

## Ofgem Consultation

### Consumer Consent Solution

#### Centre for Net Zero Response

Sent by email to: [digitalisation@ofgem.gov.uk](mailto:digitalisation@ofgem.gov.uk)

Centre for Net Zero (CNZ) is part of the world-leading group of organisations that comprise Octopus Energy Group. We are an autonomous, not-for-profit organisation that delivers pioneering research to transform future energy systems, in particular by leveraging the Octopus Energy customer dataset. A digitalised and data-driven energy sector is of central importance to our mission.

We have focused our response to align with the areas we have relevant research - namely, the use of Gen-AI to transform data-sharing. For example, we are not sharing any views on the delivery body or funding mechanism for the consent solution. We are happy to discuss our response, and share any further information directly.

We agree with the rationale behind Ofgem's work in this area. Data is a critical enabler for a smart, digital and low-carbon energy system with flexibility at its centre. There will be growing opportunities to maximise the use of energy data across all services and markets. Consumers will increasingly be asked to give, manage, and revoke consent to share their personal smart meter data with a number of actors in the energy sector (and potentially beyond). It is essential to have a transparent and robust process for that. Consumers must see who has access to their data and why, which is crucial to protect privacy and engender public trust. They should be able to opt in or out access to their assets and easily switch between suppliers and other providers, which may operate different technologies in a home. We agree that the energy industry can learn from the successes of other sectors, especially Open Banking.

While we agree with the overall direction of travel towards a less centralised data-sharing model than previously presented (i.e. a central repository), we would continue to advise against an overly centralised design (in response to consultation question 4). In Open Banking, data sharing occurs without a central platform, with the customer authenticating on both the third party app and their personal banking app. There is no additional central data-sharing service that you need to authenticate against. There are number of potential risks we encourage with a centralised design:

- a single point of failure (as Ofgem have identified), which risks inadequate protection of customer data (in terms of both privacy and cybersecurity)
- inflexible design, not agile to changes in a rapidly changing system or different use cases that develop over time.
- likely longer delivery times and higher costs (though we note that we do not yet have a sense of the costs involved with this proposal).

The risk of inadequate consumer protection should be taken seriously as it risks having the opposite effect than intended. It could add to further distrust in the energy

industry if the centralised solution allows a broader range of parties to have access to a consumer's personal data, beyond those that they have explicitly and knowingly given consent to. A consumer consent platform could also add another layer of complexity to consumer journeys (e.g. having to consent and re-consent to data-sharing), such as adopting low-carbon technologies or time-of-use tariffs, which we need to be as seamless as possible to accelerate adoption.

Furthermore, any solution to democratise data access based solely on consumer consent will be limited, given the customer has relatively little incentive to agree to share their personal data. The energy sector is different from banking in that high visibility across the system will be a critical part of grid modelling, network planning, and innovation for flexibility. In particular, as energy systems electrify and rely on renewable generation from more varied sources, in multidirectional energy flows, it is increasingly important to accurately profile and actively manage demand. Good quality and *granular* demand data is therefore highly valuable to researchers, grid operators and innovators. We need to consider specific use cases for energy data, which will determine the data-sharing solution.

**We urge Ofgem to, alongside the development of transparent and robust consumer consent processes, consider alternative ways to democratise access to smart meter data. As CNZ has raised previously, we think AI-generated synthetic data is a solution that can achieve this more easily, safely and quickly.**

Synthetic smart meter data can provide realistic profiles for consumer archetypes that cannot be attributed to individuals. CNZ has begun to generate and share this data ourselves, using an open source model - see the annex for further detail. This approach avoids privacy concerns of releasing raw data, and mitigates delivery risks around developing lengthy or complex data-sharing agreements. Even accepting the treatment of aggregated smart meter data as presumed open, GDPR principles, security controls and technical barriers are likely to continue to limit the smart meter data that suppliers and networks can share.

Given the diversity in consumers, buildings, technologies, and grid conditions, to be useful these demand profiles must be granular, i.e. household-level, with at least half-hourly resolution and combined with useful metadata such as low-carbon technology ownership. Real data, combined with useful metadata, cannot be released at this level of granularity without privacy risks.

While some uses of smart meter data are inherently specific to an individual consumer - such as billing, customer support, or a household understanding its own energy use - for many use cases we simply need realistic profiles for consumer archetypes. We have not yet explored the full potential of synthetic data, but our research already shows immense value for two broad purposes:

- a) Grid modelling and demand forecasting with accurate and granular smart meter data - from improved energy system modelling, to strategic network planning, to sizing grid connections for new housing developments.

