ORIGINAL RESEARCH **Open access**

Toward a Human-Centered AI Architecture for Smart Energy Transitions: Integrating Prosumer Behavior, Grid Intelligence, and Institutional Constraints

Hans Müller[1*], Anna Schmidt[1]

Abstract

Smart energy transitions increasingly embed artificial intelligence into grid operations, demand-side management, and prosumer coordination. However, many AI designs remain technology-centric, neglecting human behavior, user cognition, and institutional embedding. Current AI architectures for smart grids and energy systems often treat human actors as passive or perfectly rational, failing to account for prosumer heterogeneity, bounded rationality, social influence, and behavioral inertia. This gap reduces the effectiveness, acceptance, and legitimacy of AI-enabled energy services. This article proposes a human-centered AI architecture for smart energy transitions that explicitly integrates prosumer behavior, grid intelligence capabilities, and institutional constraints into a single design framework. The architecture is developed through conceptual synthesis and design-oriented reasoning, drawing on human-centered AI, behavioral energy research, smart grid informatics, and institutional analysis of energy systems. The proposed architecture organizes AI functions into three interconnected layers: a prosumer behavior and interaction layer, a grid intelligence and automation layer, and an institutional governance and constraint layer. It shows how these layers can be linked through interoperable interfaces, feedback loops, and human-in-the-loop mechanisms.A human-centered AI architecture can align technical grid optimization with the lived realities of energy users and the regulatory frameworks shaping smart energy transitions. It offers a design compass for more inclusive, adaptive, and societally robust AI-enabled energy systems.

Keywords Grid intelligence, Human-centered artificial intelligence, Smart energy transitions, Prosumer behavior, Institutional constraints, Human-AI interaction

*Correspondence:
Hans Müller
hans.mueller@rwth-aachen.de
[1] Department of Artificial Intelligence for Power System Optimization, Faculty of Electrical Engineering, RWTH Aachen University, Aachen, Germany

Introduction

Smart energy transitions are accelerating the integration of artificial intelligence into electricity systems, especially where distributed energy resources, flexible loads, smart meters, and prosumer participation create complex coordination problems. AI is increasingly positioned as a means of improving forecasting, optimization, automation, and demand-side response, as shown in reviews of AI-enabled sustainable energy systems and demand-response intelligence [1]. Yet architectures that prioritize technical optimization alone risk treating households, communities, operators, and regulators as external variables rather than central actors in the system. A human-centered AI architecture for energy transitions should therefore begin from the premise that grid intelligence is also a social, behavioral, and institutional design challenge.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

The mismatch between technology-centric AI design and real prosumer behavior is especially important because prosumers are not simply economic optimizers responding predictably to prices or automation signals. Social norms can shape electricity shifting and conservation, while feedback systems influence energy practices only when they are meaningful within everyday routines [2]. Smart meter acceptance research further shows that perceived usefulness, trust, and user understanding affect whether households accept digital energy technologies [3]. Ignoring these behavioral dimensions could lead to user disengagement, algorithmic distrust, and energy outcomes that appear technically rational but remain socially fragile.

Human-centered AI in smart energy transitions must therefore involve more than interface usability or attractive dashboards. Explainable AI methods in energy systems suggest that transparency is necessary when algorithmic recommendations affect operational and household decisions [4]. At the same time, governance-oriented work on AI in smart energy systems indicates that explainability must be connected to accountability, institutional rules, and user protection [5]. A genuinely human-centered architecture would embed behavioral models, co-adaptive learning, human oversight, and institutional awareness into the structure of AI-enabled energy services.

This article proposes a conceptual architecture that integrates prosumer behavior, grid intelligence, and institutional constraints into a coherent design framework. The architecture draws on AI-enabled demand response, smart grid optimization, and prosumer-centric market research to specify how behavioral interaction, automation, and governance layers could be linked [6]. It also builds on scholarship on collective prosumers and regulatory challenges to argue that institutional compatibility is a design parameter rather than an afterthought [7]. The article develops foundational principles, specifies architectural components, explains human-AI interaction mechanisms, and prepares the basis for governance-sensitive implementation pathways.

Background

Smart energy transitions and the prosumer shift

Smart energy transitions are transforming centralized electricity systems into decentralized, renewable-based, data-intensive systems in which consumers increasingly become prosumers who generate, store, share, and flexibly consume energy. Reviews of prosumerism show that this shift involves technological change, business model innovation, and new forms of participation in sustainable energy development [8]. Collective renewable energy prosumers further illustrate how participation can extend beyond individual households toward communities and local energy initiatives [7]. Because these systems depend on active engagement as well as digital coordination, AI architectures must connect data-driven grid intelligence with behavioral participation and institutional feasibility.

Artificial intelligence in grid intelligence and energy management

AI contributes to smart grid intelligence through load and generation forecasting, distributed energy resource coordination, anomaly detection, demand response scheduling, market operation, and automated control. Demand-side response research shows that machine learning can support flexible consumption and coordination across residential and commercial energy users [6]. AI-enabled demand response frameworks also highlight prospects for optimization in smart grid environments while acknowledging operational and implementation challenges [9]. However, these functions are often framed primarily as engineering capabilities, which means the human meaning of automation, advice, control, and consent can remain underdeveloped.

