Open Framework for Integrated Data Operations

OpenFIDO

Operated by Stanford University for the US Department of Energy under contract DE-AC02-76SF00515.

Meeting agenda

- Project description
- Goals and objectives
- Summary of major technical tasks
- Summary of results
- Conclusions and moving forward
- Q&A / Discussion

OpenFIDO

Project Description

SLAC

The Problem

- Need reliable and consistent data and model exchange
- Ingest from many sources
- Curation of heterogeneous data
- Run complex analyses on multiple tools
- Delivery of data with long-term availability

The Solution

- Identify exchange and analysis requirements for California IOUs
- Develop and test a platform to satisfy identified needs
- Demonstrate solution for primary use-cases identified
- Commercialize platform to enable large scale deployment

Goals and Objectives

SLAC

Support major use-cases

- ICA/Hosting capacity analysis
- End-use electrification analysis
- Tariff design analysis
- Resilience analysis

Other use-cases identified

- CYME to GLM conversion
- Forecasting (e.g., load weather)
- Data access (e.g., vegetation)

Validation

- Forecasting load at National Grid
- Resilience analysis at SCE
- Tariff and electrification tests

Pathways to commercialization

- Adoption by commercial entity

Summary of Major Technical Tasks

SLAC

Task 1 - Management

Task 2 - Requirements Analysis

Task 3 - Implementation

Task 4 - Testing and Validation

Task 5 - Finalize Production

Task 6 - Evaluation of Benefits

Task 7 - Technology Transfer

Summary of Major Technical Tasks

Task 1 - Management

- Subtask 1.1 - Products
- Subtask 1.2 - Kick-off Meeting
- Subtask 1.3 - CPR Meetings
- Subtask 1.4 - Final Meeting
- Subtask 1.5 - Progress Reports
- Subtask 1.6 - Final Report
- Subtask 1.7- Match Funds
- Subtask 1.8 - Permits
- Subtask 1.9 - Subcontracts
- Subtask 1.10 - TAC
- Subtask 1.11 - TAC Meetings

CALIFORNIA ENERGY COMMISSION

CALIFORNIA NATURAL RESOURCES AGENCY

FINAL PROJECT REPORT

Open Framework for Integrated Data Operations (OpenFIDO)

Agreement Number: EPC-17-047

Author(s):
David P. Chassin, Duncan Ragsdale, Alyona I. Teyber (SLAC National Accelerator Laboratory)

Kevin Rohling (Presence Product Group)

Matthew Tisdale, Katie Wu (Gridworks)

Bo Yang, Yanzhu Yu (Hitachi America Laboratories)

California Energy Commission Project Manager:
Qing Tian, Ph.D., P.E.
Energy Research and Development Division

Month Year | CEC-500-XXXX-XXX

DISCLAIMER

This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees or the State of California. The Energy Commission, the State of California, its employees, contractors, and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this report

Summary of Major Technical Tasks

SLAC

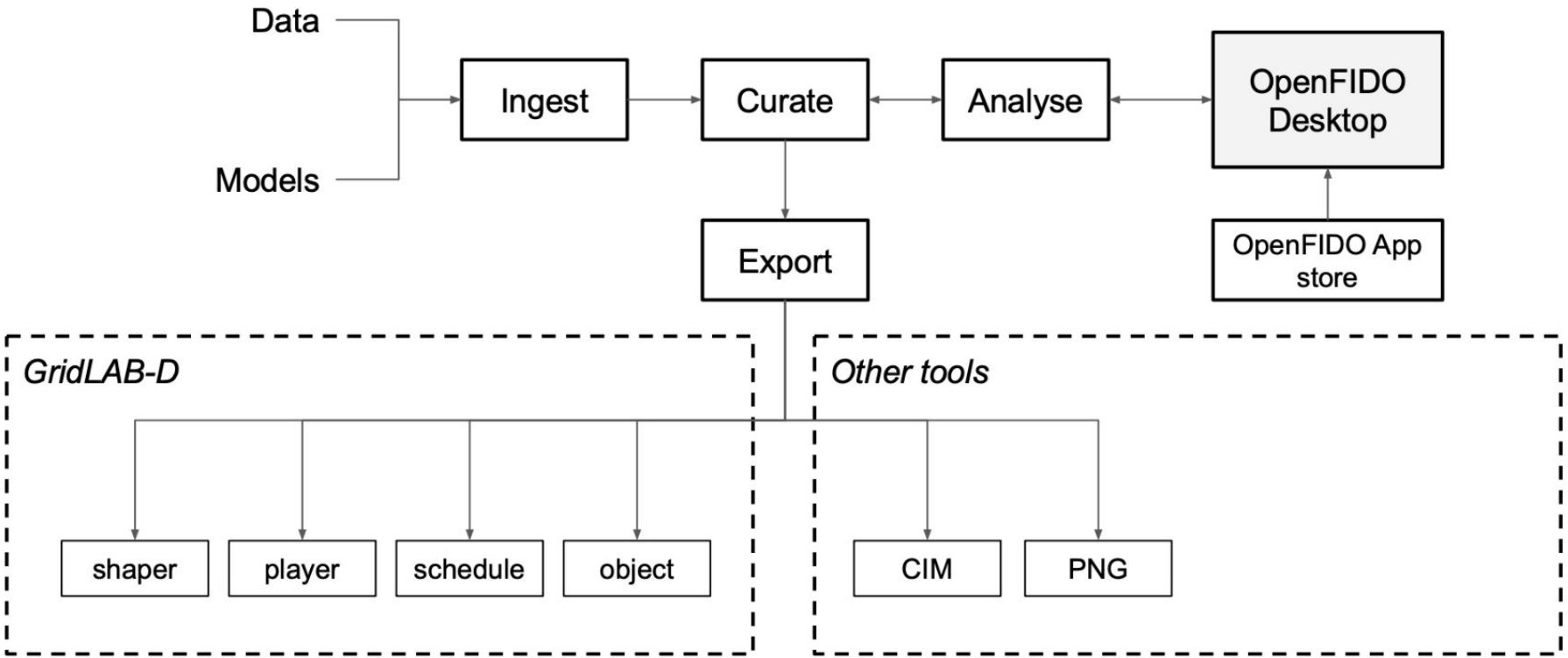
Task 2 - Requirements Analysis

Establish requirements needed for OpenFIDO extension to VADER and to assess those requirements for the data exchange implementation to VADER.

Deliverables

1. Data Exchange Requirements and Assessment Presentation
2. Data Exchange Implementation and Validation Plan
3. Data Exchange Implementation and Validation Presentation

OpenFIDO Workflow Model



User Requirements: Government Agencies

- **Tariff analysis**
  - Customer revenue and resource cost impacts at distribution level
  - Tariff parameter sensitivity analysis on revenue and costs
- **ICA analysis**
  - Implementation of a standard integrated capacity analysis method
- **LNBA analysis**
  - Implementation of a standard locational net-benefit analysis method
- **Resilience analysis**
  - Asset vulnerability assessment
  - Large-scale rare event impacts analysis

User Requirements: Utilities

- **ICA, LNBA analysis**
  - Validation of ICA and LNBA on utility test models
- **Tariff design**
  - Validation of tariff design tools on utility test models
- **Resilience planning**
  - Demonstrate asset vulnerability and extreme event analysis method
- **Resource procurement modeling**
  - Load modeling and load forecasting
  - Distributed energy resource modeling
  - Demand response modeling

User Requirements: Service/analytics Providers

SLAC

- **Custom analysis implementation and deployment**
  -
- **Automate access to public data repositories**
  - Weather data collection/processing/delivery
  - Standard grid, building, and tariff models libraries
- **Data import and export**
  - Automatic data format recognition and handling
  - Data cleaning, time-standardization, identification

