ORIGINAL RESEARCH **Open access**

Human-AI Collaboration in Energy Dispatch and Grid Control: A Critical Framework for Automation, Expertise, and Operational Responsibility

David Osei[1*], Akua Afriyie[1], Kofi Adu[2]

Abstract

Energy dispatch and grid control are transitioning from operator-centred environments to settings increasingly supported, shaped, or partially driven by artificial intelligence. While this transition offers opportunities for faster situational assessment, adaptive dispatch support, and improved resilience, it also raises critical questions about the changing role of human operators, the preservation of operational expertise, and the allocation of responsibility when decisions emerge from human–AI collaboration. Existing approaches often emphasize levels of automation but give insufficient attention to how operator cognition is sustained, how trust in AI recommendations is calibrated, and how accountability is maintained across human, algorithmic, and organizational actors. This article develops a critical conceptual framework for human–AI collaboration in energy dispatch and grid control by synthesizing insights from human–automation interaction, cognitive systems engineering, energy operations research, and organizational responsibility theory. The framework moves beyond linear automation taxonomies by defining dynamic collaboration modes that vary according to situational criticality, AI confidence, operator expertise, and operational risk. It further identifies mechanisms for preserving expert judgement, preventing skill degradation, supporting calibrated reliance, and ensuring that AI systems remain transparent and contestable within control-room practice. A responsibility architecture is proposed to clarify accountability across operators, AI developers, system owners, control-room supervisors, and utility organizations. The framework also outlines context-specific applications for routine dispatch, anomaly detection, emergency response, demand response coordination, and post-event review. By integrating automation design, expertise retention, trust calibration, and responsibility allocation, the article argues that AI-supported grid control should be designed as a socio-technical collaboration rather than a transfer of authority from humans to machines. The proposed framework can support safer, more resilient, and more accountable grid operations, but it requires empirical validation in operationally realistic control-room environments.

Keywords Human-AI collaboration, Energy dispatch, Grid control, Automation, Operational expertise, Operational responsibility

*Correspondence:
David Osei
david.osei@ug.edu.gh
[1] Department of Intelligent Energy Distribution and AI, Faculty of Engineering, University of Ghana, Accra, Ghana
[2] Department of Smart Microgrids for Rural Electrification, Faculty of Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Introduction

Energy dispatch and grid control are increasingly shaped by AI-enabled forecasting, optimization, decision support, and closed-loop control tools that alter how operators perceive and act on system conditions. The move from conventional control-room support toward AI-assisted power system decision-making is visible in the growing literature on

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.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

reinforcement learning, physics-informed learning, emergency control, and intelligent grid operations [1]. Because dispatch decisions can affect system stability, continuity of supply, market outcomes, and public safety, the integration of AI into grid control should be treated as a transformation of critical infrastructure governance rather than a narrow technical upgrade [2]. Human-AI collaboration in this setting therefore requires a framework that recognizes the high-consequence nature of operational judgement.

The central problem is that automation can encroach on core operator tasks without adequately preserving the expertise needed to recognize abnormal states, contest machine recommendations, or recover from unexpected automation behaviour. Endsley argues that lessons from human-automation research remain highly relevant to autonomy because automation can create out-of-the-loop performance problems when humans are asked to supervise systems they no longer actively control [3]. In power systems, the future role of human operators in highly automated environments has been explicitly questioned as AI and automation reshape operational authority [2]. These developments make it urgent to examine not only what AI can do, but also how operators remain competent, accountable, and meaningfully involved.

The issue is not simply whether a system is manual, automated, or autonomous, but how collaboration is structured across changing operational conditions. Human-autonomy teaming research shows that performance depends on coordination, shared situation awareness, and calibrated reliance rather than on a static allocation of tasks [4]. Trust in AI also requires careful design because both over-trust and under-trust can undermine effective decision-making when automated recommendations appear authoritative but remain fallible [5]. A critical framework for grid control must therefore address automation design, expertise preservation, and responsibility allocation as interdependent dimensions rather than isolated design concerns.

This article contributes a conceptual framework for human-AI collaboration in energy dispatch and grid control structured along three axes: automation design, expertise integration, and operational responsibility. The framework draws on human-machine teaming, cognitive systems engineering, AI governance, and power system control scholarship to conceptualize how decision authority could be dynamically allocated based on system criticality, AI confidence, and operator intervention capability [6]. It also provides analytical tools for differentiating collaboration modes, responsibility structures, and application contexts in energy operations. In doing so, it treats AI-supported grid control as a socio-technical arrangement in which operational safety depends on the quality of human-AI-organization interaction.

Background

Energy dispatch and grid control: operational realities and pressures

Energy dispatch and grid control involve continuous balancing of generation, demand, network constraints, reserve availability, voltage conditions, congestion, and contingency readiness. The operational pressures facing control rooms are intensifying as renewable variability, distributed energy resources, and inverter-based assets complicate conventional assumptions about observability and controllability [7]. AI applications in power system operation, control, and planning have expanded in response to these pressures, particularly where conventional rule-based tools struggle with high-dimensional and dynamic operating states [1]. Yet the same complexity that motivates AI adoption also increases the need for human judgement when system behaviour departs from anticipated patterns.

Automation and AI in control room environments

Control room automation has evolved from alarm processing, SCADA visualization, and energy management system support toward AI-based decision aids, reinforcement learning controllers, and autonomous grid functions. Reviews of AI in power systems show a broad transition from optimization and forecasting support toward more adaptive, learning-based control architectures [8]. However, opaque AI recommendations may create automation surprises when operators cannot easily infer why an action was recommended or why a system changed state [3]. In safety-critical energy

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

operations, such opacity can weaken operator agency unless interfaces make uncertainty, assumptions, and operational consequences visible.

