
18th March 2026

Response to Ofgem Consultation: AI Technical Sandbox

Respondent: University of Strathclyde

The University of Strathclyde, through PNDC, operates a national test and demonstration facility providing hybrid physical-digital environments for energy system validation. PNDC has the capabilities to replicate real-world operational energy environments using physical infrastructure, hardware-in-the-loop systems, and advanced digital simulation platforms. This is complemented by the University's AI Accelerator, which incorporates its AI for Energy initiative.

The Challenge

The UK energy sector faces a persistent structural barrier: AI innovations struggle to move from pilot to operational deployment due to uncertainty around:

- Safety in safety-critical infrastructure
- Interoperability across legacy and multi-vendor systems
- Scalability beyond limited pilots
- Regulatory compliance and evidencing
- Public and consumer trust.

Current approaches are insufficient:

- Purely digital simulation cannot fully replicate dynamic grid behaviour, protection coordination, or real-world communication constraints.
- Consumer-scale trials expose customers to risk before system-level validation is complete.
- Vendor-led testing lacks independent neutrality and system-wide visibility.

Energy infrastructure is inherently cyber-physical and socio-technical. Effective de-risking requires hybrid physical-digital environments capable of replicating authentic operational dynamics. Validation must therefore be:

- Operationally realistic
- Technically rigorous
- Ethically measurable
- Independently assessable
- Reproducible.

Critically, de-risking must involve coordinated participation across the whole supply chain:

- Energy companies
- Technology developers
- Consumer representatives
- Policy and regulatory stakeholders.

Without this whole-sector alignment, the industry risks fragmented innovation, duplicated investments, and inconsistent assurance standards.

A properly structured and well-implemented AI Technical Sandbox, underpinned by representative test and demonstration facilities, can:

- Reduce regulatory uncertainty
- Build sector-wide confidence
- Establish common AI validation norms
- Protect consumers and infrastructure
- Accelerate safe innovation.

Responses to Consultation Questions:

Q1. Eligibility and participation

Do you agree with the proposed eligibility criteria for lead Participants (licensees, market participants, and operators of essential services) and the encouragement of partnerships with technology providers, academia, and other innovators? Please explain your reasoning.

We agree with the proposed eligibility criteria restricting lead applicants to licensees, market participants, and operators of essential services, while encouraging structured partnerships with technology developers, academia, and innovators.

From an operational testing perspective:

- Access to live-grade operational data and infrastructure context is essential for meaningful AI validation.
- Licensees and operators of essential services bring governance maturity, regulatory accountability, and operational realism.
- Partnerships allow technical depth (AI modelling, assurance tooling, cybersecurity testing, red-teaming capability) without diluting regulatory responsibility.

In our experience operating physical and digital testbeds, **the most effective innovation programmes bring together:**

- A regulated asset owner (accountable and data-rich)
- A technology developer (innovation)
- An independent validation environment (neutral test and demonstration capability).

We support Ofgem's approach as it balances regulatory assurance with innovation.

Q2. Use case selection criteria

Are the proposed use case selection criteria (including commercial neutrality, innovation, sector impact, regulatory uncertainty, testability, governance, and data access) appropriate and sufficient to ensure a fair and transparent process? Are there other criteria, safeguards, or considerations we should include?

The proposed criteria are strong and largely sufficient. We particularly welcome:

- Emphasis on testability
- Explicit recognition of regulatory uncertainty
- Governance via Steering Group oversight
- Requirement for access to representative data.

From a test facility perspective, **we suggest strengthening three areas:**

- Operational realism criterion
Use cases should demonstrate that:
 - Scenarios replicate real-world operational stress conditions
 - AI performance can be evaluated under degraded, adversarial, or abnormal states
 - Physical system impacts (voltage stability, load management, dispatch coordination, etc.) can be meaningfully assessed
 - 'Human-in-the loop' based approaches should be incorporated; this involves constructing physical operational scenarios in a safe environment where any human interventions that may be needed can be understood.

AI tested only on static datasets does not adequately demonstrate operational safety.
- Technical assurance readiness
Applicants should provide:
 - Defined performance metrics
 - Failure mode identification plans

-
- Red-teaming and adversarial testing strategies
 - Clear rollback or containment protocols
 - An ability to test scenarios/tools for ‘traceability’, creating different operational scenarios and recording how decisions were made for example.

A sandbox must validate not just functionality, but failure containment.