User Requirements: Researchers

- **Documentation**
  - Current internals documentation is needed for all analytic methods
- **Training resources**
  - User cases with step-by-step instruction online as videos
- **Valid agents for all system components**
  - Code snippets, standardized units, parameter ranges, good defaults
  - E.g., distribution grids, households, buildings, renewable resources
- **Data-driven res/com building models**
  - CEUS/RBSA data for residential and commercial building models
- **Optimization solvers**
  - Update optimization module to support new methods

Requirements: Key Messages

- **Usability**
  - Improve GridLAB-D UI/UX and GLOW integration
- **Speed**
  - Enable multi-core/host and cloud-scale workflows
- **Flexibility**
  - Enable user-defined workflows, scenarios, and modules
- **Validation**
  - Enable user-defined model validation tools
- **Reproducibility**
  - Enable reproducible simulation and analysis capabilities

Summary of Major Technical Tasks

SLAC

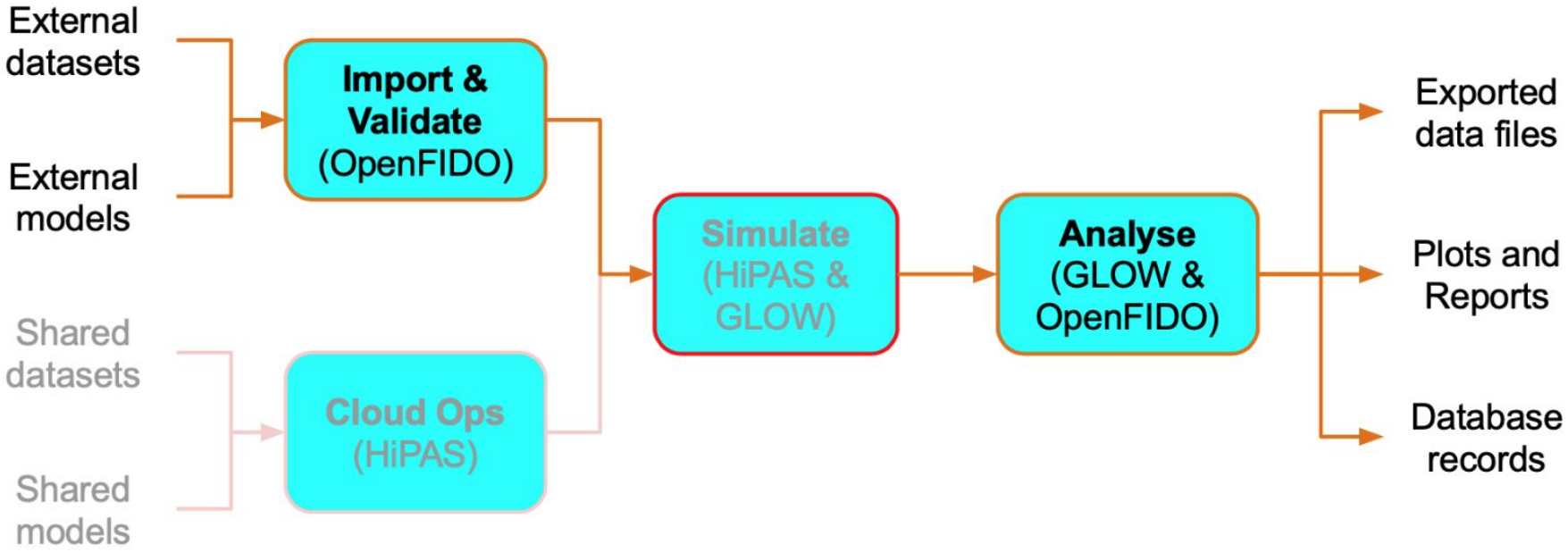
Task 3 - Implementation

Implement the data import and export codes in VADER software that meet the requirements established in Task 2.