Human-AI collaboration paradigms and their limits

Human-AI collaboration paradigms emphasize teaming, shared cognition, adaptive automation, and joint decision-making, but their transfer to grid control requires attention to the temporal urgency and institutional responsibility of power system operations. Human-autonomy teaming studies suggest that team effectiveness depends on coordination mechanisms and shared situation awareness rather than mere human presence in the loop [6]. Machines may function as teammates only when roles, expectations, and interaction protocols are explicitly designed rather than assumed [9]. In grid control, these paradigms are useful but incomplete unless they address authority, expertise retention, and accountability under operational stress.

Operator expertise, cognition, and skill degradation

Grid operator expertise is not reducible to procedural compliance because it includes pattern recognition, anomaly interpretation, anticipation of cascading consequences, and practical judgement under uncertainty. Studies of control-room cognition show that operators rely on cue utilization and cognitive resources to manage regular operational tasks, and these forms of expertise may become fragile when automation removes opportunities for active engagement [10]. Cognitive load research in energy network control rooms further suggests that interface design and information structure influence how operators understand system conditions [11]. AI-augmented control should therefore be designed to support expertise development rather than convert operators into passive monitors of algorithmic output.

Recent empirical research on trust dynamics in AI-assisted energy decision-making demonstrates that different stakeholder groups exhibit systematically different trust patterns: AI experts display skepticism due to awareness of AI limitations, energy domain experts tend toward overtrust influenced by AI's confident presentation, and non-experts show inconsistent trust behaviors [12]. This finding, derived from the VIRTSI (Variability and Impact of Reciprocal Trust States towards Intelligent Systems) methodology applied to urban energy management contexts, has direct implications for grid control room staffing and training. If energy experts operating AI-supported dispatch tools are predisposed to overtrust, then expertise preservation cannot rely solely on technical competence but must include deliberate trust calibration mechanisms that expose operators to AI failure modes and encourage systematic validation of recommendations [12]. **Figure 1** conceptualizes AI-supported grid control as a cognitive partnership in which operator expertise, active engagement, and calibrated trust must be preserved to prevent passive monitoring, skill degradation, and overreliance on confident algorithmic recommendations.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

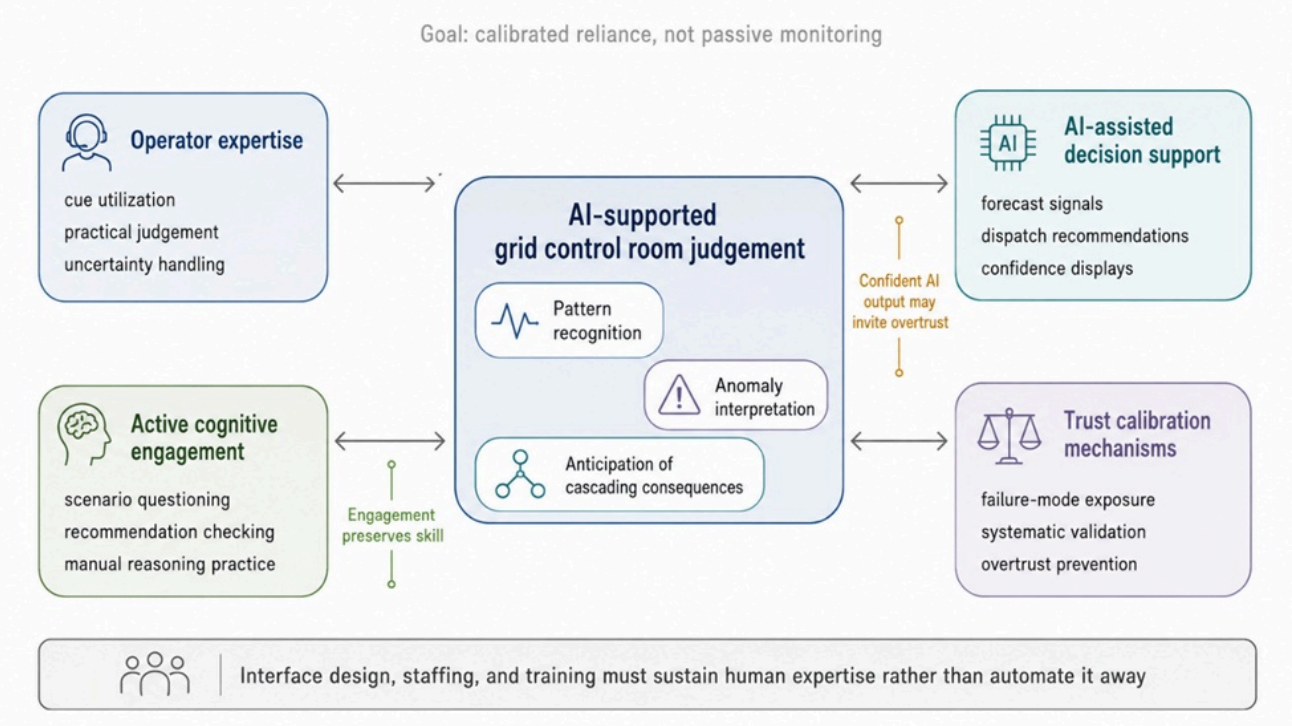


Figure 1. Preserving operator expertise and calibrating trust in AI-supported grid control rooms

Responsibility, liability, and the ‘Moral Crumple Zone’

Responsibility becomes especially difficult when AI systems recommend or initiate actions that affect grid safety, reliability, or market outcomes. The concept of the moral crumple zone warns that humans may absorb blame for failures of complex automated systems even when they had limited knowledge, control, or realistic opportunity to intervene [13]. Similar concerns appear in discussions of AI-mediated decision-making, where responsibility can be displaced onto the most visible human actor rather than the organizational or technical design chain [14]. Energy dispatch and grid control therefore require explicit responsibility architecture so that accountability is not improvised only after adverse events.

