ORIGINAL RESEARCH **Open access**

Policy-Aware Artificial Intelligence for Smart Energy Systems: Aligning Algorithmic Optimization with Public Value and Regulatory Objectives

Ahmed Ben Ali[1*], Samira Gharbi[1], Nabil Jebali[2]

Abstract

AI increasingly drives decisions in smart energy systems, from dispatch and pricing to demand response. These decisions have profound public consequences, yet algorithms often optimize for narrow private or technical objectives, misaligning with public values and regulatory goals. Existing AI design paradigms treat policy as an afterthought, leading to conflicts between algorithmic efficiency and societal objectives such as fairness, sustainability, and consumer protection. There is no systematic framework for building policy awareness into energy AI. This article develops a policy-aware AI framework that aligns algorithmic optimization in smart energy systems with public value and regulatory objectives. The framework treats public values as design inputs rather than external constraints applied after deployment. The framework is constructed through conceptual synthesis of AI alignment theory, public value management, energy regulation, and multi-objective optimization. No empirical experiments are conducted. The framework identifies key public value dimensions for energy AI, specifies mechanisms to translate regulatory objectives into algorithmic constraints, and proposes governance structures for managing trade-offs transparently. Three analytical tables map public values, regulatory translation mechanisms, and trade-off resolution strategies. Policy-aware AI offers a pathway to embed democratic and regulatory norms directly into the algorithmic core of smart energy systems, making them more accountable, legitimate, and publicly responsive. It reframes smart energy optimization as a public governance problem as well as a technical design challenge.

Keywords Optimization, Smart energy systems, Policy-aware AI, Public value, Regulatory objectives, Algorithmic alignment

*Correspondence:
Ahmed Ben Ali
ahmed.benali@utunis.tn
[1] Department of Artificial Intelligence for Smart Energy Systems, Faculty of Engineering, University of Tunis, Tunis, Tunisia
[2] Department of Intelligent Power Distribution Networks, Faculty of Science and Technology, University of Tunis El Manar, Tunis, Tunisia

Introduction

AI systems are increasingly used to coordinate smart energy functions such as demand response, forecasting, distributed energy management, and automated grid optimization. Reviews of machine learning in energy economics and demand-side response show that AI can improve operational responsiveness, yet these systems often prioritize efficiency, prediction quality, or cost reduction without fully reflecting public obligations embedded in energy systems [1]. Reinforcement learning applications in energy further illustrate how algorithmic agents may learn control strategies that optimize local rewards while leaving broader social consequences under-specified [2]. In this context, smart energy AI should be understood not only as a technical instrument but also as a governance mechanism that shapes affordability, reliability, sustainability, and access.

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.

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Policy and regulation embody collective decisions about what energy systems should provide, who should bear costs, and how risks should be distributed. Energy justice scholarship emphasizes that energy systems are judged by distributive, procedural, and recognition-based values, which cannot be reduced to operational efficiency alone. Algorithmic governance research similarly warns that automated decision systems may exercise regulatory power without adequate democratic accountability when rules are embedded invisibly into technical architectures [3]. This creates a legitimacy deficit when AI-driven energy decisions affect consumers, markets, or environmental outcomes but are not explicitly aligned with public mandates.

Policy-aware AI can be conceptualized as a design paradigm in which public value and regulatory requirements are incorporated into objective functions, constraints, training procedures, evaluation criteria, and oversight routines. Ethical AI frameworks argue that transparency, justice, responsibility, and human oversight must move from abstract principles into operational design methods [4]. The difficulty is that principles alone do not guarantee responsible AI unless they are translated into institutional practices and technical mechanisms [5]. For smart energy systems, policy awareness therefore requires a structured bridge between regulatory objectives and algorithmic optimization logic.

This article develops a conceptual framework for policy-aware AI in smart energy systems, focusing on how public values and regulatory objectives could be encoded into AI design, training, evaluation, and governance. It draws on AI governance, public sector AI, smart energy optimization, energy regulation, and algorithmic regulation literature to specify the conceptual elements of such a framework [6]. The argument proceeds by defining the public value basis of energy AI, identifying regulatory alignment mechanisms, and explaining how algorithmic systems may be governed when policy values conflict. The article is framework-oriented and does not conduct experiments, report empirical findings, or present quantitative simulations.

Background

AI in smart energy systems: functions and decisions

AI in smart energy systems performs increasingly consequential functions, including demand forecasting, demand response coordination, energy trading support, distributed resource scheduling, and real-time control. Systematic reviews of demand-side response show that AI can support flexible consumption decisions, but these decisions are shaped by incentives, tariffs, consumer behavior, and institutional rules rather than technical signals alone [7]. Reinforcement learning applications in energy systems further demonstrate how agents may be trained to optimize building control, energy use, or grid interaction through reward structures that define what the system values [8]. Because these reward structures influence who benefits, who pays, and which risks are tolerated, AI-based energy decisions have direct policy relevance.

Public value theory and energy systems

Public value in energy systems can be understood as the set of collective goods that energy infrastructures are expected to secure, including reliability, affordability, environmental sustainability, fairness, privacy, resilience, and democratic accountability. Energy justice literature has shown that energy transitions are not only technological changes but also redistributive and institutional processes that affect vulnerable groups differently [9]. AI governance research extends this concern by arguing that automated systems should be evaluated according to their social consequences, not merely their technical performance [10]. A policy-aware AI framework should therefore treat public value as a layered design requirement that informs what optimization is permitted to pursue.

