DOCKETED	
Docket Number:	25-EPIC-01
Project Title:	Electric Program Investment Charge 2026–2030 Investment Plan (EPIC 5)
TN #:	265432
Document Title:	Electric Power Research Institute, Inc. (EPRI) Comments - AI-Generated Customer Usage Data to Accelerate Building Decarbonization and Grid Modernization
Description:	N/A
Filer:	System
Organization:	Electric Power Research Institute, Inc. (EPRI)
Submitter Role:	Public
Submission Date:	8/8/2025 11:09:03 AM
Docketed Date:	8/8/2025

Comment Received From: Electric Power Research Institute, Inc. (EPRI)
Submitted On: 8/8/2025
Docket Number: 25-EPIC-01

AI-Generated Customer Usage Data to Accelerate Building Decarbonization and Grid Modernization

Additional submitted attachment is included below.

Electric Program Investment Charge 2026–2030 (EPIC 5) Research Concept Proposal Form

The California Energy Commission (CEC) is currently soliciting research concept ideas and other input for the Electric Program Investment Charge 2026–2030 (EPIC 5) Investment Plan. For those who would like to submit an idea for consideration, please complete this form and submit it to the CEC by **August 8, 2025**. More information about EPIC 5 is available below.

To submit the form, please visit the e-commenting link: https://efiling.energy.ca.gov/EComment/ECommentSelectProceeding.aspx and select the Docket **25-EPIC-01**. Enter your contact information and then use the "choose file" button at the bottom of the page to upload and submit the completed form. Thank you in advance for your input.

---

1. Please provide the name, email, and phone number of the best person to contact should the CEC have additional questions regarding the research concept:

Electric Power Research Institute, Inc. (EPRI)
Alex Summe, asumme@epri.com, (919) 244-8772
Krish Gomatom, pgomatom@epri.com, (865) 218-8070
Brandon Johnson, bjjohnson@epri.com, (865) 218-8198

Centre for Net Zero (CNZ)
Gus Chadney, gus.chadney@centrefornetzero.org
Amber Woodward amber.m.woodward@centrefornetzero.org

2. Please provide the name of the contact person's organization or affiliation:

Electric Power Research Institute, Inc. in partnership with Center for NetZero

3. Please provide a brief description of the proposed concept that you would like the CEC to consider as part of the EPIC 5 Investment Plan. What is the purpose of the concept, and what would it seek to do? Why are EPIC funds needed to support the concept?

This proposed concept outlines a data-driven framework for generating and applying synthetic customer usage data to meet California's strategic energy goals as defined in the EPIC 5 Investment Plan. The approach uses AI and machine learning to generate statistically representative, privacy-preserving building-level datasets. By allowing this data to be shared openly, researchers, policymakers and utilities can evaluate the real-world impacts of California's Building Decarbonization, electrification, flexible demand management and net-zero carbon emission investment goals.

As the electric grid becomes more dynamic, driven by the adoption of distributed energy resources (DERs), electric vehicles (EVs), and climate dependent loads, realistic, resolute, and representative customer load data becomes critical for assessing grid impacts and forecasting demand. For utility managers, energy regulators and forecasting/planners' representative baseline and technology adoption profiles by end-use, building and customer segment is critical for impact evaluation and grid investment planning for electrification, de-carbonization, and flexible demand management programs.

With the emergence of advanced AI, trained generative models can produce synthetic customer/building load data inclusive of the above-mentioned intrinsic variable effects, which can be a game changer for impact evaluation and for optimization of investments. EPIC funds are being requested through this concept paper proposal to address this critical data and informational gap to create synthetic data frameworks and models to enable generation, utilization, and collaborative enhancements to the resolute characterization of California's customer and building usage portfolios representing real-world usage and impacts.

4. In accordance with Senate Bill 96[i], please describe how the proposed concept will "lead to technological advancement and breakthroughs to overcome barriers that prevent the achievement of the state's statutory energy goals." For example, what technical and/or market barriers or customer pain points would the proposed concept address that would lead to increased adoption of clean energy technology or innovation? Where possible, please provide specific cost and performance targets that need to be met for increased industry and consumer acceptance. For scientific analysis and tools, provide more information on what data and information gaps the proposed concept would help fill, and which specific parties or end users would benefit from the results, and for what purpose(s)?

Empirical data and load shapes are difficult and expensive to obtain due to data regulations around personally identifiable information (PII)/state and local regulations as well as the expense to collect through separate metering. Load data derived from physics-based simulation/modeling, though widely used does not capture customer behavior, building envelope inefficiencies, and other intrinsic variables that are difficult to model with a high degree of certainty. A key technical barrier for enabling technology adoption is the lack of high-quality, labeled, representative, usage data to assess strategic electric program investment impacts. For instance, estimates suggest the potential costs of meeting distribution grid infrastructure needs in California to be approximately $50 billion by 2035 if EV electricity needs are met exclusively with utility distribution assets without considering the flexibility these loads can provide (Kevala Inc., 2023). In other states, utilities have proposed adding nearly 50% capacity to the distribution network to account for load growth, requiring billions of dollars of ratepayer funded investment, without accounting the flexibility heat pumps, air conditioning units, or EVs can provide (Thakore, 2024). Without real-world reflective load data and profiles, it is difficult to evaluate the true impacts of decarbonization, electrification, flexible demand management and net-zero emission goals, and to optimize investments that produce outcomes with the desired cost to benefit. Synthetic data can mitigate this pain-point to produce resolute customer- or building-level smart meter data mitigating legal, contractual, and logistical barriers/challenges around data sharing and broader collaboration.