Conceptual Foundation

Collaboration as dynamic allocation of functions

Human-AI collaboration in grid control can be conceptualized as a dynamic allocation of sensing, interpretation, recommendation, authorization, execution, monitoring, and recovery functions. Rather than treating automation as a fixed level, the framework proposes that collaboration modes could be allocated based on AI confidence, operator workload, system criticality, and reversibility of action [4]. Adaptive forms of human-autonomy teaming indicate that authority should shift as the operational context changes, but such shifts require clear role boundaries and shared situation awareness [15]. In energy dispatch, this means the human-AI relationship should be designed as a controlled coordination process rather than a one-way transfer of tasks from operator to machine.

Expertise as a resilience resource

Operator expertise should be treated as a resilience resource because it enables sensemaking when models fail, data become incomplete, or operating conditions are novel. Research on the future role of operators in automated electric power systems suggests that even highly automated grids may still depend on human understanding during abnormal and degraded states [2]. Team situation awareness research also indicates that coordination between humans and

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

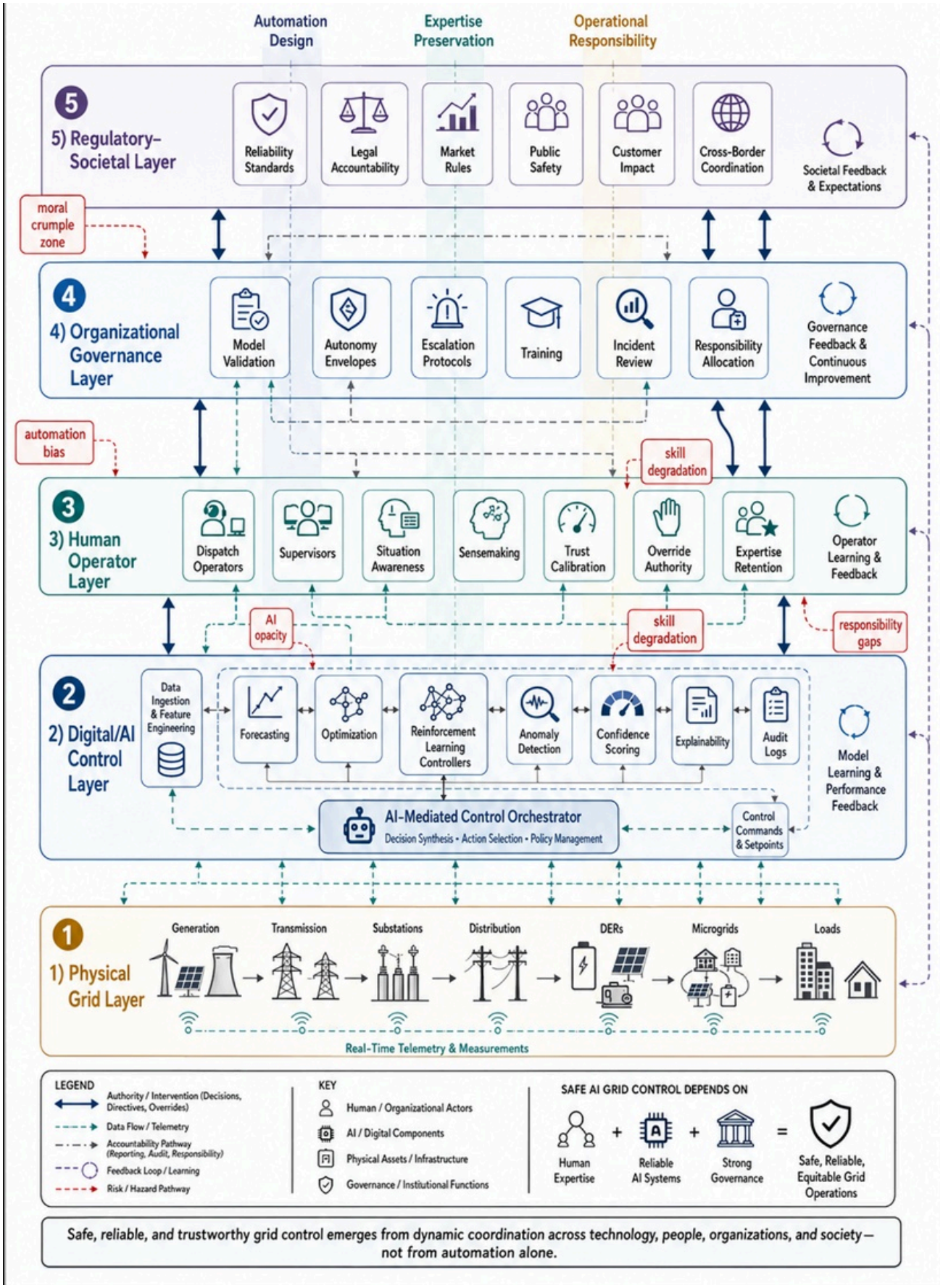
autonomous agents depends on maintaining a shared representation of task state, constraints, and likely consequences [6]. Expertise preservation is therefore not a nostalgic defense of manual control but a practical requirement for resilience in AI-mediated grid operations.

Responsibility as a designable property

Operational responsibility can be conceptualized as a designable property of the human-AI-organization system rather than as a retrospective judgment assigned only after failure. AI governance research emphasizes that human oversight must be institutionally meaningful, not merely symbolic, if accountability is to be preserved in automated decision environments [16]. Responsibility allocation in grid control should therefore be linked to foresight, control authority, intervention capability, documentation, and contestability. A collaboration architecture that does not specify these elements risks creating responsibility gaps in which AI influences high-stakes outcomes while accountability remains unclear.

Figure 2 presents the proposed multi-level socio-technical system map, showing how AI-enabled grid control emerges from interactions among the physical grid, digital control systems, human operators, organizational governance, and regulatory–societal structures. The figure highlights automation design, expertise preservation, and operational responsibility as cross-cutting dimensions that must be jointly managed to support safe, resilient, and accountable human–AI collaboration.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press



© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Figure 2. Multi-Level Socio-Technical System Map for Human–AI Collaboration in Energy Dispatch and Grid Control

Automation and Human Involvement Dimensions

Collaboration modes along the automation spectrum

The framework distinguishes collaboration modes across an automation spectrum that ranges from manual control to advisory support, consensual decision-making, supervised autonomy, and full automation with oversight. This spectrum should not be interpreted as a maturity ladder in which more automation is always better, because human-automation research shows that inappropriate autonomy can reduce situation awareness and weaken human readiness to intervene [3]. In power systems, autonomous control may be appropriate for fast, bounded, and well-validated functions, while broader dispatch and restoration choices may require explicit operator authorization [17]. The central design question is therefore which collaboration mode is appropriate for which operational condition.