Regulatory frameworks for smart energy

Smart energy systems operate within regulatory environments shaped by price controls, market rules, emission targets, consumer protection obligations, cybersecurity requirements, and infrastructure reliability mandates. Research on

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

aggregators and flexibility markets shows that new energy actors can create system value, but only when market design and regulation clarify responsibilities, incentives, and consumer protections [11]. Peer-to-peer trading and blockchain-based energy systems similarly reveal that decentralized optimization can challenge existing regulatory categories, particularly around settlement, accountability, and market participation [12]. These developments suggest that AI systems should not be designed as neutral optimizers but as policy-situated agents operating within regulatory boundaries.

AI alignment and value-sensitive design

AI alignment concerns how artificial systems can be made responsive to human goals, institutional norms, and socially legitimate constraints. Ethical AI scholarship has proposed principles such as beneficence, non-maleficence, autonomy, justice, and explicability, but practical translation remains difficult when values conflict or lack measurable definitions [4]. Society-in-the-loop theory argues that algorithmic systems should be programmed through social contracts that keep collective values in the loop rather than treating them as fixed technical parameters [13]. For energy AI, value-sensitive design would require public policy goals to influence training objectives, model evaluation, deployment thresholds, and post-deployment monitoring.

The gap between optimization and policy

The current gap between optimization and policy arises because energy AI is often designed around technically tractable objectives such as cost minimization, load reduction, prediction accuracy, or system efficiency. Reviews of AI in sustainable energy highlight substantial opportunities for intelligent control and planning, yet they also identify governance, accountability, and implementation challenges that are not solved by better algorithms alone [14]. The literature on AI in public governance similarly shows that public-sector uses of AI require legitimacy, transparency, and institutional accountability beyond technical effectiveness [15]. Policy-aware AI addresses this gap by making regulatory and public value requirements native to the algorithmic design process rather than external compliance checks.

Recent empirical research has begun to quantify how AI adoption affects energy justice outcomes, demonstrating that the relationship is neither automatic nor uniformly positive. A longitudinal study of Chinese provinces found that while AI adoption significantly enhances overall energy justice, the benefits are concentrated in advanced regions, and AI can initially exacerbate distributional injustice even as it improves procedural and recognition justice [12]. This finding challenges technological neutrality assumptions and suggests that policy-aware AI frameworks must be designed to actively counteract potential regressive distributional effects. Similarly, governance analyses of AI's expanding energy footprint emphasize that the technology creates a paradoxical trilemma across security, justice, and sustainability dimensions, requiring careful policy design to prevent efficiency gains from being offset by rising demand and inequitable infrastructure concentration [16].

Conceptual Foundation

Defining policy-aware AI

Policy-aware AI can be defined as an AI system whose objective function, constraints, learning process, evaluation metrics, and oversight routines explicitly incorporate public policy values and regulatory requirements. This definition builds on the insight that algorithmic regulation can operate both through the regulation of algorithms and through the use of algorithms as regulatory instruments [17]. In smart energy systems, policy awareness would mean that an AI controller, pricing algorithm, or forecasting tool should reflect not only technical feasibility but also affordability protections, emissions obligations, equity concerns, and reliability standards. Such systems should be designed to make policy choices visible, contestable, and revisable.

The public value stack for energy AI

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

The public value stack for energy AI can be conceptualized as a layered structure that connects foundational system obligations, social distributional values, environmental commitments, and procedural governance principles. Foundational values include reliability, safety, and security because energy infrastructures support critical social functions and cannot be governed solely through market efficiency [18]. Distributional and environmental layers include affordability, fairness, decarbonization, and sustainability, reflecting the role of AI in supporting the Sustainable Development Goals and energy transition objectives [19]. Procedural layers include transparency, explainability, accountability, and human oversight, which are necessary for perceived legitimacy in automated public decision-making [20].

Translating regulation into algorithmic form

Translating regulation into algorithmic form requires converting legal duties, policy goals, and public value commitments into computable objectives, constraints, thresholds, labels, or review triggers. The idea of regulation by design shows that rules can be embedded into technical architectures, but it also warns that coded rules may narrow legal interpretation and reduce democratic contestability [3]. In energy AI, an emission cap could be encoded as a constraint, a consumer protection rule could be encoded as an eligibility threshold, and an affordability mandate could be encoded as a penalty for disproportionate tariff impacts. However, such translation should remain auditable and revisable because public policy goals are interpretive, contested, and context-dependent.

Figure 1 illustrates the nested alignment logic through which public values, regulatory objectives, algorithmic mechanisms, and governance feedback loops are integrated into policy-aware AI design.

