Governance and Regulation

Session Chair: Dr. Jennifer Williams
University of Southampton
Assistant Professor

TAS Showcase

5th – 6th March 2022

TAS Showcase 2024
www.tas.ac.uk

Welcome

Jennifer Williams

Assistant Professor, University of Southampton

EPSRC EdgeAI Hub South England Engagement Director

- Autonomous systems will proliferate
  - Every aspect of daily life and in every sector
  - Will make us more prosperous and secure
  - Introduce new kinds of vulnerabilities and safety issues

- This is being recognized globally
  - 2023 Global AI Safety Summit at Bletchley Park
  - EU AI Act – unacceptable and high-risk applications
  - US AI Bill of Rights

- Going forward requires society-wide engagement
  - Academics, Developers, Industry, Government, Civil Society

First TAS AI Regulation Workshop – May 2023

- **45 attendees from UK universities, UK regulators**
- **Invited talks from 6 UK regulators**
- Bring together multiple stakeholders to open up communication and normalise interaction
- Examine cross-sector regulation challenges
- Regulators: NCSC, National Highways, DSIT/Office for AI, HSE, Ofcom, CMA
- Three sessions: 1) problematising, 2) solutioneering in-domain, 3) solutioneering across domains

Key takeaways:

- ➢ Regulators need cross-collaboration and cross-regulation advising to solve the biggest problems
- ➢ There is a need for more evidence-gathering and case studies
- ➢ Communication beyond policy is important for public understanding

Second TAS AI Regulation Workshop – November 2023

- **55 attendees from industry, UK universities, UK regulators, UK government**
- **Invited talks from 6 UK regulators and academics**
- Discuss challenges, opportunities and risks of AI regulation in different sectors
- Regulators: CAA, MAA/DSA, CMA, HSE, Electoral Commission, UK Parliament, DSIT/CDEI ++
- Talks and world café discussions

Key takeaways:

- ➢ Incredibly difficult to look at AI regulation across sectors
- ➢ Dfferent sectors have different appetites for AI regulation and different pressing issues
- ➢ The term "safety-critical systems" means different things to different regulators in terms of risk of harm

Overview of this session

- Two invited talks
- Discussion panel
- Audience activity session
- Getting involved

Talk: Nuala Polo

Senior Policy Advisor and AI Assurance Lead
Department for Science, Innovation, and Technology

Department for Science, Innovation & Technology

Tools for Trustworthy AI:

Implementing the UK's AI Governance Strategy

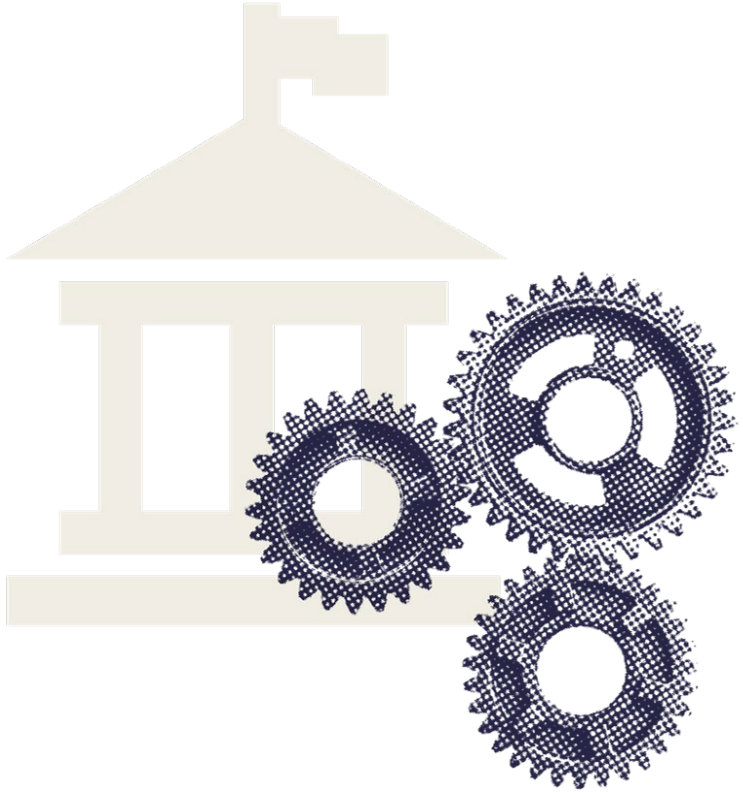
Nuala Polo – AI Assurance Lead

About the RTA

Department for Science, Innovation & Technology

- **The first of its kind in the world, the Responsible Technology Innovation Unit (RTA) leads the UK government's work to enable trustworthy innovation using data and AI.**