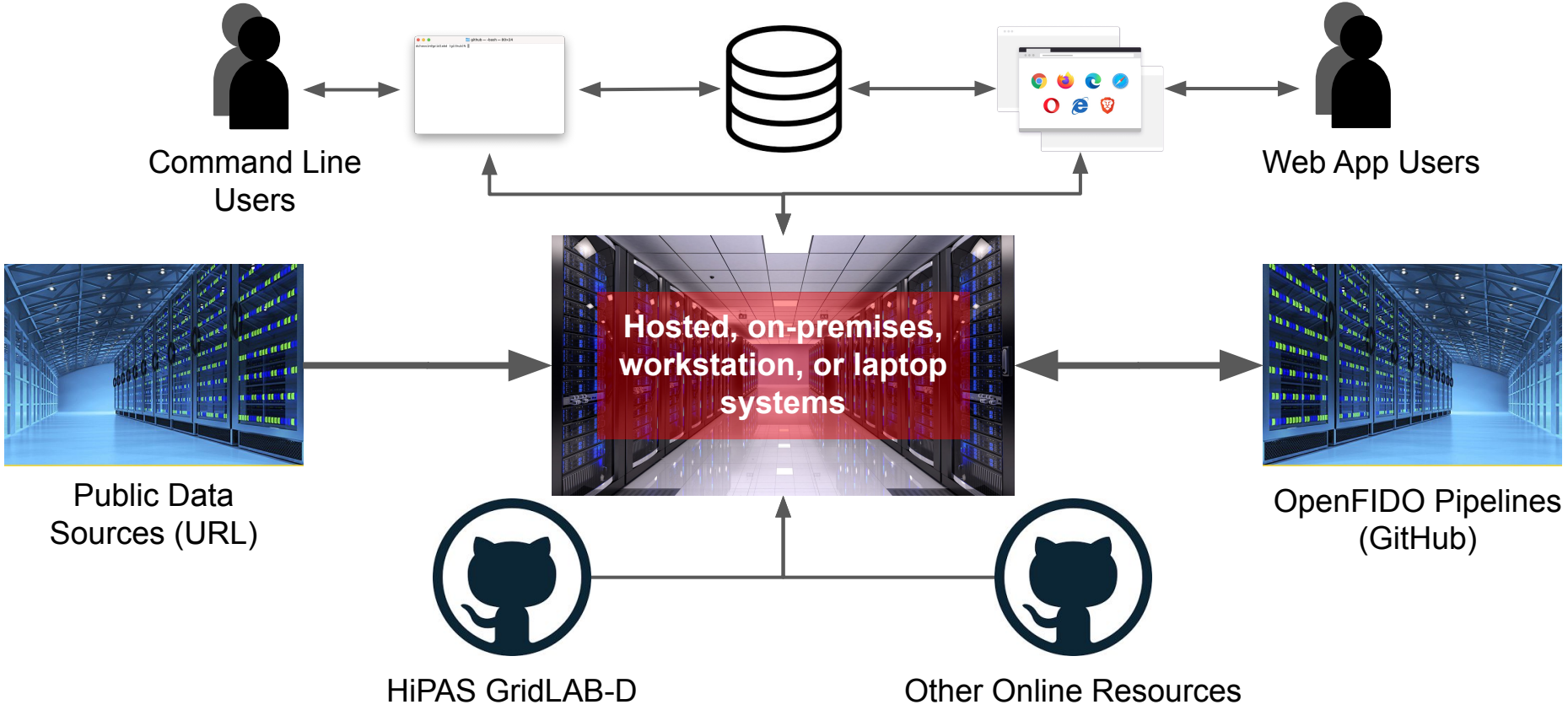
Deliverables

1. Implementation Report - CPR Report 1

Pipeline workflow support



OpenFIDO Implementation Architecture



Summary of Major Technical Tasks

SLAC

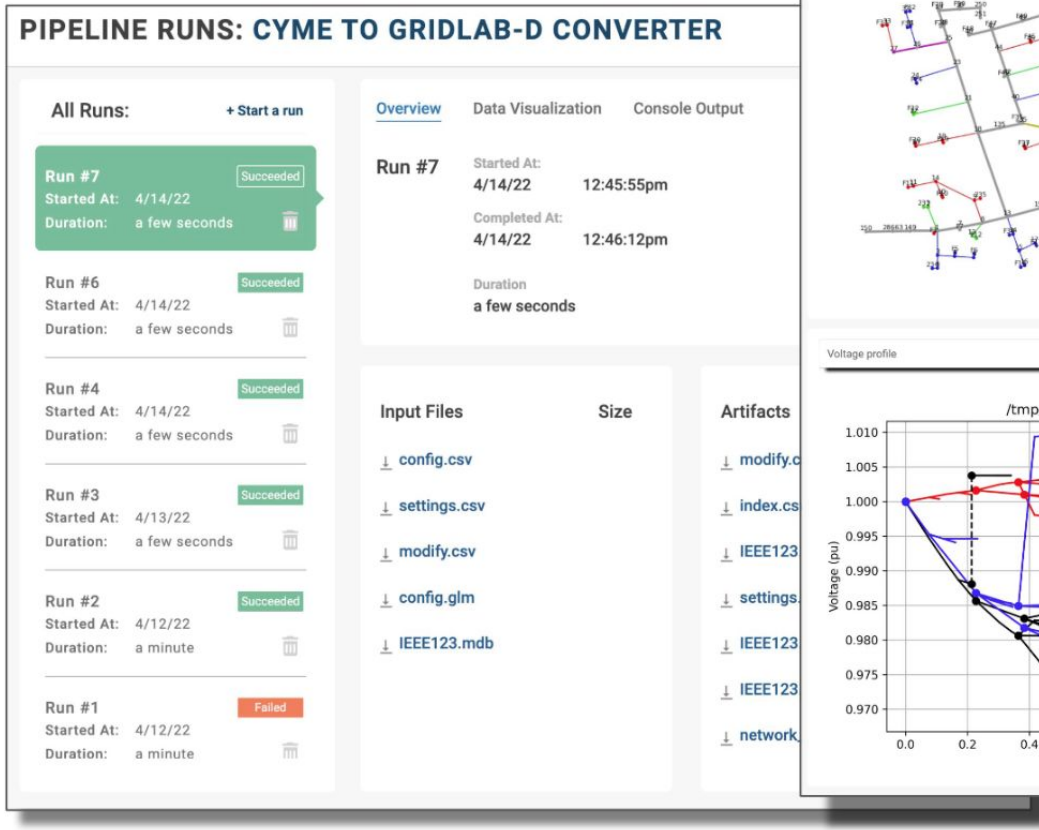
Task 4 - Testing and Validation

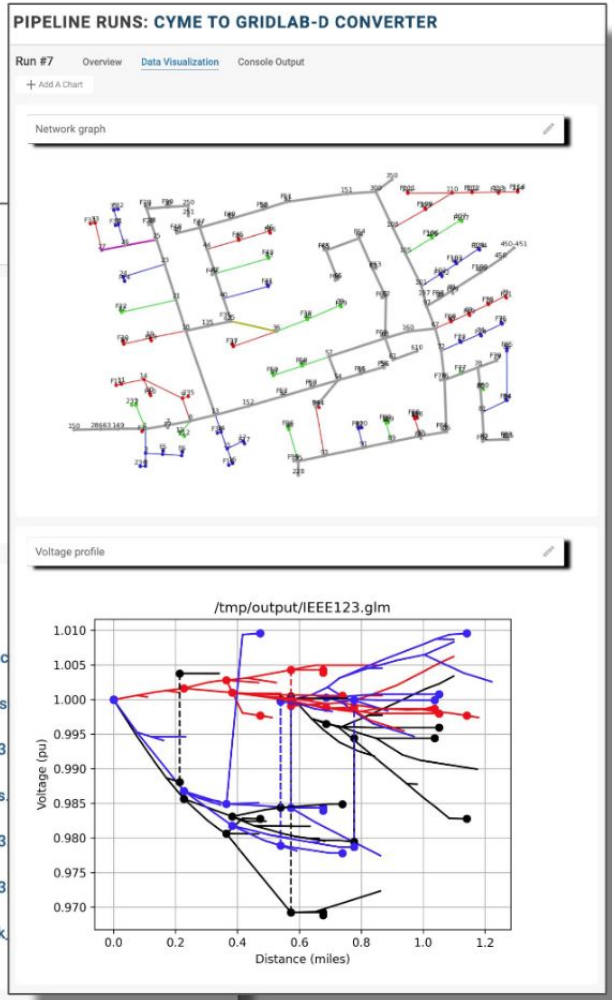
Verify the VADER code functionality and validate the data exchange capabilities between VADER software and external tools.

Deliverables

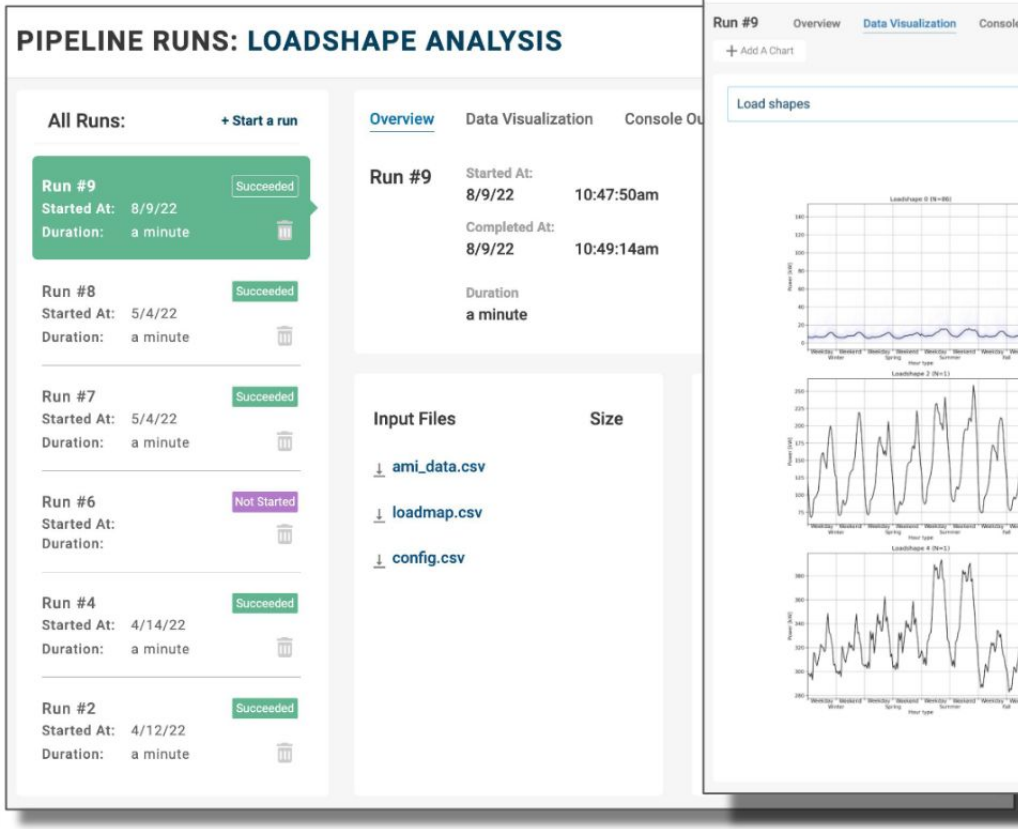
1. Testing and Validation Presentation
2. Testing and Validation Report