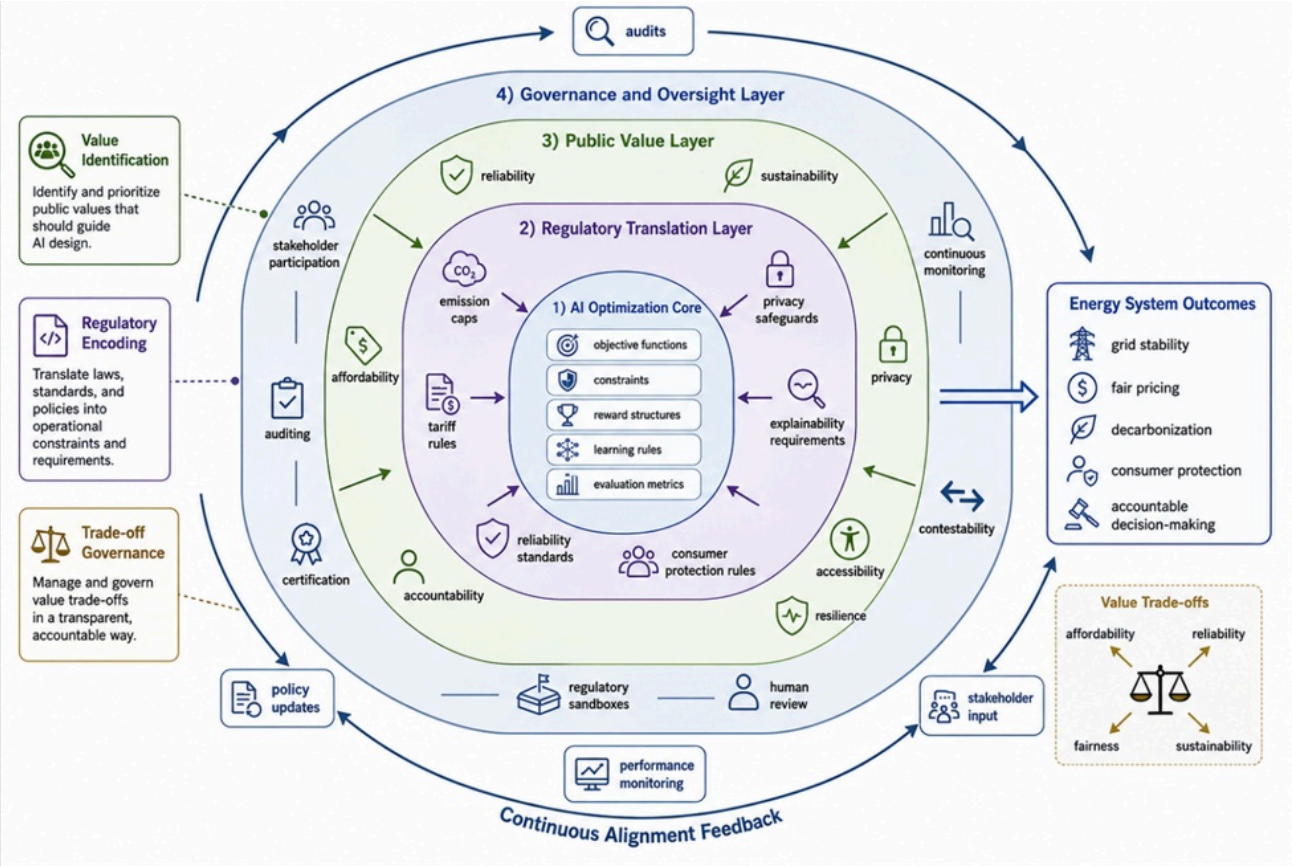


Figure 1. Nested Public-Value Alignment Model for Policy-Aware AI in Smart Energy Systems

Public Value Dimensions for Energy AI

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Mapping core public values

Core public values for energy AI include reliability, affordability, environmental sustainability, fairness, privacy, accessibility, accountability, and resilience. Energy system AI may support reliability by anticipating demand and balancing distributed resources, yet reliability-oriented optimization could impose costs on consumers or delay environmental objectives if not governed carefully [21]. Affordability and fairness are especially important where dynamic pricing, automated demand response, or flexibility markets affect households with different capacities to adapt [9]. Privacy and accountability also become central when AI systems rely on granular consumption data, household behavior patterns, or automated decisions that users may not understand or contest.

Operationalizing values as metrics

Operationalizing public values as metrics does not mean reducing policy to numbers, but it does require defining observable indicators that can guide AI design and review. For example, reliability could be represented through interruption risk, reserve adequacy, or service continuity indicators, while affordability could be reflected through bill volatility, vulnerability thresholds, or exposure to dynamic price spikes [11]. Fairness may require distributional analysis across consumer classes, regions, or vulnerable groups, while sustainability may require emissions intensity or renewable integration indicators that shape algorithmic decisions [19]. Because AI ethics tools often struggle to move from principle to practice, metric design should be paired with governance processes that keep value interpretation open to review [22]. **Figure 2** illustrates how public values in AI-enabled energy systems can be translated into observable metrics while remaining subject to governance review and value interpretation rather than being reduced to purely numerical optimization targets.

