An Overview of Tools and Data for AI

Prepared By: Theo Laughner

Presentation Outline

Artificial Intelligence

Definition
Applications
Failures
Steps/Requirements
Resources
Conclusions

AI

Definition

According to the father of Artificial Intelligence, John McCarthy, it is *"The science and engineering of making intelligent machines, especially intelligent computer programs"*.

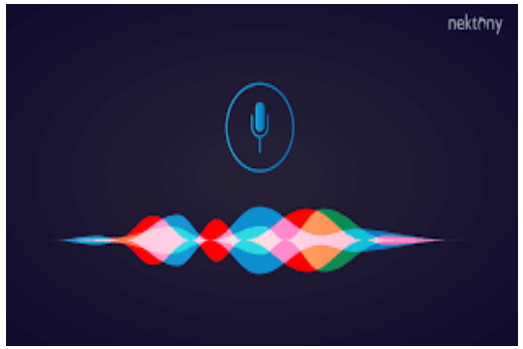
Artificial Intelligence is a way of **making a computer, a computer-controlled robot, or a software think intelligently**, in the similar manner the intelligent humans think.

Source: https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm

AI

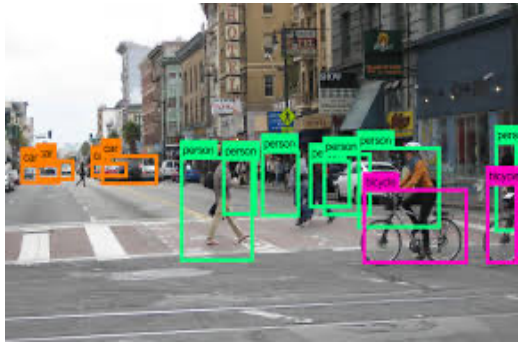
Generic Applications

Gaming



Natural Language Processing

Handwriting Recognition



Computer Vision

Robotics

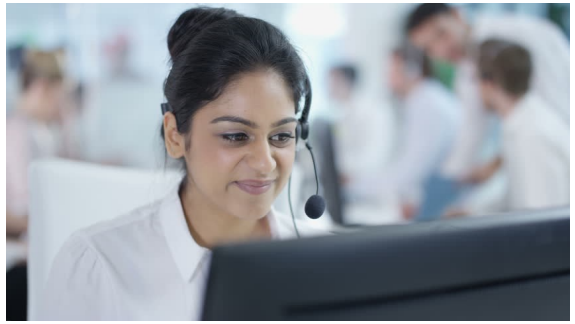
AI

Utility Applications

Asset Optimization

Operations

Maintenance



Customer Service

Cybersecurity

AI

Failures

- Amazon AI Recruitment
- OpenAI's Samantha
- Google AI Diabetic Retinopathy Detection
- Microsoft's Tay Chatbot
- IBM's Watson

Training a chatbot on Twitter users' data is probably not the safest bet. In less than 24 hours, Microsoft's Tay, an AI chatbot, started making offensive and inflammatory tweets on its twitter account. Microsoft said that as the chatbot learns to talk in a conversational manner, it can get "casual and playful" while engaging with people.

Though the chatbot did not have a clear ideology as it garbled skewed opinions from all over the world, it still raised serious questions about biases in machine learning and resulted in Microsoft deleting its social profile and suggesting that they are going to make adjustments to it.

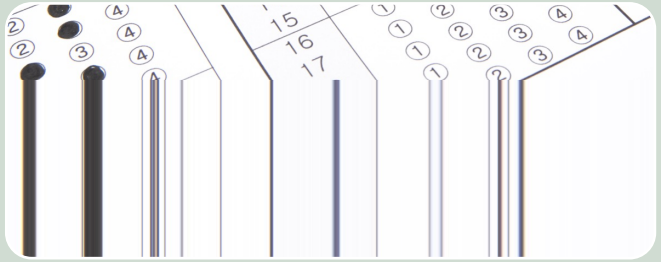
Source: https://analyticsindiamag.com/epic-ai-fails-a-list-of-failed-machine-learning-projects/#:~:text=Amazon's%20Rekognition,not%20properly%20testing%20the%20model.

