ACCELERATING GLOBAL ENERGY SYSTEMS RESEARCH WITH OPEN ACCESS TO SYNTHETIC ENERGY DATA

Centre for Net Zero

Powered by **Octopus Energy**

INTRODUCTIONS

GUS CHADNEY

Data Lead

SHENG CHAI

Senior Data Scientist

CENTRE FOR NET ZERO

An impact-driven research unit founded by Octopus Energy

THE DEMAND-SIDE CHALLENGE

DEMAND FLEXIBILITY IS ESSENTIAL

The massive uptake of intermittent renewable energy sources will result in a need of **500 GW** of **demand flexibility** globally by 2030, according to the International Energy Agency.

HOUSEHOLD CONSUMPTION IS KEY

As heat and transport electrify, we need to understand **household consumption** intimately in order to predict **usage** and optimise **flexibility**

WE NEED SMART METER DATA

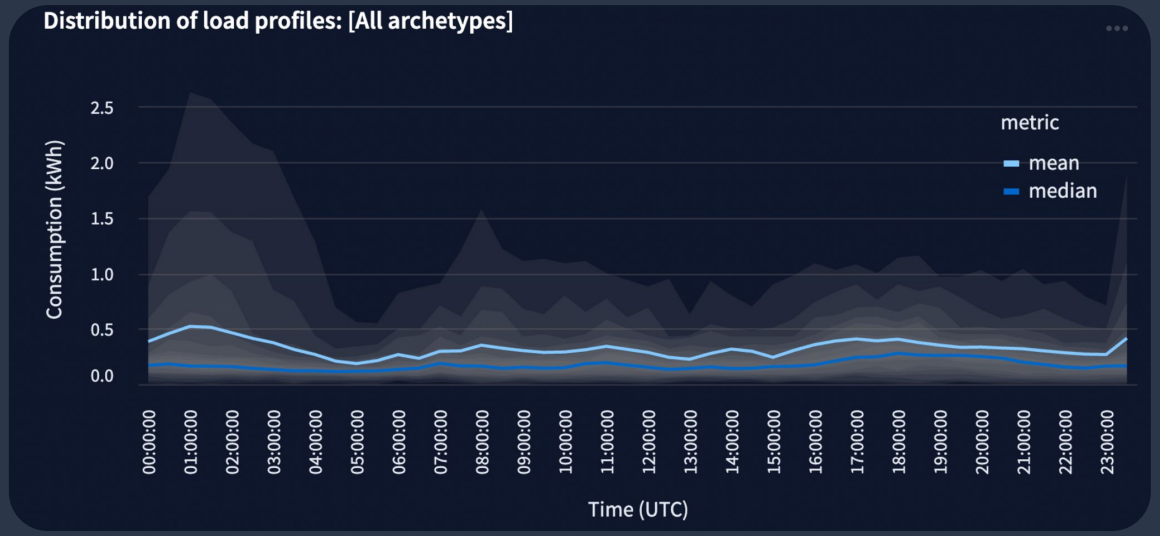
Granular **smart meter data** will unlock pioneering **research** and **innovative data products** to plan for electrification and unlock demand flexibility

ACCESS TO RAW SMART METER DATA IS ESSENTIAL FOR ENERGY RESEARCH

SYNTHETIC DATA ALLEVIATES CONSUMER PRIVACY ISSUES

AN OPEN COMMUNITY FOR SYNTHETIC SMART METER DATA WILL ACCELERATE RESEARCH EFFORTS

CNZ FARADAY



CUTTING-EDGE TECHNIQUES

Uses a combination of **Variational Autoencoders** (VAEs) and **Gaussian Mixture Model** (GMM) to provide best in class synthetic data