- b) Innovation to develop new products, services and business models using smart meter data - for example, new smart tariffs or green finance products.

On innovation, the use of synthetic data could include providing new services and advice to the individual consumer. Consider the indicative consumer journey set out in the consultation (p.28-30), which aims to provide a tailored offer to lower bills, based on the consumer's data. CNZ's model can generate daily load profiles consisting of half-hourly kWh consumption for a given set of user-specified inputs, which currently include low carbon technology, property type, and season (with other inputs such as tariff type in development). Consumer advice could therefore be provided based on different consumer archetypes, or with direct user input from customers about their home without sharing actual consumption data (e.g. LCT ownership, which over time may become more automated through DESNZ's Asset Registration workstream). A provider can show the impact on demand and therefore bills - for example, the best tariff, or the impact of adopting a certain technology. It should be noted that any bill estimates - real data or not - are highly sensitive to factors such as changes in consumption patterns and prices, and will come with a host of caveats.

At the least, synthetic smart meter data is a useful complement to address limits to the data gleaned through this consumer consent solution. including:

- Datasets need to go beyond the individual consumer, but granular enough to profile household-level demand (including by combining with metadata)
- Gaps in real data, given there is little incentives for consumers to share their data
- Biases in real data towards those more willing to share it.
- Pace - synthetic data is available today (we published a sample based on the Octopus customer dataset a month ago), whereas your consultation suggests 2027 as a potential deployment date.

We encourage Ofgem to consider synthetic data as part of broader work exploring data-sharing solutions for the energy sector, which should join up with relevant other workstreams as much as possible (e.g. asset registration). We welcome further discussion on this and look forward to doing so at the CNZ-hosted public sector roundtable shortly, including to consider where we can learn from other sectors.

## Annex: Existing work to generate and scale synthetic smart meter data

CNZ has developed a generative AI model, [Faraday](#), trained on the Octopus Energy customer dataset to produce synthetic smart meter data. Based on Conditional Variational Auto-encoder and Gaussian Mixture Model, it generates daily load profiles consisting of half-hourly kWh consumption for a given set of user-specified inputs, which currently include low carbon technology, property type, and season.

Faraday is already being used by a number of alpha testers in research:

- 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.

Faraday is freely available as open source, and is not a commercial product. We share the synthetic data it generates with third parties, and other data owners can use the Faraday algorithm to train models on their own proprietary data to produce more synthetic data. Opening up access to synthetic data generation tools will help synthetic smart meter data to be generated at scale. We would like to work towards consensus of what “good” looks like for synthetic smart meter data, ensuring quality and privacy, and ensuring it can scale quickly.

To that end, in collaboration with Linux Foundation Energy, CNZ has also established an international data community, [OpenSynth](#), to facilitate the sharing of synthetic smart meter data and algorithms so it can scale quickly, particularly in areas with limited data access. This is an open-source one-stop-shop for synthetic smart meter data, including:

- a) Model repository to host open-source algorithms for generating synthetic data
- b) Data repository to host synthetic data sets contributed by data owners (generated by models trained on their own raw proprietary data)
- c) A community for quality assurance (in particular to agree a common evaluation framework against which data is vetted to ensure privacy protection), sharing best practice and ideating use cases.

To date, Centre for Net Zero has:

- Released Faraday to alpha testers for research (e.g. [TEED Digitisation Project](#) by University of Birmingham, Better Home Leeds Project by ARUP, other academic projects by University of Manchester and King's College London)
- Published a short [paper of our own Faraday model](#) ([presented](#) at the “Tackling Climate Change with Machine Learning” Workshop at the 12th International Conference on Learning Representation 2024).

- Published "[Defining 'Good': Evaluation Framework for Synthetic Smart Meter Data](#). (jointly with leading academics from Georgia Tech, MIT, and Oxford University) proposing a common evaluation framework to benchmark algorithms for synthetic smart meter data based on three tests: fidelity, utility and privacy.
- Set up the [OpenSynth Model Repository](#) and released our own source code for Faraday's algorithm, with a basic notebook tutorial on how to use the code.
- Released a sample of our synthetic Octopus Energy dataset to the community, which is publicly available [here](#).

In the next 6 months CNZ plans to:

- Develop Faraday - continue to improve the model, including to retrain it on a national representative dataset; version 4 will be released shortly.
- Open up access - 'OpenSynth' data community will generate, share and improve more synthetic datasets
- Demonstrate real-world applications - release a paper on the potential role and specific use cases for synthetic smart meter data.