AI

Steps / Requirements

Training

- High quality data
- Accurate Annotation



Validation

- Human Review of Model

Testing

- Give the model a dataset and see how it performs

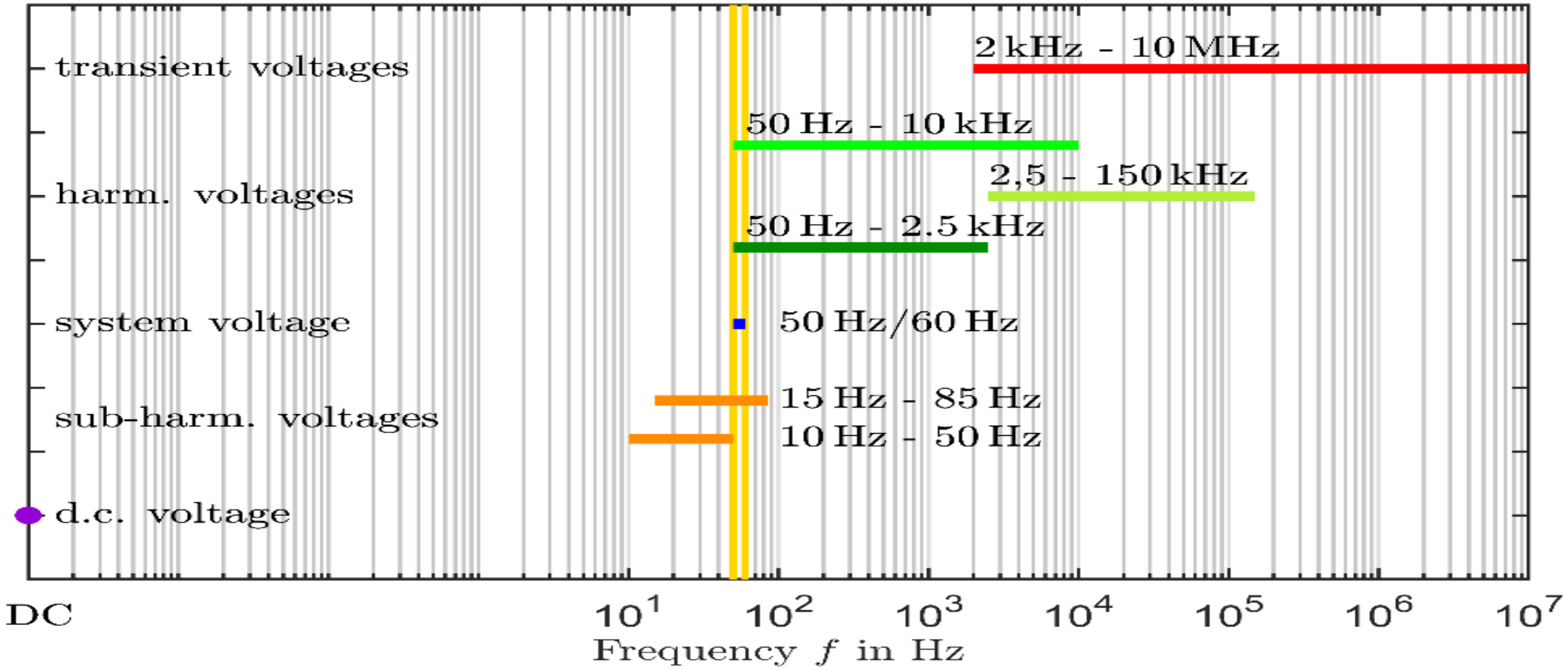
AI > Training

High Quality Data – Use Cases

Equipment	Voltage dips	Voltage swells	Harmonics	Interharmonics	Subharmonics	Supraharmonics	Slow voltage variations	Fast voltage variations	Transients	Voltage unbalance	Frequency variations	DC components
PV inverters	X											
Production units	X										X	
Active converters	X	X	X	X	X	X	X	X	X	X	X	X
LED lamps				X				X				
Power line communication						X			X			
Transformers						X			X			
Rotating machines						X			X			
Cable insulation						X						
Instrument transformers						X						
Three-phase converters										X		