Conditions triggering mode transitions

Mode transitions should be triggered by factors such as AI confidence, quality of input data, operator workload, time criticality, system stress, anomaly detection, and the reversibility of the proposed action. Reinforcement learning applications in smart grids illustrate the appeal of adaptive control but also reveal the need for safeguards when learning-based systems act in uncertain or rapidly changing environments [18]. Trust calibration research suggests that operators need experience with both AI competence and AI failure to judge when reliance is appropriate [19]. A dynamic framework should therefore make transitions visible, justified, and contestable rather than allowing automation to silently expand its authority.

Table 1 presents the proposed triggers for human-AI collaboration mode transitions in grid control operations.

Table 1. Trigger Conditions for Dynamic Human-AI Collaboration Mode Transitions in Grid Control

Trigger Dimension	Condition for Higher Autonomy (e.g., Supervised Autonomy)	Condition for Lower Autonomy (e.g., Advisory or Consensual)	Assessment Indicator	Transition Pathway
AI confidence	High confidence (e.g., >0.9) with stable uncertainty estimates	Low confidence (<0.7) or wide confidence intervals	Probabilistic output from AI model; ensemble agreement score	System proposes mode shift; operator confirms or overrides
Input data quality	Complete, validated telemetry with no known sensor faults	Missing data, telemetry gaps, or active sensor alarms	Data freshness timestamps; sensor health status flags	Automatic demotion to advisory mode until data quality restored
Operator workload	Low workload (e.g., normal conditions, multiple operators available)	High workload (e.g., cascading events, understaffing)	SCADA alarm rate; queued tasks; physiological indicators	Supervised autonomy may increase during high workload with mandatory post-hoc audit
Time criticality	High urgency (sub-second to second response window)	Low urgency (minutes to hours available for deliberation)	Time-to-action requirement for system stability	Autonomous execution within bounded envelope; full logging required
System stress state	Normal or alert state with predictable	Emergency, restoration, or black-	Contingency classification;	Mandatory return to consensual or exception-

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

	dynamics	start conditions	frequency/voltage deviation thresholds	driven collaboration
Anomaly detection	No anomalies; system behaviour matches expected patterns	Anomaly detected or model uncertainty flags raised	Statistical process control limits; reconstruction error thresholds	Hold and escalate: AI suspends action, calls operator
Reversibility of action	Actions are reversible with low consequence of error (e.g., advisory displays)	Actions are irreversible or have high consequence (e.g., breaker operation)	Consequence severity matrix pre-defined per action type	Higher autonomy only for reversible actions; irreversible actions require consensual authorization

Designing for transparency and operator agency

Collaboration modes must be supported by transparent explanations, explicit authority boundaries, and interface designs that preserve operator agency. Human-centered AI principles emphasize that reliable and trustworthy systems should keep humans informed, able to intervene, and capable of understanding system behaviour [20]. In grid control, explanation interfaces would be expected to show not only a recommended action but also its assumptions, uncertainty, constraint trade-offs, and likely operational consequences. As shown in **Table 2,** different collaboration modes place different demands on operator expertise and produce different responsibility configurations.

Table 2. Human-AI Collaboration Modes in Grid Control: Characteristics, Expertise Implications, and Responsibility Focus

Collaboration Mode	Operator Role	AI Role	Typical Situation	Expertise Demands	Primary Responsibility Locus
Advisory interpretation	Interprets AI recommendations and decides whether to act	Generates forecasts, alerts, rankings, or candidate actions	Normal dispatch planning, congestion anticipation, reserve assessment	Requires diagnostic judgement, model literacy, and ability to compare AI advice with operational experience	Operator retains decision responsibility; engineering team responsible for tool validity
Consensual authorization	Confirms or rejects AI-proposed action before execution	Proposes action with explanation, confidence indication, and constraint rationale	Voltage control, redispatch, curtailment, or switching where time allows review	Requires rapid consequence evaluation and understanding of system interdependencies	Shared responsibility between operator authorization and AI system governance
Supervised autonomy	Monitors bounded AI execution and intervenes when thresholds are crossed	Executes pre-authorized actions inside defined operational envelopes	Fast corrective actions, automated DER coordination, local frequency	Requires envelope awareness, anomaly recognition, and readiness to interrupt automation	Management and engineering hold responsibility for envelope design; operator responsible for

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

			or voltage support		feasible intervention
Exception-driven collaboration	Takes control when AI detects uncertainty, conflict, or abnormal state	Flags uncertainty, suspends action, and requests human resolution	Conflicting constraints, incomplete data, rare contingencies, degraded observability	Requires high-level sensemaking, contingency reasoning, and recovery planning	Operator leads exception response; organization responsible for escalation protocols
Full automation with oversight	Reviews logs, audits performance, and validates post-event outcomes	Executes validated routine or time-critical actions without immediate human approval	Highly bounded protection-like functions or rapid local control where delay is unsafe	Requires retrospective audit competence and understanding of automation limits	Organization, engineering, and regulator share primary responsibility for approved autonomous function

Expertise Integration in Collaborative Control

Maintaining operator competence in AI-augmented environments

Maintaining operator competence in AI-augmented environments may require simulation-based training, periodic manual operation, structured review of AI recommendations, and involvement of operators in model evaluation. Human factors research indicates that automation can weaken active control skills when operators are asked only to monitor rather than practice decision-making [3]. In grid control, training should expose operators to automation failures, degraded data, conflicting recommendations, and rare contingencies so that expertise remains available when AI support is unreliable. Operator involvement in AI review could also help translate tacit operational knowledge into design feedback for future system updates.