- Reproducibility and transferability

To maximise sector-wide learning, proposals should include:

- Defined evaluation frameworks
- Transparent metric definitions
- Potential for replication in other environments.

This strengthens commercial neutrality and accelerates standardisation

We recommend that the sandbox prioritise use cases that demonstrate clear consumer benefit, systemic value, and regulatory relevance. We would propose the following **candidate use cases for the AI Technical Sandbox**.

- AI predictive maintenance
 - Prevent unplanned outages
 - Reduce asset replacement costs passed via tariffs
 - Protect vulnerable consumers from service disruption
- Asset optimisation
 - Condition-based investment planning
 - Loss reduction
 - Improved network efficiency
- Dynamic network reconfiguration and grid balancing
 - Real-time grid balancing
 - Demand side response integration
 - Renewable variability management
 - Increased consumer participation in flexibility markets
- Real-time autonomous grid control
 - Maintain system security
 - Prevent cascading failures
 - Establish accountability for AI-driven decisions

-
- Cybersecurity and anomaly detection
 - Protect critical infrastructure
 - Safeguard smart meter and personal data
 - Prevent coordinated grid disruption
 - Automated market operations and trading
 - Maintain market integrity
 - Prevent algorithmic manipulation
 - Protect consumers from price distortion
 - Consumer-facing tariff optimisation
 - Addressing bias and fairness concerns
 - Protecting vulnerable groups
 - Ensuring data privacy compliance
 - Virtual Power Plants and Community Energy
 - Advanced dispatch optimisation
 - Automated market participation
 - Peer-to-peer trading and local energy optimisation
 - Multi-vector integration

Q3. Alignment with other initiatives

Is the proposed approach for the AI Technical Sandbox clearly distinct and complementary to other initiatives such as Ofgem's AI Reg Lab, Energy Regulation Sandbox, Future Regulation Sandbox, UKRI-funded and SIF/NIA initiatives, NESO, FCA regulatory sandbox experience, and DSIT AI Growth Lab? Are there other relevant initiatives or examples of best practice that Ofgem should consider, and if so, which ones?

The proposed sandbox is appropriately positioned as a **technical validation layer**, distinct from policy or live regulatory flexibility sandboxes.

We believe it is complementary to the referenced initiatives.

From our operational standpoint, a **critical success factor** will be to collaborate with these other initiatives for:

- Alignment on data standards
- Shared AI evaluation metrics

-
- Interoperable safety assurance frameworks
 - Consistent ethics benchmarking

We would **also suggest considering** appropriate collaboration with:

- International energy digitalisation testbeds
- IEC and ISO emerging AI governance standards
- Cross-sector cyber-physical infrastructure resilience programmes

This ensures the sandbox does not become isolated from global best practice or from learnings in other sectors.

Q4. Engagement and governance

Does the proposed governance structure (steering group, working groups, open forums) provide sufficient oversight, transparency, and opportunities for stakeholder engagement? Are there other mechanisms or safeguards that should be included to ensure effective governance and knowledge sharing?

The Steering Group, working groups, and open fora provide a **strong governance foundation**.

To enhance effectiveness, we recommend:

- Independent Technical Assurance Panel

This panel would include independent test and validation experts capable of:

- Reviewing experimental design
- Assessing robustness of simulation environments
- Evaluating hardware-in-the-loop credibility
- Challenging claims of readiness

- Structured Stage Gates

Each use case should move through:

Stage 1: Pre-test technical review

- Model architecture documentation
- Training data lineage
- Bias and fairness pre-assessment
- Defined success metrics
- Defined failure thresholds

Stage 2: Controlled functional testing

- Nominal condition performance
- System integration compatibility
- Latency and responsiveness assessment
- Interoperability with legacy systems

Stage 3: Stress and adversarial testing

- Extreme load scenarios
- Data drift simulation
- Cyber intrusion simulation
- Conflicting objective environments
- Model hallucination or instability detection

Stage 4: Human-in-the-Loop evaluation

- Operator override capability
- Explainability under time pressure
- Escalation pathways
- Cognitive load assessment

Stage 5: Post-test assurance and ethics review

- Fairness validation across user groups
- Transparency audit
- Performance variability analysis
- Consumer impact
- Residual risk documentation

Formal gates reduce ambiguity and increase transparency.

▪ Auditability

In order to enable auditability and learning, activities should be required to maintain:

- Controlled experiment logs
- Model version traceability
- Data lineage records
- Red-team documentation.