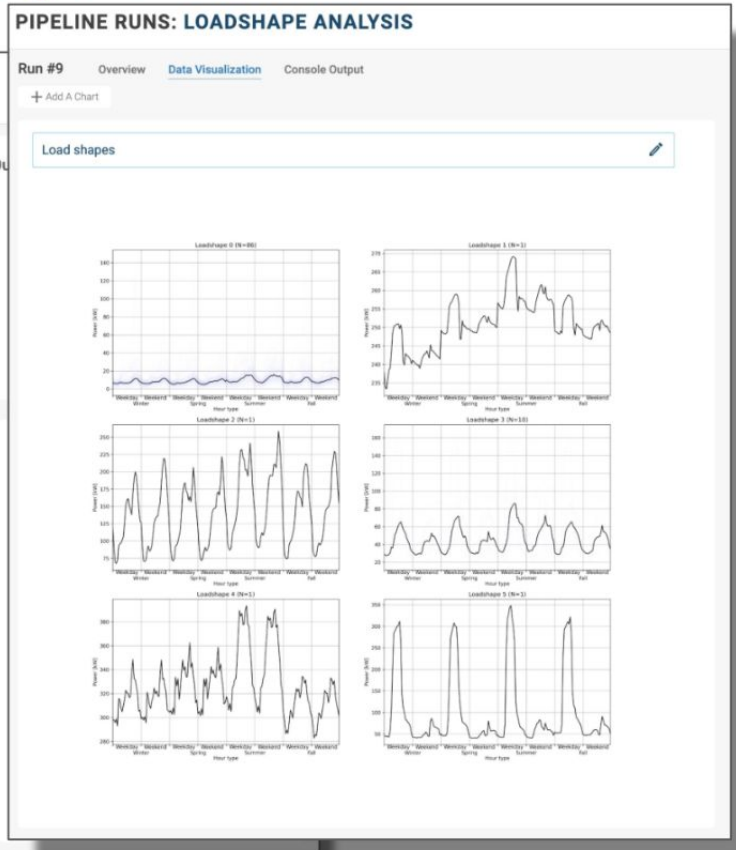
Pipeline Example 1: Cyme Converter





Pipeline Example 2: Loadshape analysis





Results from testing and validation: Inputs and Output

Pipeline inputs

Tariff design: OpenEI, GLM model files

Loadshape: AMI data, SCADA data

Weather: location, year

HiPAS GridLAB-D: GLM model files

Census: location, year

Resilience: GLM model files, time range, location

Hosting Capacity: GLM model files, load models

Electrification: GLM model files, load forecast

Address: location

Cyme converter: Cyme MDB files

Output artifacts

CSV data files

PNG charts/graph images

GLM model files

Results from testing and validation: Quality and Coverage

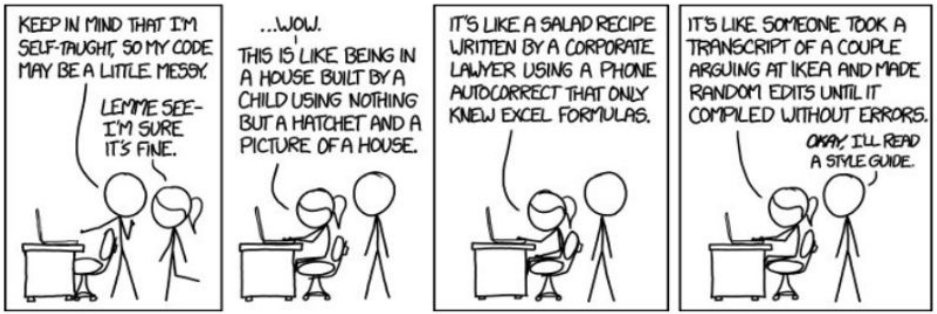
Lint code quality score

Lint score calculation

$$100.0 - 100 \times \frac{5 \times Frequency\ of\ convention\ errors}{Number\ of\ statements}$$

Score must greater than 90% to pass validation

No industry standard exists for lint score



Code testing coverage

Evaluates the fraction of tested statements

Relative to the total number of statements

Must be greater than 99% to pass validation

Google test coverage standards[1]:

- 60% is acceptable
- 75% is commendable
- 90% is exemplary

Gains are logarithmic (diminishing returns)

[1] https://testing.googleblog.com/2020/08/code-coverage-best-practices.html

Results from testing and validation: Approach

Methodology

- CI/CD-based testing/validation (GitHub)
  - Verify programmatic soundness of code
  - Coding style/conventions requirements
  - Verify code correctness
  - Verify testing coverage
- Integration testing
  - Verify that code change requests are ok
  - Verify staged system is ready to deploy
- Pipelines
  - Originate from other projects (e.g., HiPAS)
  - Tested separately and independently
  - Pipeline validation not in OpenFIDO scope

Implementation

- Phase 1 (code submittal)
  - *eslint* (code style/conventions)
  - *black* (code correctness)
  - *test* or *pytest* (unit test/validation)
  - *jest-puppeteer* (integrated test/validation)
    - Authentication
    - Account management
    - Organization management
    - Pipeline creation/running
    - Housekeeping
- Phase 2 (code review)
  - Staged deployment

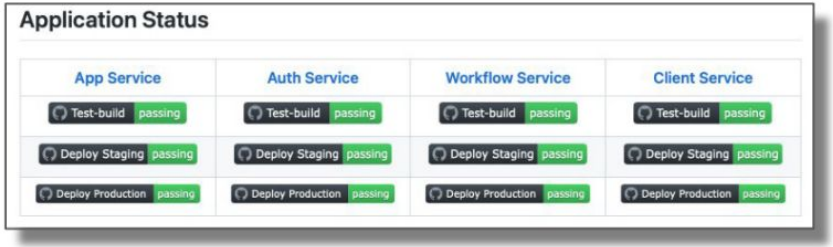
Results from testing and validation: Platform Status

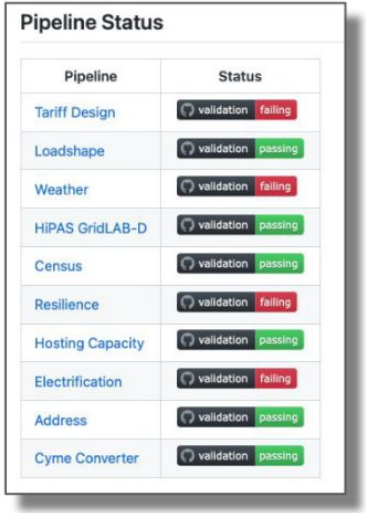
GitHub "Actions" CI/CD status reports

- Applications and pipelines status in real-time

Table A.1 - GitHub Actions

Component	Test	Result	Disposition
App	211 unit tests	>97% coverage	100% tests passing. See Table A.2
	style	100/100 ok	
	lint	>90% score	See Table A.3
Auth	136 unit tests	>92% coverage	100% tests passing. See Table A.4
	style	100% ok	
	lint	>90% score	See Table A.5
Workflow	160 unit tests	>99% coverage	100% tests passing. See Table A.6
	style	100% ok	
	lint	>90% score	See Table A.7
Client	jest-puppeteer	No issues	





Example real-time CI/CD application and pipeline status reports from GitHub.

Results from testing and validation: CI/CD Status

Table A.2 - App Missing Coverage

File location	Missed Statements	Proposed Remedy
app/pipelines/services.py	18/254 uncovered	Add additional tests for uncovered statements.
app/workflows/routes.py	4/134 uncovered	Add additional tests for uncovered statements.
app/workflows/services.py	7/279 uncovered	Add additional tests for uncovered statements.

Table A.3 - App Lint Issues

Test Result	Issue Description	Proposed Remedy
92.6% lint score	Code Style and Formatting	Run the *lint* command, and review each potential style or formatting issue listed against the linter's recommended fix.