Human-centered AI as a design paradigm

Human-centered AI can be conceptualized as a design paradigm that preserves human agency, supports understandable decisions, adapts to context, and aligns algorithmic action with social values. In energy systems, explainable AI research demonstrates that interpretability is not merely technical but also relational, because different stakeholders need different

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

forms of explanation [4]. Governance-focused AI scholarship further suggests that accountability and institutional oversight are essential when AI systems influence energy access, costs, and participation opportunities [5]. For smart energy transitions, human-centered AI should therefore be understood as an architectural commitment to agency, transparency, fairness, and contextual fit.

Behavioral models of prosumer energy decisions

Prosumer energy decisions are shaped by bounded rationality, social norms, habitual routines, mental models, temporal preferences, and feedback interpretation. Studies of electricity-related social norms show that collective expectations can influence both reduction and load shifting, indicating that behavior is socially embedded rather than purely price-driven [2]. Research on home energy management systems and real-time feedback similarly shows that feedback must be situated within domestic routines to influence consumption meaningfully [10]. A human-centered AI architecture should therefore include behavioral representations that can adapt to heterogeneous and changing patterns of energy use without assuming perfect rationality.

Institutional constraints and the governance of AI in energy

AI-enabled energy systems operate within institutional landscapes shaped by market design, utility regulation, data governance, interoperability requirements, consumer protection, and grid reliability obligations. Regulatory analysis of collective renewable energy prosumers shows that even technically feasible forms of energy sharing may be constrained by electricity market rules and legal classifications [11]. Institutional contestation around smart grids also indicates that organizations may compete over authority, data access, and control in smart energy transitions [12]. Consequently, an architecture that ignores institutional constraints would be incomplete, because permissible AI actions are shaped by rules as much as by technical capacity.

Foundational Design Principles of the Architecture

Human agency and prosumer empowerment

The first design principle is that AI should augment rather than replace human decision-making in energy contexts. Peer-to-peer and community-based market research shows that prosumers may participate in energy exchange, local coordination, and value creation when systems enable meaningful choices rather than passive enrollment [13]. Federated power plant models similarly suggest that prosumer participation can be incentivized through platforms that coordinate distributed assets while preserving user relevance and engagement [14]. The architecture should therefore support informed choices about asset operation, flexibility participation, data sharing, and automated delegation.

Contextual adaptation and behavioral sensitivity

The second design principle is contextual adaptation, meaning that AI should be sensitive to diverse prosumer preferences, routines, capabilities, and trust levels. Research on smart home energy management indicates that real-time feedback and automated systems can influence household behavior only when they resonate with lived domestic practices [10]. Studies of innovative household and community feedback approaches further emphasize that engagement depends on communication forms, social context, and interpretability [15]. A human-centered architecture should therefore avoid one-size-fits-all optimization and instead provide adaptive interaction patterns that respect behavioral heterogeneity.

Institutional compatibility and normative alignment

The third design principle is institutional compatibility, which means that AI functions should be designed to operate within regulatory frameworks governing data access, market participation, consumer rights, and grid stability. Work on regulatory challenges for collective renewable energy prosumers demonstrates that institutional design can either enable or block

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

decentralized energy participation [11]. Analyses of prosumerism after subsidy regimes further show that business models and participation opportunities depend on policy and market structures [16]. In this architecture, institutional compatibility is not a downstream compliance task but a core normative design parameter.

Architecture Components

Prosumer behavior and interaction layer

The prosumer behavior and interaction layer models, learns from, and communicates with human energy actors, including households, communities, small businesses, and local energy collectives. It could combine smart meter signals, interaction logs, preference elicitation, feedback mechanisms, and behaviorally informed recommendation engines, while recognizing that data-driven profiles must remain interpretable and contestable [3]. It may also use conceptual or agent-based representations to explore how social norms, feedback, and flexibility preferences interact without claiming empirical validation in the absence of implementation studies [2]. This layer is the primary site where human agency, consent, explanation, and behavioral adaptation are translated into AI design.

Table 1 translates the manuscript's three-layer architecture into design requirements, AI functions, human-centered safeguards, and implementation risks.

Table 1. Design Translation Matrix for Human-Centered AI in Smart Energy Transitions

Architectural domain	Core design question	Required AI capability	Human-centered requirement	Institutional constraint	Failure mode if ignored	Design implication
Prosumer behavior modeling	How should AI represent heterogeneous energy users?	Preference learning, clustering, behavioral profiling, interaction-log analysis	Users must understand and revise how they are represented	Data minimization, consent, purpose limitation	Stereotyped users; inaccurate recommendations; disengagement	Use interpretable behavioral profiles with editable assumptions
Prosumer interaction	How should AI communicate energy advice?	Personalized feedback, recommendation engines, scenario simulation	Advice must be meaningful within routines, comfort needs, and cognitive limits	Consumer-protection rules and transparency duties	Technically optimal advice is ignored or distrusted	Convert optimization outputs into household-relevant explanations
Consent and delegation	When may AI automate action?	Delegation logic, automation thresholds, preference-based control	Users must choose, limit, pause, or reverse automation	Consent management, revocation rights, liability boundaries	Hidden automation; perceived loss of control	Provide visible delegation settings and reversible automation
Grid forecasting	How should uncertainty be translated?	Load, generation, price, and	Uncertainty must be expressed	Reliability obligations and	Abstract model confidence does	Present uncertainty as human-