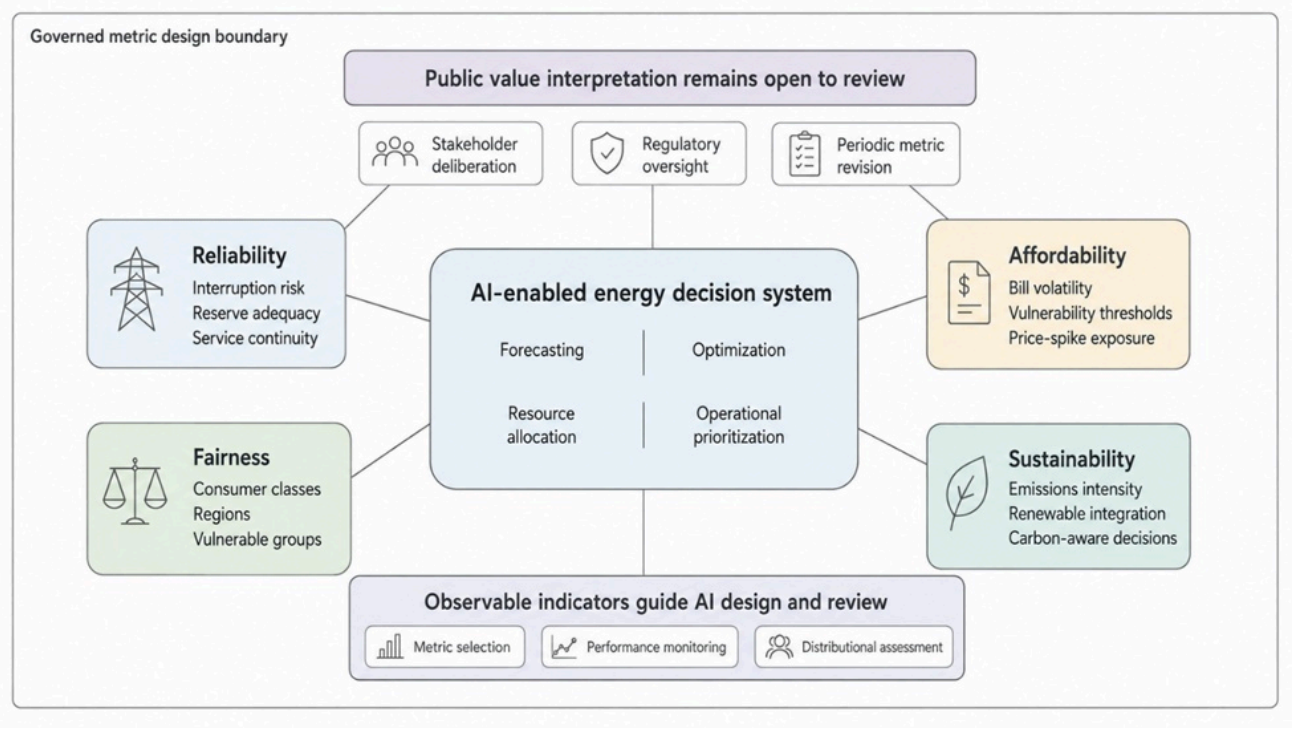


Figure 2. Translating Public Values into Governable AI Metrics for Energy Decision-Making

Policy instruments and value mandates

Policy instruments provide the institutional pathway through which public values become actionable mandates for energy AI. Price caps, tariff design, emissions standards, renewable portfolio requirements, market participation rules, data

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

protection obligations, and service quality standards each communicate a public value that could be reflected in algorithmic constraints or reward modifiers [23]. For example, consumer protection regulation may require an AI-based pricing system to avoid imposing excessive volatility on vulnerable users, while emissions regulation may require a dispatch optimizer to penalize carbon-intensive operational choices. **Table 1** maps these public values to operational definitions, policy instruments, and AI optimization implications.

Table 1. Public Value Dimensions, Operational Definitions, and Policy Instruments for Energy AI

Public Value	Operational Definition	Policy Instrument	AI Optimization Implication
Reliability	Ability of AI-supported energy decisions to maintain service continuity, grid stability, and adequate reserve margins under changing demand and supply conditions.	Reliability standards, grid codes, capacity obligations, system operator rules.	Reliability should be encoded as a non-negotiable constraint or high-priority objective that limits cost-saving actions likely to increase outage or instability risk.
Affordability	Degree to which AI-driven pricing, dispatch, and demand-response decisions avoid excessive cost burdens or bill volatility for consumers.	Price caps, regulated tariffs, social tariffs, consumer affordability mandates.	Optimization should include penalties for disproportionate bill impacts, price shocks, or strategies that shift costs onto less flexible consumers.
Environmental sustainability	Extent to which AI decisions support decarbonization, renewable integration, and reduced environmental harm.	Emission caps, carbon pricing, renewable portfolio standards, clean energy targets.	Carbon intensity and renewable utilization should be included as constraints, reward modifiers, or priority objectives alongside cost and reliability.
Fairness	Distribution of AI-generated benefits, burdens, risks, and opportunities across consumer groups, regions, and market participants.	Anti-discrimination rules, energy justice policies, targeted support schemes, equity impact assessments.	Models should be evaluated for unequal impacts, and optimization should include safeguards against systematically disadvantaging vulnerable or low-flexibility users.
Privacy	Protection of consumer data, household behavior patterns, and commercially sensitive operational information used by AI systems.	Data protection laws, consent requirements, data minimization rules, cybersecurity obligations.	Training and deployment should restrict data collection, favor privacy-preserving features, and penalize designs requiring unnecessary personal granularity.
Accessibility	Ability of diverse users and institutions to understand, use, benefit from, or contest AI-enabled energy services.	Universal service obligations, consumer information rules, accessibility standards, procedural rights.	AI interfaces and decisions should be designed with explainability, notice, and contestation pathways rather than assuming technically sophisticated users.
Accountability	Capacity to assign responsibility for AI decisions and demonstrate compliance with public mandates.	Audit requirements, certification schemes, regulatory reporting, liability frameworks.	Systems should generate traceable decision records, policy compliance logs, and explanations suitable for regulators and affected stakeholders.

