oxera

Ofgem call for input on AI assurance

—

Response by Oxera Consulting LLP
12 August 2026

Executive summary

The transition to a net-zero, highly dynamic energy sector relies on the successful integration of artificial intelligence (AI). We welcome Ofgem's call for input on AI assurance.[1] As an economic consultancy with a long history in energy market regulation and deep expertise in algorithmic auditing, we are responding to highlight the critical intersection between technical AI capabilities and real-world market dynamics. Our goal is to share our experience with Ofgem to support shaping a framework that focuses on the unique economics, innovation drivers and risks of the energy sector.

In this response, our focus spans licensed energy networks, retailers and market operators (the 'energy sector'), recognising that each faces distinct AI integration challenges and regulatory obligations. We also recognise that AI enables new tech entities to enter the market—such as smart-home ecosystems and independent load controllers—and the importance of the Energy System Value Chain (ESVC) approach to ensure comprehensive oversight.[2]

To provide market participants with clarity to drive innovation while also ensuring appropriate AI governance, we consider it beneficial for a future framework to be structured around five broad themes:

1 **Scope and proportionality (the 'additionality test')**: Sector-specific assurance requirements should focus on areas where they provide clear additional value over existing cross-economy frameworks. Any new requirements should be limited to risks arising from unique characteristics of the energy sector that are not addressed by bodies such as the ICO, NCSC and DRCF. This will focus compliance efforts where sector-specific intervention is genuinely needed.

2 **Governance and accountability**: As energy companies increasingly rely on third-party AI models and specialist software vendors, accountability can become fragmented. It would be beneficial if Ofgem's framework provided clear principles on the allocation of responsibilities between regulated firms and their AI suppliers,

[1] Ofgem (2026), 'Call for input: AI assurance', 17 June, available at: https://www.ofgem.gov.uk/sites/default/files/2026-06/ai-assurance-call-for-input.pdf.
[2] The Ofgem review (22 April 2026) proposes an ESVC approach as a way 'to close existing regulatory gaps and ensure that potential harms created by new entities are addressed in an appropriate and timely manner'. Available at: https://www.gov.uk/government/publications/ofgem-review-2026/ofgem-review-final-report.

ensuring that accountability remains transparent and contractual arrangements appropriately manage upstream risks.

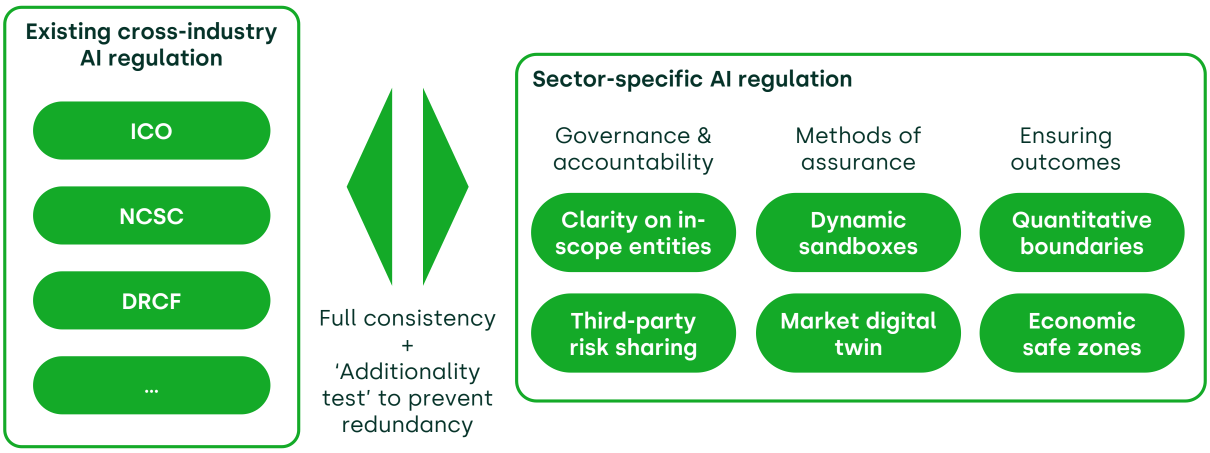
3 **Technical assurance and post-deployment monitoring**: We believe that evaluating an AI algorithm in a contained environment or relying on static documentation is insufficient. Assurance practices must include real-world and outcome-based testing and monitoring. This should include testing models against historical, messy operational data within evolving digital twins, rather than relying solely on siloed point-in-time audits.