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

		flexibility forecasting	through cost, comfort, reliability, and environmental implications	reporting requirements	not support decision-making	relevant trade-offs
Demand response coordination	How should flexibility be activated?	Demand-response optimization, scheduling, DER coordination	Participation should respect comfort, fairness, and opt-out preferences	Market rules, aggregator obligations, grid constraints	Flexibility extraction burdens users or vulnerable groups	Couple demand-response optimization with fairness and comfort thresholds
Peer-to-peer energy exchange	How should prosumers participate in local markets?	Matching, market clearing, transaction recommendation, flexibility forecasting	Users must express preferences and understand trade-offs	Market eligibility, network constraints, settlement rules	Technically possible trades become legally or socially illegitimate	Embed market rules and user preferences into transaction logic
Edge-cloud decision distribution	Where should AI decisions occur?	Edge analytics, cloud optimization, federated coordination	Local autonomy and privacy should shape computation placement	Cybersecurity, interoperability, and data-locality requirements	Over-centralization; privacy loss; weak responsiveness	Allocate functions by latency, privacy, autonomy, and explainability
Governance and auditability	How should AI decisions remain accountable?	Audit logs, decision tracing, compliance checking, model monitoring	Affected stakeholders must contest or review consequential decisions	Regulatory reporting, audit standards, institutional oversight	Opaque decisions; unclear responsibility; weak public trust	Build audit and contestability interfaces into the architecture from the start
Equity and vulnerability	Who benefits or bears burdens?	Bias detection, segmentation analysis, fairness-aware allocation	Vulnerable users must not be disadvantaged by pricing, automation, or flexibility demands	Energy justice, anti-discrimination, consumer vulnerability protections	Smart energy services privilege asset-owning or digitally skilled users	Treat inclusion as a system constraint, not an optional ethical add-on

Grid intelligence and automation layer

The grid intelligence and automation layer is responsible for technical functions such as forecasting, optimization, distributed energy resource coordination, fault detection, peer-to-peer market clearing, and automated control. Smart grid AI research indicates that such functions may support demand response, renewable integration, and operational

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

efficiency when embedded in advanced energy management systems [9]. Peer-to-peer trading and microgrid studies further show how optimization and market-clearing mechanisms can coordinate distributed prosumer resources [17]. In a human-centered architecture, however, this layer must generate outputs that are explainable, actionable, and governable rather than opaque technical commands.

Institutional governance and constraint layer

The institutional governance and constraint layer encodes regulatory rules, market boundaries, data governance policies, fairness criteria, interoperability requirements, and audit expectations. Research on blockchain in the energy sector highlights that digital platforms for energy exchange raise questions about governance, trust, regulation, and data handling [18]. Studies of decarbonization, digitalization, and decentralization also show that changing power infrastructures require institutional coordination alongside technical modernization [19]. This layer constrains the behavioral and grid layers by defining what AI systems may recommend, automate, share, optimize, or expose to human review.

Figure 1 presents the proposed human-centered AI architecture linking prosumer behavior, grid intelligence, and institutional constraints within a single sociotechnical design framework.

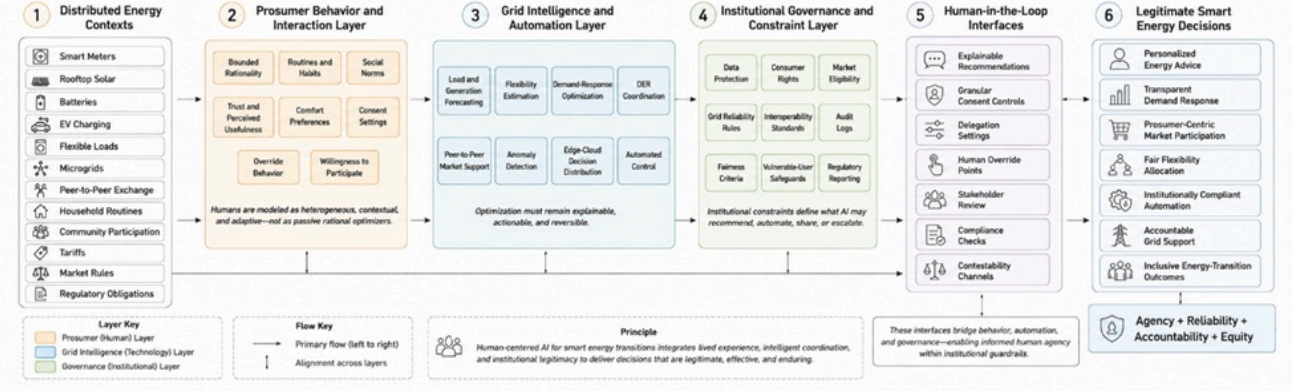


Figure 1. Human-centered AI architecture for smart energy transitions

Prosumer Behavior Integration and Human-Ai Interaction

Modeling heterogeneous prosumer behavior for AI design

The architecture can integrate heterogeneous prosumer behavior through rule-based profiles, preference models, data-driven clustering, and hybrid cognitive representations, while treating these models as design constructs rather than validated empirical findings. Prosumer pricing research shows that emergence and participation depend on conditions such as tariffs, incentives, and local system context, which implies that behavior cannot be represented through a single universal model [20]. Smart HVAC acceptance research also indicates that residential technology acceptance depends on perceived control, usefulness, and contextual fit [21]. Conceptually, behavioral modeling may support personalized recommendations, flexibility management, and more responsive interaction, provided that users can understand and revise how they are represented.