Table A.4 - Auth Missing Coverage

File location	Missed Statements	Proposed Remedy
app/auth.py	32/260 uncovered	Add additional tests for uncovered statements.
app/mail.py	13/52 uncovered	Add additional tests for uncovered statements.
app/org.py	14/221 uncovered	Add additional tests for uncovered statements.
app/services.py	3/211 uncovered	Add additional tests for uncovered statements.

Table A.5 - Auth Lint Issues

Test Result	Issue Description	Proposed Remedy
90.4% lint score	Code Style and Formatting	Run the *lint* command, and review each potential style or formatting issue listed against the linter's recommended fix.

Table A.6 - Workflow Missing Coverage

File location	Missed Statements	Proposed Remedy
app/workflows/queries.py	2/52 uncovered	Add additional tests for uncovered statements.

Table A.7 - Workflow Lint Issues

Test Result	Issue Description	Proposed Remedy
92.2% lint score	Code Style and Formatting	Run the *lint* command, and review each potential style or formatting issue listed against the linter's recommended fix.

Results from testing and validation: Pipeline Status

Pipelines status

- HiPAS GridLAB-D master/develop issues
- Once resolved CD status will update ok

Known use-case pipeline issues

- *Resilience*: pole failure accuracy uncertain, some vegetation data missing
- *Tariff design*: slow
- *Electrification*: network losses missing
- *ICA*: speed-up not meeting goals

Table B.1: Pipeline testing/validation summary

Pipeline Name	Validation status	Cause/Action
Tariff design	Failing	HiPAS GridLAB-D template download failed
Loadshape	Passing	
Weather	Failing	HiPAS NSRBD weather download failed
GridLAB-D	Passed	
Census	Passing	
Resilience	Failing	HiPAS GridLAB-D template download failed
Hosting capacity	Passed	
Electrification	Failing	HiPAS GridLAB-D template download failed
Address	Passing	
Cyme Extract	Passing	

Results of testing and validation: Cyme Converter

National Grid Load Forecast (LGF) Study

- 15 year load growth projection
- Analysis is updated annually
- 2021 LGF done w/DOE GridLAB-D
- 2022 LGF done w/HiPAS GridLAB-D
- Converted ~2000 Cyme feeders
- Used "Cyme Extract" pipeline
- Included weather and solar PV

Results of National Grid LGF Study

- Generated HiPAS GridLAB-D models
- 97.5% success unsupervised conversion
- 2.5% required manual intervention
  - Cyme network model errors
  - Cyme load model errors
  - Cyme-GridLAB model mismatches
- Validated based on energy consumption relative to 2021 within load growth

Results from testing and validation: Lessons Learned

Requirements

- Data exchange needs evolve quickly
- Data formats change without warning
- Vendors can be uncooperative/resistant
- Need standard approach to validation
- Need legal framework for data sharing

Implementation

- Dependencies across tools is challenging
- Python changes/updates frequent issue
- Utilities need rapid pipeline deployment
- Utilities need time to validate tools
- Security compliance not part of CI/CD

Summary of Major Technical Tasks

SLAC

Task 5 - Finalize Production

Prepare and release the final OpenFIDO extensions of VADER software.

Deliverables

1. Developer and User Training Documentation
2. Final Product Release Report
3. Production Report (CPR Report #2)

Task 5: Key Findings

SLAC

Requirements

Open Source Software

- Users want open-source software
- Utilities wary of open-source support

Deployment Flexibility

- Utility migration to cloud is going slow
- On-premise servers still preferred
- Local workstation/laptop still desired

Result Reproducibility

- Must retrieve/reproduce old results

Implementation

Deployment Platform

- Python and React preferred
- Python language stability concerns

CI/CD Modernization

- Older CI/CD tools unstable
- GitHub is preferred

Commercializability

- Limited open-source licenses (no GPL)
- LF Energy application pending

Summary of Major Technical Tasks

Task 6 - Evaluation of Benefits

Report the benefits resulting from this project.

Deliverables

1. Kick-off Meeting Benefits Questionnaire
2. Final Meeting Benefits Questionnaire

Summary of Major Technical Tasks

Task 7 - Technology Transfer

Develop a plan to make the knowledge gained, experimental results and lessons learned available to the public and key decision-makers.

Deliverables

1. Initial Fact Sheet
2. Final Project Fact Sheet
3. Final Presentation Materials
4. Technology/Knowledge Transfer Plan
5. Technology/Knowledge Transfer Report

Summary of Results

Capabilities delivered in pipelines

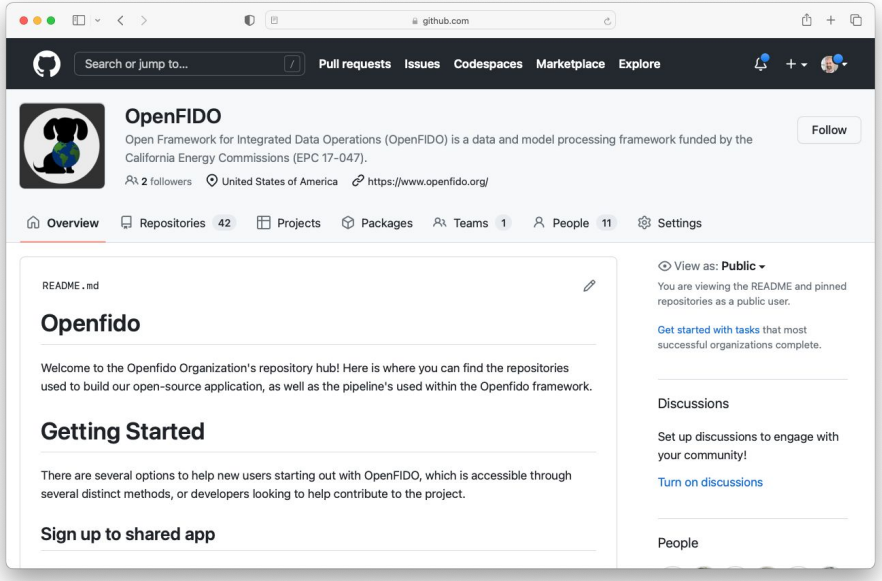
Data/Model Handling

1. Weather
2. Census
3. Address
4. CYME Converter

Analysis Tools

1. Hosting capacity analysis
2. Electrification impacts analysis
3. Resilience analysis
4. Tariff design analysis

Commercialization by LF Energy



https://source.openfido.org/

Summary of Results

Capabilities delivered in pipelines

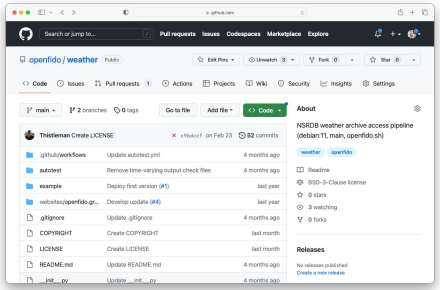
Data/Model Handling

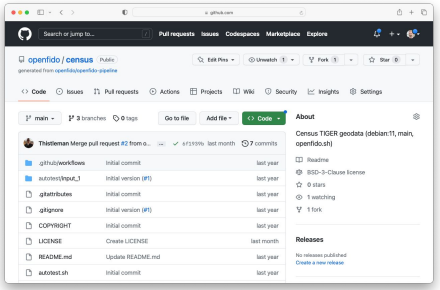
1. **Weather**
2. **Census**
3. **Address**
4. **CYME Converter**

