-30-what-is-the-angle/) and [Part 2](https://blog.ni4ai.org/post/2020-07-29-what-is-the-angle-part-2/).

Phasors can be calculated using curve fitting, fourier transforms, and other techniques. These different approaches have implications on information loss, and on downstream applications for the data. Here, we encourage you to explore different algorithms for calculating phasors. What metrics would you use to compare algorithms? How much historical data is needed to calculate the phasor? What is the computational burden?

# NI4AI Links

* Get a login: ni4ai.org
* [Blog](http://blog.ni4ai.org)
* [Github tutorials](https://github.com/PingThingsIO/ni4ai-notebooks/tree/main/tutorials)
* [Platform demo](https://blog.ni4ai.org/post/2020-11-15-demo-video/)
* [Data sets](https://blog.ni4ai.org/post/2100-01-01-datasets/)
* [Workshop videos](https://blog.ni4ai.org/post/2020-10-31-workshops/)

# Other Energy Data Sets

**OE 417 data**. This is a survey on outages, cyber threats, and other types of emergency events. Huge shout out to whoever cleaned up this dataset and put it into a table.

clean data table <http://insideenergy.org/2014/08/18/data-explore-15-years-of-power-outages/>

raw data <https://www.oe.netl.doe.gov/oe417.aspx>

**EIA 930 data**. This survey tracks hourly load by balancing authority area. Might be pretty cool to dig into. I want to say they tracked interchange as well? I'm not sure!

real-time data map <https://www.eia.gov/realtime_grid/#/status?end=20201020T16>

There used to be archived data available too. Might take some poking around to find it. Maybe it's even somewhere in this API? <https://www.eia.gov/opendata/>

**CAISO local marginal price data.** This data is stored in a system called OASIS which is notoriously difficult to use. That said, there's lots of interesting stuff to look. It's also possible that they've made the portal easier to use.

<http://www.caiso.com/TodaysOutlook/Pages/prices.html>

**Wildfire mitigation data release**. This data was just recently released and gives the location, age, and failure record for grid assets -- poles, transformers, not sure what else. It also gives the location of ignitions and wire down incidents.

<https://www.cpuc.ca.gov/wildfiremitigationplans/>