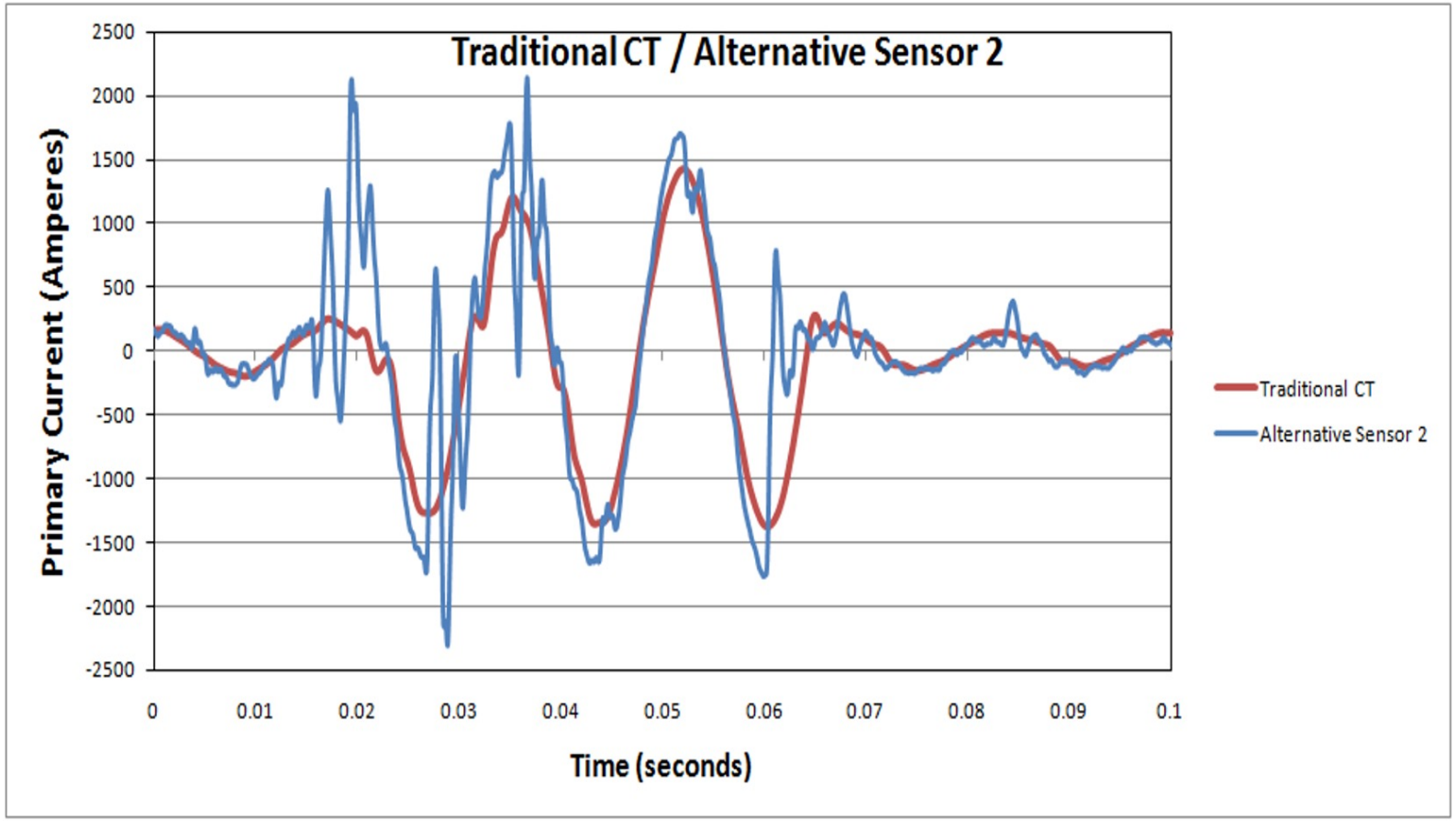
AI > Training

High Quality Data - Sensors



AI > Training

High Quality Data - Annotation



These two measurements were recorded at the same sampling rate.

However, one was recorded with a traditional CT.

One was recorded with a sensor that has better frequency response.

The blue trace more accurately captures the phenomenon.

AI

Resources for the Electric Power Industry

File Formats –

- PQDIF
- COMTRADE
- PQds

Data –

- EPRI Disturbance Library
- ORNL Disturbance Library

Tools –

- PingThings – NI4AI

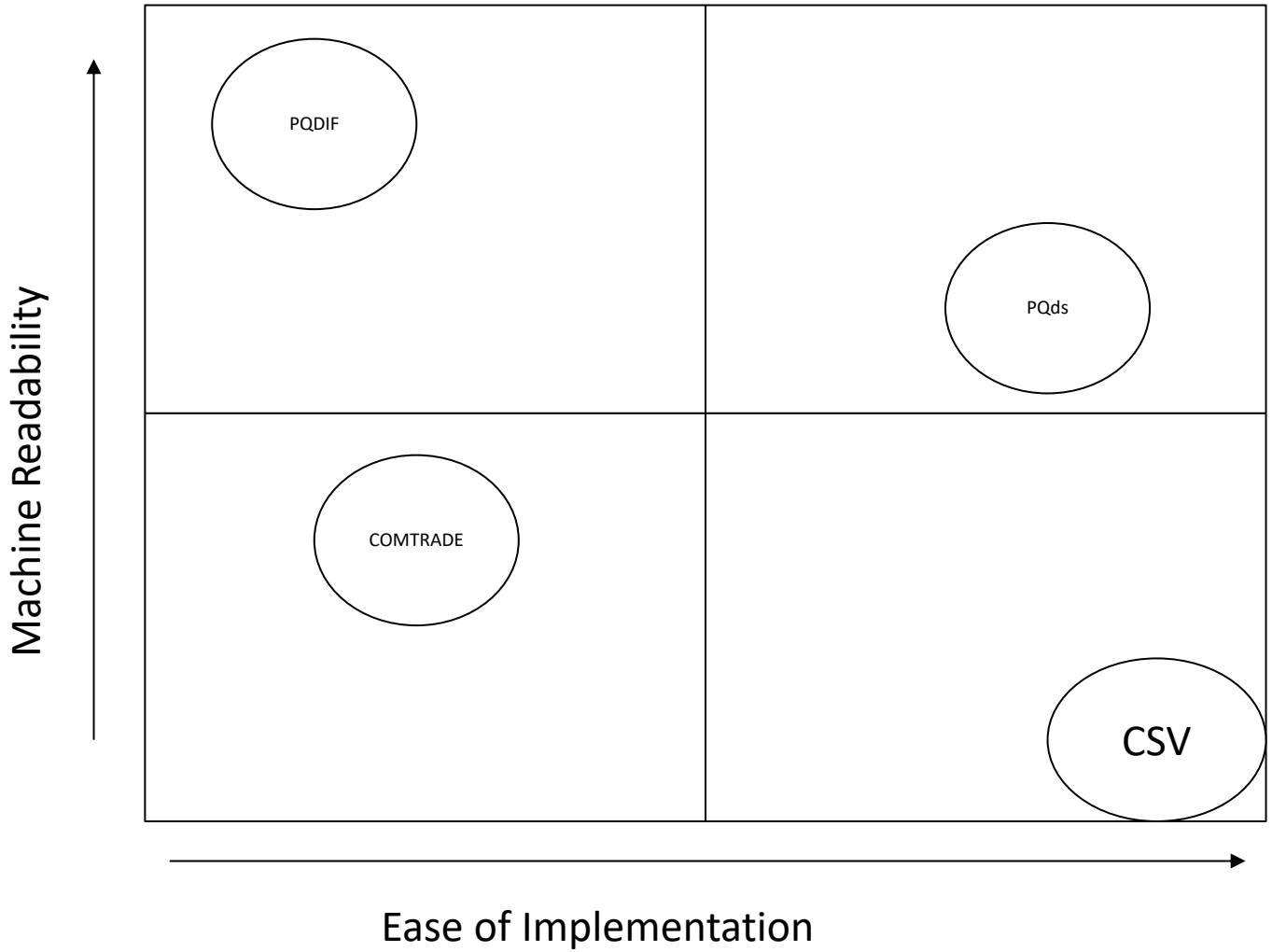
Benchmarks/Outage Info –

- DOE – Outage Data Initiative
- Grid Metrics – Power Outages
- Whisker Labs – CPQI

AI > Resources

File Formats

- **IEEE PQDIF (1159.3)**
- **IEEE COMTRADE (37.111)**
- **CSV (RFC 4180)**
- **PQds (In Development)**

