ORIGINAL RESEARCH **Open access**

Digital Twin Governance for AI-Based Energy Systems: Transparency, Reliability, and Decision Accountability

Youssef Hariri[1*], Hassan Nasser[1], Fatima Zahra[2]

Abstract

Digital twins, coupled with AI, are becoming integral to energy systems for monitoring, simulation, and autonomous decision-making. Their governance, however, lags behind technical deployment, raising concerns about opaque twin-driven decisions. Existing digital twin implementations lack coherent governance frameworks addressing transparency, reliability, and decision accountability simultaneously. This gap exposes energy systems to risks from inaccurate, unexplainable, or unaccountable twin outputs that influence critical grid and market decisions. This article proposes a governance framework for digital twins in AI-based energy systems that systematically integrates transparency, reliability, and decision accountability. The framework is constructed through conceptual synthesis of digital twin literature, AI governance, energy system governance, and cyber-physical accountability. No empirical data are used. The framework comprises three intertwined governance dimensions: transparency, reliability, and decision accountability. These dimensions include model explainability, data lineage, twin-to-physical documentation, continuous validation, uncertainty management, robustness criteria, auditability, responsibility allocation, human oversight, and contestation. A governance framework for digital twins can transform them from opaque technical mirrors into accountable socio-technical decision-support tools. Such a framework aligns twin capabilities with the safety, fairness, and trust expectations of energy system stakeholders.

Keywords Digital twin, Transparency, Governance, AI-based energy systems, Reliability, Decision accountability

*Correspondence:
Youssef Hariri
youssef.hariri@um5.ac.ma
[1] Department of Artificial Intelligence for Renewable Energy Systems, Faculty of Sciences and Engineering, Mohammed V University, Rabat, Morocco
[2] Department of Smart Grid Machine Learning, Faculty of Engineering, University of Casablanca, Casablanca, Morocco

Introduction

Digital twins are increasingly positioned as core digital infrastructures within AI-based energy systems, supporting applications that range from building energy management and smart campuses to urban multi-energy systems and grid-facing optimization [1]. In energy contexts, these twins no longer function only as static replicas of physical assets but are increasingly designed to support simulation, real-time monitoring, and anticipatory decision-making [2]. As AI is embedded into these cyber-physical representations, digital twins may influence decisions about flexibility, maintenance, dispatch, and operational control [3]. This shift from passive mirroring toward active decision influence makes governance a central requirement rather than an external compliance concern.

The governance challenge emerges from three interconnected deficits. First, many digital twin systems provide limited transparency into the data streams, model structures, synchronization assumptions, and AI reasoning processes that shape their outputs [4]. Second, the reliability of twin-generated recommendations may be difficult to establish when

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

models operate under changing physical, market, climatic, or behavioral conditions [5]. Third, accountability can become ambiguous when twin developers, AI modelers, data providers, system operators, and asset owners all contribute to a decision chain without clearly assigned responsibility [6]. These deficits reinforce one another because opaque models are difficult to validate, unreliable outputs are difficult to defend, and untraceable decisions are difficult to contest.

A dedicated governance framework is therefore needed because transparency, reliability, and accountability cannot be treated as isolated technical add-ons. Broader explainable AI literature shows that interpretability and explanation are necessary but insufficient unless they are connected to institutional review and operational responsibility [7]. Similarly, AI governance scholarship emphasizes that accountable systems require layered oversight, ethical principles, and organizational procedures rather than model-level transparency alone [8]. For energy digital twins, these issues are amplified because twin outputs can affect critical infrastructure, market participation, and the distribution of risks among consumers, operators, and communities.

The objective of this article is to develop a conceptual governance framework for digital twins in AI-based energy systems structured around transparency, reliability, and decision accountability. Building on digital twin scholarship that characterizes twins as data-rich, lifecycle-oriented cyber-physical systems [9], the article conceptualizes governance as a design requirement that spans technical architecture, operational use, and institutional embedding. The framework specifies how transparency artifacts, reliability assurance mechanisms, and accountability chains could be governed through mutually reinforcing procedures. It contributes to digital twin and responsible AI debates by repositioning the energy digital twin as a socio-technical governance object rather than only a technical decision-support tool.

Background

Digital twins in energy systems: architectures and applications

Digital twins in energy systems can be conceptualized across several architectural types, including asset twins, process twins, system twins, and operational twins. Energy applications include building energy twins, renewable plant twins, smart campus twins, urban multi-energy twins, and grid-facing decision-support twins [1]. Studies of smart-building and urban energy twins show that these systems depend on continuous integration of sensor data, simulation models, and operational logic to maintain a meaningful link between physical and virtual environments [10]. As digital twins become more deeply integrated with AI, their architectures increasingly support learning, prediction, anomaly detection, and optimization rather than simple visualization.