Co-adaptive human-AI interaction loops

Co-adaptive human-AI interaction loops allow AI recommendations and prosumer behavior to influence each other over time. Human-in-the-loop energy flexibility research suggests that flexibility integration at neighborhood scale requires systems that can combine data management with human participation and oversight [22]. In the proposed architecture, AI

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

could learn from user feedback, overrides, declined recommendations, and evolving preferences, while prosumers could learn from explanations, simulations, and energy feedback. This reciprocal learning process should be designed to support alignment rather than manipulation, ensuring that adaptation remains compatible with agency and consent.

Trust, transparency, and user acceptance mechanisms

Trust, transparency, and acceptance mechanisms are essential because AI-mediated energy decisions may affect comfort, cost, privacy, and participation in local energy markets. Smart meter acceptance research shows that trust and perceived usefulness influence whether households accept digital energy technologies [3]. Explainable AI research in energy systems further indicates that explanations should be tailored to stakeholder needs, including prosumers, operators, aggregators, and regulators [4]. The architecture should therefore include transparent explanations, granular consent controls, human override options, and clear feedback about the likely consequences of AI-supported energy decisions.

Table 2 summarizes the major human-centered interaction requirements and associated AI design responses within smart energy systems.

Table 2. Human-Centered Interaction Requirements for AI-Enabled Smart Energy Systems

Human-centered interaction issue	AI-related challenge	Human-centered architectural response	Expected sociotechnical benefit
User trust in automation	Opaque algorithmic recommendations and low interpretability	Explainable AI interfaces, transparent decision summaries, and user-readable feedback	Increased trust and informed participation
Perceived loss of control	Excessive or hidden automation of household energy assets	Human override mechanisms and adjustable delegation settings	Preservation of user agency and autonomy
Behavioral heterogeneity	Standardized optimization assumptions ignore diverse routines	Adaptive behavioral modeling and preference-sensitive recommendations	Better contextual fit and user acceptance
Cognitive overload	Complex energy data and technical terminology	Simplified explanations and routine-oriented visualization tools	Improved comprehension and engagement
Privacy concerns	Extensive smart meter and interaction-data collection	Consent management, data minimization, and privacy-preserving analytics	Greater confidence in digital participation
Participation inequality	Vulnerable or digitally excluded users benefit less from AI services	Inclusive design standards and fairness-aware optimization	More equitable energy transition outcomes
Distrust in local energy markets	Unclear transaction logic in peer-to-peer trading	Transparent market rules and auditable transaction histories	Enhanced legitimacy of decentralized coordination

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Grid Intelligence and AI Decision Layers

Forecasting and optimization under human-relevant uncertainty

The grid intelligence layer uses AI for load, generation, flexibility, and price forecasting, but a human-centered architecture should translate uncertainty into terms that matter to users and institutions. Reviews of AI in sustainable energy systems show that forecasting and optimization are central to intelligent energy management, yet these functions often emphasize technical metrics rather than human consequences [1]. Explainable AI techniques may help express uncertainty through expected bill effects, comfort implications, reliability risks, and environmental trade-offs rather than abstract model confidence scores [4]. This design orientation would allow prosumers, operators, and regulators to interpret AI outputs in relation to everyday choices and institutional responsibilities.

Table 3 presents a comparative mapping between conventional grid-centric AI optimization and the proposed human-centered AI approach.

Table 3. Comparison between Conventional Grid-Centric AI and Human-Centered AI Architectures in Smart Energy Systems

Design dimension	Conventional grid-centric AI approach	Proposed human-centered AI approach
Primary optimization goal	Technical efficiency and grid stability	Joint optimization of efficiency, agency, fairness, and institutional legitimacy
Representation of users	Passive consumers or rational actors	Heterogeneous prosumers with evolving preferences and routines
Human involvement	Minimal intervention after automation	Continuous human-in-the-loop interaction and oversight
Explainability	Technical outputs aimed at operators	Stakeholder-specific explanations for households, operators, and regulators
Treatment of uncertainty	Statistical confidence metrics	Human-relevant trade-offs involving comfort, cost, and reliability
Governance integration	External compliance checks after deployment	Embedded regulatory constraints and auditability mechanisms
Equity considerations	Often absent or secondary	Treated as a core architectural requirement
Data governance	Centralized data extraction and processing	Consent-based, privacy-aware, and contestable data management
Market participation	Efficiency-driven transaction matching	Preference-sensitive and institutionally constrained participation
Adaptation logic	System adapts mainly to operational conditions	Co-adaptive learning between AI systems and human actors

Automated coordination with human override points

Automated coordination of distributed energy resources, demand response, and virtual power plants can be designed with explicit human override points rather than full delegation by default. AI-enabled demand response research indicates that automation can improve coordination in smart grids, but such coordination also raises questions about user control, consent, and acceptable intervention [9]. Prosumer scheduling approaches based on intelligent edge computing further

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

show that automation may be distributed across household and grid interfaces, making override design especially important [23]. In this architecture, meaningful human control would require visible delegation settings, reversible automation, and escalation pathways when AI actions affect comfort, cost, or participation rights.