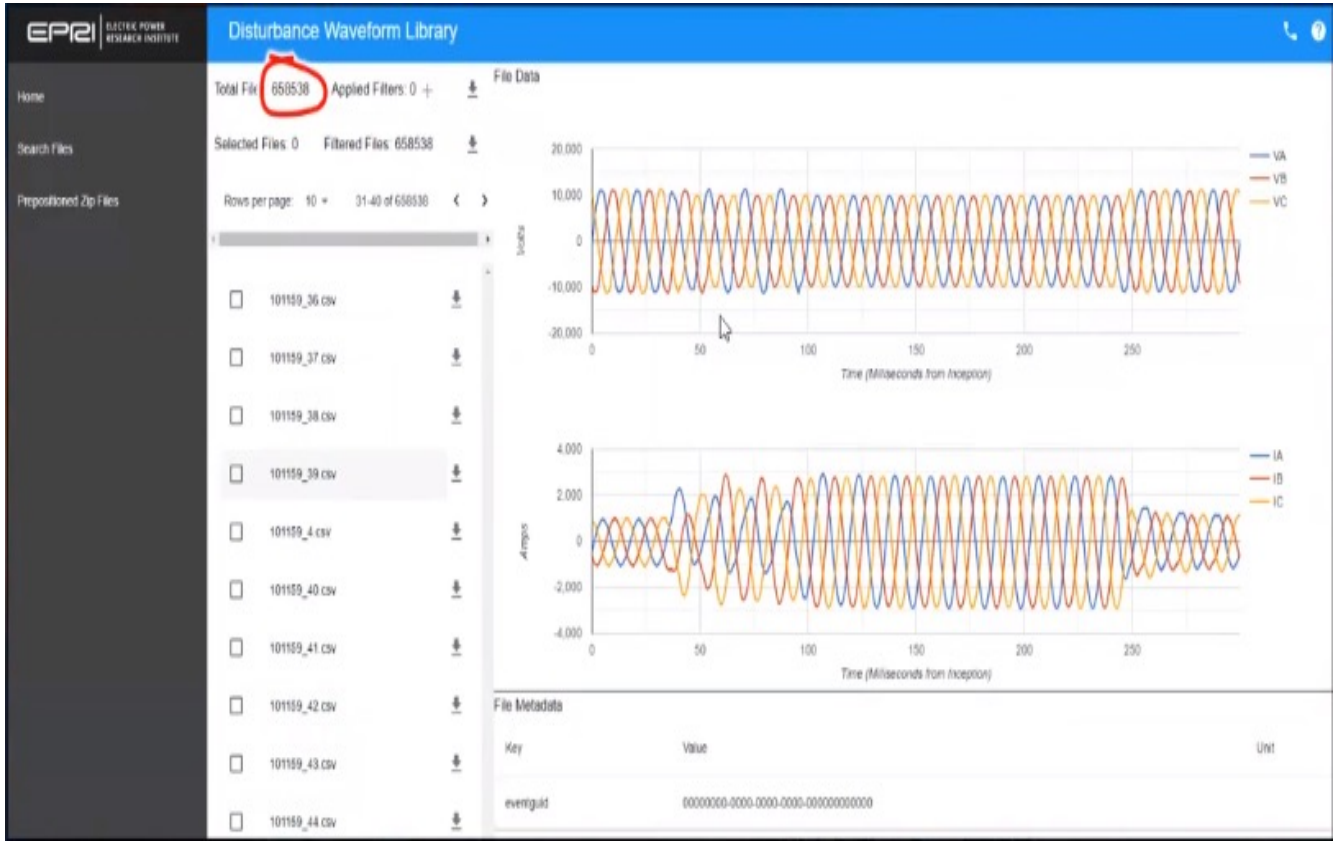


AI > Resources

Data

EPRI Disturbance Library

- Objective – Construct a large library of power quality disturbance data with and without metadata describing cause and event information.