Trust calibration and automation bias mitigation

Trust calibration is central because operators may either over-rely on AI recommendations or reject useful AI support when explanations are poor or prior failures remain unresolved. Empirical reviews of trust in AI show that transparency, reliability, task characteristics, and user experience shape whether people rely appropriately on automated systems [5]. Automation bias is especially concerning in grid control because an AI recommendation framed as optimal may suppress operator questioning during time pressure [21]. The framework therefore requires uncertainty communication, explanation of assumptions, safe exposure to AI errors, and institutional permission for operators to challenge automation.

Table 3 summarizes recommended interventions for calibrating operator trust across different trust states and risk patterns.

Table 3. Trust Calibration Interventions for Grid Control Operators Working with AI Decision Support

Trust State	Observable Indicators	Operational Risk	Cognitive Mechanism	Recommended Intervention	Measurement Approach
Distrust (under-	Operator rejects AI recommendations	Loss of efficiency;	Negative prior experience; lack	Exposure to AI successes with	Rejection rate tracking; response

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

reliance)	despite demonstrated accuracy; verifies all outputs manually	operator fatigue; slower response times	of transparency; perceived AI incompetence	explainable outputs; structured trust-building scenarios in simulation	time measurement; post-shift confidence surveys
Calibrated trust (appropriate reliance)	Operator accepts AI recommendations when confidence is high; questions when uncertainty is present	Low risk; optimal human-AI team performance	Accurate mental model of AI capabilities and limitations; experience with both successes and failures	Reinforce with feedback on calibration accuracy; maintain through periodic exposure to AI edge cases	Agreement rate vs. AI accuracy comparison; override justification quality assessment
Overtrust (over-reliance)	Operator accepts incorrect AI recommendations; fails to monitor or question; automation bias evident	Automation bias; missed anomalies; skill degradation	Authority bias; cognitive ease; lack of exposure to AI failures	Deliberate exposure to AI errors in simulation; forced verification at critical decision points; uncertainty visualisation emphasis	Override rate (too low); eye-tracking for monitoring patterns; post-event error attribution
Complacent trust	Operator ignores AI outputs or disengages from monitoring	Loss of situation awareness; delayed intervention	Low workload; high reliability leading to vigilance decrement	Adaptive automation varying AI authority; periodic unscheduled automation failures in training; secondary task loading	Monitoring dwell time; response to infrequent alarms; physiological vigilance measures

Designing for joint learning and adaptation

Joint learning in collaborative control can be conceptualized as a bidirectional process in which AI systems learn from operator feedback while operators learn the strengths, limits, and failure modes of AI tools. Human-autonomy teaming research suggests that coordination improves when human and autonomous agents develop shared expectations and adaptive interaction patterns over time [22]. In grid control, this may require mechanisms for annotating AI recommendations, documenting overrides, feeding post-event lessons into model governance, and updating training scenarios. Such adaptation should be governed carefully, because learning systems that change without sufficient explanation or validation could undermine both trust and accountability.

Operational Responsibility Architecture

Mapping responsibility across the decision chain

Operational responsibility in human-AI grid control should be mapped across the full decision chain, including data acquisition, model design, validation, recommendation generation, operator interpretation, action authorization, execution, monitoring, and post-event review. The operator may hold responsibility for situated judgement when sufficient information, time, and intervention authority are available, but responsibility should not be assigned to the operator when the AI system, interface, or organizational protocol makes meaningful oversight unrealistic [13]. AI developers and

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

engineering teams would be expected to bear responsibility for model validity, testing, boundary definition, and failure-mode disclosure, while control-room managers and utility executives should be accountable for deployment conditions and staffing policies [16]. Regulators also play a secondary but important role by defining oversight expectations, auditability requirements, and acceptable autonomy boundaries for safety-critical energy operations [7].

The Role of organisational culture and governance

Responsibility allocation is not only a technical matter because the same AI recommendation can be interpreted differently depending on organizational culture, training norms, and incident-review practices. A just culture would avoid treating the operator as a convenient liability sink when failures arise from poor interface design, inadequate testing, management pressure, or ambiguous escalation rules [14]. Teaming research suggests that human-machine collaboration depends on role clarity and mutual coordination, which are organizational achievements as much as interface properties [23]. In grid control, governance should therefore define how operators may challenge AI recommendations, how disagreements are escalated, and how accountability is distributed across the socio-technical system.

Accountability mechanisms: audit trails, explanation logs, and review

Accountability mechanisms should make the human-AI decision process reconstructable after both successful and adverse events. Explanation logs could document the AI recommendation, uncertainty estimate, input data state, violated or satisfied constraints, operator response, override rationale, and final executed action, while audit trails would allow post-event reviewers to distinguish human judgement errors from design, data, validation, or governance failures [20]. AI-supported emergency control and explainable reinforcement learning research show that interpretability becomes particularly important when automated recommendations affect stability-related decisions under time pressure [24]. As shown in **Table 4**, accountability should be differentiated by decision category rather than imposed as a single generic requirement.