Peer-to-peer trading and prosumer-centric market facilitation

AI can support peer-to-peer energy trading by matching local generation and demand, clearing transactions, forecasting available flexibility, and helping prosumers express preferences. Peer-to-peer and community-based market research shows that such systems require careful market design as well as technical coordination [13]. Grid-influenced peer-to-peer trading further demonstrates that local transactions must remain compatible with network constraints and system reliability [24]. The proposed architecture therefore links the behavioral layer, where users define preferences and acceptable trade-offs, with the institutional layer, where fairness, market rules, and regulatory limits shape what transactions may occur.

Edge-cloud continuum and real-time decision distribution

The edge-cloud continuum determines where AI decisions are processed, from household devices and local gateways to aggregator platforms and cloud-based grid intelligence systems. Prosumer scheduling research suggests that intelligent edge computing may support more responsive and privacy-sensitive decisions close to the user [23]. Microgrid market design studies also show that local coordination mechanisms can support distributed energy exchange while interacting with wider system needs [25]. A human-centered architecture should therefore place functions according to latency, privacy, autonomy, explainability, and behavioral responsiveness rather than purely computational efficiency.

Institutional Constraints and Governance Interfaces

Regulatory embeddedness of AI functions

The architecture must embed regulatory constraints directly into AI processes so that automated recommendations and controls remain auditable and institutionally legitimate. Regulatory research on collective renewable energy prosumers shows that energy sharing, aggregation, and community participation depend on legal classifications and market rules [26]. Institutional contestation around smart grids further indicates that different organizations may dispute authority over data, infrastructure, and customer relationships [12]. Regulatory embeddedness should therefore be explicit in the architecture through constraint libraries, compliance logs, audit interfaces, and traceable decision rules.

Data governance, privacy, and prosumer agency

Data governance is central because human-centered AI in energy systems depends on flows of smart meter data, device data, preference data, transaction data, and interaction records. Blockchain energy research highlights both the promise of decentralized data-enabled coordination and the governance challenges associated with trust, privacy, and regulatory oversight [18]. AI governance work in smart energy systems similarly suggests that transparency and accountability are necessary when data-driven decisions affect households and energy markets [5]. The institutional layer should therefore manage consent, purpose limitation, anonymization, data portability, and revocation rights, while the behavioral layer provides controls that prosumers can understand and use.

Interoperability and multi-stakeholder alignment

Interoperability is necessary because AI-enabled smart energy services must connect devices, meters, market platforms, aggregators, utilities, grid operators, and regulatory reporting systems. Research on decarbonization, digitalization, and decentralization shows that key power infrastructures are being reshaped by interacting technical and organizational changes [19]. Conceptual work on smart energy systems further emphasizes the need for theoretical frameworks that connect emerging technologies, system integration, and digital coordination [27]. The proposed architecture treats

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

interoperability as both a technical and institutional problem, requiring shared data models, governance interfaces, and negotiated responsibilities among stakeholders.

Equity, vulnerability, and inclusive energy transitions

Institutional constraints related to equity and consumer vulnerability must be reflected directly in human-centered AI design. Normative analysis of prosumerism asks what decentralized energy participation is for, showing that prosumer systems may reproduce inequalities if they privilege technically capable or asset-owning households [28]. Collective prosumer research also indicates that community energy can broaden participation, but only when institutional conditions support inclusion rather than exclusion [7]. The architecture should therefore allow the behavioral layer to accommodate vulnerable users and require the grid layer to avoid discriminatory pricing, flexibility allocation, or automation burdens.

Recent scholarship further reinforces the importance of embedding justice, agency, and accountability into AI-enabled smart energy systems. Research on AI and energy justice argues that algorithmic management of electricity systems may unintentionally reproduce inequalities through opaque decision-making, discriminatory optimization, and uneven distribution of digital benefits unless governance safeguards are integrated into system design [29]. Systematic reviews of smart local energy systems similarly demonstrate that procedural justice, inclusion, and equitable participation are critical for ensuring that digital energy transitions remain socially legitimate and accessible beyond technologically privileged groups [30]. These findings strengthen the rationale for embedding fairness constraints, participatory oversight, and contestability mechanisms directly into the proposed human-centered AI architecture. **Figure 2** illustrates how equity, vulnerability, participatory oversight, and contestability can be embedded directly into human-centered AI architecture for inclusive prosumer energy systems.