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Regulatory Alignment Mechanisms

Encoding constraints and objectives

Regulatory alignment can be achieved by encoding policy requirements as hard constraints, soft penalties, reward modifiers, optimization priorities, or post-decision review triggers. In reinforcement learning for energy, the reward function defines the behavior that the agent is expected to learn, so policy-aware design would require public objectives to be represented within the learning environment rather than evaluated only after deployment [2]. A hard constraint may prohibit unsafe operating states, while a soft penalty may discourage carbon-intensive dispatch or excessive bill volatility. Multi-objective optimization can then be conceptualized as a policy translation method that makes trade-offs explicit rather than allowing technical convenience to determine them implicitly.

Regulatory sandboxes and iterative alignment

Regulatory sandboxes may support policy-aware AI by allowing utilities, regulators, technology providers, and public stakeholders to test algorithmic behavior under controlled conditions before full deployment. This approach is particularly relevant where AI systems interact with emerging flexibility markets, aggregators, or decentralized energy trading arrangements that challenge established regulatory categories [9]. Sandboxes could evaluate whether AI systems respect consumer protection rules, market integrity safeguards, emissions obligations, and explainability requirements under plausible operating scenarios. However, sandbox governance should avoid treating experimentation as deregulation, because the purpose of policy-aware testing is to refine alignment with public mandates rather than suspend them.

Certification and auditing for policy compliance

Certification and auditing mechanisms are needed to verify that AI systems satisfy policy requirements throughout design, deployment, and adaptation. Work on AI explainability and governance in smart energy systems suggests that transparency, documentation, and oversight are essential for maintaining trust when algorithms influence infrastructure decisions [16]. Compliance with emerging AI regulation may also require evidence about risk classification, human oversight, documentation, and monitoring, especially where AI systems affect critical infrastructure or consumer welfare [24]. **Table 2** illustrates how regulatory objectives could be translated into algorithmic constraints or modifications and how those translations may be verified.

Table 2. Regulatory Objectives and Their Translation into AI Constraints

Regulatory Objective	Regulatory Instrument	Algorithmic Constraint/Modification	Verification Method
Limit carbon emissions from AI-supported dispatch	Emission caps, carbon pricing, clean energy standards.	Add carbon-intensity penalties to objective functions or impose hard limits on emissions-intensive actions when cleaner feasible alternatives exist.	Audit dispatch recommendations against emissions accounting records and review whether carbon penalties affected model decisions.
Protect consumers from excessive energy prices	Retail price caps, tariff regulation, social tariff schemes.	Constrain dynamic pricing outputs within approved price bands and penalize bill volatility for vulnerable consumer categories.	Compare algorithmic pricing decisions with tariff rules, consumer bill simulations, and vulnerability safeguards.
Maintain system reliability	Grid codes, reliability standards, reserve requirements.	Prohibit control actions that violate reserve margins, voltage limits, frequency stability thresholds, or service continuity standards.	Conduct compliance review using operational logs, system-state records, and exception reports for constraint violations.

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Prevent discriminatory distribution of burdens	Equity mandates, anti-discrimination rules, energy justice assessments.	Add fairness constraints that monitor unequal cost, curtailment, or service-quality impacts across consumer groups and regions.	Perform distributional impact audits and require documented justification for materially unequal outcomes.
Protect consumer data and privacy	Data protection regulation, consent requirements, data minimization rules.	Restrict input features, use aggregation or privacy-preserving learning, and penalize unnecessary personal data dependency.	Review data lineage, feature inventories, consent records, and privacy impact assessments.
Safeguard market integrity	Market rules, anti-manipulation provisions, aggregator regulation.	Prevent bidding, trading, or dispatch strategies that exploit market design weaknesses or create artificial scarcity signals.	Audit transaction logs, bidding behavior, and model incentives against market surveillance rules.
Ensure explainability and contestability	AI transparency requirements, administrative due process, regulatory reporting.	Require interpretable outputs, reason codes, decision traces, and escalation triggers for high-impact recommendations.	Examine explanation records, appeal pathways, human review logs, and regulator-facing documentation.

Embedding Policy into AI Design and Training

Policy-aware objective functions

Policy-aware objective functions should be designed so that AI systems optimize within the boundaries of public value rather than treating public obligations as external constraints. In smart energy applications, a loss function could include cost, reliability, carbon intensity, consumer vulnerability, and regulatory compliance terms, with each term reflecting a policy-relevant dimension of system performance [8]. Work on trustworthy and responsible AI shows that ethical objectives must be operationalized through concrete design tools, otherwise fairness, transparency, and responsibility remain abstract principles without implementation force [22]. For energy AI, this means that affordability, emissions reduction, and reliability should not be added after training but should shape the learning target from the beginning.

