RESPONSE TO OFGEM'S CALL FOR INPUT ON AI ASSURANCE

AUGUST 2026

WWW.FRONTIER-ECONOMICS.COM

Contents

1 Executive summary

AI could create substantial benefits for energy consumers and the wider system. It can expand access to advice, improve customer service, support forecasting and network operation, strengthen asset management, improve trading and help firms make faster and better-informed decisions. Some AI applications, however, have a higher risk profile because AI behaviour can be uncertain, decisions can be made at speed or scale, and systems may take consequential actions without prior human approval.

Existing governance frameworks provide an essential foundation. They identify the outcomes firms should protect, allocate responsibility and establish the legal and regulatory constraints within which firms must operate. The next question is how firms should translate those principles into the design and operation of particular AI applications.

Our central proposition is that deployment optimisation provides the practical link between governance principles and those operational decisions. By deployment optimisation, we mean comparing feasible ways of configuring and controlling an AI application and selecting the approach expected to produce the greatest consumer and system value within existing legal and regulatory constraints.

Firms already make trade-offs between benefits, risks and controls. Much of that judgement is embedded in professional experience and established processes rather than formally specified. AI makes these choices harder to leave implicit. A firm must decide what an application should achieve, what information it may use, what actions it may take, how it should respond when objectives conflict, when a person should intervene, and how performance should be monitored. If these choices are not made deliberately, the application and its surrounding processes will still reflect defaults that may not produce the firm's intended outcomes. Deployment optimisation makes those choices explicit and provides a disciplined basis for making them.

Deployment optimisation does not mean seeking to eliminate every conceivable risk. That may be impossible and, where additional controls materially reduce access, speed, accuracy or other benefits, may not produce the best outcome for consumers. Nor does aggregate benefit provide a licence to disregard legal obligations or expose individuals to foreseeable or serious harm. The task is to select the best feasible approach within those constraints and to be explicit about who receives the benefits, who bears the risks, and how those risks will be managed.

The distribution of benefits and costs also matters. An application can improve average outcomes while producing poorer results for a small group, including consumers in vulnerable circumstances. Firms should therefore examine the likelihood, severity and distribution of adverse outcomes, not only overall performance. They should also distinguish three questions: whether the original deployment decision was reasonable; whether the application continues to perform consistently with that decision; and what correction or redress is owed to a person who experiences a poor outcome.

Assurance should enable firms, boards and Ofgem to understand and challenge those decisions. For higher-impact uses, firms should be able to explain the basis for their deployment decision. This should include:

- what benefits the application is expected to deliver, for whom, and compared with the current process or another credible option;
- the relevant constraints and potential harms;
- the alternative configurations and controls considered;
- why the selected approach was chosen;
- how benefits and harms are distributed; and
- what evidence would trigger intervention or a review of the decision.

Ofgem can help secure the benefits of responsible AI adoption as well as protect against its risks. A clear signal that beneficial deployment is supported, accompanied by practical clarity about good decision-making and the evidence Ofgem would expect, would reduce regulatory uncertainty that might otherwise delay or narrow adoption. Ofgem also has a sector-wide role that no individual firm can perform: identifying patterns that arise across several firms because they use common models, data sources or suppliers, or because the same subset of consumers is adversely affected in ways that no firm can see alone.

We do not suggest that Ofgem should configure individual applications, prescribe universal error rates or establish a new assurance regime before the evidence warrants it. We suggest engagement with firms, consumer representatives and technical experts to develop a common language, proportionate evidential expectations, worked examples and channels for identifying cross-sector risks. Current frameworks may then be supplemented where that engagement reveals a material gap.

The remainder of this response explains why AI makes operational choices more important; develops deployment optimisation as an economic framework; identifies the evidence Ofgem could expect firms to produce; tests the approach against the difficult case of wider benefit and individual detriment; and considers how Ofgem can enable responsible adoption. It does not attempt to answer every consultation question separately.