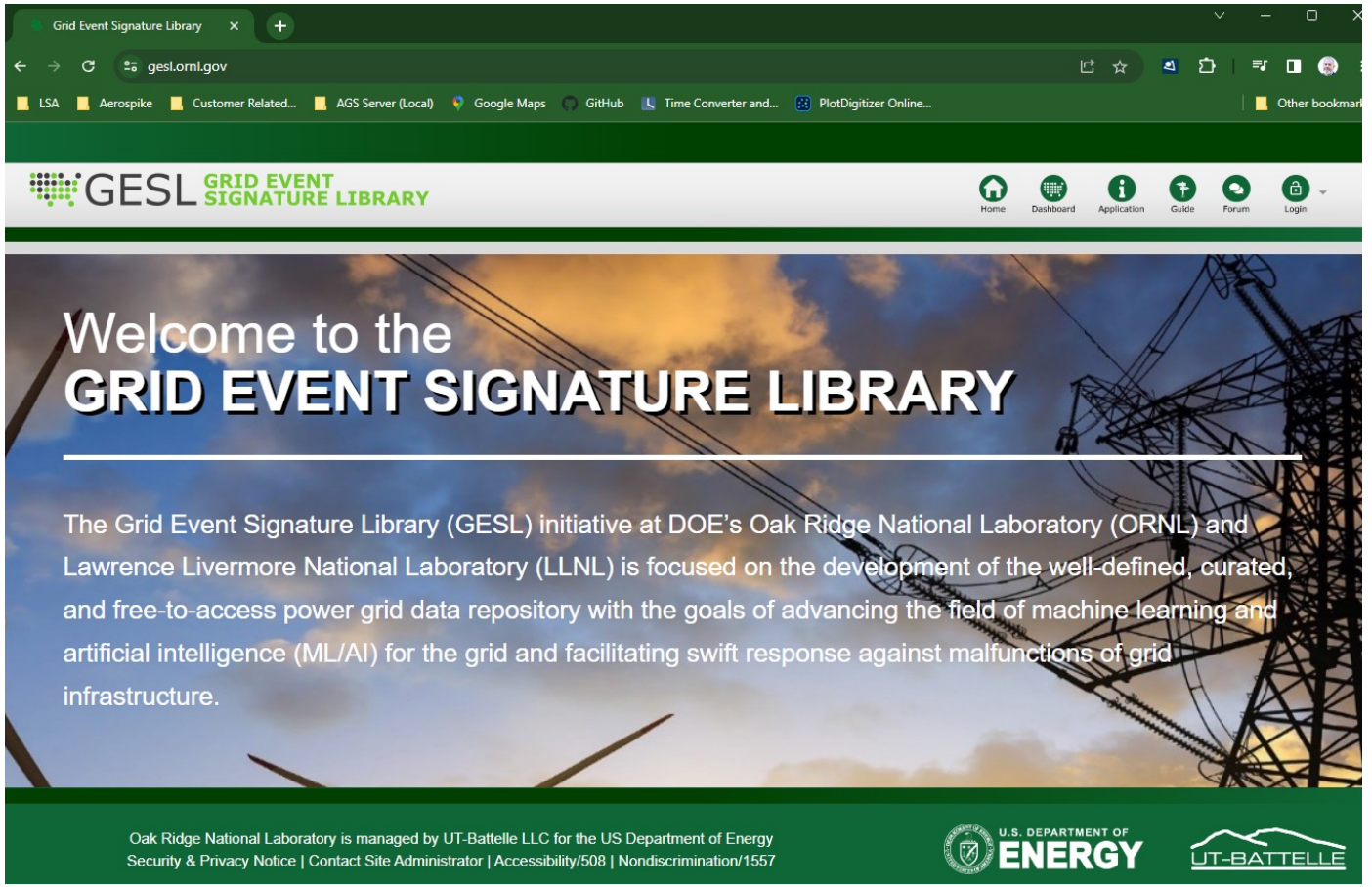


AI > Resources

Data

ORNL/LLNL Grid Event Signature Library

- Objective – Development of the well-defined, curated, and free-to-access power grid data repository with the goals of advancing the field of machine learning and artificial intelligence (ML/AI) for the grid and facilitating swift response against malfunctions of grid infrastructure.



AI > Resources

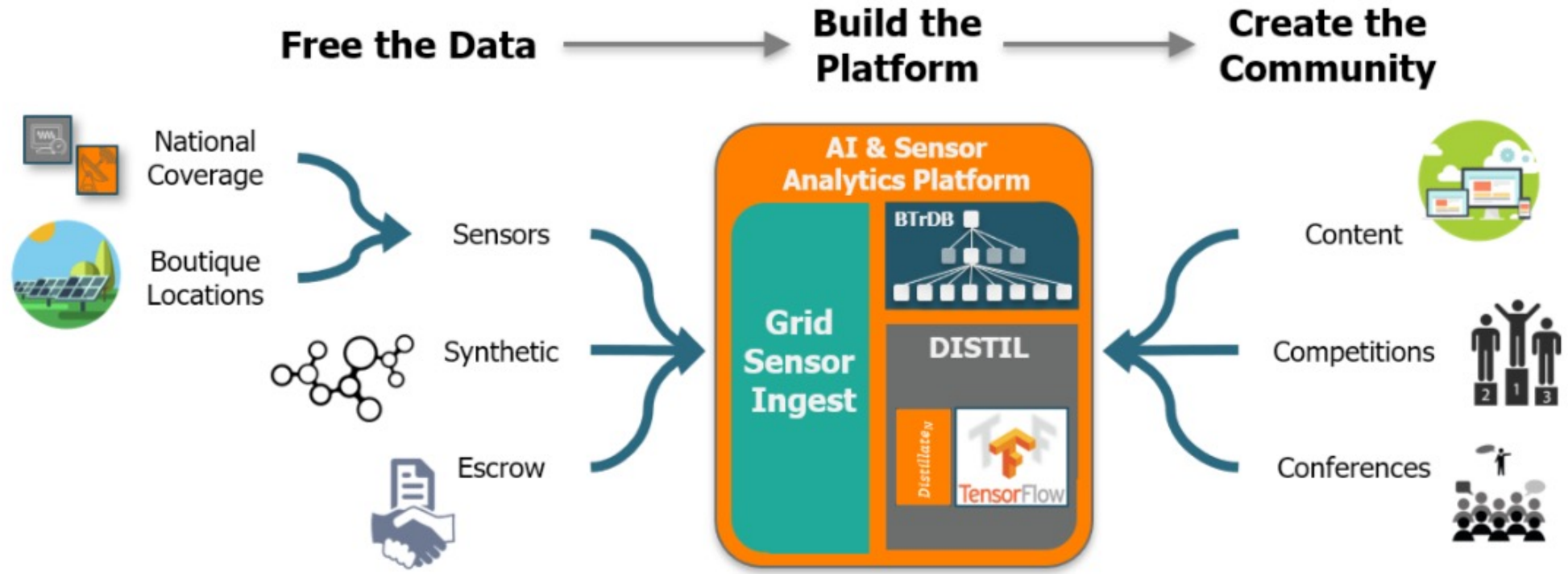
Tools

NI4AI Objectives

1. Remove obstacles to developing new AI use cases
2. Host open access data to train algorithms to study specific problems on the grid
3. Train analysts to work with next-generation grid data
4. Build a community to exchange expertise between engineers and analysts
5. Connect analysts with stakeholders in industry who could benefit from the tools they develop