This strengthens defensibility and sector trust.

Q5. Timelines and next steps

Are the proposed next steps for developing and launching the pilot clear, and is there anything further we should consider as we refine the timeline?

The proposed phased development **approach is clear**.

We recommend:

- Early publication of technical testing expectations
- Clear articulation of minimum data maturity requirements
- Early engagement workshops to shape high-quality applications.

From experience, applicants often underestimate the time required for:

- Data preparation
- Model validation
- Integration into representative operational environments.

Allowing preparatory phases will improve pilot quality.

Q6. Ethics and responsible AI

Does the consultation and proposed pilot sufficiently address ethical considerations (fairness, transparency, responsible use, consumer trust) in line Consultation AI Technical Sandbox with Ofgem's AI guidance? Are there further steps we should take to embed ethics and safety in the sandbox?

We **strongly support** the integration of ethics and safety checkpoints.

To further embed responsible AI we would propose:

- Ethics-by-Design templates
Provide structured templates covering:
 - Fairness assessment
 - Bias detection metrics
 - Transparency documentation
 - Explainability approaches
 - Consumer impact modelling

-
- Adversarial and stress testing

Ethical AI must be evaluated under:

- Data drift scenarios
- Adversarial manipulation
- Extreme operational events
- System conflict conditions

- Human-in-the-Loop evaluation

This evaluation should assess:

- Operator override mechanisms
- Decision explainability under pressure
- Escalation pathways

AI safety in energy systems is socio-technical, not purely algorithmic.

Q7. Stakeholder support

Do you have suggestions for how Ofgem can best support stakeholders throughout the pilot and beyond?

To **maximise uptake and quality participation**, Ofgem could provide:

- Pre-application clinics or technical briefings
- Published exemplar use cases
- Standardised evaluation metric libraries
- Model documentation guidance
- Shared ethical assessment checklists

We **also recommend** facilitating shared access to **neutral test and demonstration facilities** capable of:

- Hardware-in-the-loop simulation
- Digital twin integration
- Cyber-physical attack simulation
- Real-time performance benchmarking.

This lowers barriers to rigorous testing.

Q8. General feedback

Do you have any other comments, suggestions, or concerns regarding the proposed pilot, the consultation process, or the expected outcomes? Please provide evidence, examples, or reasoning to support your responses wherever possible.

We **strongly support the creation of an AI Technical Sandbox** that:

- Tests AI under representative operational conditions
- Integrates ethical assurance with technical safety
- Generates reproducible, sector-wide learning
- Maintains commercial neutrality
- Protects consumers and market integrity.

From our experience operating real-world representative test environments, **we emphasise AI credibility in the energy sector will depend not on innovation alone, but on demonstrable safety under realistic operational stress.**

If structured effectively, this sandbox can:

- Define the operational “art of the possible”
- Establish common AI validation standards
- Reduce regulatory uncertainty
- Build trust across operators, regulators, and consumers.

We would welcome continued engagement to support development of robust technical validation frameworks that ensure AI systems are proven safe, resilient, fair, and fit for deployment in critical energy infrastructure.

ANNEX

Response to Ofgem Consultation: AI Technical Sandbox

Respondent: University of Strathclyde

Role of Representative Test & Demonstration Facilities in the AI Technical Sandbox

Why Representative Testing Matters

AI systems deployed in energy networks interact with:

- Physical infrastructure (generation, storage, networks, protection systems)
- Human operators
- Market signals
- Cybersecurity layers
- Safety-critical control logic.

Testing AI only on historical datasets is insufficient. A robust sandbox must replicate:

- Dynamic system behaviour
- Fault conditions
- Adversarial events
- Multi-system interactions
- Human-machine decision loops.

The University of Strathclyde PNDC facility enables controlled experimentation in environments that mirror real-world operational conditions using physical equipment, hardware-in-the-loop platforms, digital twins and advanced simulation software.

Unique Advantages of Incorporating a Representative Test and Demonstration Facility

Integrating a facility capable of replicating real-world operational conditions using physical infrastructure, hardware-in-the-loop systems, and advanced simulation tools fundamentally strengthens the AI Technical Sandbox. It moves the sandbox from a conceptual or data-only exercise to a credible, operationally grounded validation environment.