1.1 How this response relates to the consultation questions

THEME IN THIS RESPONSE	OFGEM QUESTIONS ADDRESSED
Implementation challenge and distinctive AI risks	Q1(d); Q2(b)-(d); Q3(a)-(b)
Deployment optimisation and proportionality	Q2(b)-(c); Q4(a)-(b); Q6(a)-(b)
Evidence Ofgem could expect firms to produce	Q1(a), (d)-(e); Q6(a)-(b); Q8(b); Q9(b)
Wider benefit, distributional effects and individual detriment	Q2(b)-(c); Q4(a)-(b); Q6(a)
Guidance, communication and sector-wide monitoring	Q1(c); Q2(a), (d); Q7(a); Q8(a)-(c); Q9(a)-(b)

2 The implementation challenge

2.1 AI brings significant benefits, but some uses have higher risk profiles

AI applications in energy range from low-impact tools that summarise documents or help employees draft material to systems that interact directly with consumers, support control-room decisions or execute actions in markets and operational systems. The benefits may include improved access and service, more accurate forecasting, more efficient use of assets and networks, reduced cost and better-informed decisions.

The risk profile also varies. A drafting tool whose output is checked by an employee is different from a consumer-facing service that provides personalised advice. Both differ from an application permitted to switch a tariff, alter a payment arrangement, trade within specified limits or take a network-management action without prior approval. The relevant question is therefore not whether AI is risky in the abstract, but what combination of uncertainty, authority, scale and consequence a particular use creates.

Existing frameworks identify important principles: accountability, fairness, transparency, safety, human oversight, security and organisational capability. They establish the ends that firms should protect. They do not, and cannot, determine every operational choice needed to deliver those ends in a particular application.

A requirement to deliver fair outcomes does not determine how to configure a service that expands access to advice but retains a risk of occasional error. A requirement for human oversight does not determine which cases require review, what the reviewer needs to see or how much delay it is reasonable for review to introduce. A requirement for transparency does not determine when a less explainable model is justified by a material improvement in performance. Those decisions require a way to compare the available options.

2.2 AI makes implicit trade-offs explicit

AI does not invent trade-offs. Human advisers already balance speed, accuracy, caution and the circumstances of the individual. Operations teams decide which cases require review and which can proceed. Traders work within limits and escalation arrangements. Control-room operators combine forecasts, engineering standards and judgement. Much of this is embedded in experience, professional norms and established processes.

An AI application needs these choices to be expressed explicitly in its design and operating environment. Its behaviour is shaped by the model and data used, the objectives and instructions it is given, the information and tools it can access, the actions it is permitted to take, the circumstances in which it must defer to a person, and the controls around it. Silence is itself a choice: if a firm does not specify how competing objectives should be balanced or unusual cases handled, the application will still respond, but according to patterns and defaults that the firm may not have intended.

The specification also needs to cover more of the decision. A human can recognise an unusual case, ask another question or depart from normal practice using context that was never written down. An AI application has no dependable organisation-specific frame beyond the objectives, information, instructions and permissions it has been given. No specification will anticipate every circumstance, which makes testing, monitoring and revision part of the original design rather than an afterthought.

Good governance should make the resulting choices visible. For each higher-impact use, a firm should be able to explain:

- the value it expects the application to create for consumers, the energy system and the firm, and the alternative against which that value is measured;
- the outcomes, obligations and limits on the level and distribution of risk that constrain the design;
- the principal configuration and control choices, including instructions, permissions, escalation and human involvement;
- who approved those choices, who can change them and who owns the resulting outcomes; and
- how performance will be monitored and when the choices will be revisited.

2.3 The level of rigour required depends on uncertainty and consequence

Two dimensions are particularly important.

Uncertainty of behaviour. A relatively deterministic tool can be tested against known inputs and expected outputs. A generative system may respond differently to materially similar cases, and its behaviour may change when the model, data or connected tools change. Assurance must then consider the distribution of performance, including rare but serious outcomes and differences between groups, rather than rely only on a set of pass-or-fail tests.

Consequence and reversibility. A tool that recommends an answer to an employee creates a different exposure from one that advises a consumer directly or can take an action without prior review. The more consequential and difficult to reverse the action, the stronger the case for evidence before deployment, limits on authority, effective intervention arrangements and intensive monitoring.

This applies beyond retail. An AI tool supporting demand forecasting or network constraint management creates different risks depending on whether it provides information to a control-room operator, recommends an action requiring approval, or can take defined actions without prior approval. The greater its authority, the stronger the case for rigorous testing, clear intervention arrangements and a reliable fallback if it behaves unexpectedly.