- It is vital that the public can **trust** innovation in data and AI. To earn that trust, RTA works with **partners** across the public sector, industry and academia, in the UK and internationally, to identify and tackle barriers to responsible innovation.



What is AI governance

AI governance refers to mechanisms including **laws, regulations, policies, institutions, and norms** that set out processes for making decisions about AI.

The goal of AI governance is to maximise the benefits of AI systems, while mitigating potential risks and harms.

Department for Science, Innovation & Technology

Department for Science, Innovation & Technology

Key elements of our pro-innovation framework

Cross-sectoral principles

Our framework will be underpinned by a set of cross-sectoral principles including concepts such as transparency, safety and security, to guide how actors in the AI ecosystem approach responsible AI and AI risk

Leveraging existing regulator expertise

We will leverage the sector expertise of our world-class regulators, focusing on outcomes rather than the technology itself. We balance the economic and societal potential benefits of AI against its risks.

Context-specific

We acknowledge that AI is a dynamic, general-purpose technology and that the risks arising from it depend principally on the context of its application. The same AI used in different context may need regulating differently

Central functions to drive coherence

To ensure that the overall framework offers a proportionate, coherent and effective response to risk while promoting innovation across the regulatory landscape

The proposed cross-cutting principles

Department for Science, Innovation & Technology

Existing regulators will be expected to implement the framework underpinned by 5 values-focused cross-sectoral principles, based on **OECD AI Principles**

Safety, Security & Robustness	Ai systems should function in a robust , secure and safe way throughout the AI life cycle, and risks should be continually identified, assessed and managed.
Appropriate Transparency & Explainability	AI systems should be appropriately transparent and explainable
Fairness	AI systems should not undermine the legal rights of individuals or organisations, discriminate unfairly against individuals or create unfair market outcomes. Actors involved in all stages of the AI life cycle should consider definitions of fairness that are appropriate to a system's use, outcomes and the applicant of relevant law
Accountability & Governance	Governance measures should be in place to ensure effective oversight of the supply and use of AI systems, with clear lines of accountability established across the AI life cycle. AI life cycle actors should take steps to consider, incorporate and adhere to the principles and introduce measures necessary for the effective implementation of the principles at all stages of the AI life cycle.
Contestability & Redress	Where appropriate, users, impacted third parties and actors in the AI life cycle should be able to contest an AI decision or outcome that is harmful or creates material risk of harm.

