












***Intervention: Operationalising Ethics for AI regulation and governance***

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

**3<sup>rd</sup> December 2024**

 <p><b>Professor Sandra Wachter</b> <i>Professor of Technology and Regulation</i></p> <p>Professor Sandra Wachter is Professor of Technology and Regulation focusing on law and ethics of AI, Big Data, and robotics as well as Internet regulation at the Oxford Internet Institute at the University of Oxford</p> <p><a href="#">VIEW PROFILE</a></p>	 <p><b>Professor Brent Mittelstadt</b> <i>Director of Research, Associate Professor, Senior Research Fellow</i></p> <p>Professor Brent Mittelstadt is a data ethicist and philosopher specializing in AI ethics, professional ethics, and technology law and policy. He is the founder/coordinator of the Governance of Emerging Technologies (GET) programme.</p> <p><a href="#">VIEW PROFILE</a></p>	 <p><b>Professor Chris Russell</b> <i>Dieter Schwarz Associate Professor, AI, Government &amp; Policy, Research Associate</i></p> <p>Chris Russell is the Dieter Schwarz Associate Professor, AI, Government and Policy. Dr Russell's work lies at the intersection of computer vision and responsible AI.</p> <p><a href="#">VIEW PROFILE</a></p>
 <p><b>Dr Johann Laux</b> <i>British Academy Postdoctoral Fellow</i></p> <p>Johann Laux works at the intersection of law and the social sciences. His current research is interested in the governance of emerging technologies as well as the design of institutions.</p> <p><a href="#">VIEW PROFILE</a></p>	 <p><b>Professor Netta Weinstein</b> <i>Research Associate</i></p> <p>Netta Weinstein is a professor of clinical and social psychology at the University of Reading. She studies human motivation, behaviour and well-being in collaboration with Professors Wachter and Mittelstadt.</p> <p><a href="#">VIEW PROFILE</a></p>	 <p><b>Dr Daria Onitiu</b> <i>Postdoctoral Researcher</i></p> <p>Daria Onitiu is a Postdoctoral Researcher within the Trustworthiness for AI Auditing project. She is interested in the facets and boundaries of the law to address pressing societal problems caused by novel technologies.</p> <p><a href="#">VIEW PROFILE</a></p>
 <p><b>Dr Lizzie Barclay</b> <i>PhD Student, University of Reading</i></p> <p>Lizzie is Associate Medical Director at Annalise AI, a former NHS doctor and also a PhD student at the University of Reading. She has a particular interest in the intersection of law and emerging technologies.</p>	 <p><b>Fyonna Boateng</b> <i>Placement Student, University of Reading</i></p> <p>Fyonna Boateng is a Psychology undergraduate at the University of Reading. With Dr. Netta Weinstein, she is investigating public opinions of AI and its potential biases.</p>	 <p><b>Eoin Delaney</b> <i>Postdoctoral Researcher</i></p> <p>Eoin is a postdoctoral researcher on the Trustworthiness Auditing in AI project. His research interests include explainable and trustworthy AI, interpretable machine learning and responsible AI.</p>

## 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.*

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**



– ICS ← 35 ← 35.020

ISO/IEC TR 24028:2020

Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence

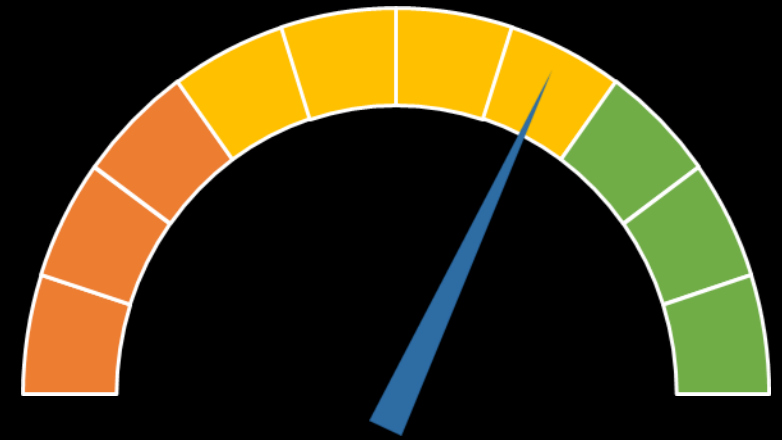
# 2

## 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](mailto:daria.onitiu@oii.ox.ac.uk)