AI-based decision-making in energy twins

AI is often embedded within digital twins to support forecasting, demand-side response, anomaly detection, optimization, and what-if simulation. Reviews of machine learning in energy systems show that AI models are widely used to represent nonlinear relationships, forecast energy demand, and guide operational decisions under uncertainty [11]. When these models are placed inside digital twins, their opacity may propagate through the twin interface, making recommendations appear objective even when they depend on contestable assumptions or incomplete data [12]. Governance must therefore address not only the AI model itself but also the way twin interfaces translate AI outputs into operationally persuasive recommendations.

Governance of cyber-physical and digital systems

Governance of cyber-physical and digital systems traditionally draws on IT governance, data governance, safety management, and organizational accountability. However, digital twins present a distinctive challenge because they combine data infrastructures, simulation models, AI components, physical asset interfaces, and human decision pathways within one evolving system [13]. Layered AI governance approaches suggest that oversight should operate across technical, organizational, and societal levels rather than focusing only on algorithmic performance [14]. Energy digital

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

twins require such adaptation because they interact with critical infrastructure, market rules, public-service obligations, and safety-sensitive operational environments.

Transparency, reliability, and accountability as interlocking governance pillars

Transparency refers to the ability to understand and inspect the data, model, assumptions, and decision logic of the digital twin. Reliability refers to the ability of the twin to produce dependable outputs under defined conditions, including through validation, robustness assessment, and uncertainty communication [15]. Accountability refers to the capacity to trace decisions, assign responsibility, and enable review or contestation when twin-informed actions affect stakeholders [8]. These pillars are interlocking because reliability without transparency cannot generate warranted trust, while accountability without reliability reduces responsibility to procedural formality rather than substantive assurance.

Recent research on socio-technical energy transitions emphasizes that governance gaps in digital infrastructures often arise not from missing technical features but from misaligned institutional responsibilities and accountability vacuums . Similarly, studies of AI-based energy management systems show that operational trust depends less on algorithmic accuracy alone than on transparent error handling and clear fault attribution when systems deviate from expected behavior [16].

Existing governance gaps in digital twin energy deployments

The digital twin literature in energy has developed strong technical accounts of architecture, integration, monitoring, and optimization, yet governance remains comparatively underdeveloped [1]. Reviews of digital twin concepts and applications emphasize enabling technologies, interoperability, and implementation challenges, but they rarely specify accountability structures for decisions shaped by twin outputs [4]. The responsible AI literature provides relevant principles of explainability, fairness, human oversight, and accountability, yet these principles are often not tailored to the lifecycle and cyber-physical coupling of energy twins [17]. The resulting gap is an absence of an integrated framework that connects twin transparency, reliability assurance, and decision accountability within the specific institutional setting of AI-based energy systems.

Conceptual Foundation of Digital Twin Governance

The digital twin as a socio-technical governance object

A digital twin in an AI-based energy system should be conceptualized not merely as a technical artifact but as a socio-technical governance object. It mediates between physical energy assets, data infrastructures, AI models, operational interfaces, and institutional decisions [18]. This mediation is especially significant because digital twin outputs can shape how operators perceive system states, how markets value flexibility, and how organizations allocate responsibility for action. Governance must therefore supervise the interfaces between the physical system, the computational representation, and the human or institutional decision environment.

Figure 1 illustrates the proposed multi-layer governance architecture for AI-enabled energy digital twins, showing how power system operations, human decision-making, AI system capabilities, and governance and institutional structures are vertically integrated through bidirectional information, decision, and feedback flows.