Data governance for policy alignment

Policy-aware AI requires data governance structures capable of measuring public values without creating new risks of surveillance, exclusion, or bias. Smart energy systems increasingly rely on granular consumption, device, market, and network data, but privacy and accountability concerns arise when such data are used to infer household behavior or automate consumer-facing decisions [12]. AI governance literature emphasizes that public trust depends on transparency, legitimacy, and responsible institutional use of data, especially when automated systems affect public goods or regulated services [25]. Therefore, policy-aligned datasets should include not only technical variables such as load, price, and voltage, but also carefully governed indicators of consumer vulnerability, service quality, environmental impact, and distributional fairness.

Learning from regulatory signals

Policy-aware AI systems could learn from regulatory signals by treating policy updates, enforcement decisions, audit findings, and stakeholder feedback as inputs into model refinement. In electricity systems, emerging AI regulation and sector-specific rules indicate that compliance is not static, because obligations may evolve as technologies, risks, and institutional expectations change [24]. A policy-aware learning process would therefore require monitored adaptation rather than unrestricted online optimization, especially where AI systems affect prices, access, reliability, or market

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

participation. Such adaptation may be achieved by creating regulatory feedback loops that update constraints, revise reward modifiers, and trigger human review when algorithmic behavior drifts away from approved policy objectives.

Explainability as a policy requirement

Explainability should be treated as a policy requirement because regulators, operators, and affected consumers need reasons for AI-supported decisions that influence public infrastructure and household welfare. Research on transparency in AI-based public decision-making shows that explanations are central to perceived legitimacy, particularly when automated decisions affect rights, benefits, or access to essential services [20]. In smart energy systems, explainability could support tariff justification, demand-response participation, dispatch accountability, and consumer appeals by making visible why an AI system prioritized one objective over another [16]. Policy-aware AI should therefore provide explanations that are not merely technical descriptions but regulator-facing accounts of how public values and legal constraints shaped the decision.

Trade-off Governance and Contestability

Inevitable trade-offs among public values

Policy-aware AI must recognize that public values in energy systems frequently conflict rather than align automatically. A demand-response algorithm may reduce system costs while shifting inconvenience to households with less flexibility, and a carbon-oriented dispatch strategy may increase short-term affordability pressures if distributional safeguards are absent [7]. Energy justice scholarship shows that sustainability transitions can generate uneven benefits and burdens, making fairness and recognition central to energy governance. Therefore, policy-aware optimization should not claim to eliminate trade-offs but should make them explicit, governed, and open to institutional review.

Transparent trade-off resolution mechanisms

Transparent trade-off resolution mechanisms are needed when AI systems must choose among efficiency, equity, sustainability, reliability, and consumer protection. Algorithmic regulation scholarship warns that automated systems can hide normative choices inside technical architectures, making political decisions appear as neutral optimization outcomes [17]. One response is to use lexicographic ordering for non-negotiable obligations, stakeholder-weighted optimization for contestable priorities, and constraint-based approaches for legal minimum standards. These mechanisms would be expected to make value prioritization visible and to prevent private or technical objectives from silently overriding public mandates.

Stakeholder participation and contestation

Stakeholder participation is essential because the relative importance of energy values cannot be determined by developers, utilities, or regulators alone. Society-in-the-loop thinking argues that algorithmic systems should remain connected to collective judgment when they govern socially consequential domains [13]. In policy-aware energy AI, participation could include regulatory hearings, consumer panels, public value workshops, municipal consultation, and expert review of proposed value weights or constraints. **Table 3** presents conceptual mechanisms for resolving public value conflicts and identifies how stakeholder involvement changes the governance implications of each mechanism.

Table 3. Trade-off Resolution Mechanisms for Public Value Conflicts in Energy AI

Trade-off Example	Resolution Mechanism	Stakeholder Involvement	Governance Implication
Affordability versus decarbonization	Constraint-based carbon reduction with affordability safeguards	Regulators, consumer advocates, environmental	Prevents decarbonization algorithms from shifting

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

		agencies, and utilities define acceptable cost exposure limits.	excessive transition costs onto vulnerable consumers.
Reliability versus market efficiency	Lexicographic ordering that treats reliability as a priority constraint before cost optimization	System operators and regulators define minimum reliability conditions before market optimization is permitted.	Clarifies that efficiency gains cannot justify decisions that undermine critical infrastructure obligations.
Privacy versus demand-response precision	Data minimization and privacy-preserving optimization	Consumers, data protection authorities, utilities, and technology providers review acceptable data granularity.	Ensures that more accurate control does not automatically justify intrusive household-level data collection.
Equity versus aggregate welfare maximization	Stakeholder-weighted optimization with distributional impact review	Community representatives, vulnerable consumer groups, and regulators participate in setting equity weights.	Makes distributive choices explicit and contestable rather than hidden within aggregate performance indicators.
Innovation versus regulatory certainty	Regulatory sandbox with staged deployment conditions	Regulators, utilities, innovators, and public interest representatives jointly define testing boundaries.	Allows experimentation while preserving public protection and creating evidence for rule refinement.
Automation efficiency versus human accountability	Human-in-the-loop escalation for high-impact or ambiguous decisions	Operators, regulators, and affected stakeholders identify decisions requiring human review or appeal.	Prevents accountability gaps by linking automated recommendations to responsible institutional actors.