4 **Consumer outcomes and energy-system risks**: Qualitative mandates for concepts such as 'algorithmic fairness' can create operational uncertainty. To provide regulatory certainty, we recommend establishing clear, risk-based quantitative boundaries for acceptable outcomes, including for pricing, service differentials, outage prioritisation and access to financial support.

5 **Innovation and experimentation**: To enable rapid but safe modernisation, an AI assurance framework should act as an enabling mechanism. By establishing shared testing environments and providing a layer of regulatory protection for responsible experimentation, Ofgem has an opportunity to give energy companies clear operational boundaries within which they can innovate confidently.

By combining thorough, risk-based testing of proposed AI systems with an outcome-focused supervisory approach, we believe that Ofgem could establish an assurance framework that protects consumers, ensures market integrity and provides structural certainty for private investment in the net-zero transition.

Figure 1.1 Proposed core elements of sector-specific AI regulation



Source: Oxera.

1 Introduction

The transition of the UK energy sector towards a net-zero, highly dynamic system relies on the successful integration of advanced computational models, including AI. As the deployment of AI moves from enabling back-office efficiency to critical infrastructure operations—such as system balancing, demand forecasting and dynamic pricing—the framework for regulatory assurance must evolve accordingly.

Below, we set out our evidence and reasoning in response to the questions raised in Ofgem's call for input, followed by additional considerations we believe would be beneficial for Ofgem to incorporate before issuing formal guidance.

2 Question 1: Current AI assurance practices

a. How do organisations evidence that AI systems are operating as intended and delivering safe, fair and effective outcomes? & d. How are AI governance frameworks translated into operational practice?

Currently, many organisations provide evidence of AI safety primarily through procedural documentation, such as model cards,[3] risk logs and point-in-time audits. Our view is that these tools are structurally mature within the technology sector, but that their translation into energy market economics requires special consideration to capture how an algorithm behaves when interacting with volatile grid physics or coupled wholesale markets.

This reliance on static documentation alone would create a 'reality gap'. Static compliance assumes that if code is audited, the outcomes are predictable. However, for non-deterministic models (e.g. reinforcement learning for grid balancing), the model may encounter new inputs and adapt based on them. Therefore, solely relying on 'checkbox compliance' cannot guarantee operational safety, nor does it protect a firm from retrospective regulatory action if the model drifts.

For organisations to provide evidence that systems are operating as intended, assurance practices must include testing under real-world conditions—including 'extreme' values, outliers and corrupted data—and continuous, empirical observation.

b. What AI assurance approaches are currently used or under development, including in-house and third-party?

Drawing on insights from our work with utility companies and their data, we anticipate that the most significant operational AI risks in utility networks do not stem from sophisticated algorithmic flaws, but from messy, incomplete or inconsistent real-world data and differences in interpretation of pre-set guardrails. To assess these risks, we observe a

[3] Model cards are structured overviews of how an advanced AI model was designed and evaluated.

growing interest in sandbox environments in order to prioritise testing models against historical, imperfect operational data rather than sanitised synthetic datasets.[4]

Demonstrating that an algorithm fails safely under 'dirty data' conditions offers far more robust commercial and regulatory assurance than perfect performance in a sterile environment. Having open conversations between firms and Ofgem about what is meant by terms such as 'reasonable', 'proportionate' and 'fair' in specific contexts will help ensure that there is a shared understanding of when guardrails are expected to kick in.

c. What tools and technical capabilities are available to support AI assurance in practice, how mature and effective are they, and where are there gaps or opportunities for shared or sector-wide approaches?