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

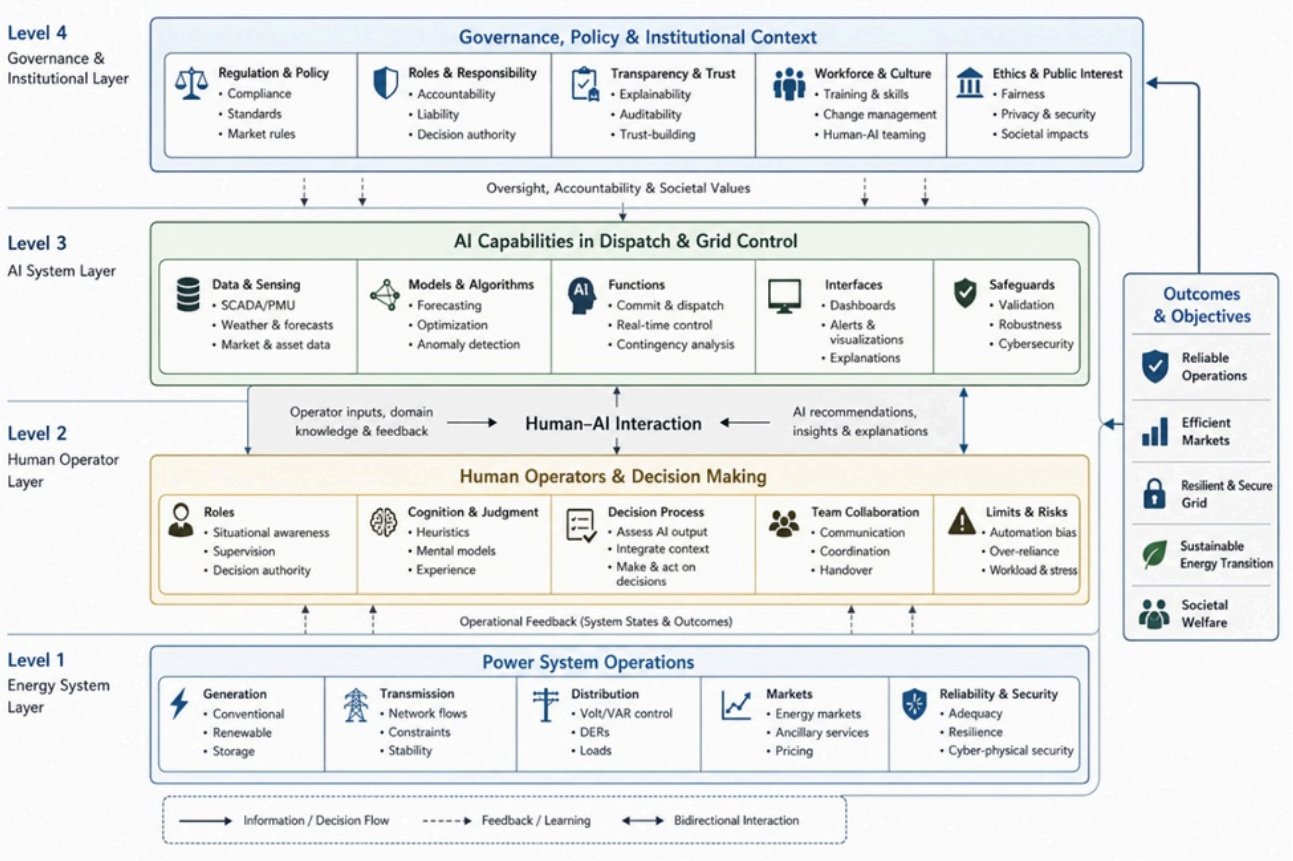

Figure 1. Multi-layer governance architecture for AI-enabled energy digital twin systems spanning power system operations, human decision-making, AI system functions, and governance-policy context

The Triple governance mandate: transparency, reliability, accountability

The proposed framework is organized around a triple governance mandate: transparency, reliability, and accountability. Explainable AI scholarship shows that interpretability techniques can help reveal how model outputs are generated, but such techniques must be connected to decision contexts and user needs [16]. Reliability assurance ensures that twin outputs remain credible under known operational conditions, while accountability mechanisms ensure that decisions informed by those outputs can be traced, reviewed, and challenged. None of the three mandates can be sacrificed because transparent but unreliable twins mislead users, reliable but opaque twins undermine trust, and transparent reliable twins without accountability leave affected parties without recourse.

Table 1 summarizes the distinct governance objectives, operational mechanisms, and failure consequences associated with each of the three pillars.

Table 1. Governance Pillars for Digital Twins in AI-Based Energy Systems: Definitions, Mechanisms, and Failure Modes

Governance Pillar	Definition in Energy Twin Context	Operational Mechanisms	Failure Consequence
Transparency	Ability to inspect data provenance, model architecture, synchronization assumptions, and AI reasoning that shape twin outputs	Model cards, data sheets, twin passports, synchronization logs, explainability interfaces	Opaque recommendations that cannot be validated or contested

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

Reliability	Capacity to produce dependable outputs under defined operational conditions with quantified uncertainty	Continuous validation, uncertainty quantification, robustness testing, drift monitoring, recertification protocols	Confident but incorrect twin outputs leading to operational errors
Decision Accountability	Ability to trace decisions, assign responsibility, and enable review or contestation of twin-informed actions	Responsibility allocation matrices, audit trails, decision provenance records, contestation channels	No clear recourse or responsibility when twin-informed decisions cause harm

Lifecycle perspective: from twin design to decommissioning

Digital twin governance should follow a lifecycle perspective that spans design, development, commissioning, operation, adaptation, and decommissioning. Lifecycle-oriented digital twin scholarship emphasizes that twins evolve with physical systems, data streams, model updates, and operational objectives rather than remaining fixed after deployment [9]. This means that governance requirements should also evolve, with design-stage documentation, commissioning-stage validation, operation-stage monitoring, update-stage recertification, and decommissioning-stage archival responsibility. A lifecycle approach prevents governance from being reduced to one-time approval and instead treats it as continuous stewardship of a changing cyber-physical system.

Framework Components: Transparency and Explainability

Model and data transparency in energy digital twins

Model and data transparency require digital twin developers and operators to document AI model architectures, training data sources, data lineage, synchronization methods, feature transformations, and assumptions about physical system behavior. Smart-building twin research highlights the importance of integrating building information models, Internet of Things data, and operational data, but each integration point can introduce errors or hidden assumptions if not documented [10]. In AI-based energy systems, transparency should therefore extend from raw data acquisition to model inference and from twin synchronization to human-facing recommendations. This framework treats transparency as a layered documentation obligation rather than a single explanation attached to a final output.