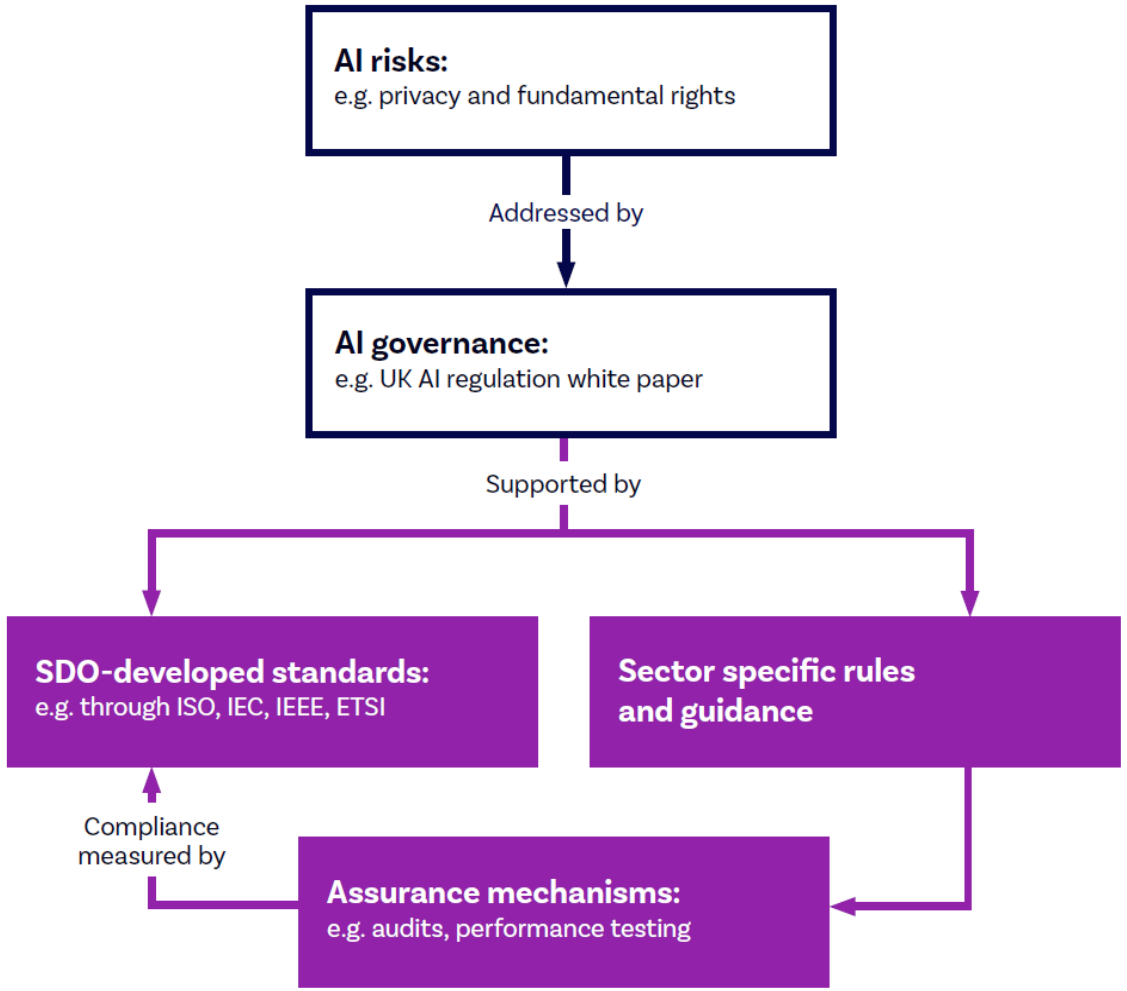
Tools for trustworthy AI

Department for Science, Innovation & Technology

Tools for trustworthy AI will play a critical role in enabling the responsible adoption of AI by **supporting the implementation of regulatory framework** and **boosting international interoperability.**

These tools include:

1. **Assurance Mechanisms**
2. **SDO-developed standards**



Assurance techniques

Department for Science, Innovation & Technology

The goal of assurance techniques, is to measure, evaluate and communicate whether AI systems are trustworthy.

There are a range of different techniques for assuring AI systems, that can be used in combination with one another across the AI lifecycle.

Technical standards

Department for Science, Innovation & Technology

There are many different types of standards, including:

Foundational standards build common understanding around definitions and terminology

Process standards universalise best practice in organisational management and governance

Measurement standards define metrics and methods for quantitative measurement

Performance standards set specific performance thresholds for acceptability

Without standards of some kind, we have **advice**, rather than **assurance.**

Department for Science, Innovation & Technology

Thank you

For more information please contact rtau@dsit.gov.uk

Talk: Richard Sturman

Defence Safety Authority
Military Aviation Authority
Futures Strategic Development - Maritime