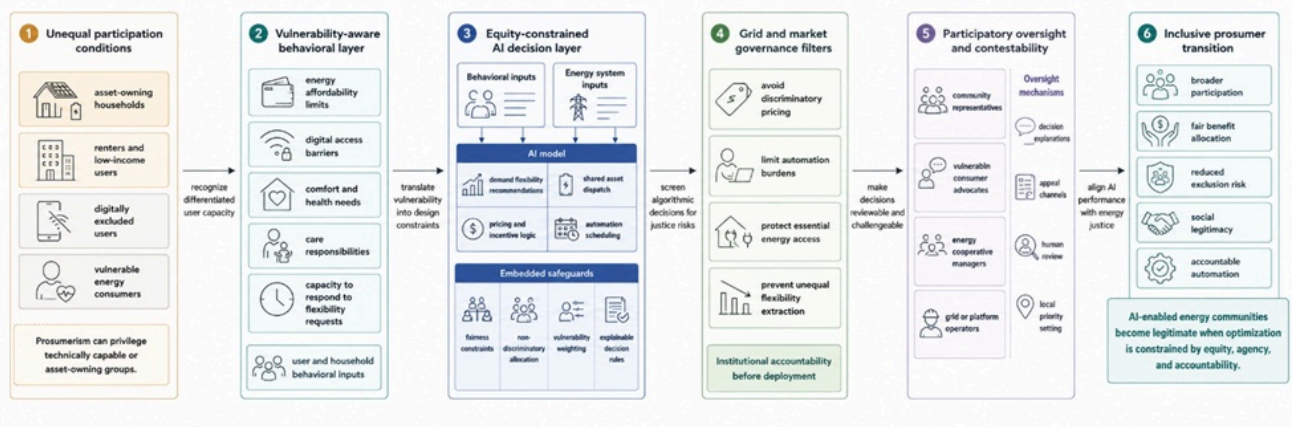


Figure 2. Equity-by-Design Architecture for Inclusive AI-Enabled Prosumer Energy Transitions

Architecture Integration and Validation Pathways

Conceptual integration and consistency

The three layers of the architecture are integrated through shared data models, interaction protocols, feedback loops, and governance constraints that connect human behavior with grid intelligence and institutional legitimacy. Peer-to-peer energy sharing simulations show that distributed energy coordination requires mechanisms for linking user actions, local exchange, and system-level constraints [31]. Demand-response AI research similarly implies that flexibility cannot be treated only as a technical resource because it depends on user participation and operational coordination [6]. Conceptually, the architecture is consistent because each layer performs a distinct function while constraining and informing the others.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Stakeholder-based validation and design workshops

Because this article proposes a conceptual architecture rather than an implemented system, validation should proceed through expert review, participatory design workshops, regulatory sandboxes, and co-creation with prosumers, operators, utilities, aggregators, and policymakers. Community and household feedback research suggests that user engagement methods can reveal how people interpret energy information and whether feedback supports meaningful action [15]. Institutional smart grid research further indicates that stakeholder contestation must be surfaced rather than hidden during design [12]. Such validation pathways would not prove system performance, but they could test conceptual coherence, practical relevance, governance feasibility, and acceptability before technical implementation.

Table 4 outlines a validation framework for assessing the conceptual architecture before technical deployment in real-world smart energy systems.

Table 4. Governance-Sensitive Validation Framework for the Proposed Human-Centered AI Architecture

Validation dimension	What must be tested	Suitable validation approach	Key stakeholders	Evidence generated	Success criterion	Why it strengthens the architecture
Conceptual coherence	Whether the three layers perform distinct but connected functions	Expert review, architecture walkthrough, design critique	Energy AI researchers, systems engineers, sociotechnical scholars	Structured feedback on completeness, overlap, and missing interfaces	Reviewers can trace how behavior, automation, and governance interact	Prevents the framework from becoming a loose list of components
Behavioral realism	Whether prosumer assumptions reflect bounded rationality, routines, trust, and heterogeneity	Participatory design workshops, household scenario testing, persona validation	Households, community energy members, small businesses, behavioral researchers	User interpretations, preference conflicts, perceived burden, comprehension gaps	Users recognize the scenarios as plausible and can revise system assumptions	Ensures that the architecture does not reproduce technology-centric rational-user assumptions
Human-AI interaction quality	Whether explanations, consent controls, and override options are usable and meaningful	Interface prototyping, think-aloud testing, explanation comprehension tasks	Prosumers, aggregators, UX researchers, consumer advocates	Comprehension scores, usability problems, trust concerns, override behavior	Users can understand, accept, reject, or modify AI-supported decisions	Links explainability to agency rather than treating it as a technical add-on
Operational feasibility	Whether AI functions can support forecasting, coordination, and	Simulation, digital sandboxing, operator review, edge-cloud	Grid operators, utilities, aggregators, DER platform providers	Feasibility assessment of latency, reliability, scalability, and	Technical functions can operate without removing human	Tests whether human-centered requirements remain viable

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

	automation under grid constraints	architecture testing		control boundaries	oversight or regulatory traceability	in operational contexts
Institutional compatibility	Whether AI recommendations comply with market rules, consumer protection, and regulatory duties	Regulatory sandbox, compliance mapping, policy stress testing	Regulators, utilities, legal experts, market operators, policymakers	Rule-conflict logs, compliance gaps, liability concerns, audit requirements	High-risk AI actions can be traced to explicit legal and institutional constraints	Makes institutional legitimacy a design property rather than a post-deployment correction
Data governance robustness	Whether data collection, sharing, retention, and revocation are transparent and controllable	Privacy impact assessment, consent-flow testing, data lineage review	Data protection officers, prosumers, platform operators, regulators	Consent risks, data-flow maps, portability gaps, revocation problems	Users and auditors can identify what data are used, why, and under what authority	Protects the architecture from becoming dependent on opaque data extraction
Equity and vulnerability protection	Whether AI-enabled services distribute benefits and burdens fairly	Equity impact assessment, vulnerable-user scenario testing, fairness audit	Low-income households, renters, elderly users, disability advocates, consumer regulators	Differential impact patterns, exclusion risks, burden allocation concerns	Flexibility, pricing, and automation do not systematically disadvantage vulnerable users	Aligns smart energy AI with energy justice and inclusive transition goals
Accountability and contestability	Whether affected actors can challenge, review, or escalate AI-mediated decisions	Redress pathway testing, audit-trail inspection, stakeholder governance review	Prosumers, regulators, utilities, ombuds bodies, community representatives	Contestability gaps, unclear responsibility chains, escalation failures	Consequential decisions have identifiable responsible parties and usable challenge routes	Converts trust and accountability from abstract values into operational mechanisms
Transferability	Whether the architecture can be adapted across different regulatory and grid contexts	Comparative case mapping, regional adaptation workshops, policy-context analysis	Local authorities, utilities, community energy groups, researchers	Context-specific modifications, non-transferable assumptions, adaptation requirements	The architecture can be customized without losing its core three-layer logic	Supports use as a flexible design template rather than a universal blueprint