The advantages fall into six key categories:

True Operational Realism (Not Just Model Validation)

Most AI testing environments rely solely on static datasets. A representative facility enables:

- Real-time system dynamics
- Physical equipment interaction

-
- Protection system behaviour testing
 - Network constraints and latency modelling
 - Operator interface interaction

AI systems are validated against the dynamic, non-linear behaviour of real infrastructure, not simplified abstractions. This exposes emergent risks that do not appear in offline analysis.

Safe Testing of High-Risk and Edge Scenarios

A physical-digital hybrid environment allows controlled simulation of:

- Frequency instability events
- Voltage collapse conditions
- Extreme weather disruption
- Market volatility spikes
- Communication loss
- Coordinated cyber-physical incidents

These scenarios cannot ethically or safely be reproduced in live networks.

The sandbox can evaluate AI behaviour under stress conditions that are rare but system-critical, improving resilience assurance before live deployment.

Cyber-Physical Risk Validation

Energy infrastructure is cyber-physical. AI interacts with:

- Sensors
- Control signals
- Protection relays
- Distributed assets
- Human operators

A representative facility enables:

- Hardware-in-the-loop validation
- Signal corruption testing
- Latency injection
- Adversarial telemetry manipulation
- Protection mis-coordination testing

The sandbox can detect cascading failure modes, oscillatory instability, and unsafe control behaviours that would not be visible in purely digital simulation environments.

Measurable Ethical and Consumer Impact Testing

Ethics cannot remain theoretical in operational AI.

A representative environment allows:

- Fairness testing across consumer segments
- Bias amplification under stress conditions
- Explainability testing under time pressure
- Human override validation
- Consumer outcome scenario modelling

Ethical assurance becomes observable and measurable, not just policy-aligned documentation. This materially strengthens consumer trust and regulatory credibility.

Independent, Commercially Neutral Validation

A neutral test facility embedded in the sandbox:

- Applies standardised evaluation frameworks
- Uses consistent stress-test libraries
- Produces comparable performance metrics
- Enables anonymised benchmarking

Reduces risk of competitive distortion from opaque vendor claims. Evidence is generated under controlled, transparent, and repeatable conditions.

This directly supports Ofgem's objectives around market integrity and commercial neutrality.

Faster Translation from Innovation to Regulatory Insight

By generating structured, reproducible evidence, a representative facility:

- Identifies regulatory friction points
- Quantifies system-level impact
- Clarifies where existing rules suffice
- Highlights where guidance gaps exist
- Supports evidence-based policy updates

It complements initiatives involving:

- National Energy System Operator
- UK Research and Innovation
- Energy Systems Catapult

The sandbox becomes a technical evidence engine for regulatory evolution not just an experimentation forum.

Failure Containment Validation (Not Just Performance Optimisation)

In critical infrastructure, one important question is:

“Does the AI optimise well?”

But a more important question is:

“How does the AI fail?”

A representative facility enables:

- Controlled degradation testing
- Safe-mode validation
- Override path verification
- Escalation protocol assessment
- Recovery behaviour analysis

This directly strengthens system resilience and reduces systemic risk exposure.

Building Sector-Wide Confidence

Confidence in AI deployment depends on:

- Demonstrable safety
- Transparent evaluation
- Reproducible testing
- Independent oversight
- Evidence under stress

A representative test facility provides tangible proof that:

- AI has been challenged rigorously
- Risk boundaries are understood
- Human control remains viable
- Consumer protections remain intact

It transforms the sandbox from a policy initiative into a credibility-building instrument for the entire sector.

Summary

Incorporating a representative physical and digital test and demonstration facility provides advantages that cannot be replicated by:

- Policy sandboxes alone
- Data-only environments
- Academic modelling exercises
- Vendor self-testing.

It enables:

- Realistic system interaction
- Safe stress experimentation
- Independent assurance
- Ethical validation in context
- Failure containment verification
- Reproducible regulatory evidence.

For a sector as safety-critical and infrastructure-dependent as energy, these advantages are not incremental — they are foundational.

PNDC complements other facilities in that it offers a controllable live-grid environment. The PNDC difference is the physical infrastructure: real switchgear, protection relays, grid emulation equipment, and lab-scale network hardware that can actually fail, trip, or respond in real time to what an AI model tells it to do. This enables network and system operators to test AI on real equipment with no risk of impacting customers.

Power network AI risks include cascading failures, protection system miscoordination, and grid instability — risks that simply *cannot* be safely tested on real consumers and real infrastructure at early technology readiness stages.

Evidence generated at the PNDC - real-world hardware test data at scale - would fill the need for empirical evidence that regulatory mechanisms require.