UKRI Trustworthy Autonomous Systems Programme

TAS Showcase – 06 Mar 24

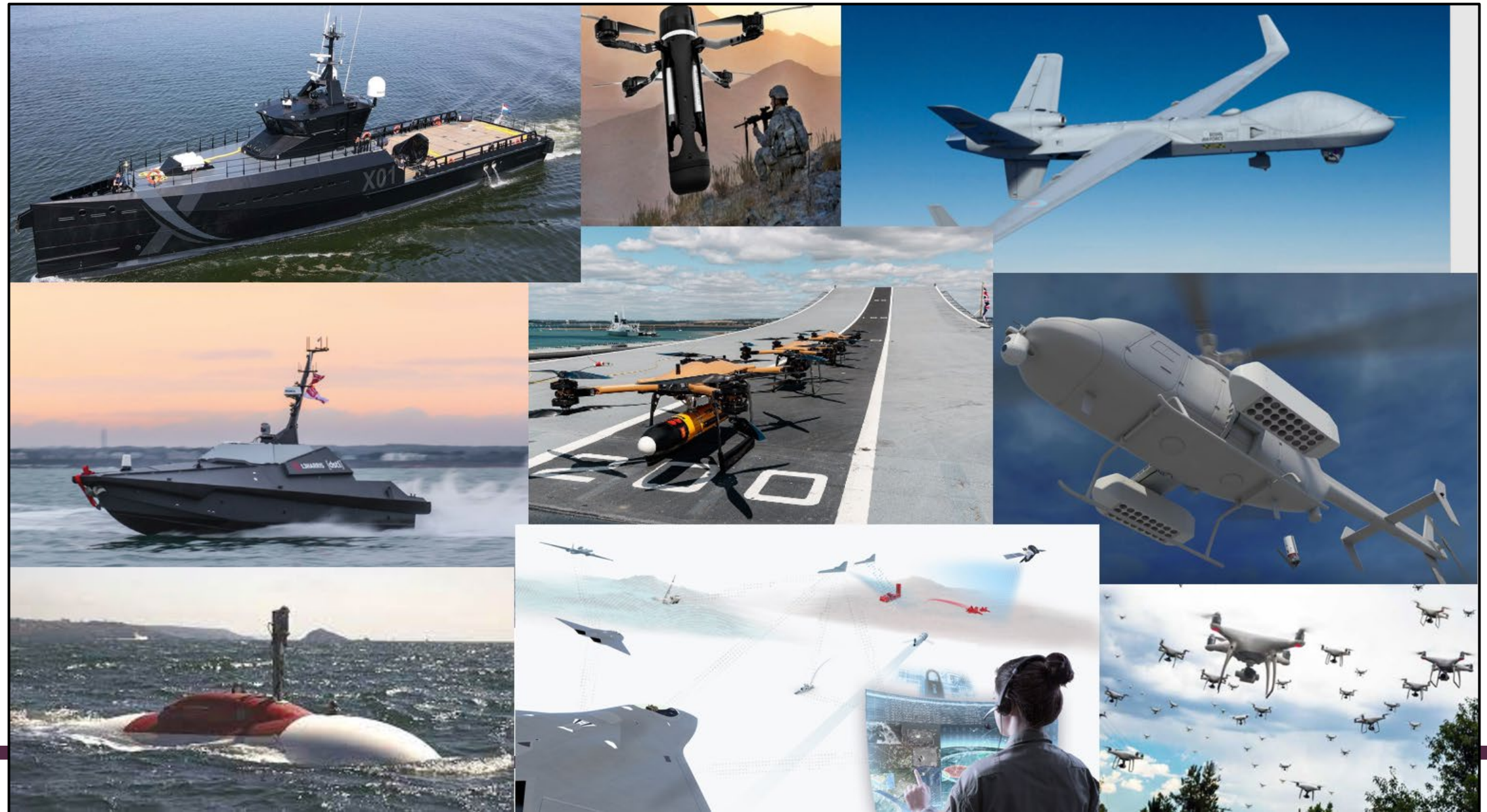
Defence Safety Authority

Cdr Richard Sturman

- **"We are at a crossroads in human history" (Michelle Donelan, DSIT, 27 Oct 23)**
- **I believe nothing in our foreseeable future will be more transformative for our economy, our society, and all our lives, than this technology."**

(PM Rishi Sunak announces new AI safety Institute - statement 26 Oct 23)

Opportunity?