Table 4. Operational Responsibility Matrix for Human-AI Collaboration in Grid Control

Decision Category	Primary Responsibility	Secondary Responsibility	AI Role	Accountability Mechanism	Contestability Pathway
Routine economic dispatch	Operator or dispatch supervisor for final authorization	Market operations team and model engineering group	Recommends feasible schedules, congestion options, reserve adjustments, and forecast-informed alternatives	Dispatch log linking AI recommendation, operator approval, market constraint set, and deviation rationale	Operator review, market audit, and internal dispatch governance panel
Renewable balancing and curtailment	Control-room operator within approved operating policy	Forecasting team, asset management, and regulator	Estimates variability, proposes balancing actions, and ranks curtailment or flexibility options	Forecast provenance record, curtailment justification note, and uncertainty trace	Asset owner appeal, regulatory review, and post-event fairness assessment
Emergency switching or	Operator when feasible intervention time	Protection engineering, reliability	Detects instability, proposes or executes	High-resolution event recorder, AI confidence log,	Independent reliability review, operator debrief,

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

remedial action	exists; automation governance body when action is pre-authorized	management, and executive safety governance	bounded emergency action, and records triggering conditions	override window record, and emergency review file	and engineering failure-mode analysis
Restoration after outage	Restoration coordinator or senior operator	Field crews, distribution management, engineering support, and customer operations	Suggests restoration sequence, prioritizes constraints, and updates situational model as field data arrive	Restoration decision chronology, field confirmation record, and exception log	Restoration review board, affected-customer complaint pathway, and regulator inquiry
Market participation through automated bids	Market operations management	AI governance lead, compliance office, and regulator	Generates bid strategies, identifies arbitrage opportunities, and monitors risk limits	Bid rationale archive, model version record, compliance screen, and human approval trace	Compliance escalation, market monitor investigation, and executive accountability review
Decentralized DER or microgrid coordination	Local operating authority or aggregation manager	Platform provider, community governance body, and distribution utility	Coordinates local flexibility, peer-to-peer exchange, or islanding support within defined limits	Local transaction log, autonomy boundary record, and participant-impact assessment	Participant dispute process, platform audit, and utility-regulator mediation

Legal and regulatory dimensions of responsibility

Legal and regulatory dimensions of responsibility become more complex when AI systems contribute to decisions but cannot be treated as moral or legal agents in the same way as human organizations. Recent discussions of human oversight under AI governance regimes emphasize that oversight must be operationally feasible, documented, and connected to real intervention capacity rather than being a formal checkbox [16]. In energy operations, regulators may need to distinguish between advisory AI, bounded autonomy, and fully automated action because each mode creates different expectations for operator review and organizational liability [7]. Responsibility architecture should therefore align technical autonomy boundaries with legal accountability, licensing conditions, and reliability obligations.

Application Contexts and Implementation Scenarios

Transmission system operation and wide-area control

In transmission system operation, human-AI collaboration would be expected to prioritize reliability, situational awareness, and rapid coordination across wide-area contingencies. AI may support contingency screening, remedial action selection, frequency-security assessment, and emergency control, but transmission operators must remain able to interpret whether recommendations are valid under changing topology and uncertain telemetry [25]. Physics-informed neural networks and trajectory-constraint approximation methods illustrate how AI can support power system stability assessment, yet these tools should be embedded within transparent operating envelopes and reviewable control protocols [26]. In this context,

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

the framework favours consensual authorization or supervised autonomy for high-speed bounded functions, with strong post-event audit and operator override provisions.

Distribution network management with high DER penetration

Distribution management creates different collaboration challenges because distributed energy resources, flexible loads, field crews, and prosumers introduce a more fragmented operational environment. AI-based Volt-VAR control and microgrid reinforcement learning show how automated systems can coordinate voltage, flexibility, and local constraints, but responsibility may become distributed across utilities, aggregators, vendors, and customers [27]. Human-AI collaboration in this context should be designed around boundary management, exception handling, and clear escalation from automated local control to human distribution operators. Because distribution systems often have less complete observability than transmission networks, AI confidence and data-quality indicators should strongly influence whether automation remains advisory or becomes supervised autonomy.

Microgrids, virtual power plants, and decentralised control

Microgrids, virtual power plants, and peer-to-peer coordination environments change the meaning of human-AI collaboration because the relevant human actor may be a community operator, facility manager, aggregator, or platform supervisor rather than a professional transmission dispatcher. Reinforcement learning for grid-connected microgrids illustrates the potential for automated local optimization, but such systems may affect resilience, fairness, and participant expectations in ways that exceed technical control objectives [17]. In decentralized settings, responsibility should be anchored in platform governance, autonomy envelopes, participant consent, and transparent dispute pathways. **Table 5** compares how the framework varies across transmission, distribution, and microgrid or community control environments.

Table 5. Contextual Variation of Human-AI Collaboration, Expertise, and Responsibility Across Grid Operational Environments

Grid Operational Environment	Typical Collaboration Mode	Expertise Profile Required	Key Responsibility Challenges	Recommended Governance Adaptation
Transmission control	Consensual authorization for major actions; supervised autonomy for bounded fast-response functions	Senior system-operations expertise, contingency reasoning, stability awareness, and cross-area coordination judgement	Short intervention windows, cascading-risk attribution, multi-actor reliability obligations, and difficulty distinguishing operator error from automation-design failure	Require strict autonomy envelopes, real-time explanation displays, reliability-grade audit trails, and independent post-contingency review
Distribution management	Advisory and exception-driven collaboration, with supervised autonomy for localized voltage or DER coordination	Hybrid operational expertise combining network constraints, field operations, DER behaviour, and customer-impact interpretation	Fragmented responsibility among utility operators, field crews, DER aggregators, vendors, and prosumers	Establish escalation protocols, data-quality thresholds, vendor accountability clauses, and customer-impact contestability processes
Microgrid or community control	Supervised autonomy and full automation with oversight for routine local optimization;	Local resilience knowledge, platform literacy, community-governance	Ambiguous accountability when AI decisions affect local resilience, participant equity, or	Define community-facing autonomy boundaries, participant consent mechanisms, platform

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

	human intervention for islanding, restoration, or participant disputes	awareness, and basic electrical operations competence	peer-to-peer exchange outcomes	auditability, and utility-regulator mediation pathways

Cross-border and interconnected system challenges

Cross-border and interconnected systems complicate human-AI collaboration because operational decisions may affect neighbouring control areas operating under different reliability rules, data-sharing arrangements, and regulatory cultures. The ENTSO-E Transparency Platform illustrates the importance and difficulty of shared electricity data infrastructures across jurisdictions, even before AI-mediated control is added to the operational chain [28]. When AI supports redispatch, balancing, or emergency coordination across interconnected systems, responsibility may be divided among multiple system operators, market actors, and regulators. The framework therefore requires interoperable audit trails, shared explanation standards, and pre-agreed escalation rules for AI-influenced decisions that cross organizational boundaries.