AI > Resources

Tools > NI4AI



AI > Resources

Tools > NI4AI

Time Period (Years)	Sample Rate (Hz)	Number of Rows	Data Volume (number of data points)			
			100 PMUs (2,000 Columns)	200 PMUs (4,000 Columns)	500 PMUs (10,000 Columns)	1000 PMU (20,000 Columns)
1	30	0.95 G*	1.89 T*	3.78 T	9.46 T	18.9 T
1	60	1.89 G	3.78 T	7.57 T	18.9 T	37.8 T
3	30	2.84 G	5.68 T	11.4 T	28.4 T	56.8 T
3	60	5.68 G	11.4 T	22.7 T	56.8 T	114 T

*G (giga) stands for 10^9, T (tera) stands for 10^{12}

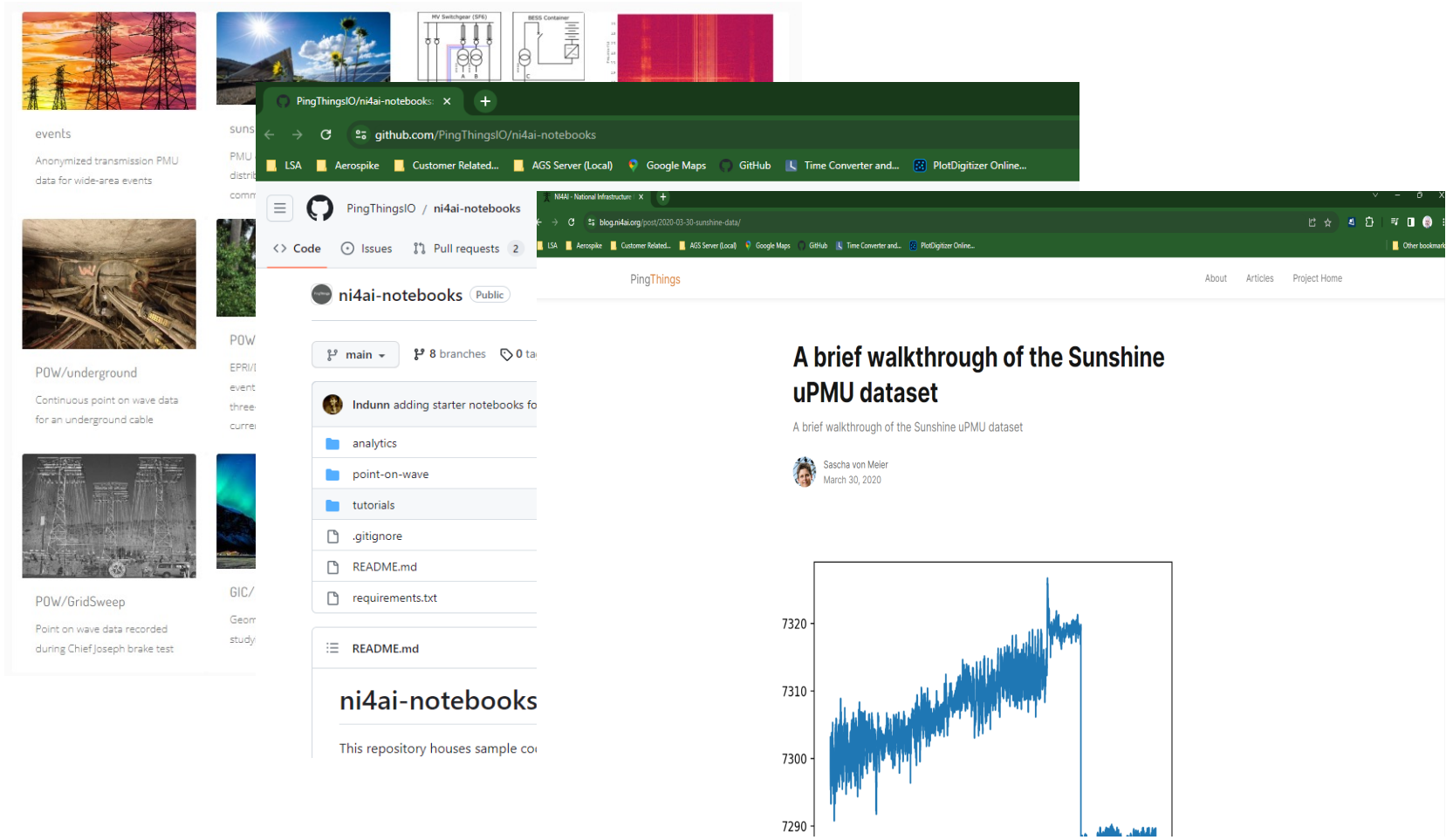
AI > Resources

Tools > NI4AI

Read Speed (data points per sec)	Time to Read 1.89 Trillion Data Points
10,000	6 years
100,000	7.3 months
1,000,000	3.125 weeks
10,000,000	2.1875 days
100,000,000	5.25 hours
1,000,000,000	31.5 minutes

AI > Resources

Tools > NI4AI



www.ni4ai.org

Plotter
PingThings data visualization tool for seeing and analyzing data streams.

NI4AI Blog
News and information about sensor data research and analytics.

Data sets
Find out more about open access data sets hosted in NI4AI.

API tutorials
Learn how to get started using the BTrDB API.

Github repo
Explore open source code other NI4AI users have built.