Policy and Regulatory Implications

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Encouraging human-centered AI through energy regulation

Regulators could encourage human-centered AI by requiring AI-enabled energy services to demonstrate explainability, user control, consent management, and consumer impact assessment before deployment. Regulatory challenges for collective prosumers show that policy frameworks can strongly shape whether decentralized participation becomes feasible or constrained [11]. AI explainability and governance research suggests that smart energy systems require oversight mechanisms that connect technical transparency with institutional accountability [5]. Certification schemes, regulatory sandboxes, and grid code modifications could therefore help embed human-centered design into AI-enabled energy services.

Organizational governance for utilities and aggregators

Utilities, aggregators, and platform operators can support human-centered AI through internal governance structures that define responsibilities for model oversight, user communication, ethical review, and continuous monitoring. Studies of prosumer business models after subsidy regimes show that organizational strategies shape how prosumers participate in emerging energy markets [16]. Blockchain and platform-based energy research also indicates that new intermediary roles may create governance questions around trust, accountability, and market power [18]. Organizations should therefore establish human-AI collaboration protocols, stakeholder advisory panels, escalation procedures, and regular reviews of algorithmic impacts on users.

Aligning digital and energy justice agendas

Human-centered AI for smart energy transitions should align digital governance with energy justice by ensuring fair access, procedural participation, and equitable distribution of benefits. Normative work on prosumerism demonstrates that decentralized energy systems should be evaluated not only by efficiency but also by their social purposes and distributive consequences [28]. Collective renewable energy prosumer research shows that community-based models may support broader participation when institutional frameworks enable inclusion [7]. The proposed architecture can contribute to this alignment by making consent, explanation, fairness, vulnerability, and regulatory accountability visible design requirements rather than optional features.

Limitations

Conceptual nature and absence of empirical implementation

The proposed architecture is conceptual and has not been implemented, tested, or validated in a real-world energy system. Although AI demand response, peer-to-peer trading, and smart home energy management studies provide relevant design knowledge, this article does not report technical performance, user testing, or empirical behavioral outcomes [6]. Future work should instantiate the architecture in pilot projects and evaluate it through human-centered metrics such as perceived agency, trust, comprehension, fairness, comfort impact, and institutional auditability. Such work would need to distinguish between the conceptual promise of the architecture and the practical challenges of deployment.

Contextual variability and transferability challenges

The architecture also faces transferability limits because prosumer behavior, grid infrastructure, market design, and institutional culture vary substantially across regions. Prosumer pricing research shows that participation conditions depend on local policy and economic settings [20], while regulatory studies of collective prosumers show that legal frameworks shape what forms of coordination are possible [26]. Smart energy system conceptualization further suggests that digital energy architectures must be adapted to emerging trends and contextual configurations rather than applied uniformly [27]. Therefore, the architecture should be treated as a design template requiring customization for specific regulatory cultures, grid maturity levels, and social contexts.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Conclusion

This article has proposed a human-centered AI architecture for smart energy transitions that systematically integrates prosumer behavior, grid intelligence, and institutional constraints. Its main contribution is to conceptualize AI-enabled energy systems as sociotechnical architectures in which human actors are not peripheral users but central participants in design, operation, and governance.

Technology-centric AI risks undermining user trust, behavioral engagement, and institutional legitimacy when it optimizes grid functions without accounting for lived energy practices and regulatory realities. A human-centered architecture can better align technical optimization with human agency, social values, and the practical conditions under which smart energy transitions unfold.

The architecture emphasizes the heterogeneity of prosumers, the importance of meaningful human oversight, and the need for transparent operation within regulatory guardrails. By linking behavioral interaction, grid automation, and institutional governance, it provides a conceptual basis for AI systems that are more explainable, accountable, adaptive, and socially responsive.

Future research should move from conceptual specification toward interdisciplinary implementation, stakeholder co-design, policy experimentation, and real-world piloting. Such efforts will be essential for evolving human-centered AI from an architectural principle into a practical foundation for inclusive and trustworthy smart energy transitions.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 02 Jul 2022 Revised: 08 Sep 2022 Accepted: 23 Sep 2022
Published online: 15 June 2023

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

References

1. Ahmad T, Zhang D, Huang C, Zhang H, Dai N, Song Y, et al. Artificial intelligence in sustainable energy industry: status quo, challenges and opportunities. J Clean Prod. 2021;289:125834.

2. Horne C, Kennedy EH. The power of social norms for reducing and shifting electricity use. Energy Policy. 2017;107:43-52.

3. Chen CF, Xu X, Arpan L. Between the technology acceptance model and sustainable energy technology acceptance model: investigating smart meter acceptance in the United States. Energy Res Soc Sci. 2017;25:93-104.