Explainability mechanisms for twin-driven decisions

Explainability mechanisms for twin-driven decisions should connect physical system states, AI inferences, simulated scenarios, and recommended actions. Methods such as feature attribution, local explanations, counterfactual reasoning, and rule-based summaries can help users understand why a twin recommends a fault response, dispatch adjustment, or maintenance action [19]. However, explanations must be adapted to digital twin contexts because the relevant question is not only why an AI model produced an output but also how the virtual representation corresponds to the physical system. A governance framework should therefore require explanations that make visible the relationship between twin state, data uncertainty, model reasoning, and decision consequence.

Transparency documentation and auditability

Transparency should be operationalized through standardized documentation artifacts such as model cards, data sheets, twin passports, synchronization logs, change histories, and decision trace reports. These artifacts would be expected to support internal audit, external regulatory review, and post-incident investigation by preserving information about data provenance, model purpose, update history, and known limitations. Digital twin surveys emphasize that definitions and

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

design implications vary widely across systems, which makes structured documentation especially important for comparability and oversight [6]. As shown in **Table 2**, different types of energy digital twins require distinct transparency, reliability, and accountability controls because their decision roles and risk profiles differ.

Table 2. Digital Twin Types and Their Governance Requirements for Transparency, Reliability, and Accountability

Digital Twin Type	Key Energy Use Case	Primary Transparency Needs	Primary Reliability Requirements	Accountability Focus
Asset Twin	Monitoring a wind turbine, transformer, battery unit, inverter, or building system component	Clear documentation of sensor provenance, asset model assumptions, maintenance history, and component-level state estimation logic	Calibration against physical inspection records, sensor fault detection, degradation tracking, and bounded uncertainty for asset condition outputs	Assignment of responsibility for maintenance recommendations, deferred repairs, and escalation when twin evidence conflicts with field inspection
Process Twin	Representing energy workflows such as demand response, maintenance scheduling, charging coordination, or building energy optimization	Visibility into workflow rules, AI optimization objectives, operational constraints, and human override points	Stress-testing under workflow disruption, missing data, changed schedules, and nonstandard operating conditions	Traceability of process adjustments, operator overrides, and automated prioritization decisions affecting service delivery
System Twin	Coordinating grid zones, virtual power plants, multi-energy districts, or interconnected renewable portfolios	Disclosure of system boundaries, aggregation assumptions, interdependency models, and cross-asset data transformations	Validation of coupling logic across assets, robustness under cascading disturbances, and consistency between simulated and observed system behavior	Responsibility allocation across asset owners, aggregators, platform operators, and system operators when system-level recommendations redistribute risk
Operational Twin	Supporting near-real-time dispatch, fault response, congestion management, or emergency control	Real-time visibility into twin state, data freshness, confidence indicators, model version, and automated decision rules	Continuous divergence monitoring, real-time anomaly detection, fail-safe thresholds, and recertification after model or infrastructure updates	Immediate accountability for automated or operator-approved actions, including audit trails for time-critical interventions
Planning Twin	Evaluating infrastructure investment, network expansion, flexibility procurement, or long-term decarbonization pathways	Documentation of scenario assumptions, demand projections, technology cost assumptions, and policy constraints	Scenario plausibility review, sensitivity testing, and transparent treatment of uncertainty in long-horizon projections	Accountability for how twin outputs inform capital allocation, regulatory justification, and distribution of long-term system benefits and burdens

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

Framework Components: Reliability and Robustness

Verification, validation, and uncertainty quantification

Reliability governance begins with verification, validation, and uncertainty quantification. Verification asks whether the digital twin has been built according to its design specifications, while validation asks whether the twin adequately represents the relevant physical or operational system for its intended decision use [15]. Continuous calibration research in smart-building twins illustrates that twin models may require repeated adjustment as operating conditions, occupancy patterns, equipment behavior, and data streams change [5]. For governance purposes, every decision-relevant twin output should therefore be accompanied by a statement of uncertainty, domain of applicability, and validation status.

Data quality and twin synchronization governance

Data quality and synchronization governance are essential because digital twins depend on continuous flows of sensor, operational, market, weather, and asset-condition data. If data feeds are delayed, corrupted, biased, incomplete, or poorly synchronized with physical system states, the twin may create a misleading representation of reality [20]. Governance mechanisms should therefore define minimum requirements for sensor reliability, data freshness, missing-data handling, synchronization frequency, and alerting when twin-to-physical divergence exceeds acceptable limits. In AI-based energy systems, these controls are especially important because AI models may amplify poor data quality by producing confident but unreliable recommendations.

