

Opacity as a feature, not a flaw: Role-sensitive explainability, institutional trust, and the LoBOX ethics governance framework for AI

F. Herrera Triguero; M.R. Calderón Cuadrado

Abstract-

This paper introduces the LoBOX (Lack of Belief: Opacity & eXplainability) ethics governance framework, a governance-centric approach for managing artificial intelligence (AI) opacity when full transparency is infeasible. While transparency-centric approaches treat transparency as the social/ideal goal and therefore opacity as a design flaw, LoBOX suggests opacity is a condition which should be ethically governed through role-sensitive explanation and institutional accountability. The LoBOX framework comprises a three-stage pathway: reduce accidental opacity, bound irreducible opacity, and delegate trust through institutional oversight. Integrating the stakeholder-sensitive explanation described in the RED/BLUE XAI model, which is aligned with emerging legal instruments such as the EU AI Act, LoBOX offers a scalable and context-aware alternative to transparency-centric approaches. LoBOX reframes trust as an outcome of institutional credibility, structured justification, and stakeholder-sensitive accountability, and it is designed to remain aligned with evolving technological contexts and stakeholder expectations while ethically governing opacity. In the end, to ensure responsible AI systems, LoBOX moves from transparency ideals to ethical governance, emphasizing that trustworthiness in AI must be institutionally grounded and contextually justified.

Index Terms- Explainable artificial intelligence (XAI)Algorithmic opacityAI governanceInstitutional trustRole-sensitive explainabilityAccountability

Due to copyright restriction we cannot distribute this content on the web. However, clicking on the next link, authors will be able to distribute to you the full version of the paper:

[Request full paper to the authors](#)

If your institution has an electronic subscription to Technology in Society, you can download the paper from the journal website:

[Access to the Journal website](#)

Citation:

Herrera, F.; Calderón, R. "Opacity as a feature, not a flaw: Role-sensitive explainability, institutional trust, and the LoBOX ethics governance framework for AI", Technology in Society, , .