Table 6 identifies key responsibility and accountability challenges specific to cross-border and interconnected system operations with AI support.

Table 6. Responsibility and Accountability Challenges in Cross-Border AI-Mediated Grid Operations

Challenge Category	Specific Issue	AI Contribution	Responsibility Fragmentation	Required Governance Mechanism	Example Scenario
Divergent reliability standards	Adjacent control areas operate under different reliability rules and autonomy limits	AI recommends redispatch or balancing actions affecting neighbouring area	System operator A, system operator B, market operator, national regulator	Inter-operator service-level agreements specifying AI authority limits across boundaries; joint validation of AI recommendations	AI proposes counter-trade between Norway and Sweden; Swedish rules permit supervised autonomy while Norwegian rules require consensual authorization
Data sharing asymmetry	One control area has higher telemetry quality or model fidelity than its neighbour	AI confidence and recommendations degrade when input data quality differs asymmetrically	Data provider, data consumer, both system operators, market surveillance	Mandatory data quality disclosure; AI confidence thresholds for cross-border actions; fallback to manual coordination	German TSO has high-quality phasor measurements; Dutch TSO has degraded telemetry; AI confidence for coordinated action differs unsafely
Regulatory oversight mismatch	One jurisdiction permits bounded AI autonomy; another requires human-in-the-loop for same function	Cross-border emergency control AI proposes action acceptable on one side but prohibited on the other	Both TSOs, both national regulators, regional coordination body (e.g., ENTSO-E,	Pre-agreed minimum common autonomy standard; escalation to human coordinators when jurisdictional conflict	AI proposes automatic load shedding for frequency control; permitted in Finland but

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

			regional reliability council)	detected; mutual recognition agreements	requires control-room authorisation in Estonia
Liability and redress gaps	Failure caused by AI recommendation affects consumers or market participants across border	AI decision log attribution ambiguous; which operator's governance approved the action?	Liability insurers, both TSOs, affected parties, multiple national courts	Cross-border liability protocol; joint AI audit trail with tamper-evident logging; international arbitration pathway	AI-curtailed wind generation in Denmark affects balancing market prices in Germany; Danish wind farm seeks compensation but cannot identify responsible party
Language and explanation barriers	AI explanations generated in one language; operators in neighbouring area lack fluency	Explanation quality degrades in translation; technical terms misinterpreted	Both control rooms, AI vendor, training organisations	Standardised technical vocabulary for cross-border AI explanations; visual explanation formats (non-linguistic); joint simulation training	French TSO AI explains voltage support action in French; Belgian operators misinterpret confidence indication, leading to inappropriate override

Governance, Policy, and Organizational Implications

Embedding the framework into operational codes and standards

The proposed framework could be embedded into operational codes, operator licensing, reliability procedures, procurement requirements, and AI governance standards for energy control environments. Human-centered AI design principles indicate that trustworthy AI requires reliability, safety, transparency, and meaningful human control, but these values must be translated into concrete operating rules for dispatch and grid control [20]. Smart grid intrusion-detection work also shows that explainable AI may be needed not only for operational control but also for cybersecurity-related decision support in critical infrastructure [29]. Embedding the framework into codes and standards would help prevent collaboration practices from being determined solely by vendor defaults or short-term efficiency pressures.

Organisational readiness and change management

Adopting human-AI collaboration in grid control requires organizational readiness, including training redesign, new governance roles, revised incident-review practices, and explicit policies for operator challenge and override. Research on machines as teammates suggests that AI collaboration alters team structures and work processes rather than simply adding another tool to existing workflows [9]. Utilities may therefore need AI operations stewards, model governance committees, human factors specialists, and cross-functional review boards that connect engineering, control-room practice, cybersecurity, compliance, and executive responsibility. Without such organizational change, AI deployment may increase technical capability while leaving expertise preservation and accountability underdeveloped.

Evaluation Strategy for the Framework

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Conceptual soundness and human factors alignment

The framework can be evaluated by assessing whether its collaboration modes align with human factors principles such as situation awareness, trust calibration, workload, explainability, and meaningful human control. Conceptual soundness depends on whether automation, expertise, and responsibility reinforce one another rather than create contradictions.

Expert review and simulation-based validation

Future validation could use expert panels, control-room simulations, training exercises, and field studies. These should examine whether operators understand AI explanations, recognize automation limits, challenge inappropriate recommendations, and use accountability logs for post-event learning.

Transferability across grid contexts and cultures

The framework should be tested across different grid sizes, regulatory systems, market structures, infrastructure conditions, and organizational cultures. Core principles such as auditability and meaningful oversight may be general, while autonomy limits, licensing expectations, and escalation procedures require local adaptation.

Limitations

Conceptual and unvalidated nature

The framework remains conceptual and has not yet been tested in operational control rooms, simulations, or utility deployments. It should therefore be treated as a structured conceptual contribution rather than a proven implementation model.

Generality versus specificity in diverse grid environments

Grid environments differ in infrastructure, regulation, staffing, data quality, market design, and organizational culture. The framework cannot prescribe one universal automation mode or responsibility structure, and should be adapted through local risk assessment, stakeholder review, and empirical testing.

Conclusion

This article developed a three-dimensional framework for human–AI collaboration in energy dispatch and grid control, linking automation design, expertise preservation, and operational responsibility. Its central contribution is to show that grid control should be understood as collaborative, cognitive, and accountable work rather than as a simple transfer of tasks from humans to AI.

As grid control becomes more AI-mediated, meaningful human involvement remains essential for safety, resilience, and public trust. AI can support dispatch and control when it strengthens operator judgement, clarifies uncertainty, and operates within transparent boundaries. It becomes problematic when it erodes expertise, obscures authority, or diffuses accountability.