The same logic applies to energy trading. An application might produce a forecast, recommend a bid or trade for approval, or execute transactions within specified limits. The deployment decision then includes the markets and products it may access, its exposure limits, how it should respond to unusual conditions or conflicting signals, and when a person must intervene. Correlated strategies or common data and model dependencies may also affect market functioning even where each firm's individual exposure appears limited.

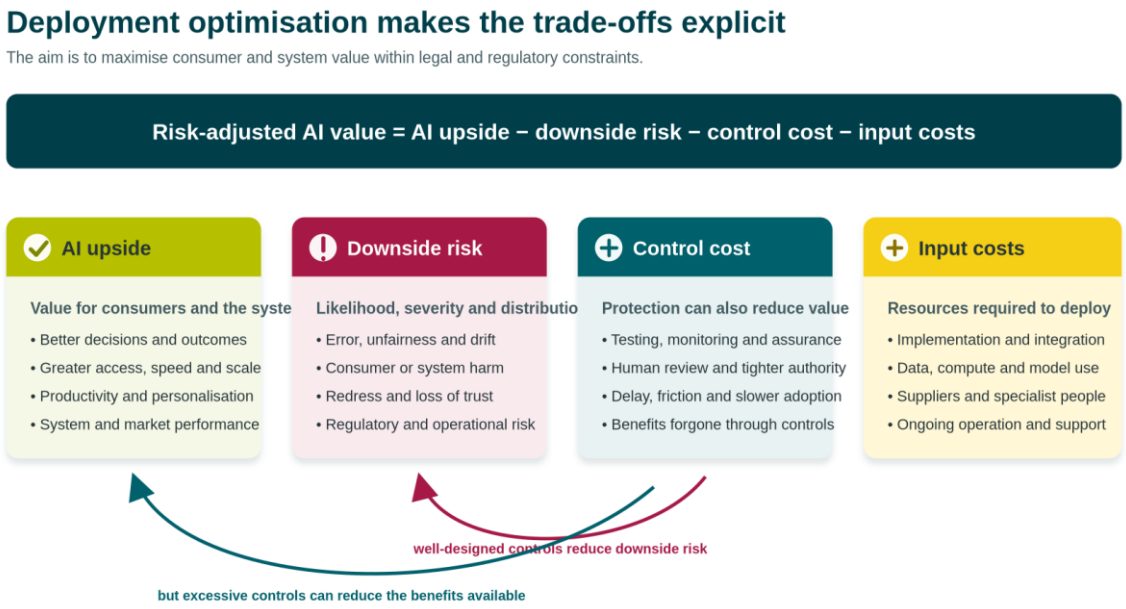
Speed and scale increase the consequences of getting these choices wrong. Shared reliance on common models, data sources or suppliers can also produce correlated risks. An issue affecting a small subset of one firm's consumers may appear immaterial; the same issue affecting that subset across many firms may be significant. Each firm remains responsible for its deployment, but Ofgem may be better placed to identify a pattern visible only across the sector.

3 Deployment optimisation

3.1 Making the value equation explicit

Deployment optimisation is the process of comparing feasible ways to configure and control an AI application and selecting the approach expected to produce the greatest consumer and system value within existing legal and regulatory constraints. Economics provides a disciplined way to make that comparison.

Figure 1. The components of risk-adjusted AI value



Source: Frontier Economics.

AI upside includes value for consumers and the system as well as commercial return: improved access and service, better decisions, productivity, speed, scale, personalisation, forecasting, asset use and market or system performance. It should be measured against what would happen without the application, normally the existing process or another feasible configuration. The question is whether the application improves on what would otherwise exist.

Downside risk includes the likelihood, severity and distribution of error, unfairness, drift, consumer or system harm, redress, operational failure, loss of trust and regulatory challenge. Averages are insufficient where a low-frequency event could cause serious harm or poorer outcomes are concentrated among a particular group.

Control cost includes both expenditure and value forgone when risk is reduced. Testing, monitoring, human review, tighter permissions and slower processes may all be justified. They also consume resources and can reduce access, speed or other benefits. This interaction is why an instruction to minimise risk does not by itself identify the right configuration.

Input costs include implementation, integration, data, computing, model use, suppliers and human resources. The relevant calculation is wider than the firm's profit and loss account. Licence obligations and consumer protection, safety, equality and data-protection requirements constrain the feasible options. Consumer harm and effects on the energy system matter even where the firm does not bear their full financial cost. Not every effect needs to be monetised: quantitative evidence can be combined with qualitative judgement, scenarios and behavioural testing.