When considering sector-wide approaches, we believe it is essential to take into account the regulatory baseline to avoid additional complexity from redundancy. For example, cross-sector guidance on data privacy and general model safety already exists (e.g. via the ICO, the NCSC, and the foundational work of the DRCF on AI auditing). We would therefore advise that any new sector-wide tools or guidance pass a strict additionality test, asking: 'What specific energy-market risk or market failure does a new regulation address that existing cross-sector guidance does not?' This would focus additional requirements strictly on the gaps unique to the energy sector.

f. What skills, expertise and resources are required for effective AI assurance, and where are the main capability gaps?

Effective AI assurance requires a mix of technical, empirical and sector-specific knowledge. Where any one of these is missing, the risk will be higher. We would point out that a significant capability gap is likely to exist within regulatory bodies themselves. Effective AI assurance will generate large quantities of complex, unstructured data. Without internal engineering and data pipelines capable of evaluating algorithmic architectures, regulators risk becoming a structural bottleneck, causing significant product approval delays for the industry. A natural extension of Ofgem's supervisory role therefore seems to be to build internal capabilities to use AI-enabled pipelines to ingest and compare these industry submissions. Building the regulators' own AI capability is a practical necessity to turn complex assurance information into swift, consistent regulatory judgements, thereby reducing approval delays for industry products.

3 Question 2: Risks and challenges

a. What are the main barriers to implementing effective AI assurance? & d. Where are current AI assurance approaches most limited in practice?

[4] Oxera (2026), 'The reality gap: testing AI from compliance to enforcement', *Agenda*, April. Available at: https://www.oxera.com/insights/agenda/articles/the-reality-gap-testing-ai-from-compliance-to-enforcement/

We find that a primary barrier to effective AI assurance today is the lack of system-level observability. Current assurance approaches, derived largely from IT auditing, focus on whether a specific algorithm performs adequately under isolated test conditions. This approach is limited because it ignores market dynamics, real-world data conditions and the interaction of multiple algorithms, which are particularly pronounced in the energy sector.

To catch these compound effects, assurance testing must move to evolving 'digital twins' that simulate both upstream generation constraints and downstream consumer response profiles. We believe that evaluating an algorithm within an environment that models multi-agent behaviours is the only viable way to identify emergent systemic risks before they manifest in production.

Furthermore, a growing challenge in AI adoption is managing accountability across complex supply chains. Energy firms increasingly license, fine-tune or integrate third-party AI models from technology providers. Ofgem's framework would benefit from explicitly recognising that while the licensing entity remains ultimately accountable to the consumer, technical risk apportionment should be governed by standardised industry risk-sharing templates. This ensures that third-party vendors face commercial incentives to maintain the integrity of the AI models they supply, empowering energy companies to enforce rigorous assurance standards upstream where the technical risks are best managed.

b. What are the key risks associated with AI use in the energy system (including system reliability, market functioning and consumer outcomes)? & c. Which risks are most difficult to assess, evidence, or link to real-world outcomes?

We believe that the most difficult key risks to evidence and assess are systemic and cumulative, particularly the following three.

1 **Algorithmic bias**: Fundamentally, we believe there is no such thing as a 'bias-free algorithm': an algorithm needs to make choices, and this will always entail trade-offs. For example, if an algorithm's objective function is to improve efficiency, it may do so at the cost of fairness. Or, if the training data contains human bias, it is prone to propagating that bias. This is why, on the one hand, clarity is required on what constitutes an 'acceptable' bias versus a bias to be avoided. On the other hand, continuous monitoring of outcomes is key for the success and credibility of any AI system.

2 **Algorithmic coordination**: If multiple market participants deploy AI for wholesale trading or dynamic retail pricing, there is a risk that these algorithms could interact in undesirable ways. For example, they could independently learn to coordinate to achieve commercial targets ('tacit coordination'), leading to sub-optimal market outcomes or elevated prices without any intent to collude. This represents a novel frontier of antitrust risk where the algorithm acts 'rationally' from the individual firm's perspective, meaning standard internal IT audits would be unlikely to detect a market-wide failure. Firms should have access to central, market-wide simulators to test for these unintended systemic interactions.