Impact

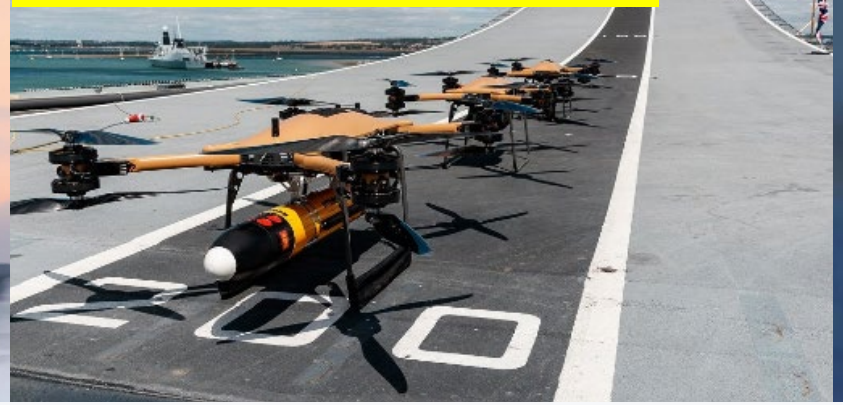
- Reduced risk to 1st party operators
- Autonomous operations (many to one)
- Greater endurance/range/ops in austere conditions
- Increased ‘footprint’ / ‘mass’
- Reduced operating costs
- Political freedoms (grey zone/thresholds)
- Improved safety (?)
- What else???

Risks

- Safety concerns (lack of assurance tools/standards)
- Rogue systems
- Societal concern (LAWS)
- Lack of transparency/explainability/understanding
- Lack of SQEP operators/maintainers (?)
- Over-reliance on tech!
- What else??









Questions?

AI Regulation Panel Discussion

Stuart Anderson

Sana Khareghani

Nuala Polo

Justyna Lisinska

Ram Ramamoorthy

Rich Sturman

Socially Assistive Robots

Thinking about regulations for future technologies

Marta Romeo and Miranda Addey

Robot & Frank (2012) movie by Jake Schreier

Internet of Robotic Things (IoRT)

- Ambient Assistive Living
- Intelligent devices that monitor and fuse sensor data from a variety of sources
- Local and distributed intelligence to determine best course of action
- Socially assistive robot

Use Case

Each morning the robot reminds its user to step **on the scale** to check their weight. During the day, the user uses a **smart fork** to eat all their meals. After a month of weight checking, the **robot** detects an increase in weight and checks the records of the smart fork and **smartwatch**. The robot notices an increase intake of carbs and fat and a lower physical activity so it looks for its user to **recommend** a more balance diet and propose a 30 minutes walk every day, which the user has **to log onto the robot** to confirm having complete.

The Regulating Game

Reflect on the system described:

1)Who should regulate it?

2)What part of the system should your first choice regulating agency regulate?

3)What is important when considering data ownership?

4)Who should be held responsible for the system's mistakes?

5)Who pays for Socially Assistive Robots and their ecosystem if they become essential?

The Regulating Game

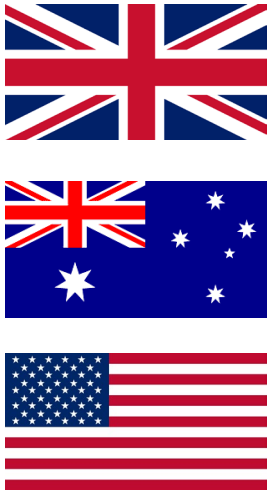
Slido Session: Governance and Regulation

AI Regulation Engagement

RAiUK

New RAI UK International Partnership
AI Regulation Assurance for Safety-Critical Systems