As of date, no AI-generated synthetic customer load data exists for U.S customer usage data. Synthetic load data frameworks and analytics are not just a technical advancement but a strategic investment for a prudent and responsible energy future. Various theoretical studies have applied generative AI technologies to create synthetic AMI data (Zhang et al., 2019; Liang and Wang, 2022; Gu et al., 2019) and various techniques have been used in the UK and Europe to create and openly share high quality time series data for energy systems (Centre for Net Zero, 2024; Lin, N., Palensky, P. and Vergara, P.P., 2024).

5. Please describe the anticipated outcomes if this research concept is successful, either fully or partially. For example, to what extent would the research reduce technology or ratepayer costs and/or increase performance to improve the overall value proposition of the technology? What is the potential of the innovation at scale? How will the innovation lead to ratepayer benefits in alignment with EPIC's guiding principles to improve safety,[ii] reliability,[iii] affordability,[iv] environmental sustainability,[v] and equity?[vi]

If Synthetic Data Architectures/Analytics are implemented at scale they offer a range of benefits across the economy, including but not limited to the following:

- Transport Electrification: Synthetic data enables detailed simulation of EV charging demand across household types, tariff structures, and control strategies. When combined with distribution network data, it helps identify where infrastructure upgrades are needed and where flexible charging can reduce peak loads and minimize costly grid reinforcements.

- Distributed Energy Resource Integration: By modelling households with solar, batteries, and EVs, synthetic data improves forecasting and real-time grid operations, while reducing the cost of data collection and management. It also supports financial institutions to evaluate payback periods for DER investments, helping to design sustainable lending products that broaden access to clean technologies.
- Building Decarbonization: Synthetic smart meter data allows for accurate modelling of how electrified buildings interact with the grid, supporting more precise connection requests and system planning. It can quantify both the demand and the flexibility potential of low-carbon technologies, helping prioritize upgrades and maximize consumer and system benefits.
- Achieving 100 Percent Net-Zero Carbon Emissions: Synthetic data can inform the design and evaluation of policies, standards, and regulatory interventions. It enables granular distributional analysis to ensure that decarbonization strategies equitably support the households and communities most in need.
- Consumer Products and Services: Synthetic data enables assessments for customer benefit and regulatory cost-effectiveness studies for integration into consumer-facing products to accelerate behind the meter technology adoption, conservation efficiency and flexible demand management providing consumers actionable tools and information feedback to manage their consumption and reduce energy/demand bills.
- Policy and Rate Making: Synthetic data improves distributional analysis when designing energy policies and ruling on rate cases, ensuring interventions target, or supporting those who most need it. An example application is scenario analyses of Building Decarbonization in housing development planning with electrification measures to characterize grid impact through electricity system and network modeling. Consumer products/services that enable flexible grid management is another scenario for analysis of electricity system and network modeling to assess regulatory design, net zero policy and rate making.
- Financial Products and Services: Alternative data sources can support risk assessment and portfolio resilience and enhance the development of sustainable lending products and services so that more people can afford to adopt Distributed Energy Resources.

6. Describe what quantitative or qualitative metrics or indicators would be used to evaluate the impacts of the proposed research concept.

   This research proposal will have a three-stage process involving metrics at each stage:

   1. Training and validation of an AI model using empirical data from the three California IOUs to generate household consumption profiles, evaluating for **fidelity**, **utility,** and **privacy** (Centre for Net Zero, 2024).
   2. Further validation of California IOU customer data against substation telemetry data or alternative simulated data, generating estimates for differences in electricity consumption (**kWh**), peak demand (**kW**)
   3. Estimation of grid and customer impacts for scenarios covering Building Decarbonization, electrification, DER integration and net-zero carbon emission goals for California's regions and climate zones (**$**)**.**

Once all stages have been completed, we would propose sharing the model algorithm, and the synthetic datasets openly, through public platforms to allow for researchers, policymakers, and utilities in California to benefit and enable further innovation for the benefit of ratepayers and the economy of California.

7. Please provide references to any information provided in the form that supports the research concept's merits. This can include references to cost targets, technical potential, market barriers, equity benefits, etc.