3.2 Comparing feasible configurations

The relevant choice is rarely simply to deploy or not deploy. Depending on the use, a firm might change:

- the model, data or information sources;
- the objective and instructions given to the application;
- the actions, tools, markets or systems it may access;
- confidence, referral, exposure and escalation thresholds;
- the circumstances requiring human approval or review;
- warnings, explanations and consumer communications; or
- fallback arrangements, alternative channels and routes to correction.

Each configuration changes the expected benefits, residual risks and costs. Optimisation should compare credible alternatives rather than justify a preferred design after the event. It should identify uncertainty and allow authority to increase as evidence develops. A controlled trial, recommendation-only role or limited set of actions may be justified before deployment at greater scale or with wider authority.

Current legal and regulatory requirements are constraints on the firm's choice, not factors that can simply be traded off against wider benefits. New technology may reveal that an existing boundary no longer produces the best outcome, but whether to change that boundary is a matter for government or the regulator. Until it changes, the firm must choose among options that comply with it.

3.3 Monitoring, review and revision are part of the decision

The original deployment decision rests on assumptions about benefits, risks, controls and behaviour. Some will prove wrong. Others will cease to hold as models are updated, consumer behaviour changes, markets evolve, new tools are connected or evidence accumulates.

Monitoring asks whether the application is operating as expected. Review asks whether the chosen operating position remains the right one. A system may remain within its original tolerance while technological or market changes make a different configuration preferable. An apparently isolated incident may reveal a failure mode that was not considered at deployment.

The initial decision should therefore specify what will be monitored, who will review the evidence and what will trigger intervention or revision of the deployment design. For higher-impact uses, triggers

may include a material model update, a change in permissions, evidence of poorer outcomes for a group, repeated consumer corrections, unusual trading or system behaviour, a change in the relevant alternative, or a common failure affecting other firms using the same provider.

This economic framing leads directly to an evidential question for Ofgem: what should a firm be able to show to demonstrate that it made, implemented and revisited these choices responsibly?

4 What Ofgem could expect firms to evidence

Governance identifies the responsibilities, outcomes and constraints that apply. Deployment optimisation is the process of choosing how a particular application should operate within them. Assurance tests and substantiates that decision before deployment and as evidence accumulates. It should reuse existing governance and compliance evidence where that is sufficient, with the depth of evidence proportionate to the uncertainty and consequences of the use.

For higher-impact applications, Ofgem could expect the evidence to allow the firm, its board and the regulator to understand four things.

4.1 The expected value and alternative

What benefit is the application expected to create, for whom and compared with what? A claim that AI improves accuracy, efficiency or access is incomplete without describing the existing service or process, the consumers or system users who benefit, and the value of speed, scale, personalisation or improved decisions.

This prevents optimisation from becoming a risk-only exercise. Controls that remove most of the benefit may be no more consistent with good outcomes than weak controls that leave foreseeable harm unaddressed.

4.2 The alternatives and basis for the choice

What material adverse outcomes were considered? Which people or parts of the system might bear them? What legal and regulatory constraints apply? Which alternative configurations and controls were examined, and how did their benefits, risks and costs differ? Who approved the selected position?

The answer need not be a single quantitative model. It should be a coherent record of the decision, supported by testing, economic analysis, technical evidence and behavioural research where appropriate.

4.3 Performance, review and revision

Is the application delivering the benefits expected? How often do adverse outcomes occur, how severe are they and how are they distributed? Does performance vary across repeated interactions or groups? Are permissions, escalation and human involvement operating as intended? Has a model, supplier, data or market change altered the underlying risk?

The existence of human review is not itself evidence that the control is effective. Reviewers may defer to the AI output, share its underlying biases, or lack the information, time or authority needed to intervene. Assurance should test what human involvement changes in practice.

The evidence should set out use-case-specific measures and tolerances. For a consumer service, these might include the accuracy and appropriateness of responses, unsafe or misleading advice, successful escalation, unresolved contacts and variation across groups. For a system or trading use, they might include forecast error under different conditions, actions outside expected parameters, near misses, limit breaches, override frequency, market impacts and successful fallback. Monitoring should distinguish expected variation that can be managed routinely from evidence requiring intervention or revision of the deployment decision.