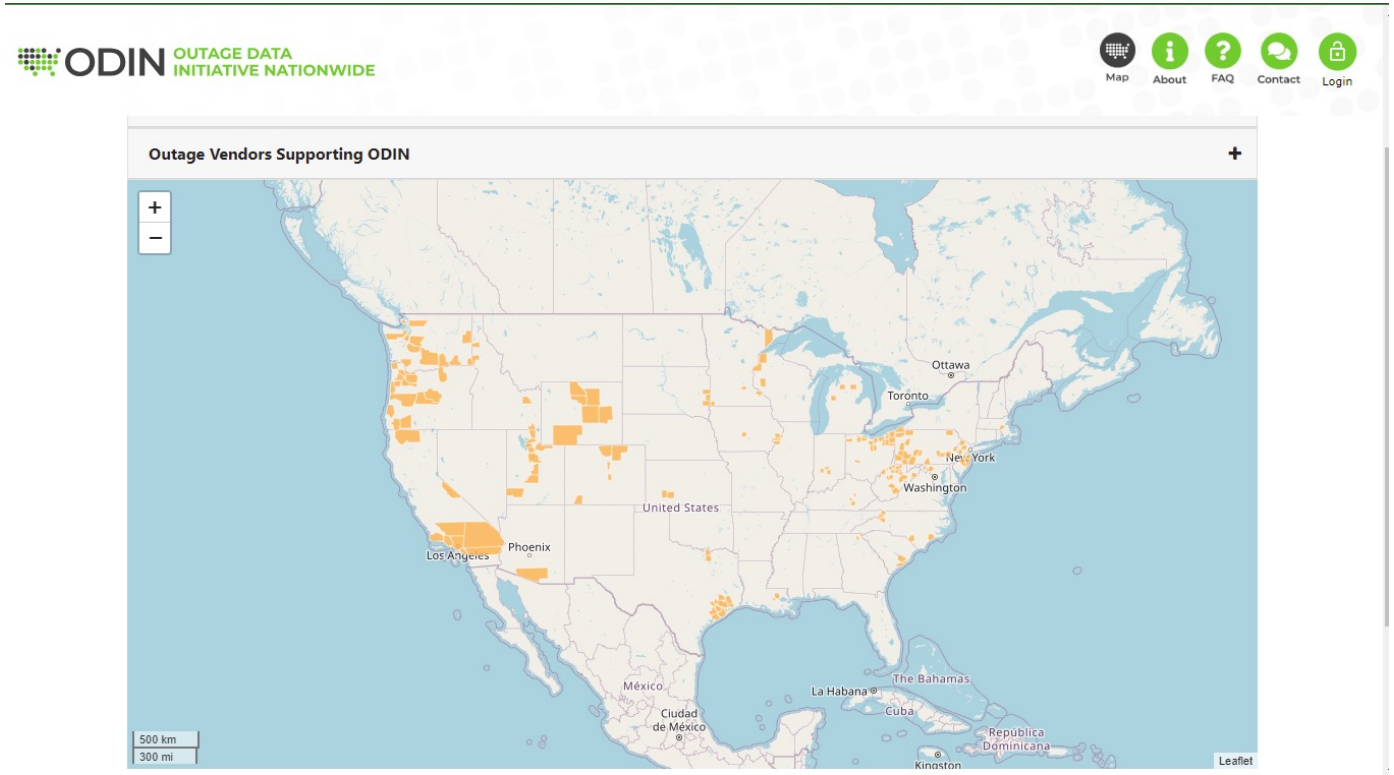
API docs
Access data programmatically using our Python library.

Videos
Watch recordings of past workshops and webinars.

API key
Get your unique key to query the data using our API.