1) Centre for Net Zero (2024) Faraday: Synthetic Smart Meter Generator For the Smart Grid : Tackling Climate Change with Machine Learning, ICLR 2024 https://arxiv.org/abs/2404.04314

   Available at: https://www.centrefornetzero.org/papers/faraday-synthetic-smart-meter-generator-for-the-smart-grid (Accessed on: 7 August 2025)

2) Centre for Net Zero (2025) Policy paper: Synthetic Data for Smart Energy Synthetic Data for Smart Energy: Applications for AI-Generated Smart Meter Data Available at: https://www.centrefornetzero.org/papers/synthetic-data-for-smart-energy-applications-for-ai-generated-smart-meter-data

   (Accessed on: 7 August 2025)

3) Chai, S., Chadney, G., Avery, C., Grunewald, P., Van Hentenryck, P., & Donti, P. L. (2024). Defining'Good': Evaluation Framework for Synthetic Smart Meter Data. Available at: *arXiv preprint arXiv:2407.11785* (Accessed on: 7 August 2025)

4) Linux Foundation Energy (2025) OpenSynth. Available at: https://lfenergy.org/projects/opensynth/
5) Lin, N., Palensky, P., Vergara, P.P.P. (2024) EnergyDiff: Universal Time-Series Energy Data Generation using Diffusion Models https://arxiv.org/abs/2407.13538 (Accessed on: 24 February 2025)
6) CPUC Data Access Working Group Workshop 6 - 1-27-2025 Customer Usage Data and Customer Privacy https://ucla.app.box.com/s/nrhfwa9otsvjpi0acyhe5pz46x3ifs3c
7) Thakore, I. (2024) Xcel Energy wants a nearly $5 billion grid upgrade to make room for heat pumps and EVs. CPR News. Available at: https://www.cpr.org/2024/12/17/xcel-energy-proposal-5-billion-gridupgrade/ (Accessed: 12 March 2025)
8) Kevala Inc., "Electrification Impacts Study Part 1: Bottom-Up Load Forecasting and SystemLevel Electrification Impacts Cost Estimates," 9 May 2023. [Online]. Available: https://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M508/K423/508423247.PDF.

8. The EPIC 5 Investment Plan must support at least one of five Strategic Goals:[vii]
    a. Transportation Electrification
    b. Distributed Energy Resource Integration
    c. Building Decarbonization
    d. Achieving 100 Percent Net-Zero Carbon Emissions and the Coordinated Role of Gas
    e. Climate Adaptation

    Please describe in as much detail as possible how your proposed concept would support these goals.

As outlined above, the development and use of synthetic data frameworks for assessment and scenario-based analyses of electric program investments has a broad range of potential benefits, in alignment with the EPIC 5 strategic goals. Further, there is alignment with the CPUC's DER Action Plan 2.0 which includes the vision that "data from smart meters and other ratepayer-funded "smart" devices is available for research purposes while retaining privacy protections and is used to improve program design and marketing."

About EPIC

The CEC is one of four EPIC administrators, funding research, development, and demonstrations of clean energy technologies and approaches that will benefit electricity ratepayers of California's three largest investor-owned electric utilities.

EPIC is funded by California utility customers under the auspices of the California Public Utilities Commission.

To learn more about EPIC, visit: https://www.energy.ca.gov/programs-and-topics/programs/electric-program-investment-charge-epic-program

EPIC 5 documents and event notices will be posted to: https://www.energy.ca.gov/proceeding/electric-program-investment-charge-2026-2030-investment-plan-epic-5

Subscribe to the EPIC mailing list to stay informed about future opportunities to inform the development of EPIC 5: https://public.govdelivery.com/accounts/CNRA/signup/31897

---

i See section (a) (1) of Public Resources Code 25711.5 at: https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=PRC&sectionNum=25711.5.

ii EPIC innovations should improve the safety of operation of California's electric system in the face of climate change, wildfire, and emerging challenges.

iii EPIC innovations should increase the reliability of California's electric system while continuing to decarbonize California's electric power supply.

iv EPIC innovations should fund electric sector technologies and approaches that lower California electric rates and ratepayer costs and help enable the equitable adoption of clean energy technologies.

v EPIC innovations should continue to reduce greenhouse house gas emissions, criteria pollutant emissions, and the overall environmental impacts of California's electric system, including land and water use.

vi EPIC innovations should increasingly support, benefit, and engage disadvantaged vulnerable California communities (DVC). (D.20-08-046, Ordering Paragraph 1.) DVCs consist of communities in the 25 percent highest scoring census tracts according to the most recent version of the California Communities Environmental Health Screening Tool (CalEnviroScreen), as well as all California tribal lands, census tracts with median household incomes less than 60 percent of state median income, and census tracts that score in the highest 5 percent of Pollution Burden within CalEnviroScreen, but do not receive an overall CalEnviroScreen score due to unreliable public health and socioeconomic data.

vii In 2024 the CPUC adopted five Strategic Goals to guide development of the EPIC 5 Investment Plan. A description of the goals can be seen in Appendix A of CPUC Decision 24-03-007 available at: https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M527/K228/527228647.PDF