Accountability for algorithmic value decisions

Accountability for algorithmic value decisions requires identifying who is responsible when an AI system prioritizes one public value over another. Public-sector AI research shows that automated decision-making can blur responsibility across developers, vendors, public agencies, and operational users unless governance roles are specified in advance [6]. In smart energy systems, accountability should attach not only to technical failure but also to the design of reward functions, constraint hierarchies, data choices, explanation practices, and deployment thresholds. A policy-aware framework therefore requires documented responsibility chains connecting algorithm designers, utilities, regulators, auditors, and human operators.

Institutional Integration and Regulatory Oversight

Embedding policy-aware AI into utility operations

Embedding policy-aware AI into utility operations would require changes in procurement, model governance, operational protocols, and staff responsibilities. Utilities adopting AI for forecasting, flexibility management, customer interaction, or market participation should require model documentation that explains how regulatory duties and public values are represented in the system design [26]. Research on the lack of alignment between AI strategies and electricity regulation suggests that energy organizations may adopt AI faster than regulatory frameworks adapt, creating gaps between innovation and oversight [26]. Policy-aware implementation should therefore involve interdisciplinary teams that include engineers, compliance officers, public policy specialists, consumer protection experts, and operational decision-makers.

Regulatory oversight and continuous alignment

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Regulatory oversight should move from one-time approval toward continuous alignment because AI systems may change through retraining, updated data streams, revised market conditions, or shifting policy priorities. Recent work on AI regulation in the energy sector highlights the need to connect general AI governance requirements with sector-specific energy law, infrastructure risk, and market regulation [16]. Continuous oversight could include model registration, periodic audits, incident reporting, public value impact assessments, and mandatory review when AI behavior materially affects consumers or grid operations. This approach would conceptualize regulators not as external observers but as institutional actors in the ongoing alignment of algorithmic optimization with public purpose.

Evaluation Strategy for the Framework

Conceptual adequacy and policy relevance

The framework can be evaluated conceptually by asking whether it adequately connects AI design elements with public administration, regulatory economics, and energy governance concerns. Public governance literature shows that AI adoption raises questions of legitimacy, collaboration, institutional capacity, and accountability that cannot be answered through technical performance alone [15]. A conceptually adequate framework should therefore explain how public values enter the AI lifecycle, how regulatory objectives become computable, and how trade-offs remain contestable. Its policy relevance depends on whether it helps regulators and utilities identify where algorithmic optimization may support, ignore, or undermine public mandates.

Expert and stakeholder validation

Expert and stakeholder validation would be appropriate because policy-aware AI involves interpretive judgments about values, legal obligations, and acceptable trade-offs. Cross-sector collaboration in public-sector AI demonstrates that technical experts, public agencies, and institutional stakeholders must work together when automated systems are used for public purposes [18]. Validation could be conducted through policy workshops, regulator-developer review panels, consumer advocacy consultations, and scenario-based exercises that examine how the framework handles conflicting objectives. Such validation would not produce empirical performance claims but would test whether the framework is understandable, institutionally usable, and normatively credible.

Transferability and scalability

The transferability of the framework depends on whether its design logic can be adapted across different regulatory regimes, market structures, and energy transition pathways. AI governance principles vary across jurisdictions, and global reviews of AI ethics guidelines show both convergence around high-level values and divergence in implementation mechanisms [10]. In energy systems, this means that a policy-aware AI framework should be modular enough to accommodate different emissions rules, tariff designs, market models, and consumer protection regimes. Scalability would require standardized documentation, audit templates, and regulatory translation procedures while still allowing local policy priorities to shape algorithmic objectives.

Limitations

The challenge of quantifying public values

A central limitation of policy-aware AI is that public values cannot always be quantified without distortion. Principles such as fairness, legitimacy, affordability, and accountability may require interpretation, deliberation, and institutional judgment rather than simple metric construction [5]. Translating legal and ethical concepts into algorithmic constraints may create useful discipline, but it may also narrow public values into what is easiest to measure. Therefore, the framework should be understood as a governance aid rather than a complete solution to the normative complexity of energy policy.

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

Dynamic regulatory and political landscapes

A second limitation is that regulatory and political landscapes are dynamic, meaning that policy-aware AI cannot rely on fixed value specifications indefinitely. Energy regulation changes in response to decarbonization targets, market reforms, consumer protection concerns, technological innovation, and geopolitical pressures [26]. AI systems that are aligned at deployment may become misaligned if laws, tariffs, emissions rules, or public expectations change. The framework therefore requires further policy design work to specify how alignment should be updated, who has authority to revise value weights, and how affected stakeholders can challenge outdated algorithmic rules.

Conclusion

Policy-aware AI offers a conceptual pathway for aligning smart energy optimization with public value and regulatory objectives. Rather than treating AI as a neutral technical instrument, the framework developed in this article positions algorithmic design as part of the governance architecture of energy systems. It shows that objective functions, constraints, training data, evaluation methods, and oversight structures can all carry policy significance.

The core argument is that policy awareness should be treated as a design imperative, not a compliance burden added after deployment. Smart energy AI systems influence affordability, reliability, sustainability, fairness, privacy, and accountability, so their design must reflect the public responsibilities of the infrastructures they help govern. This requires developers and institutions to treat regulation as a source of design intelligence rather than a barrier to innovation.

