

# AI Assurance Use Case – Ofgem Sandbox (Summary)

## AI Assurance Use Case – Ofgem Sandbox (Summary)

Lantrix Limited — March 2026

---

### 1. The Problem

AI is increasingly being introduced into operational energy systems, influencing real-world decisions across grid operations, asset management, and incident response.

However, current assurance approaches focus primarily on model performance, rather than how AI systems behave within real workflows, data flows, and control environments.

This creates a gap:

- Limited visibility of decision pathways
- No consistent mapping from decisions to risks and controls
- Insufficient evidence to support safe deployment

As a result, organisations face regulatory uncertainty when introducing AI into critical systems.

---

### 2. The Approach

This use case proposes a structured, system-level approach to AI assurance:

#### **Represent the system**

Architecture, workflows, and decision paths

#### **Identify risks**

Map decision points to potential failure and harm scenarios

#### **Apply controls**

Technical and process controls mapped explicitly to risks

#### **Generate and execute tests**

Realistic scenarios, including edge cases and failure modes

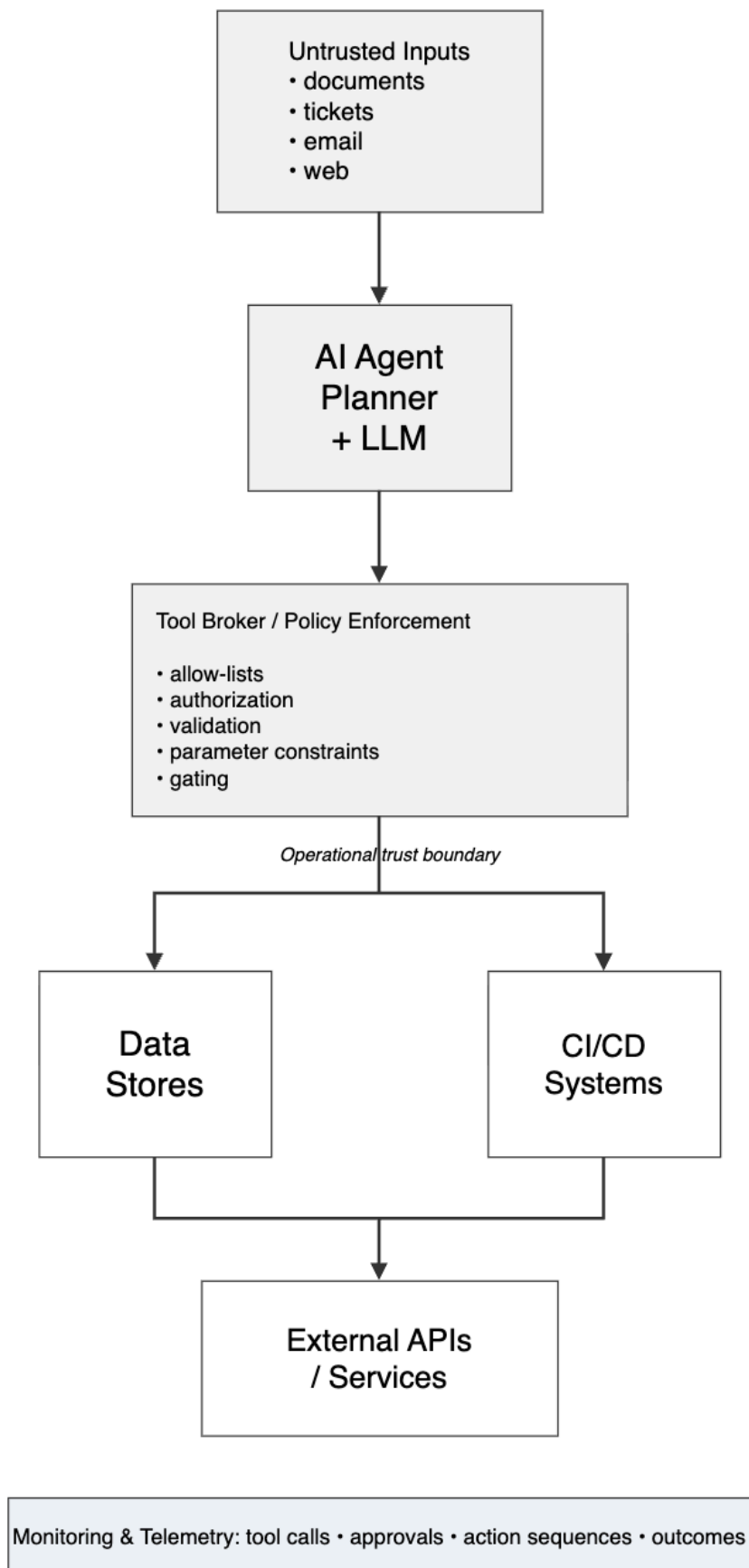
#### **Produce evidence**

Traceable, repeatable, and auditable outputs

This enables a continuous, evidence-driven assurance process rather than static validation.

---

## AI Agent as a System Actor Orchestrating Tools Across Trust Boundaries



### 3. Proposed Use Case

The use case focuses on evaluating an AI-assisted operational decision workflow within a controlled sandbox environment.

The objective is to demonstrate how such systems can be safely integrated into critical infrastructure with:

- clear visibility of decision pathways
  - explicit mapping of risks and controls
  - repeatable validation under realistic conditions
  - auditable evidence of system behaviour
- 

The following is an example risk in AI-assisted workflows where decision approval and execution can diverge:

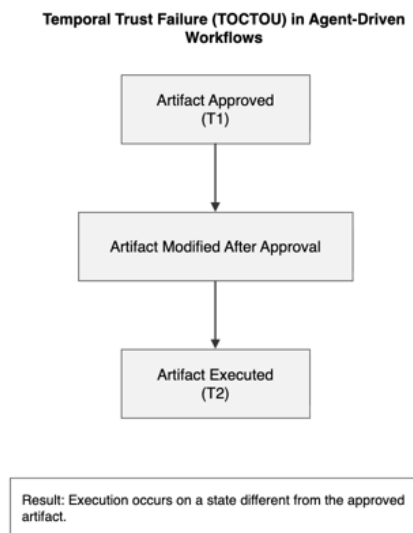


Figure 2 – Temporal trust failure (TOCTOU) in automation workflows where approval and execution occur at different states.

---

### 4. Expected Outputs

- Structured system model of an AI-assisted workflow
- Explicit mapping of risks and controls
- Defined test scenarios and results
- Evidence pack demonstrating system behaviour

- Recommendations for regulatory guidance
- 

## **5. Why This Fits the Ofgem Sandbox**

This use case directly supports the sandbox objectives:

- Controlled testing of AI in realistic operational scenarios
- Risk-based evaluation of system behaviour
- Generation of structured, auditable evidence
- Improved understanding of safe AI deployment

It provides a practical method to evaluate AI systems as part of operational infrastructure, not in isolation.

---

**Lantrix Limited**