Robustness to changing conditions and adversarial inputs

Digital twin governance must also address robustness under changing conditions, including distribution shifts, extreme weather, asset degradation, cyber-physical disruption, and adversarial manipulation. Energy demand-side response and machine learning reviews show that AI-based energy decisions depend on assumptions about users, system states, and operating environments that may change over time [21]. A reliable digital twin should therefore be subject to continuous stress-testing, scenario review, adversarial-risk assessment, and monitoring for model drift. Robustness governance does not imply that a twin can eliminate uncertainty; rather, it requires that uncertainty, failure modes, and safe operating boundaries are explicitly managed before twin outputs are used in critical decisions.

Reliability Assurance and Verification

Twin performance monitoring and anomaly detection

Reliability assurance for energy digital twins should include continuous monitoring of the divergence between the virtual representation and the physical energy system. In smart grid, transportation, and smart city digital twin applications, divergence can arise from sensor drift, delayed data ingestion, topology changes, equipment degradation, or incomplete cyber-physical integration [22]. Governance structures should therefore require automated alerts when the twin's representation no longer corresponds to observed physical behavior within predefined decision-use boundaries. These alerts should not merely flag technical malfunction but should trigger review of whether the twin remains authorized to support operational decisions.

Recertification and periodic validation protocols

Digital twins should be periodically recertified, especially after major asset replacement, network reconfiguration, AI model retraining, data pipeline redesign, or changes in operational objectives. Building energy flexibility research shows that digital twins can combine data-driven models and physical logic to support dispatch decisions, but such hybrid systems may require renewed validation when the operating environment changes [23]. Recertification could be governed through protocols similar to physical asset inspection, where the twin's assumptions, synchronization quality, model behavior, and

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

decision permissions are reviewed at defined intervals. This approach would help prevent outdated twins from retaining decision authority after their physical or institutional context has changed.

Reliability transparency: communicating limitations to users

Reliability information must be communicated in ways that human operators, managers, regulators, and affected stakeholders can understand. Digital twin interfaces that use extended reality, dashboards, or immersive visualization may improve operational awareness, but they can also create a sense of certainty if they do not display confidence boundaries, data freshness, known limitations, and failure modes [24]. Governance should therefore require reliability transparency as a user-facing obligation, not only as an internal engineering record. A twin recommendation should be accompanied by clear information about whether the twin is operating inside its validated domain, whether data streams are current, and whether uncertainty is material for the decision at hand.

Regulatory standards and industry-grade reliability benchmarks

Reliability governance should also connect energy digital twins to emerging industry-grade benchmarks and standardization practices, even where energy-specific digital twin governance standards remain incomplete. Cyber-physical building energy management research suggests that digital twins and AI can support operational intelligence, but their use in critical infrastructure requires stronger assurance of safety, robustness, and organizational control [25]. Existing digital twin standardization efforts from industry and cyber-physical systems could be adapted to define documentation requirements, validation protocols, version-control expectations, and auditability criteria for energy applications. Such standards would not replace contextual governance but would provide a shared baseline for regulators, utilities, vendors, and asset owners.

Framework Components: Decision Accountability

Responsibility allocation across twin-actor networks

Decision accountability in digital twin governance requires explicit allocation of responsibility across twin developers, data providers, AI modelers, asset owners, operators, platform vendors, market participants, and regulators. Semantic digital twin research in construction shows that twins are not isolated models but structured information environments that connect multiple professional roles, data standards, and lifecycle responsibilities [26]. In energy systems, this means that accountability should follow the decision chain rather than rest only with the operator who sees the final recommendation. Governance rules should specify who is responsible for data quality, model updates, validation decisions, operational approval, override authority, and post-incident review.

Audit trails and decision provenance

Audit trails should record the twin state, input data, model version, AI recommendation, uncertainty indicators, human interventions, override decisions, and final action for every critical twin-enabled decision. Predictive maintenance digital twins illustrate the importance of connecting condition monitoring, model outputs, and maintenance choices into a traceable decision pathway [27]. In AI-based energy systems, auditability should also include decision provenance, meaning that reviewers can reconstruct how a recommendation moved from data ingestion through model inference to operational execution. This requirement supports safety investigation, regulatory reporting, liability assessment, and organizational learning after failures or contested decisions.

Contestability and redress for twin-driven decisions

Contestability mechanisms are needed when digital twin–enabled decisions affect consumers, communities, market actors, or infrastructure users. Algorithmic governance scholarship cautions against assuming that complex computational systems are inherently inscrutable, because accountability can be strengthened through documentation, procedural

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

access, and reviewable decision records [28]. Energy digital twins should therefore include channels through which affected parties can challenge decisions such as load shedding, flexibility dispatch, tariff-related optimization, maintenance deferral, or investment prioritization. **Figure 2** illustrates decision accountability as a traceable governance structure in which responsibility follows the full twin-enabled decision chain, from data provision and model development to operational approval, override, audit review, and redress. The figure emphasizes that accountability should not rest only with the final operator but should be distributed across all actors whose inputs, models, approvals, and interventions shape the decision outcome.