Future work should deepen the interdisciplinary collaboration required to operationalize policy-aware AI in the energy transition. Engineers, regulators, public policy scholars, utilities, consumer advocates, and communities all have roles in defining what energy AI should optimize and what it must never compromise. Policy-aware AI therefore represents both a technical design agenda and a democratic governance agenda for smart energy systems.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 24 Aug 2023 Revised: 30 Oct 2023 Accepted: 15 Nov 2023
Published online: 15 June 2025

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press

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/.

References

1. Ghoddusi H, Creamer GG, Rafizadeh N. Machine learning in energy economics and finance: a review. Energy Econ. 2019;81:709-27. https://doi.org/10.1016/j.eneco.2019.05.006.

2. Vázquez-Canteli JR, Nagy Z. Reinforcement learning for demand response: a review of algorithms and modeling techniques. Appl Energy. 2019;235:1072-89. https://doi.org/10.1016/j.apenergy.2018.11.002.

3. Yeung K. 'Hypernudge': big data as a mode of regulation by design. In: The social power of algorithms. Routledge; 2019. p.118-36.

4. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. 2018;28(4):689-707. https://doi.org/10.1007/s11023-018-9482-5.

5. Mittelstadt B. Principles alone cannot guarantee ethical AI. Nat Mach Intell. 2019;1(11):501-7. https://doi.org/10.1038/s42256-019-0114-4.

6. Wirtz BW, Weyerer JC, Geyer C. Artificial intelligence and the public sector—applications and challenges. Int J Public Adm. 2019;42(7):596-615.

7. 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. https://doi.org/10.1016/j.rser.2020.109899.

8. Perera AT, Kamalaruban P. Applications of reinforcement learning in energy systems. Renew Sustain Energy Rev. 2021;137:110618. https://doi.org/10.1016/j.rser.2020.110618.

9. Schittekatte T, Meeus L. Flexibility markets: Q&A with project pioneers. Util Policy. 2020;63:101017. https://doi.org/10.1016/j.jup.2020.101017.

10. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1(9):389-99. https://doi.org/10.1038/s42256-019-0088-2.

11. Burger S, Chaves-Ávila JP, Batlle C, Pérez-Arriaga IJ. A review of the value of aggregators in electricity systems. Renew Sustain Energy Rev. 2017;77:395-405. https://doi.org/10.1016/j.rser.2017.04.014.

12. 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. https://doi.org/10.1016/j.rser.2018.10.014.

13. Rahwan I. Society-in-the-loop: programming the algorithmic social contract. Ethics Inf Technol. 2018;20(1):5-14.

14. 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.

15. Zuiderwijk A, Chen YC, Salem F. Implications of the use of artificial intelligence in public governance: a systematic literature review and a research agenda. Gov Inf Q. 2021;38(3):101577. https://doi.org/10.1016/j.giq.2021.101577.

16. Alsaigh R, Mehmood R, Katib I. AI explainability and governance in smart energy systems: a review. Front Energy Res. 2023;11:1071291. https://doi.org/10.3389/fenrg.2023.1071291.

17. Ulbricht L, Yeung K. Algorithmic regulation: a maturing concept for investigating regulation of and through algorithms. Regul Gov. 2022;16(1):3-22.

18. Mikhaylov SJ, Esteve M, Campion A. Artificial intelligence for the public sector: opportunities and challenges of cross-sector collaboration. Philos Trans R Soc A. 2018;376(2128):20170357.

19. Vinuesa R, Azizpour H, Leite I, Balaam M, Dignum V, Domisch S, et al. The role of artificial intelligence in achieving the Sustainable Development Goals. Nat Commun. 2020;11(1):233. https://doi.org/10.1038/s41467-019-14108-y.

20. de Fine Licht K, de Fine Licht J. Artificial intelligence, transparency, and public decision-making: why explanations are key when trying to produce perceived legitimacy. AI Soc. 2020;35(4):917-26.

21. Mocanu E, Mocanu DC, Nguyen PH, Liotta A, Webber ME, Gibescu M, et al. On-line building energy optimization using deep reinforcement learning. IEEE Trans Smart Grid. 2019;10(4):3698-708.

22. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. In: Ethics, governance, and policies in artificial intelligence. Cham: Springer; 2021. p.153-83.

23. Zhang C, Wu J, Long C, Cheng M. Review of existing peer-to-peer energy trading projects. Energy Procedia. 2017;105:2563-8.

24. Niet I, van Est R, Veraart F. Governing AI in electricity systems: reflections on the EU artificial intelligence bill. Front Artif Intell. 2021;4:690237. https://doi.org/10.3389/frai.2021.690237.

25. Cath C, Wachter S, Mittelstadt B, Taddeo M, Floridi L. Artificial intelligence and the ‘good society’: the US, EU, and UK approach. Sci Eng Ethics. 2018;24(2):505-28. https://doi.org/10.1007/s11948-017-9901-7.

26. Niet I. Between vision and practice: lack of alignment between AI strategies and energy regulations in the Dutch electricity sector. Discov Artif Intell. 2022;2(1):24.

© 2025 The Author(s). | Published by Institute of Smart Energy Systems Press