The framework offers diagnostic and design tools for distinguishing collaboration modes, assigning responsibility across decision types, and adapting governance to transmission, distribution, and microgrid contexts. Future research should use simulations, expert review, field pilots, and regulatory engagement to test the framework under realistic uncertainty, time pressure, and organizational constraints.

Acknowledgements

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 11 Mar 2023 Revised: 17 May 2023 Accepted: 02 Jun 2023
Published online: 15 June 2024

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. Pandey U, Pathak A, Kumar A, Mondal S. Applications of artificial intelligence in power system operation, control and planning: a review. Clean Energy. 2023;7(6):1199-218.

2. Prostejovsky AM, Brosinsky C, Heussen K, Westermann D, Kreusel J, Marinelli M. The future role of human operators in highly automated electric power systems. Electr Power Syst Res. 2019;175:105883. https://doi.org/10.1016/j.epsr.2019.105883.

3. Endsley MR. From here to autonomy: lessons learned from human-automation research. Hum Factors. 2017;59(1):5-27.

4. O'Neill T, McNeese NJ, Barron A, Schelble B. Human-autonomy teaming: a review and analysis of the empirical literature. Hum Factors. 2022;64(5):904-38. https://doi.org/10.1177/0018720820960865.

5. Glikson E, Woolley AW. Human trust in artificial intelligence: review of empirical research. Acad Manag Ann. 2020;14(2):627-60.

6. Demir M, McNeese NJ, Cooke NJ. Team situation awareness within the context of human-autonomy teaming. Cogn Syst Res. 2017;46:3-12.

7. Biagini V, Subasic M, Oudalov A, Kreusel J. The autonomous grid: automation, intelligence and the future of power systems. Energy Res Soc Sci. 2020;65:101460.

8. Heymann F, Quest H, Lopez-Garcia TB, Ballif C, Galus MD. Reviewing 40 years of artificial intelligence applied to power systems: a taxonomic perspective. Energy AI. 2024;15:100322. https://doi.org/10.1016/j.egyai.2023.100322.

9. Seeber I, Bittner E, Briggs RO, De Vreede T, De Vreede GJ, Elkins A, et al. Machines as teammates: a research agenda on AI in team collaboration. Inf Manag. 2020;57(2):103174.

10. Sturman D, Wiggins MW, Auton JC, Loft S, Helton WS, Westbrook JI, et al. Control room operators' cue utilization predicts cognitive resource consumption during regular operational tasks. Front Psychol. 2019;10:1967.

11. Afzal U, Prouzeau A, Lawrence L, Dwyer T, Bichinepally S, Liebman A, et al. Investigating cognitive load in energy network control rooms: recommendations for future designs. Front Psychol. 2022;13:812677.

12. Sola A, Corchero C, Salom J, Sanmarti M. Multi-domain urban-scale energy modelling tools: a review. Sustain Cities Soc. 2020;54:101872.

13. Elish MC. Moral crumple zones: cautionary tales in human-robot interaction. Engag Sci Technol Soc. 2019;5:40-60. https://doi.org/10.17351/ests2019.260.

14. Hohenstein J, Jung M. AI as a moral crumple zone: the effects of AI-mediated communication on attribution and trust. Comput Hum Behav. 2020;106:106190.

15. Demir M, Cooke NJ, Amazeen PG. A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. Cogn Syst Res. 2018;52:497-507. https://doi.org/10.1016/j.cogsys.2018.07.029.

16. Laux J. Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act. AI Soc. 2024;39(6):2853-66.

17. Arwa EO, Folly KA. Reinforcement learning techniques for optimal power control in grid-connected microgrids: a comprehensive review. IEEE Access. 2020;8:208992-9007. https://doi.org/10.1109/ACCESS.2020.3038735.

18. Zhang D, Han X, Deng C. Review on the research and practice of deep learning and reinforcement learning in smart grids. CSEE J Power Energy Syst. 2018;4(3):362-70.

19. Lebiere C, Blaha LM, Fallon CK, Jefferson B. Adaptive cognitive mechanisms to maintain calibrated trust and reliance in automation. Front Robot AI. 2021;8:652776. https://doi.org/10.3389/frobt.2021.652776.

20. Shneiderman B. Human-centered artificial intelligence: reliable, safe & trustworthy. Int J Hum Comput Interact. 2020;36(6):495-504.

21. Ruschemeier H, Hondrich LJ. Automation bias in public administration—an interdisciplinary perspective from law and psychology. Gov Inf Q. 2024;41(3):101953.

22. Demir M, Likens AD, Cooke NJ, Amazeen PG, McNeese NJ. Team coordination and effectiveness in human-autonomy teaming. IEEE Trans Hum Mach Syst. 2019;49(2):150-9.

23. Walliser JC, de Visser EJ, Wiese E, Shaw TH. Team structure and team building improve human-machine teaming with autonomous agents. J Cogn Eng Decis Mak. 2019;13(4):258-78.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

24. Zhang K, Zhang J, Xu PD, Gao T, Gao DW. Explainable AI in deep reinforcement learning models for power system emergency control. IEEE Trans Comput Soc Syst. 2022;9(2):419-27. https://doi.org/10.1109/TCSS.2021.3096824.

25. Glavic M, Fonteneau R, Ernst D. Reinforcement learning for electric power system decision and control: past considerations and perspectives. IFAC-Pap. 2017;50(1):6918-27.

26. Huang B, Wang J. Applications of physics-informed neural networks in power systems—a review. IEEE Trans Power Syst. 2023;38(1):572-88.

27. Wang W, Yu N, Gao Y, Shi J. Safe off-policy deep reinforcement learning algorithm for volt-var control in power distribution systems. IEEE Trans Smart Grid. 2020;11(4):3008-18.

28. Hirth L, Mühlenpfordt J, Bulkeley M. The ENTSO-E Transparency Platform—a review of Europe's most ambitious electricity data platform. Appl Energy. 2018;225:1054-67.

29. Yayla A, Haghnegahdar L, Dincelli E. Explainable artificial intelligence for smart grid intrusion detection systems. IT Prof. 2022;24(5):18-24.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press