4.4 Treatment of affected consumers

For consumer-facing uses, can a person understand the role AI played, obtain human consideration, correct an error and access redress where appropriate? Are individual incidents used to test the assumptions supporting the wider deployment decision?

The treatment owed to an affected consumer is separate from whether the original deployment decision was reasonable. A firm may need to correct an error or provide redress even where the application remains beneficial and responsibly governed overall. Conversely, the fact that correction or redress is owed does not by itself establish that the original deployment decision was unreasonable.

5 Stress test: wider benefit and individual detriment

The value of clearer expectations is most apparent where an application improves outcomes overall but cannot eliminate every poor individual outcome without losing much of the benefit. This is not the only difficult deployment decision, but it exposes a distinction that an outcomes-based framework needs to handle.

Consider an AI service that helps consumers understand bills, compare tariffs and decide whether to switch. If the alternative is that many receive no advice, remain disengaged and stay on a less suitable tariff, the service may create substantial consumer benefit. Yet a system whose responses can vary may still give an incorrect or inappropriate answer in some cases despite reasonable testing and controls.

The benefit would need evidence. The firm should be able to show who received support who would not otherwise have done so, whether understanding or decisions improved, which consumers switched, and whether the benefits persisted after errors and subsequent corrections were included. It should also identify whether poor outcomes are random or concentrated among consumers with particular characteristics or circumstances.

5.1 The trade-off is familiar

Fraud detection in banking provides a useful analogy. A very sensitive threshold may catch more fraud but block or delay many legitimate payments. A permissive threshold reduces friction but allows more fraud to pass. Eliminating every missed fraud would not be optimal if it made the payment system unusable. Banks are nevertheless expected to prevent and monitor fraud, investigate failures and treat affected customers appropriately.

Energy network regulation contains related logic. The regime seeks to reduce the frequency and duration of interruptions and provides protections for affected customers. A single interruption does not by itself establish that the network was inefficiently designed or operated; the assessment also considers its cause and severity, wider performance, prior decisions and the operator's response.

These are analogies rather than regulatory precedents. Their value is that they separate three questions:

1. Was the application designed and deployed reasonably on the evidence then available?
2. Is its performance over time consistent with the assumptions supporting that decision?
3. What correction or redress is owed to the person who experienced the adverse outcome?

How redress is funded raises a further economic question. If a firm captures only part of the wider benefit from an application but bears the full cost of every adverse outcome, it may avoid uses that would create net consumer benefit. Insulating firms from those costs altogether would weaken incentives to optimise and monitor properly. How consumer protection, fault and the allocation of

redress costs should interact is therefore a policy and regulatory design question, not one that an individual firm can settle through its deployment decision alone.

5.2 A poor outcome is evidence, not the whole assessment

Suppose a firm restricted the information sources used by an advice tool, prohibited it from taking certain actions, tested a wide range of consumer circumstances, established escalation routes and provided ready access to human consideration. A poor response in an unusual case may require correction and may reveal a new risk. It does not necessarily establish that the original decision to deploy was unreasonable.

The conclusion would be different if the firm ignored evidence of poor performance, failed to test foreseeable circumstances, selected weak controls to preserve benefits, or continued deployment after its assumptions ceased to hold. In those cases, the adverse outcome points to a deficient deployment decision or deficient governance.

Aggregate benefit cannot excuse a breach of a current obligation or provide a blanket defence for individual harm. Nor should every adverse output automatically be treated as proof that the firm should not have deployed the application. Firms and Ofgem need to distinguish the quality of the original decision, performance in use, the foreseeability, severity and distribution of harm, and the firm's response to the affected consumer.

Without clarity about that distinction, firms may restrict AI to low-value uses, require human approval in every case or avoid applications that could expand access and improve outcomes. The result could be less consumer and system benefit, not simply less risk. The converse risk is that firms invoke innovation and aggregate benefit to defend poor controls. Clearer expectations would help avoid both outcomes.

6 How Ofgem can enable responsible adoption

Regulatory clarity is part of the enabling framework for responsible AI adoption. Ofgem can help in three connected ways: signal support for beneficial uses, clarify what good deployment decisions and evidence look like, and use its sector-wide perspective to identify risks that individual firms cannot see.