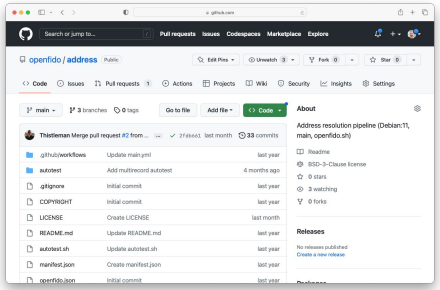
Analysis Tools

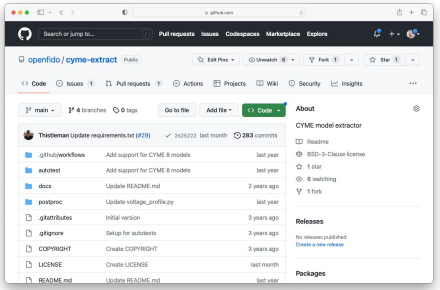
1. Hosting capacity analysis
2. Electrification impacts analysis
3. Resilience analysis
4. Tariff design analysis

Commercialization by LF Energy









https://source.openfido.org/

Summary of Results

Capabilities delivered in pipelines

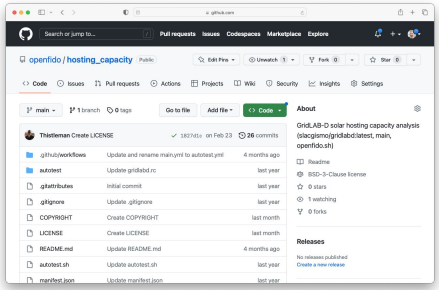
Data/Model Handling

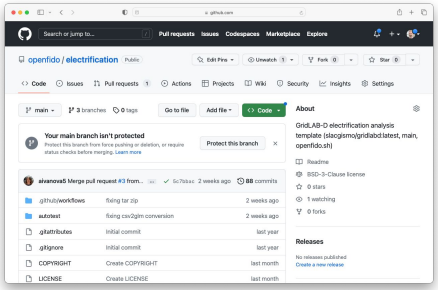
1. Weather
2. Census
3. Address
4. CYME Converter

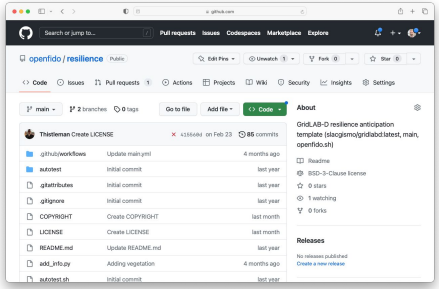
Analysis Tools

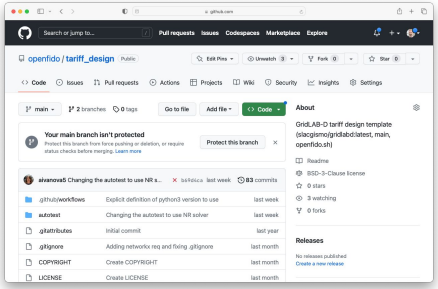
1. **Hosting capacity analysis**
2. **Electrification impacts analysis**
3. **Resilience analysis**
4. **Tariff design analysis**

Commercialization by LF Energy







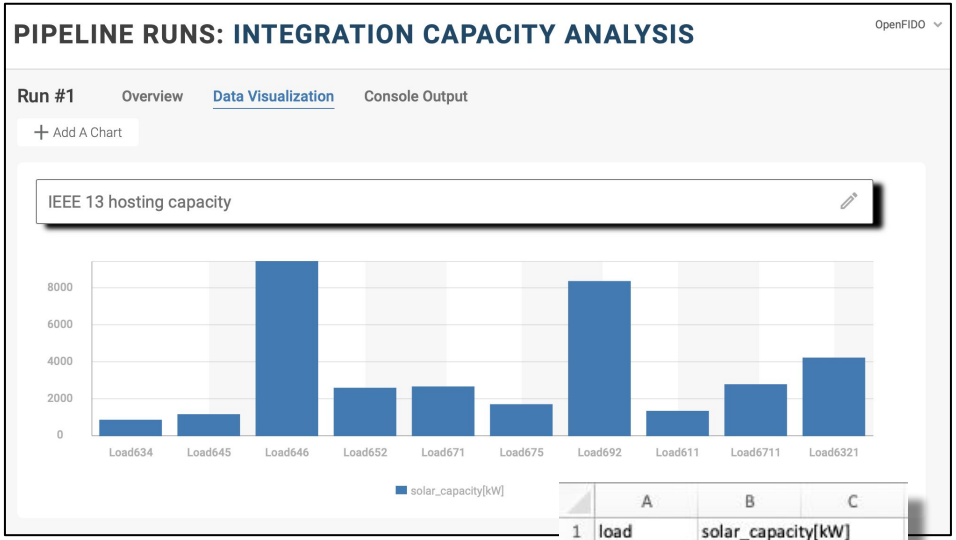


https://source.openfido.org/

Results: Hosting Capacity

Finds maximum PV capacity

- ANSI voltage limits
- 3% voltage fluctuations
- Over-current limits
- Probes PV levels at all loads
- Runs time-series if enabled
- Includes weather, DR, etc.
- Valid method for other DERs

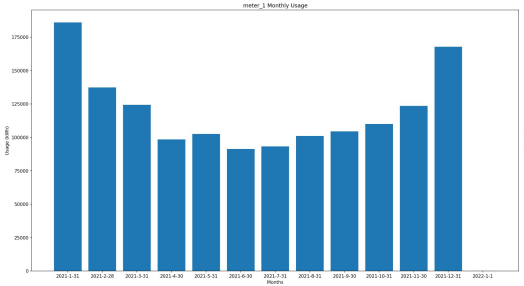


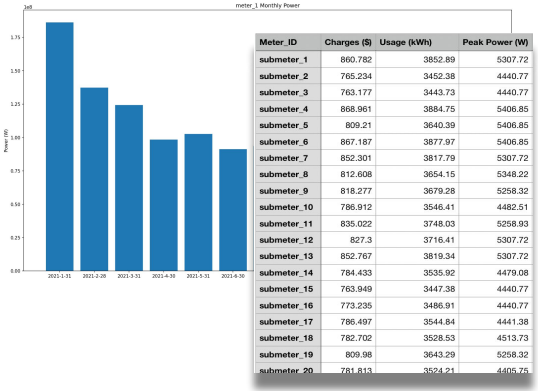
	A	B	C
1	load	solar_capacity[kW]	
2	Load634	840	
3	Load645	1140	
4	Load646	9420	
5	Load652	2570	
6	Load671	2640	
7	Load675	1680	
8	Load692	8340	
9	Load611	1320	
10	Load6711	2760	
11	Load6321	4200	

Results: Tariff Design

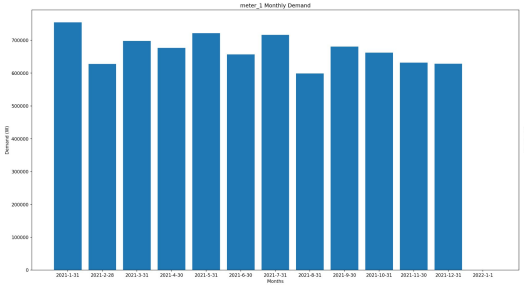
Monthly power, energy, and costs aggregated to feeder level.

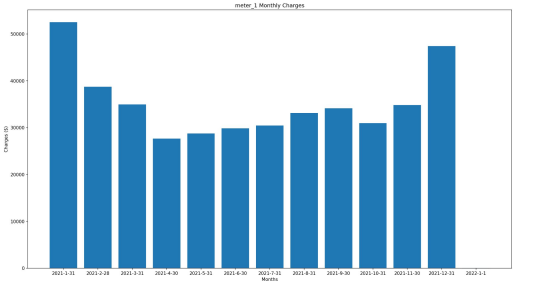
- Data from OpenEI utility rate database (NREL)
- Uses `create_meters` tool to generate meters for models, as needed.
- Validated on IEEE-13 and DOE taxonomy feeders.
- 99% accuracy between manual calculation and template.