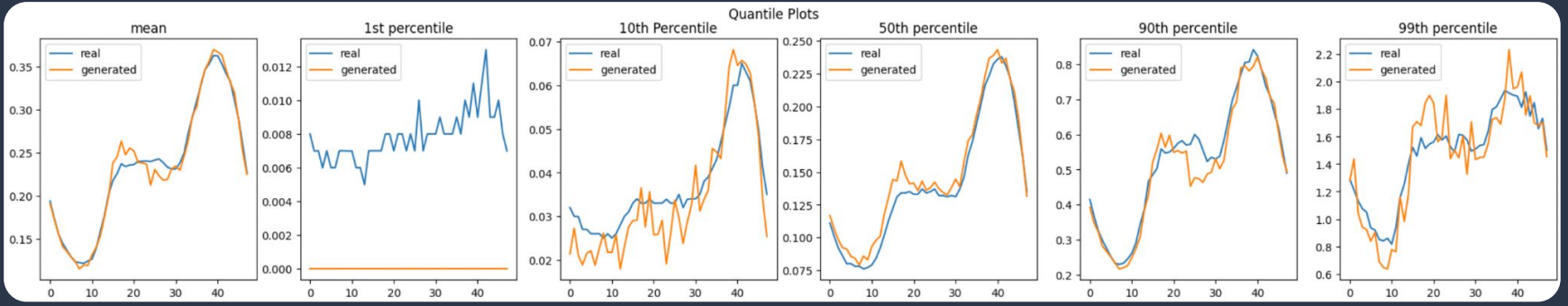
TRAINED ON REAL-WORLD DATA

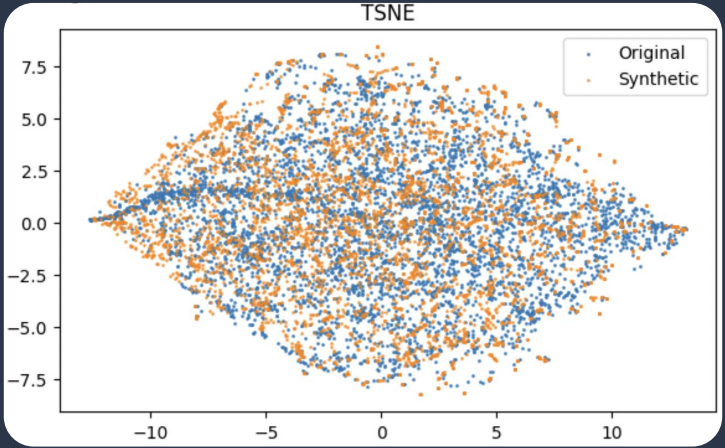
Faraday was trained on **7 million** day profiles over a **1 year** period from **20K** Octopus Energy UK households.

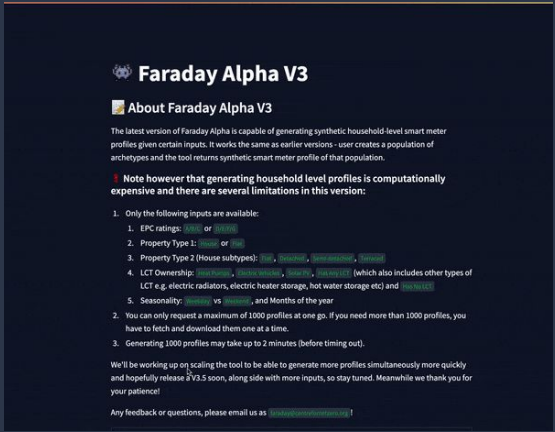
SUPPORTS ARCHETYPES

Household profiles can be generated with different **LCT** mixtures, **seasonality** and **EPC** ratings

CNZ FARADAY







Faraday Alpha V3

About Faraday Alpha V3

The latest version of Faraday Alpha is capable of generating synthetic household-level smart meter profiles given certain inputs. It works the same as earlier versions - user creates a population of archetypes and the tool returns synthetic smart meter profile of that population.

Note however that generating household level profiles is computationally expensive and there are several limitations in this version:

1. Only the following inputs are available:
   1. EPC ratings: [redacted] or [redacted]
   2. Property Type 1: [redacted] or [redacted]
   3. Property Type 2 (House subtypes): [redacted]
   4. LCT Ownership: [redacted] (which also includes other types of LCT e.g. electric radiators, electric heater storage, hot water storage etc) and [redacted]
   5. Seasonality: [redacted] vs [redacted], and Months of the year
2. You can only request a maximum of 1000 profiles at one go. If you need more than 1000 profiles, you have to fetch and download them one at a time.
3. Generating 1000 profiles may take up to 2 minutes (before timing out).

We'll be working up on scaling the tool to be able to generate more profiles simultaneously more quickly and hopefully release a V3.5 soon, along side with more inputs, so stay tuned. Meanwhile we thank you for your patience!

Any feedback or questions, please email us as [redacted]!

USE CASES

CURRENT

- TEED Digitisation Project by University of Birmingham
- Better Home Leeds Project by ARUP
- Commercial research projects by industry consultancies such as Parity Projects and Turley
- Other academic research projects by PhDs and Postdocs from University of Manchester and King's College London

POTENTIAL

- Regional, national and global grid “digital twins”
- Future energy system simulations
- Designing smart tariffs
- Greenfield grid design
- Extreme weather resilience planning
- Scenario planning

WHY AN OPEN COMMUNITY?