3 **Cumulative operational risk**: As AI is integrated into thousands of micro-processes across a company's operations, it creates an unmapped web of dependencies. This decentralised accumulation of risk is difficult to quantify in traditional risk registers

and requires—just like the risk of algorithmic coordination in the market—a systems-level approach.

4 Question 3: Critical infrastructure considerations

a. How should AI assurance reflect the criticality of energy systems as national infrastructure? & b. What level of rigour is appropriate for high-impact or safety-critical AI use cases?

As critical national infrastructure, the energy system cannot rely only on ex post correction. At the same time, pre-deployment assurance requirements must not stifle grid innovation. Assurance should therefore aim to reflect a proactive, supervisory approach that protects industry investments while ensuring the stability of critical infrastructure.

We recommend that Ofgem—potentially leveraging the proposed sandbox powers in the government's Regulating for Growth Bill and operating via the AI Growth Lab framework—invests in creating a phased, industry-collaborative roadmap to develop a macro-level digital twin of the UK energy market, serving as a shared testing environment that integrates specific trading dynamics, network constraints and balancing models.

Crucially, this digital twin model should operate as a mechanism for regulatory certainty. If an energy firm voluntarily submits its trading or balancing algorithm to multi-agent testing in Ofgem's realistic, 'messy data' sandbox, that firm should gain a degree of protection against retrospective ex post enforcement, provided the algorithm operates within the tested parameters and governance. This mechanism explicitly de-risks innovation for the industry while giving the regulator systemic oversight.

5 Question 4: Consumer protection and fairness

a. How can AI assurance support fair treatment of consumers, including vulnerable groups? & b. What risks arise from AI-driven pricing, segmentation or prioritisation, e.g. fairness, transparency, consumer outcomes?

The industry faces significant uncertainty when designing innovative, AI-driven dynamic tariffs, or when using AI to manage debt recovery and prioritise outage responses for vulnerable consumers. Guidelines demanding that models simply be 'fair' or interventions be 'reasonable' lack precision and are easily contested. For example, research indicates that mathematically optimising for identical group outcomes, or demographic parity, can conflict with cost-reflective pricing. [5] Furthermore, AI creates novel economic vulnerabilities by accurately identifying and pricing inelastic demand.

[5] Kleinberg, J., Mullainathan, S. and Raghavan. M. (2017), 'Inherent Trade-Offs in the Fair Determination of Risk Scores', in 8th Innovations in Theoretical Computer Science Conference (ITCS 2017), *Leibniz International*

To unlock industry innovation without harming vulnerable groups, future guidance should provide objective mathematical boundaries. We recommend that Ofgem establish econometric 'safe zones' that explicitly insulate highly inelastic, essential-use demand profiles—such as those linked to medical equipment or baseline low-income heating—from dynamic price optimisation. Defining these quantitative threshold boundaries allows algorithms to optimise for grid flexibility safely, giving the industry the confidence it needs to deploy advanced tariffs at scale.

6 Question 5: Cyber security and resilience

a. How should AI assurance align with existing cyber and operational security frameworks, e.g. NIS Regulations? & b. How can AI assurance support system resilience, including identifying and mitigating cyber, operational and AI-specific risks?

To be effective, Ofgem's AI assurance should in our view strictly apply the additionality test, ensuring it does not seek to replicate baseline IT security (e.g. Cyber Essentials) but focuses exclusively on AI-specific failure modes (like data poisoning and adversarial machine learning attacks) in the energy sector. We also believe that fully consistent alignment with existing and forthcoming cyber frameworks (such as the expansion of the NIS Regulations under the Cyber Security and Resilience Bill) is critical.

Additionally, believe it's crucial that Ofgem should explicitly acknowledge and provide guidance regarding the trade-off between efficiency, effectiveness and explainability of AI models. It is common to find, for instance, that there are trade-offs between model explainability and predictive accuracy, particularly where more complex models are used. A deep neural network may predict complex grid behaviour with significantly higher precision than a simple regression model, but by design operates as a 'black box'. Effective innovation will rely on clear, specific guidance on how to balance such trade-offs.