Meter_ID	Charges ($)	Usage (kWh)	Peak Power (W)
submeter_1	860.782	3852.89	5307.72
submeter_2	765.234	3452.38	4440.77
submeter_3	763.177	3443.73	4440.77
submeter_4	868.961	3884.75	5406.85
submeter_5	809.21	3640.39	5406.85
submeter_6	867.187	3877.97	5406.85
submeter_7	852.301	3817.79	5307.72
submeter_8	812.608	3654.15	5348.22
submeter_9	818.277	3679.28	5258.32
submeter_10	786.912	3546.41	4482.51
submeter_11	835.022	3748.03	5258.93
submeter_12	827.3	3716.41	5307.72
submeter_13	852.767	3819.34	5307.72
submeter_14	784.433	3535.92	4479.08
submeter_15	763.949	3447.38	4440.77
submeter_16	773.235	3486.91	4440.77
submeter_17	786.497	3544.84	4441.38
submeter_18	782.702	3528.53	4513.73
submeter_19	809.98	3643.29	5258.32
submeter_20	781.813	3524.21	4405.75

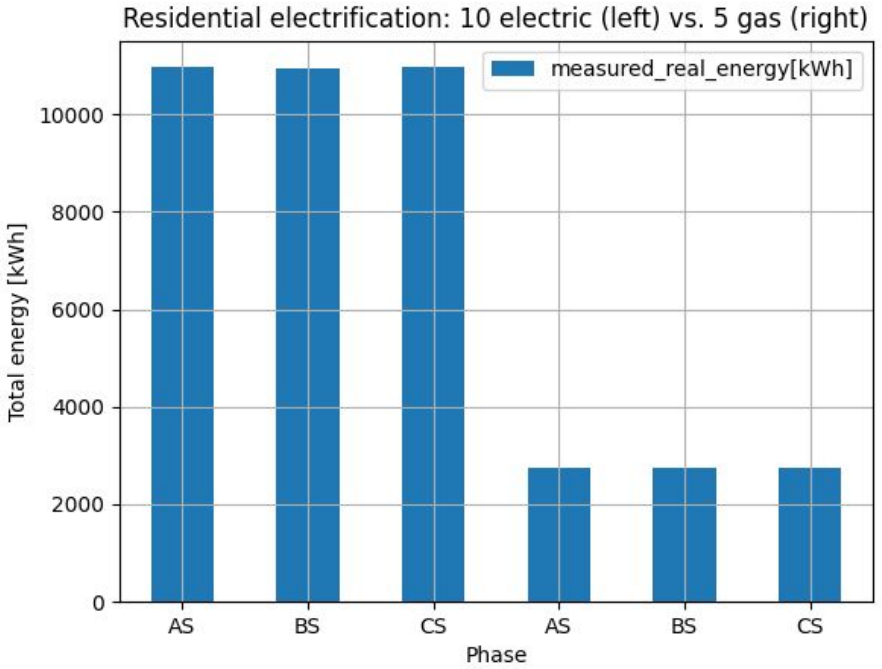




Results: Electrification

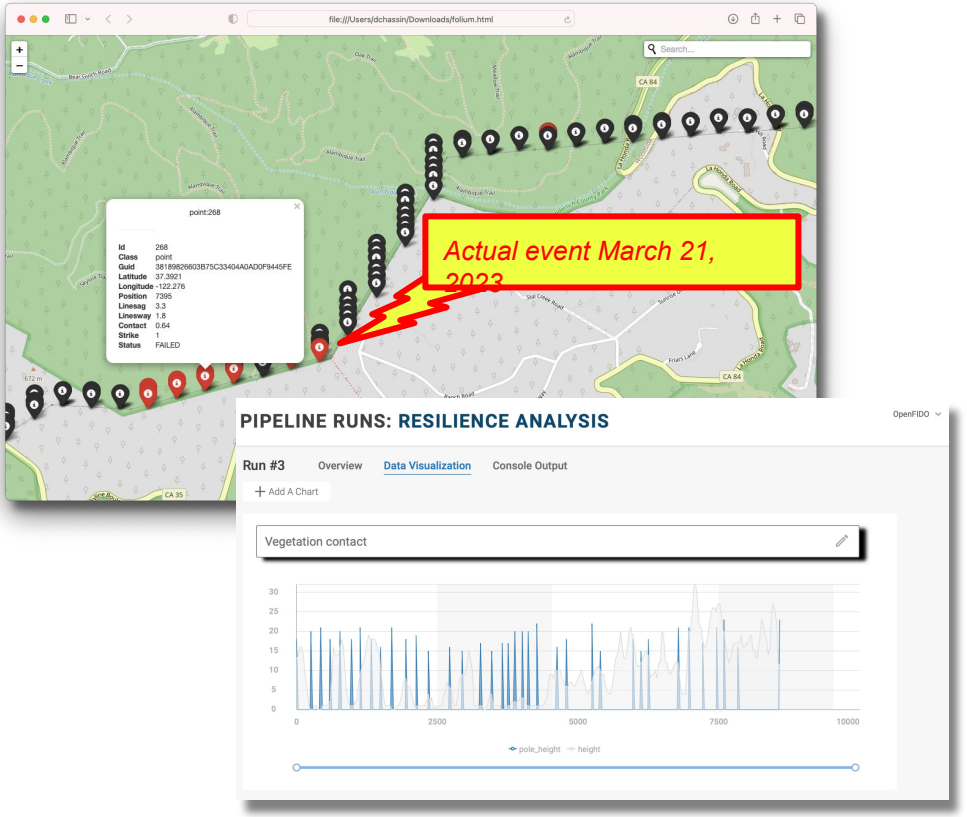
Determines the increased load at substation due to residential electrification

- Validated via manual calculations using component validation.
- Available for different climate regions
- Additional summer peaks are exposed due to availability of A/C under electrification



Results: Resilience

- Deployed at SCE
- Vegetation analytics tested using PG&E 230kV line
- Pole Analysis validated using manual calculations.
  - Automated skeleton grid formulation for bulk pole analysis
  - Integration with utility distribution grid available
  - March 21, 2023 outage matched prediction for similar weather event