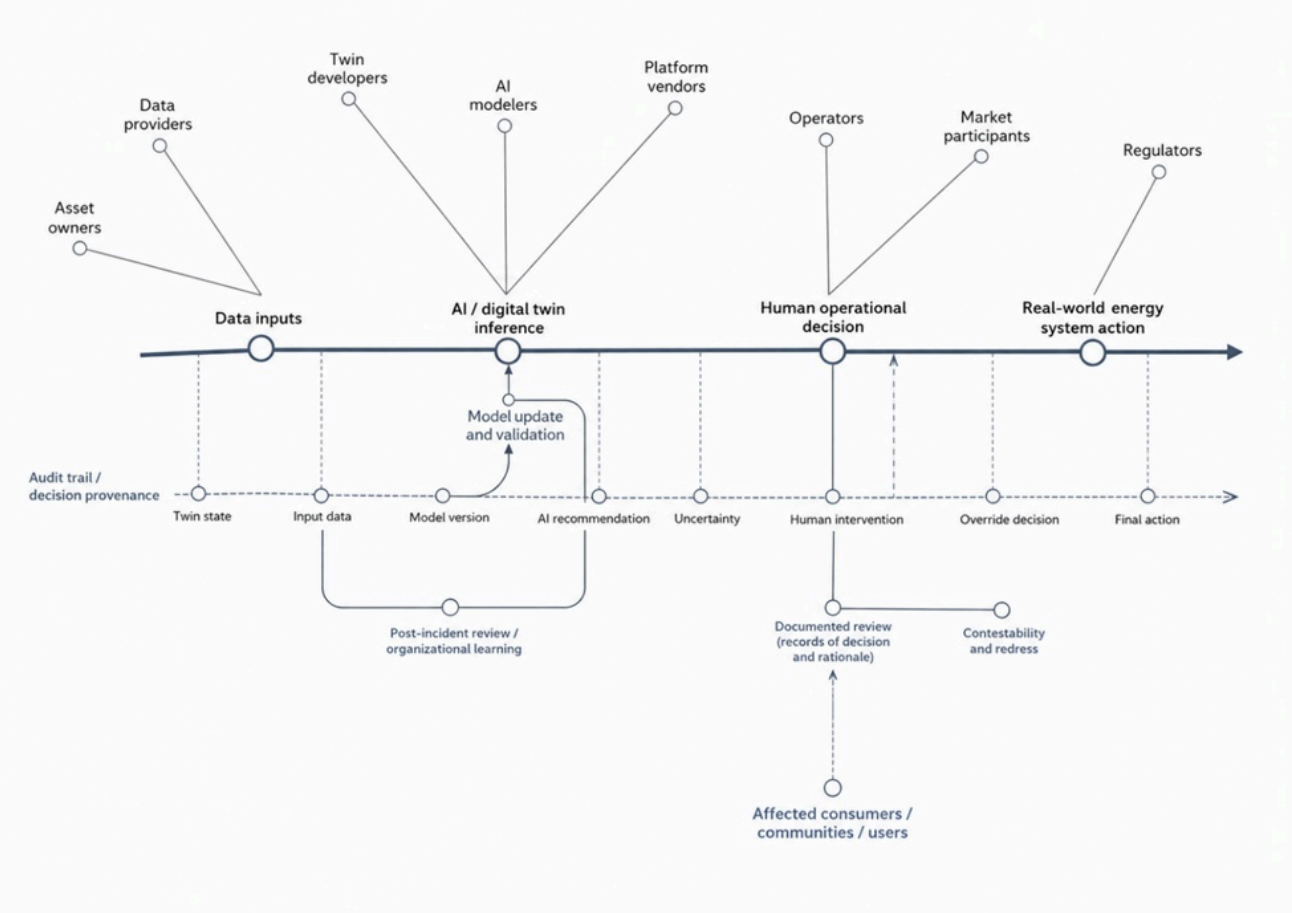

Figure 2. Decision accountability across twin-actor energy digital twin governance

Governance Integration and Institutional Embedding

Embedding digital twin governance into energy utility and market structures

The proposed framework should be embedded into existing utility governance, market rules, regulatory reporting, procurement requirements, and operational risk-management systems. Research on fairness in machine learning shows that governance cannot rely only on collecting more data or improving models, because institutional decisions about objectives, constraints, and responsibility shape whether algorithmic systems produce acceptable outcomes [29]. For energy digital twins, this means that governance requirements should be included in vendor contracts, model approval procedures, market participation rules, grid-code compliance, and internal operational authorization. Digital twin outputs

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

should not become decision infrastructure unless the organization can demonstrate transparency documentation, reliability assurance, and accountability assignment.

Multi-stakeholder governance mechanisms

Multi-stakeholder governance mechanisms could include digital twin oversight committees, external technical review panels, consumer advisory groups, regulator-supervised audit boards, and cross-functional reliability councils. Urban multi-energy digitalization research shows that digital twins can operate across life-cycle phases and institutional boundaries, which increases the need for governance arrangements that connect technical experts, planners, operators, regulators, and users [2]. Such bodies would be expected to review high-risk twin use cases, approve updates, examine incidents, assess contestation claims, and advise on whether twin outputs should retain decision authority. Multi-stakeholder oversight is particularly important where digital twin decisions distribute costs, risks, or reliability burdens unevenly across communities.