4. Machlev R, Heistrene L, Perl M, Levy KY, Belikov J, Mannor S, et al. Explainable artificial intelligence (XAI) techniques for energy and power systems: review, challenges and opportunities. Energy AI. 2022;9:100169.

5. Alsaigh R, Mehmood R, Katib I. AI explainability and governance in smart energy systems: a review. Front Energy Res. 2023;11:1071291.

6. Antonopoulos I, Robu V, Couraud B, Kirli D, Norbu S, Kiprakis A, et al. Artificial intelligence and machine learning approaches to energy demand-side response: a systematic review. Renew Sustain Energy Rev. 2020;130:109899.

7. Horstink L, Wittmayer JM, Ng K. Pluralising the European energy landscape: collective renewable energy prosumers and the EU's clean energy vision. Energy Policy. 2021;153:112262.

8. Parra-Dominguez J, Sanchez E, Ordonez A. The prosumer: a systematic review of the new paradigm in energy and sustainable development. Sustainability. 2023;15(13):10552.

9. Khan MA, Saleh AM, Waseem M, Sajjad IA. Artificial intelligence enabled demand response: prospects and challenges in smart grid environment. IEEE Access. 2023;11:1477-505.

10. Nilsson A, Wester M, Lazarevic D, Brandt N. Smart homes, home energy management systems and real-time feedback: lessons for influencing household energy consumption from a Swedish field study. Energy Build. 2018;179:15-25.

11. Leal-Arcas R, Lesniewska F, Proedrou F. Prosumers: new actors in EU energy security. In: Netherlands Yearbook of International Law 2017. The Hague: TMC Asser Press; 2018. p. 139-72.

12. Rohde F, Hielscher S. Smart grids and institutional change: emerging contestations between organisations over smart energy transitions. Energy Res Soc Sci. 2021;74:101974.

13. Sousa T, Soares T, Pinson P, Moret F, Baroche T, Sorin E. Peer-to-peer and community-based markets: a comprehensive review. Renew Sustain Energy Rev. 2019;104:367-78.

14. Morstyn T, Farrell N, Darby SJ, McCulloch MD. Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants. Nat Energy. 2018;3(2):94-101.

15. Gupta R, Barnfield L, Gregg M. Exploring innovative community and household energy feedback approaches. Build Res Inf. 2018;46(3):284-99.

16. Brown D, Hall S, Davis ME. Prosumers in the post subsidy era: an exploration of new prosumer business models in the UK. Energy Policy. 2019;135:110984.

17. Zhang C, Wu J, Zhou Y, Cheng M, Long C. Peer-to-peer energy trading in a microgrid. Appl Energy. 2018;220:1-2.

18. Andoni M, Robu V, Flynn D, Abram S, Geach D, Jenkins D, et al. Blockchain technology in the energy sector: a systematic review of challenges and opportunities. Renew Sustain Energy Rev. 2019;100:143-74.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

19. Di Silvestre ML, Favuzza S, Sanseverino ER, Zizzo G. How decarbonization, digitalization and decentralization are changing key power infrastructures. Renew Sustain Energy Rev. 2018;93:483-98.

20. Kuznetsova E, Anjos MF. Prosumers and energy pricing policies: when, where, and under which conditions will prosumers emerge? A case study for Ontario (Canada). Energy Policy. 2021;149:111982.

21. Tak AN, Becerik-Gerber B, Soibelman L, Lucas G. A framework for investigating the acceptance of smart home technologies: findings for residential smart HVAC systems. Build Environ. 2023;245:110935.

22. Zeiler W, Labeodan T. Human-in-the-loop energy flexibility integration on a neighbourhood level: small and big data management. Build Serv Eng Res Technol. 2019;40(3):305-18.

23. Slama SB. Prosumer in smart grids based on intelligent edge computing: a review on artificial intelligence scheduling techniques. Ain Shams Eng J. 2022;13(1):101504.

24. Tushar W, Saha TK, Yuen C, Morstyn T, Poor HV, Bean R. Grid influenced peer-to-peer energy trading. IEEE Trans Smart Grid. 2020;11(2):1407-18.

25. Mengelkamp E, Gärttner J, Rock K, Kessler S, Orsini L, Weinhardt C. Designing microgrid energy markets: a case study: the Brooklyn Microgrid. Appl Energy. 2018;210:870-80.

26. Inês C, Guilherme PL, Esther MG, Swantje G, Stephen H, Lars H. Regulatory challenges and opportunities for collective renewable energy prosumers in the EU. Energy Policy. 2020;138:111212.

27. Darani ZH, Taheri Demne M, Zanjirani DM, Zackery A. Conceptualization of a new generation of smart energy systems and the transition toward them using anticipatory systems. Eur J Futures Res. 2021;9(1):15.

28. Brown D, Hall S, Davis ME. What is prosumerism for? Exploring the normative dimensions of decentralised energy transitions. Energy Res Soc Sci. 2020;66:101475.

29. Noorman M, Espinosa Apráez B, Lavrijssen S. AI and energy justice. Energies. 2023;16(5):2110.

30. Knox S, Hannon M, Stewart F, Ford R. The (in)justices of smart local energy systems: a systematic review, integrated framework, and future research agenda. Energy Res Soc Sci. 2022;83:102333.

31. Zhou Y, Wu J, Long C. Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework. Appl Energy. 2018;222:993-1022

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press