7 Question 6: Proportionality

a. What does proportionate AI assurance look like across different use cases and risk levels? & b. How can assurance approaches be tailored while remaining effective and practical, including for smaller organisations?

True 'proportionality' in AI assurance must not be confused with simply demanding more or fewer documents. Rather, it requires linking the compliance burden directly to the systemic risk posed by using models, ensuring the reporting framework itself generates actionable insight without dragging down commercial viability. We therefore would encourage Ofgem to build three economic considerations into its reporting framework.

Proceedings in Informatics, **67**, pp. 43:1–43:23, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, available at: https://doi.org/10.4230/LIPIcs.ITCS.2017.43.

1 **Actionable oversight**: Reporting requirements should be strictly limited to actionable regulatory insights. An overly burdensome framework creates a compliance overhead that crowds out smaller innovators and diverts capital from product development towards administrative box-ticking.
2 **Granularity of the attribution of benefits**: In complex operations, AI benefits are often difficult to attribute cleanly across interdependent applications.
3 **Accounting for enabling costs**: Assessments must account for enabling costs (e.g. data standardisation, infrastructure upgrades) alongside gross benefits.

8 Question 7: External assurance and standards

a. Are existing frameworks and standards sufficient, or is sector-specific AI assurance guidance needed? & b. What role should independent assurance (e.g. audits, certification) play? & c. What are the benefits and risks of external assurance approaches?

While external standards (such as ISO/IEC 42001) provide a useful structural baseline, traditional independent assurance poses significant commercial and operational risks if applied uncritically.

As highlighted in the DRCF's foundational paper on auditing algorithms[6]—which has informed our own algorithmic audit frameworks—the third-party AI assurance ecosystem remains in a developmental stage and lacks a consensus around sector-specific methodologies.

Building on the DRCF's insights, we identify two primary risks if standardised external audits are applied to the energy sector without appropriate adaptation.

1 A standardised external audit typically will only provide a snapshot in time. While this is critical for validating initial code implementations and major architectural changes, it is not sufficient for continuous market protection. For adaptive models, a static certification is limited as models ingest new data and the algorithms evolve. Regular outcome-based product reviews are therefore needed to ensure ongoing consumer protection.
2 There is a lack of energy-specific expertise within the general auditing market. Many third-party AI auditors will not have deep expertise in energy markets and regulation. Consequently, an external audit might, for example, certify an algorithm as technically secure and mathematically unbiased, while failing to recognise that its pricing logic creates systemic market distortions. Sector-specific guidance therefore seems essential to ensure independent assurance synthesises data science with rigorous energy market economics.

[6] Digital Regulation Cooperation Forum (2022), 'Auditing algorithms: the existing landscape, role of regulators and future outlook', 23 September, available at: https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/auditing-algorithms-the-existing-landscape-role-of-regulators-and-future-outlook.

© Oxera 2026

9 Question 8: Future guidance

a. What would be most useful in future AI assurance guidance for the energy sector? & b. What types of evidence are most useful in demonstrating outcomes in practice? & c. What examples or case studies would be valuable?

We believe that the most useful feature of future AI assurance guidance will be strict discipline in its scope. We would therefore reiterate our recommendation that Ofgem should apply a formal additionality test to any new regulatory requirements, protecting the industry from regulatory duplication. As we outline in this response, future guidance should focus on providing a regulatory sandbox for testing, clear contractual models for AI supply chain liability, operational guidelines for assessing multi-agent interactions, and establishing explicit econometric limits for qualitative requirements.

Regarding evidence, we believe that Ofgem should look further than expecting perfection in controlled environments. The ability to fail safely under real-world, degraded data constraints is valuable evidence of system resilience and should be recognised as a vital part of an assurance 'gold standard'.

10 Question 9: Communication

a. How should organisations communicate AI assurance to different audiences? & b. What information is most useful for consumers, boards, senior management and affected groups?

Communication of AI assurance should generally be stratified, recognising that protecting proprietary algorithmic logic is vital for market competition.