First, Ofgem can make clear that responsible AI adoption is compatible with, and can contribute to, its consumer and system objectives. The relevant aim is not adoption for its own sake. It is to avoid regulatory uncertainty causing firms to delay, narrow or abandon applications that could improve outcomes, while maintaining accountability for poor decisions and breaches of existing obligations.

Second, Ofgem can provide practical decision infrastructure for firms and their assurance providers. A common language, a set of questions that higher-impact deployments should answer, proportionate evidential expectations and worked examples would make existing principles easier to apply. This need not prescribe a single methodology or standard threshold: different uses present different choices.

Issues for engagement and possible guidance include:

- the questions firms should answer about the value sought, relevant constraints and alternative configurations;
- how the depth of evidence should vary with behavioural uncertainty, authority, scale, consequence and reversibility;
- the types of testing and monitoring evidence that can support a deployment decision;
- how firms should respond to drift, model updates and emerging risks;
- how overall performance and outcomes for particular groups or individuals should be considered together;
- how correction, redress and regulatory accountability should operate where benefits and harms fall on different consumers, including how the allocation of costs affects deployment incentives; and
- how firms should identify and report common model, data or supplier dependencies.

Third, Ofgem can strengthen cross-sector visibility. It could use existing supervisory engagement, incident information and appropriately designed channels for firms and assurance providers to identify repeated failure modes, correlated dependencies and harms affecting the same subset of consumers across several firms. This is not a substitute for firms' own monitoring. It addresses information that no single firm is well placed to assemble.

Engagement should precede detailed guidance. Discussions with firms, consumer representatives and technical experts could test where uncertainty is affecting adoption, what evidence firms can produce, where existing arrangements already answer the question and where common expectations would add value. Frontier is exploring these questions with companies considering how to govern and deploy AI; we do not present those discussions as evidence of established industry practice.

Worked examples could make any subsequent guidance tangible without prescribing a technical solution. Examples could compare an application that recommends tariff information to an employee with one that advises a consumer or executes a switch; trace the increasing authority of a forecasting or network-management tool; and examine a trading application moving from forecast to recommendation to execution within limits. They could show how expected value, constraints, configuration choices, evidence and monitoring change with uncertainty and consequence.

More detailed intervention would be justified only where engagement reveals a material gap. The objective is enough certainty to support beneficial deployment and effective challenge, without creating a new assurance regime before the evidence warrants it.

7 Conclusion

AI offers important consumer and system benefits, but some uses have higher risk profiles. Existing principles identify the outcomes to protect; the implementation challenge is how firms should translate them into choices about the objectives, information, permissions, controls and acceptable performance of particular applications.

Deployment optimisation provides an organising discipline. It requires firms to compare the consumer and system benefits, downside risks, control costs and input costs of feasible configurations; select a position within existing legal and regulatory constraints; examine how benefits and harms are distributed; and explain why the choice is preferable to the alternatives. Assurance should test and substantiate that decision and show when it will be revisited.

Zero risk is not always achievable or desirable, because controls can remove the benefits an application was intended to create. Equally, aggregate benefit cannot excuse a breach or make serious individual harm irrelevant. A sound outcomes-based approach needs to distinguish the original deployment decision, the application's performance over time and the treatment owed to an affected consumer.

Ofgem can help secure the benefits of responsible adoption by signalling support for uses that improve consumer and system outcomes, clarifying the reasoning and evidence it expects from firms, and monitoring patterns that are visible only across the sector. This would reduce regulatory uncertainty while strengthening accountability.

8 Sources

- Ofgem, *AI assurance: call for input*, 17 June 2026.
- Ofgem, *Ethical AI use in the energy sector*, updated May 2026.
- Ofgem, *RIIO-2 Electricity Distribution: Annual Report 2024 to 2025*, January 2026.
- Ofgem, *Guide to the RIIO-ED1 electricity distribution price control*, January 2017.
- Payment Systems Regulator, materials on APP scams, reimbursement requirements and performance reporting.
- Department for Science, Innovation and Technology, *Trusted third-party AI assurance roadmap*, 3 September 2025.
- National Cyber Security Centre, *Guidelines for secure AI system development*, version 1.0.
- Information Commissioner's Office, guidance on AI and data protection and AI audit framework materials.