

Field	Value
Document ID	GOV-EXT-001
Version	1.0
Effective Date	December 15, 2024
Document Owner	Chief AI Officer
Classification	Internal Use
Review Cycle	Annual

Table of Contents

Explainability & Transparency Standard..... 1

1. Purpose..... 1

2. Scope..... 1

2.1 Systems Coverage..... 1

2.2 Documentation Requirements..... 2

2.3 Stakeholder Coverage.....2

3. Requirements Framework..... 2

3.1 Risk-Based Transparency Levels.....2

3.2 Explanation Method Requirements.....2

3.3 Model Card Standards..... 3

3.4 User Notice Requirements..... 4

3.5 Limitation Documentation..... 5

4. Implementation Requirements..... 5

4.1 Explanation Infrastructure..... 5

4.2 Documentation Management..... 5

4.3 Training and Competency..... 6

5. Governance and Oversight..... 6

5.1 Roles and Responsibilities.....6

5.2 Monitoring and Metrics..... 7

6. Compliance and Standards Alignment..... 8

6.1 Regulatory Mapping..... 8

6.2 Industry Standards Integration..... 8

7. Implementation Examples..... 8

7.1 Credit Decision System Example..... 8

7.2 Content Recommendation Example.....9

7.3 Healthcare Diagnostic Support Example..... 9

8. Supporting Artifacts..... 9

8.1 Document Templates.....9

8.2 Technical Implementation Guides..... 10

8.3 Training Materials..... 10

9. Continuous Improvement..... 10

9.1 Feedback Integration..... 10

Explainability & Transparency Standard

1. Purpose

This standard establishes comprehensive requirements for AI system explainability and transparency to ensure stakeholders can understand AI decision-making processes, limitations, and appropriate usage contexts. The standard mandates right-sized interpretability measures proportional to system risk and impact, while ensuring clear, accessible communication to all user populations through standardized notices and documentation.

The standard addresses regulatory compliance requirements under the EU AI Act Article 13 (transparency obligations), OECD AI Principles on transparency and explainability, and ISO 42001:2023 information management requirements. Organizations implementing this standard will achieve consistent, defensible transparency practices across their AI portfolio while maintaining operational efficiency through risk-proportionate approaches.

2. Scope

2.1 Systems Coverage

This standard applies to all AI systems within scope of organizational AI governance, including:

- Machine learning models in production environments
- Decision support systems incorporating AI components
- Automated decision-making systems affecting individuals
- AI-powered recommendation engines
- Natural language processing applications
- Computer vision systems
- Predictive analytics platforms
- AI systems integrated within third-party applications

2.2 Documentation Requirements

The standard encompasses the following transparency artifacts:

- Model cards providing system overview and performance characteristics
- Technical explainability documentation for internal stakeholders
- User-facing explanations and notices
- Algorithmic impact assessments where required
- Audit trails for significant decisions
- Training data documentation and lineage