- For consumers, the focus should be on outcomes and redress, not technical explainability. Providing a consumer with the mathematical weights of an algorithm is unhelpful and risks exposing commercial IP. Communication should clearly state the outcome determined and the precise mechanism for human intervention and appeal.
- For boards and senior management, communication should translate algorithmic performance into enterprise risk and liability protection. Boards require assurance on the economic trade-offs the firm has made and clarity on supply chain liability (who bears the penalty if a vendor's AI fails).
- For Ofgem, communication should move away from market participants submitting static risk registers towards providing Ofgem with dynamic data on performance thresholds and evidence of safe operation within shared digital twin environments.

© Oxera 2026

11 Additional considerations: structural gaps in the consultation framework

While the call for input covers the foundational elements of AI governance, we believe there are two areas that Ofgem might seek to incorporate to fully protect the energy market's integrity and investments.

11.1 Jurisdictional arbitrage and the cross-sector boundary problem

While we recognise that Ofgem is in the process of establishing the new load controller licensing regime—designed to bring organisations that remotely control consumer energy smart appliances under regulatory oversight—under the ESVC approach, the rapid evolution of AI warrants continued attention on the cross-sector regulatory boundaries and clarity on how far specific AI regulation extends across these boundaries. For example, applying AI assurance mandates exclusively to currently licensed utilities could create a structural disadvantage and a potential blind spot for system operators.

To maintain market fairness, we recommend that Ofgem continues to coordinate with the DRCF to explore assurance principles that are function-based rather than strictly entity-based. Tying assurance expectations to the function and systemic impact of an AI tool, rather than solely the regulatory status of the firm deploying it, will help ensure that any algorithmic activity impacting grid balancing, wholesale market functioning or consumer outcomes remains subject to appropriate and proportionate oversight.

11.2 The temporal gap in cyber assurance

Question 5: Cyber security and resilience asks how AI assurance aligns with existing cyber frameworks. However, it overlooks a unique temporal threat of AI. As the findings of the UK AI Security Institute (AISI, formerly the 'AI Safety Institute') have highlighted, frontier AI models are increasingly capable of sophisticated, autonomous cyber offence at machine speed. Procedural governance cannot defend against algorithmic attacks on critical energy infrastructure. Assurance must acknowledge whether a firm has deployed active, AI-driven defensive architecture capable of countering AI-driven attacks in real time.

12 Conclusion

We welcome Ofgem's proactive approach to AI assurance. To ensure the energy sector benefits from AI innovation while maintaining system integrity, governance must match the speed and complexity of the technology. By moving from static compliance to empirical supervision and providing clear quantitative guidance, Ofgem has the opportunity to establish a regulatory framework that operates as an enabler of safety and innovation alike.

Providing the industry with objective, quantitative certainty across consumer performance metrics, protecting licensees from upstream tech liability, and offering robust digital twin testing environments will unlock the private investment required to deliver a smarter, net-zero grid.

We would welcome the opportunity to discuss these economic and governance frameworks further with the Emerging Technologies team.

Contact

Sebastian Wernicke
Partner
+49 (0) 30 1663 57063
sebastian.wernicke@oxera.com

oxera.com

Oxera Consulting GmbH is registered in Germany under the commercial register number HRB 148781 B (Charlottenburg District Court) with its registered office at Rahel-Hirsch-Straße 10, Berlin 10557, Germany.

Although every effort has been made to ensure the accuracy of the material and the integrity of the analysis presented herein, Oxera accepts no liability for any actions taken on the basis of its contents. With regard to our services to you, in the absence of any other signed agreement between you and us, you agree to be bound by our standard Terms of Engagement, which can be found https://www.oxera.com/wp-content/uploads/2025/03/ToE-GER-en-GB.pdf.

No Oxera entity is either authorised or regulated by any Financial Authority or Regulation within any of the countries within which it operates or provides services. Anyone considering a specific investment should consult their own broker or other investment adviser. Oxera accepts no liability for any specific investment decision, which must be at the investor's own risk.

© Oxera 2026. All rights reserved. Except for the quotation of short passages for the purposes of criticism or review, no part may be used or reproduced without permission.