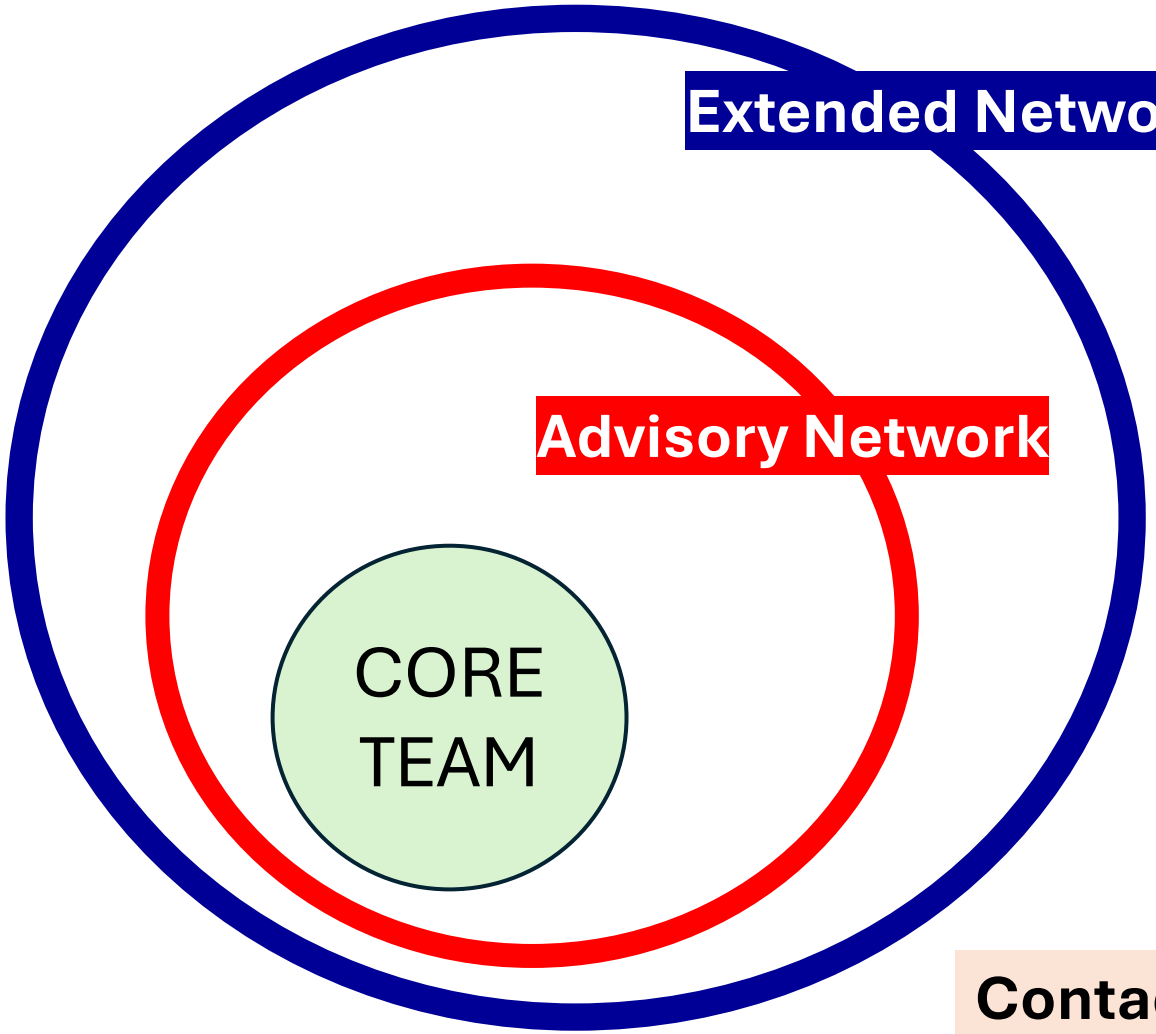
Communications | Maritime | Aerospace



PI Jennifer Williams (University of Southampton, UK)
Co-I Peng Wei (George Washington University, USA)
Co-I Zena Assad (Australia National University, Australia)

- Develop a toolkit for AI regulation safety assurance
- Deep dive into specific technologies in the three sectors
- Explore cross-regulation
- Identify technical and regulatory gaps

How to engage



Advisory Network

Industry, government, third-sector

Extended Network

TAS Community, RAI Community
Special interest groups, academics, Researchers

Contact: Jennifer Williams j.williams@soton.ac.uk

Questions

Get in Touch

contact@tas.ac.uk
@tas_hub
https://www.linkedin.com/groups/8966227