STANDARDISATION

We would like to drive consensus on what “good” looks like for synthetic smart meter data, ensuring quality and privacy

COMPETITION

The performance and ability of the generative algorithms will increase massively if contribution is open

VARIABILITY

Consumption profiles vary globally, multiple contributors will ensure we capture all edge cases for research

VOLUME

Synthetic smart meter data needs to be generated at scale, open-sourcing the algorithms will encourage all holders of real data to do this

SYNTHETIC DATA ECOSYSTEM

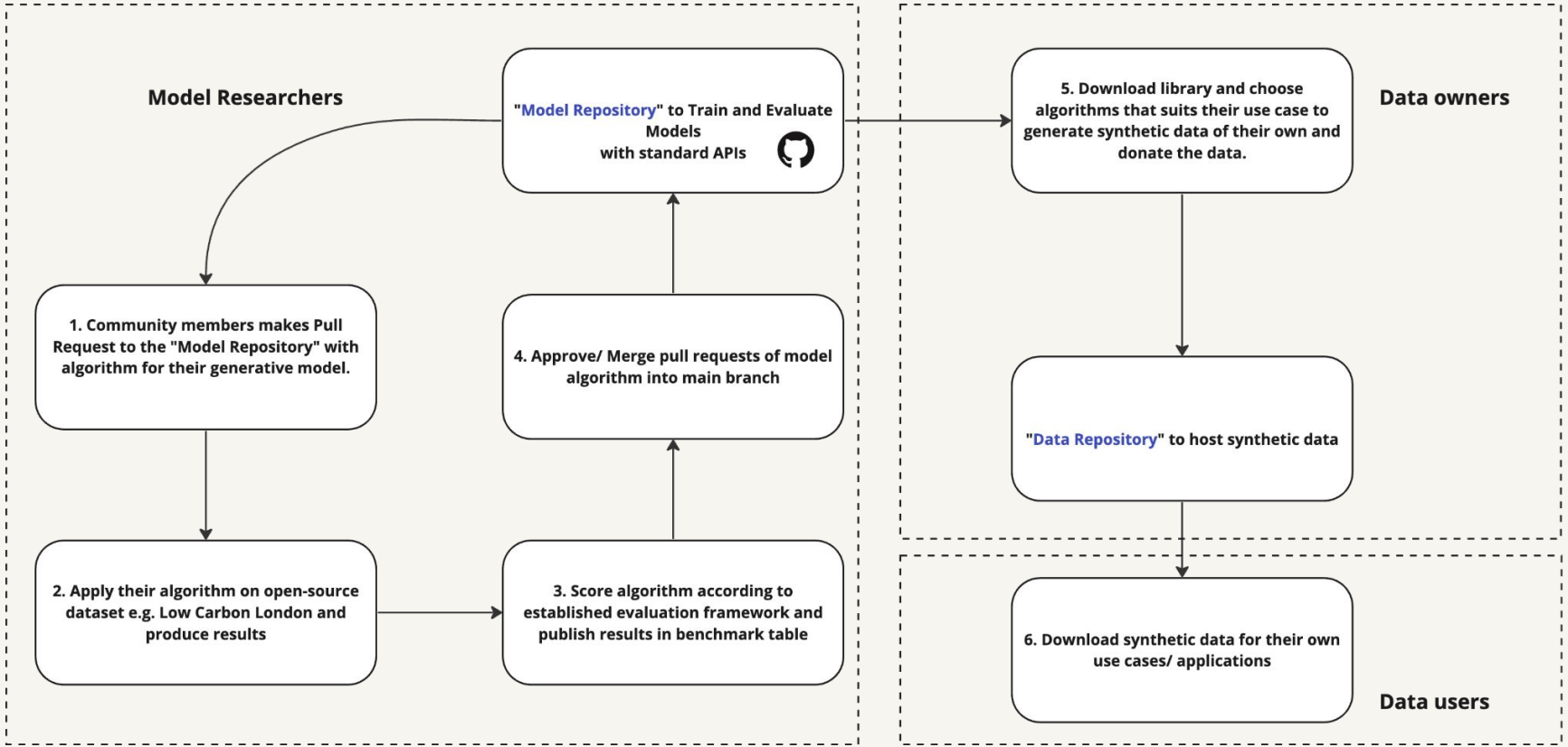
MODEL REPOSITORY

- Standardised APIs / framework to enable:
  - Model training with varied algorithms on arbitrary data sets
  - Evaluation of models to benchmark consistently and ensure quality
- Host algorithm / code for generative models that are vetted against a common evaluation framework
- Community can contribute towards algorithm / evaluation framework as research in the area progresses

DATA REPOSITORY

- Data owners can download algorithm/ code from “Model Repository” to train on their proprietary data to generate synthetic data
- Data owners can donate synthetic data to a Data Repository

SYNTHETIC DATA ECOSYSTEM



WHY LF ENERGY?

COMMUNITY

Leverage LF Energy's expertise creating and growing open-source communities

GOVERNANCE

Build the management framework to ensure quality controls and instil confidence

LICENSING

Navigate and implement correct licenses for usage of software and data

MARKETING

Outbound marketing support to grow community with workshops and events

NEXT STEPS

DEFINITION OF GOOD

We will be publishing a **technical paper** defining the **definition of good** that looks at fidelity, utility and privacy metrics

CONTINUED DEVELOPMENT

We will continue to improve our own generator **Faraday**, as well as lay the groundwork for the **synthetic data ecosystem**

OUTREACH

We will be building up our contact book of **interested parties**, and plan small, focused **workshops**

THANK YOU

Centre for Net Zero

Powered by **Octopus Energy**