Evaluation Strategy for the Governance Framework

Conceptual coherence and alignment with governance principles

The framework can first be evaluated conceptually by assessing whether its components are internally coherent, nonredundant, and aligned with recognized principles of responsible AI and digital system governance. AI ethics scholarship identifies recurring principles such as transparency, accountability, fairness, non-maleficence, responsibility, and human oversight, but these principles require operational translation for specific technical domains [17]. In this article, conceptual coherence would mean that transparency artifacts support reliability assurance, reliability evidence supports accountability claims, and accountability mechanisms create incentives for better transparency and validation. The framework's value lies in connecting these principles to the specific lifecycle, cyber-physical coupling, and decision authority of energy digital twins.

Expert and stakeholder validation

The framework could be further assessed through expert and stakeholder validation, such as Delphi studies, regulatory workshops, utility review panels, vendor consultations, and consumer-representative deliberation. AI4People emphasizes that ethical AI governance should connect abstract principles to practical mechanisms capable of guiding real institutional behavior [8]. For energy digital twins, validation should examine whether the proposed governance dimensions are understandable, implementable, proportionate to risk, and compatible with existing operational responsibilities. Such validation would not be intended to generate performance metrics but to refine the framework's categories, clarify accountability roles, and identify implementation barriers.

Adaptability to different energy digital twin architectures

The framework should also be evaluated for adaptability across centralized, distributed, federated, asset-specific, and system-level digital twin architectures. Digital twin literature shows wide variation in definitions, technical configurations, data models, and application domains, which means governance must be modular rather than tied to one architecture [6]. In a centralized operational twin, accountability may focus on system operator authorization and real-time audit trails, whereas in a distributed virtual power plant twin it may focus on platform governance, participant rights, and market transparency. The proposed framework can therefore be conceptualized as a governance architecture that adjusts its controls according to decision risk, system scale, and the degree of automation.

Limitations

Conceptual abstraction and the need for implementation pilots

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

This article remains conceptual and does not test the proposed framework in a deployed energy digital twin environment. Digital twin research across industry emphasizes that practical implementation depends on domain-specific data quality, system integration, organizational capacity, and lifecycle management conditions that vary substantially across sectors [9]. As a result, the framework requires further validation through implementation pilots in settings such as building energy management, virtual power plants, distribution networks, renewable asset management, and multi-energy districts. Such pilots would help refine the framework's documentation requirements, accountability roles, and reliability thresholds without converting the article into an empirical performance study.

Evolving technology and regulatory landscapes

A second limitation is that digital twin and AI technologies are evolving quickly, while regulatory expectations for explainability, safety, accountability, and critical infrastructure governance remain unsettled. Explainable AI literature demonstrates that explanation methods, interpretability taxonomies, and responsible AI practices continue to develop, which means governance frameworks must remain adaptable rather than fixed [7]. Energy digital twins may also incorporate new forms of reinforcement learning, generative AI interfaces, automated control, and market optimization that create governance questions not fully anticipated here. The framework should therefore be treated as a revisable conceptual architecture that requires continuous updating as technologies, standards, and institutional expectations change.

Conclusion

This article has proposed an integrated governance framework for digital twins in AI-based energy systems centered on transparency, reliability, and decision accountability. Its central contribution is to connect these three dimensions as mutually reinforcing governance requirements rather than separate technical or organizational concerns. The framework positions the digital twin as a socio-technical decision infrastructure whose legitimacy depends on how it is documented, validated, monitored, audited, and contested. In doing so, it shifts attention from what digital twins can technically optimize to how their decision influence should be governed.

Governance should be designed into the digital twin lifecycle from the beginning rather than retrofitted after operational failure, regulatory scrutiny, or stakeholder contestation. Design-stage decisions about data lineage, model architecture, synchronization logic, uncertainty communication, and human oversight shape the later accountability of the entire system. A twin that is not transparent at design time may be difficult to validate during operation and nearly impossible to audit after a disputed decision. Lifecycle governance therefore treats responsible twin deployment as continuous stewardship rather than one-time technical commissioning.

Future progress will require collaboration among researchers, utilities, regulators, technology vendors, market operators, consumer representatives, and infrastructure owners. Such collaboration should focus on piloting the framework, refining energy-specific governance standards, and developing shared expectations for documentation, validation, auditability, and redress. Digital twins can support more adaptive, intelligent, and resilient energy systems only if their decision power is matched by transparent and accountable governance. The next phase of digital twin development should therefore treat governance not as a constraint on innovation but as a condition for trustworthy energy system transformation.

Acknowledgements

None

Conflict of interest

None

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

Financial support

None

Ethics statement

None

Received: 26 Jan 2023 Revised: 31 Mar 2023 Accepted: 17 Apr 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. do Amaral JVS, dos Santos CH, Montevechi JAB, de Queiroz AR. Energy digital twin applications: a review. Renew Sustain Energy Rev. 2023;188:113891. https://doi.org/10.1016/j.rser.2023.113891.