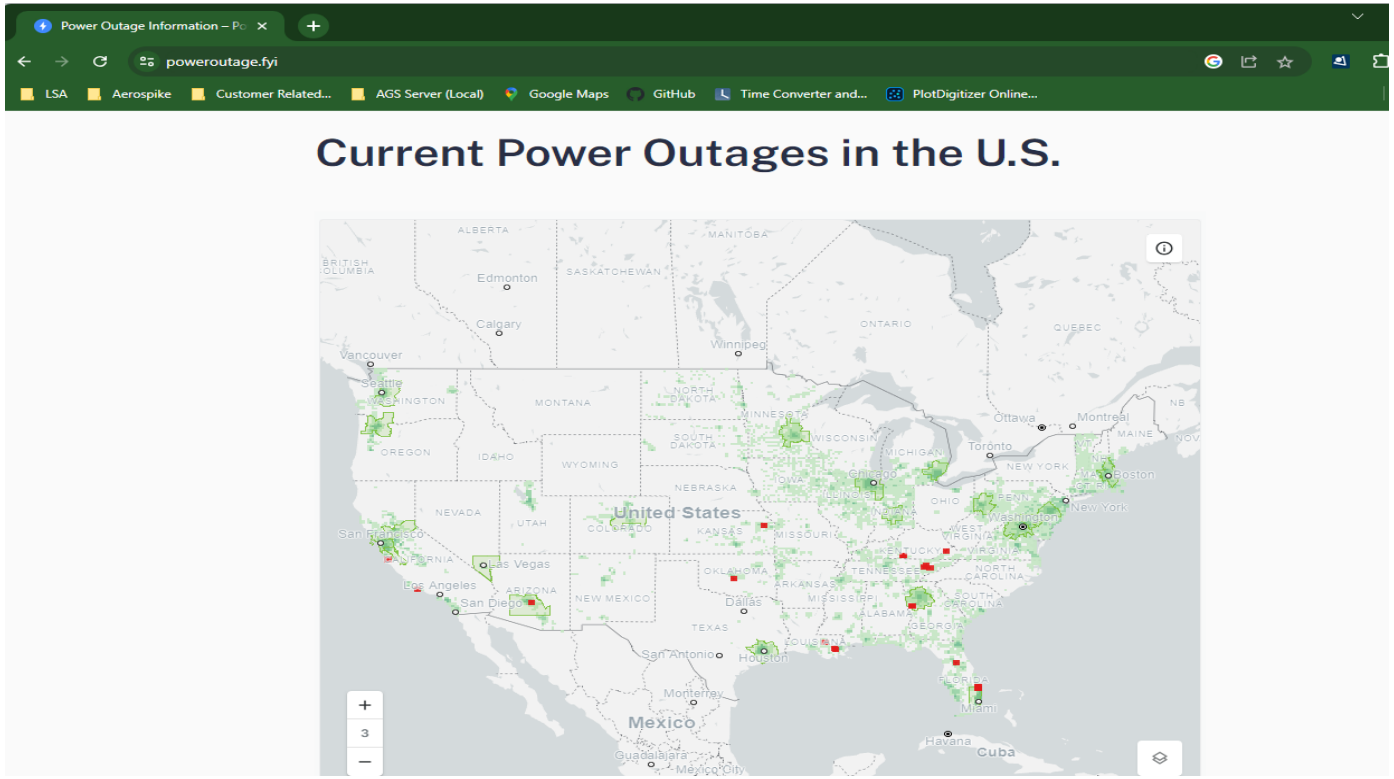
AI > Resources

Outage Info – http://odin.ornl.gov/



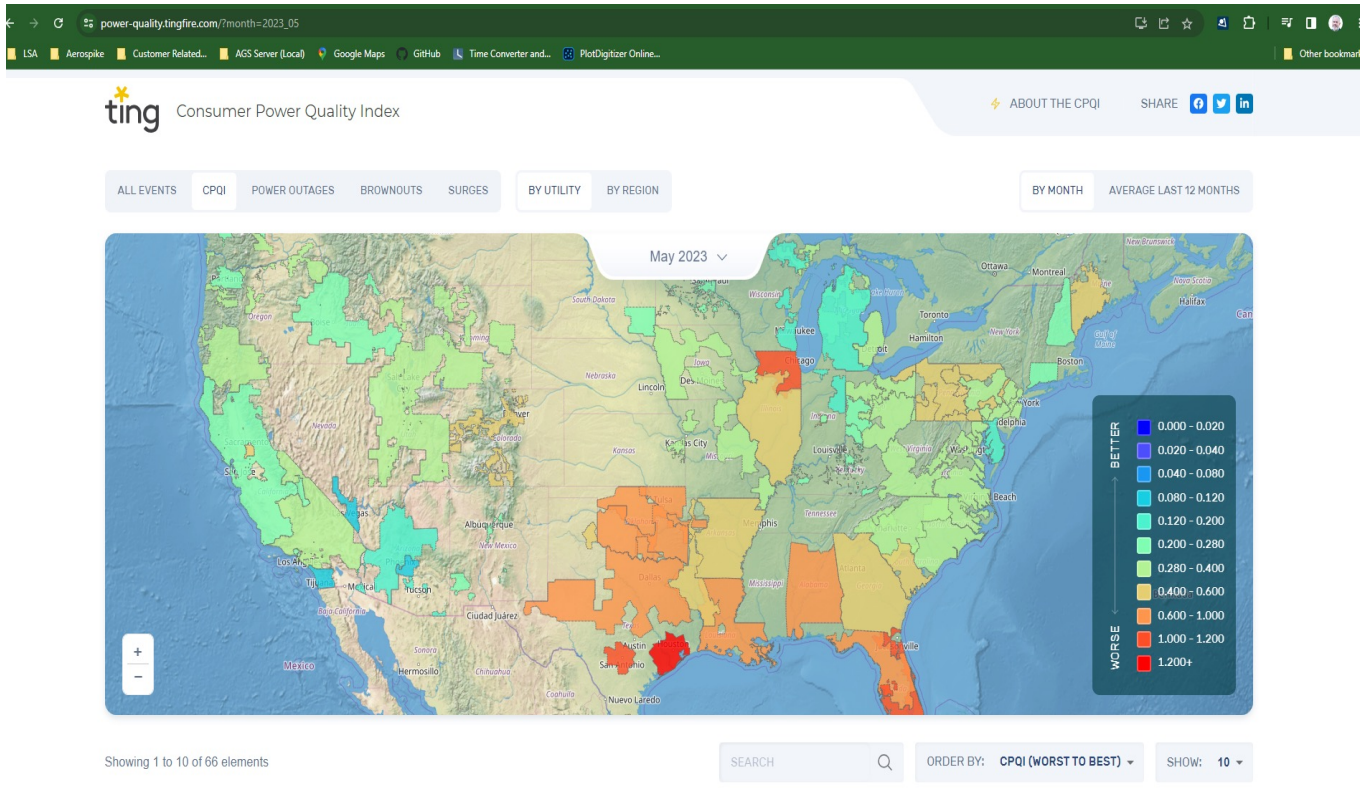
AI > Resources

Outage Info – http://poweroutage.fyi/



AI > Resources

Benchmarks – http://power-quality.tingfire.com



UTILITY	CPQI ↓	ALL EVENTS INDEX	POWER OUTAGE INDEX	BROWNOUT INDEX	SURGES INDEX
NATIONAL AVERAGE	0.47	0.39	0.26	0.11	0.02
CENTERPOINT ENERGY	1.25	1.23	0.78	0.44	0.01
CITY OF SAN ANTONIO - (TX)	1.18	1.12	0.66	0.45	0.01
FLORIDA POWER & LIGHT CO	1.12	0.64	0.36	0.16	0.12
COMMONWEALTH EDISON CO	1.03	0.49	0.16	0.19	0.14
TAMPA ELECTRIC CO	1.01	1.01	0.86	0.15	0
DUKE ENERGY FLORIDA, LLC	0.94	0.72	0.48	0.19	0.05

AI

Conclusions

- Know the use case
  - Identify the question being answered.
- Have good data
  - There are already some free data sources.
  - If generating your own data – make sure it will meet your needs.
- Have the right tools for the job
  - There are already tools that are available (for free).