Summary of Results

Capabilities delivered in pipelines

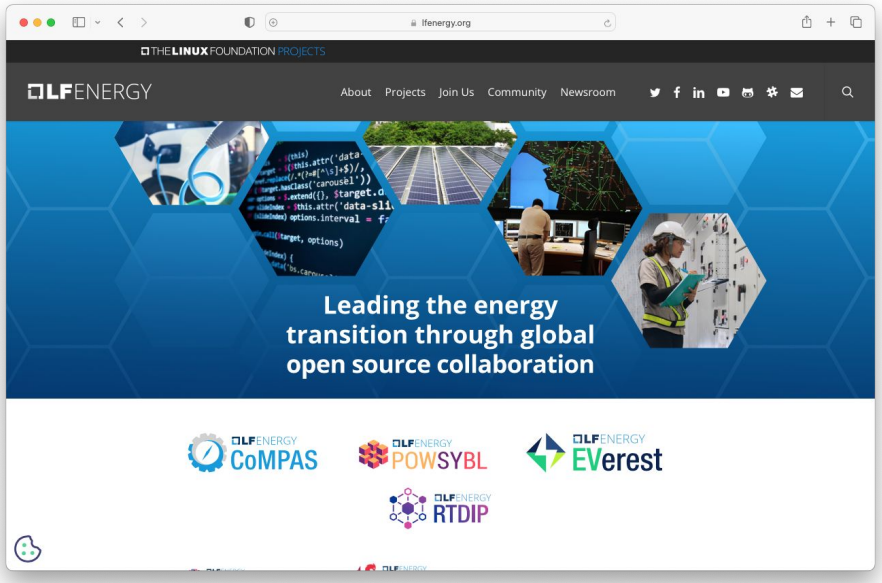
Data/Model Handling

1. Weather
2. Census
3. Address
4. CYME Converter

Analysis Tools

1. Hosting capacity analysis
2. Electrification impacts analysis
3. Resilience analysis
4. Tariff design analysis

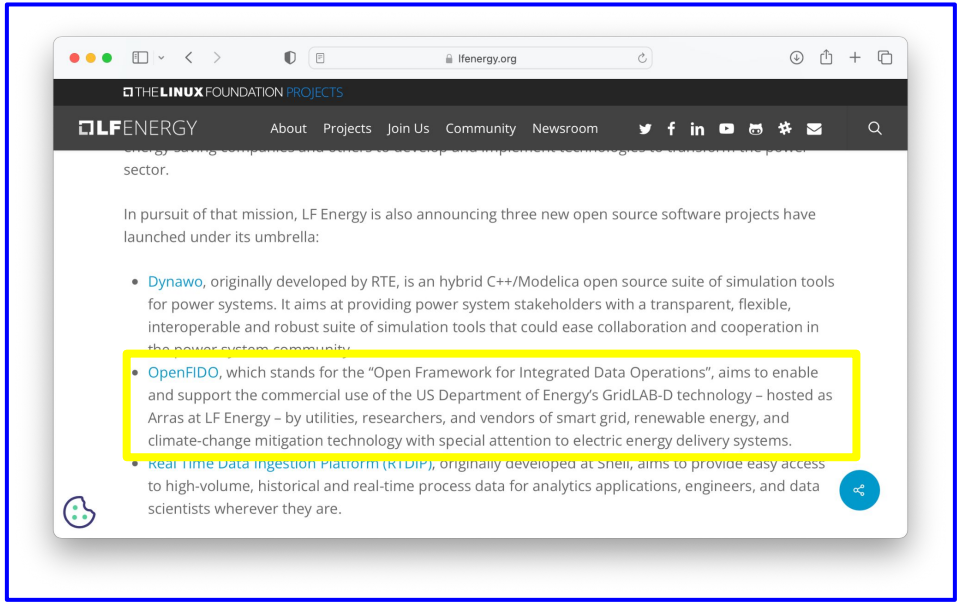
Commercialization by LF Energy



https://www.lfenergy.org/

Linux Foundation Energy Commercialization

- Adopted January 2023
- Open-source release
- Support for ongoing AWS operations
- Initial Technical Steering Committee meeting TBD
- TSC positions open to all interested parties



https://www.lfenergy.org/

Summary of Results

Activities

1. Operational support
2. Open issues/resolution
3. LF Energy commercialization

Online documentation

- http://help.openfido.org/
- https://auth.openfido.org/apidocs/
- https://api.openfido.org/apidocs/

Online tutorials

1. Getting Started
2. Use Case Tutorial
3. New Developer Build
4. Loadshape Analysis
5. Loadshape Synthesis
6. Weather Forecast
7. Resilience
8. Electrification
9. Tariff Design
10. Cyme Converter

Summary of Results

Online/cloud platform (AWS)

- Registered/invited users only
- https://app.openfido.org/

On-premise/private cloud platform

- Testing in progress at SCE

Local/private workstation platform

- Deployed via OpenFIDO CLI
- Available to general public
- **https://source.openfido.org/cli**

Use-case Pipelines

1. CYME model extractor
2. GridLAB-D (converters, geodata, tools)
3. Integration Capacity Analysis
4. Electrification
5. Tariff analysis
6. Resilience analysis (DOE/GMLC)
7. Wildfire analysis (DOE/CESER/DHS)
8. NERC load composition (DOE/OE)

Summary of Results

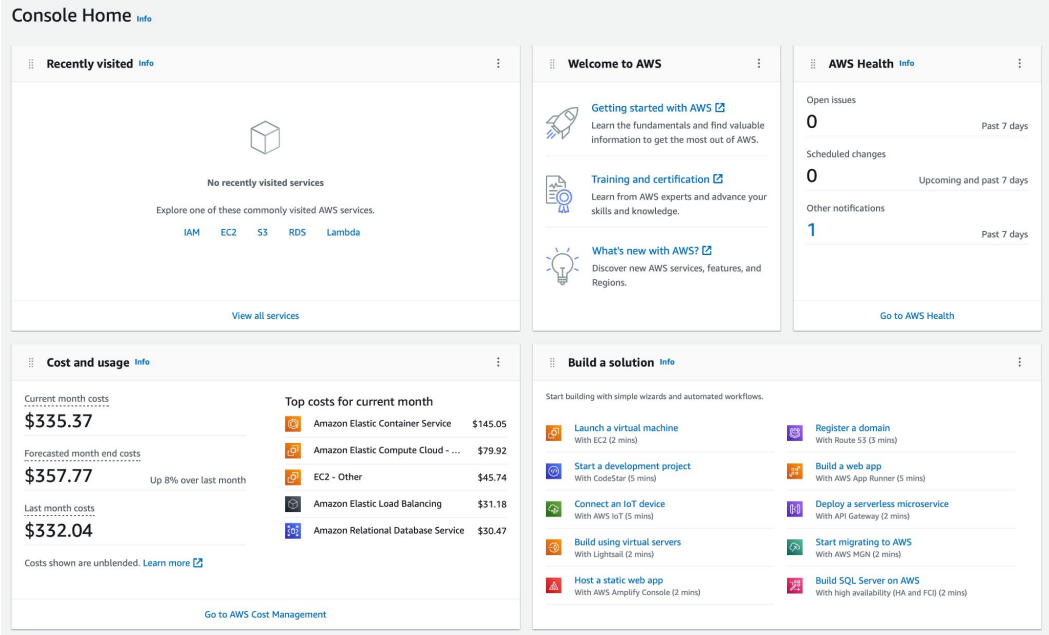
Two deployments on AWS

- Production (for users)

  app.openfido.org

- Staging (for developers)

  app-staging.openfido.org



Conclusions and Moving Forward

Support for major use-cases

- ICA/Hosting capacity analysis
- End-use electrification analysis
- Tariff design analysis
- Resilience analysis

Other use-cases supported

- CYME to GLM conversion
- Forecasting (weather, load)
- Data access (vegetation)

Validation completed

- Forecasting load at National Grid
- Resilience analysis at SCE
- Tariff and electrification

Pathway to commercialization

- Adopted by Linux Foundation
- First TSC to be held summer '23

Conclusions and Future Work

SLAC

Identified issues

1. Pipeline start dialog

   Users need more guidance about data/files required to perform complex analyses

2. Visualization dialog

   Lacks options to diverse CSV outputs (e.g., transpose, data format)

3. Lack of clear diagnostic outputs

   Output messages are often cryptic and lack details on how to fix the problem found

Resolution status

→ Form-based pipeline start dialog

Pipelines present input form that generates files needed to run

→ Upgraded visualization dialog

Visualization dialog will need to support more options and formats

→ Improved diagnostic output

Need pipelines to follow error message troubleshooting guides

Conclusions and Future Work

Load decomposition

End-use loadshapes from SCADA and AMI data

Census data

Customer demographic, economic, and population data for load modeling

Advanced building loads

Satellite imagery to generate/link building load models to network

Grid resilience analysis

Standardize PSPS optimization method,

Support long term climate impacts analysis

Advanced workflows

Capability built-in but not publicly available needs to be connected to standard tools such as Apache BEAM

Shared data/model security

Security enhancements to share artifacts with state agencies and utility partners

Legacy data handling/aging

Access to long-term historical data repositories for weather, assets, usage, prices, etc.

Thank you!

SLAC

Questions and Discussion

Contact	David Chassin (dchassin@slac.stanford.edu)
Website	https://github.com/openfido
Online	https://app.openfido.org/