2. Koirala B, Cai H, Khayatian F, Munoz E, An JG, Mutschler R, et al. Digitalization of urban multi-energy systems: advances in digital twin applications across life-cycle phases. Adv Appl Energy. 2024;16:100196. https://doi.org/10.1016/j.adapen.2024.100196.

3. Shen Z, Arraño-Vargas F, Konstantinou G. Artificial intelligence and digital twins in power systems: trends, synergies and opportunities. Digital Twin. 2024;1(3):11.

4. Fuller A, Fan Z, Day C, Barlow C. Digital twin: enabling technologies, challenges and open research. IEEE Access. 2020;8:108952-71. https://doi.org/10.1109/ACCESS.2020.2998358.

5. Eneyew DD, Capretz MA, Bitsuamlak GT. Continuous model calibration framework for smart-building digital twin: a generative model-based approach. Appl Energy. 2024;375:124080. https://doi.org/10.1016/j.apenergy.2024.124080.

6. Barricelli BR, Casiraghi E, Fogli D. A survey on digital twin: definitions, characteristics, applications, and design implications. IEEE Access. 2019;7:167653-71. https://doi.org/10.1109/ACCESS.2019.2953499.

7. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82-115.

8. 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.

9. Tao F, Zhang H, Liu A, Nee AYC. Digital twin in industry: state-of-the-art. IEEE Trans Ind Inform. 2019;15(4):2405-15. https://doi.org/10.1109/TII.2018.2873186.

10. Eneyew DD, Capretz MA, Bitsuamlak GT. Toward smart-building digital twins: BIM and IoT data integration. IEEE Access. 2022;10:130487-506. https://doi.org/10.1109/ACCESS.2022.3229370.

11. Mosavi A, Salimi M, Faizollahzadeh Ardabili S, Rabczuk T, Shamshirband S, Varkonyi-Koczy AR. State of the art of machine learning models in energy systems: a systematic review. Energies. 2019;12(7):1301.

12. Ahmad T, Chen H, Guo Y, Wang J. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: a review. Energy Build. 2018;165:301-20.

13. Minerva R, Lee GM, Crespi N. Digital twin in the IoT context: a survey on technical features, scenarios, and architectural models. Proc IEEE. 2020;108(10):1785-824. https://doi.org/10.1109/JPROC.2020.2998530.

14. Gasser U, Almeida VA. A layered model for AI governance. IEEE Internet Comput. 2017;21(6):58-62.

15. Jones D, Snider C, Nassehi A, Yon J, Hicks B. Characterising the digital twin: a systematic literature review. CIRP J Manuf Sci Technol. 2020;29:36-52. https://doi.org/10.1016/j.cirpj.2020.02.002.

16. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access. 2018;6:52138-60.

17. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1(9):389-99.

18. Semeraro C, Lezoche M, Panetto H, Dassisti M. Digital twin paradigm: a systematic literature review. Comput Ind. 2021;130:103469. https://doi.org/10.1016/j.compind.2021.103469.

19. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv. 2018;51(5):1-42.

20. Zaballos A, Briones A, Massa A, Centelles P, Caballero V. A smart campus' digital twin for sustainable comfort monitoring. Sustainability. 2020;12(21):9196. https://doi.org/10.3390/su12219196.

21. 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.

22. Jafari M, Kavousi-Fard A, Chen T, Karimi M. A review on digital twin technology in smart grid, transportation system and smart city: challenges and future. IEEE Access. 2023;11:17471-84.

23. Song Y, Xia M, Chen Q, Chen F. A data-model fusion dispatch strategy for the building energy flexibility based on the digital twin. Appl Energy. 2023;332:120496. https://doi.org/10.1016/j.apenergy.2022.120496.

24. Yeom S, Kim J, Kang H, Jung S, Hong T. Digital twin (DT) and extended reality (XR) for building energy management. Energy Build. 2024;323:114746. https://doi.org/10.1016/j.enbuild.2024.114746.

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

25. Agostinelli S, Cumo F, Guidi G, Tomazzoli C. Cyber-physical systems improving building energy management: digital twin and artificial intelligence. Energies. 2021;14(8):2338.

26. Boje C, Guerriero A, Kubicki S, Rezgui Y. Towards a semantic Construction Digital Twin: directions for future research. Autom Constr. 2020;114:103179.

27. Aivaliotis P, Georgoulias K, Chryssolouris G. The use of digital twin for predictive maintenance in manufacturing. Int J Comput Integr Manuf. 2019;32(11):1067-80.

28. Kroll JA. The fallacy of inscrutability. Philos Trans R Soc A Math Phys Eng Sci. 2018;376(2133):20180084.

29. Veale M, Binns R. Fairer machine learning in the real world: mitigating discrimination without collecting sensitive data. Big Data Soc. 2017;4(2):2053951717743530. https://doi.org/10.1177/2053951717743530.

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