Intervention: Operationalising Ethics for AI regulation and governance

Dr **Daria Onitiu**, Postdoctoral researcher, Oxford Internet Institute
Email: daria.onitiu@oii.ox.ac.uk

3rd December 2024

Trustworthiness Auditing for AI project & Governance of Emerging Technologies (GET) programme

Multidisciplinary perspectives for the governance of emergent tech (law, AI ethics, computer science and social psychology)

Real-world deployment / implementation challenges of AI in medicine and health.
(Large) Generative AI, the legal and technical constraints.

2

1. **How to balance risks and benefits of emerging technologies under conditions of acceleration, and demand in adoption?**
2. **How are risks and benefits balanced in practice; who has a say, and for whom?**

Specifically relevant when:

[OECD Framework for Anticipatory Governance of Emerging Technologies, 2024, p. 12](#):
Agile regulation: Experimentation and testing under regulatory supervision should be encouraged to foster innovation, reduce uncertainty, and ensure that governance systems remain relevant and effective. Policy makers should also explore the potential of non-binding governance approaches such as high-level principles, technical standards and codes of conduct.

① *Making safety and benefits explicit*

Risk framing and narrow safety and usability claims ([Onitiu, Wachter and Mittelstadt, 2024,, p.3](#))

For example, examining the capacity of an AI-imaging tool using pictures from hospital A or publicly available images to claim that it performs better than a healthcare professional reading.

Deflating **overstated claims**

European Parliament

Artificial intelligence in healthcare

Applications, risks, and ethical and societal impacts

STUDY
Panel for the Future of Science and Technology

EPRS | European Parliamentary Research Service

ICS ← 35 ← 35.020

ISO/IEC TR 24028:2020
Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence

② *Weighing priorities*

From rhetoric to **real-world impact:**

"liver allocation algorithm used by the UK's National Health Service (NHS) has been found to discriminate by age. No matter how ill, patients under the age of 45 seem currently unable to receive a transplant, due to the predictive logic underlying the algorithm" (AI Accountability Lab, 2024)

Providers weighing priorities, such as fairness will entail some value-judgements, such as who is included and excluded from a device's intended use (Onitiu, Wachter, Mittelstadt, 2024).
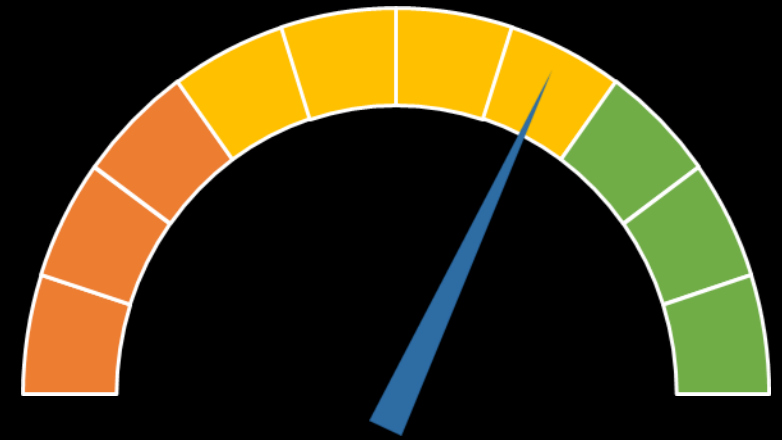
Technical safeguards
Demonstrating safety and effectiveness
Promoting confidence

③ *Common methodologies*

[Practical Guidance on Agile Regulatory Governance to Harness Innovation, 2021, p. 2.](#)
More holistic, open, inclusive, adaptive and better-co-ordinated governance models to enhance systemic resilience by enabling the development of agile, technology neutral and adaptive regulation that upholds fundamental rights, democratic values and the rule of law. It will also involve acknowledging that innovation ecosystems and related value chains tend to cut across national or jurisdictional boundaries and thus require concerted governance approaches.

Specifically, stakeholder involvement and collaboration between regulators, Industry, civil society, as well as users of technology; international coordination and shared understanding of innovation and legal certainty, capacity building and/or institutional mandates needed to oversee and monitor technological developments.

THANK YOU

Dr Daria Onitiu

daria.onitiu@oii.ox.ac.uk