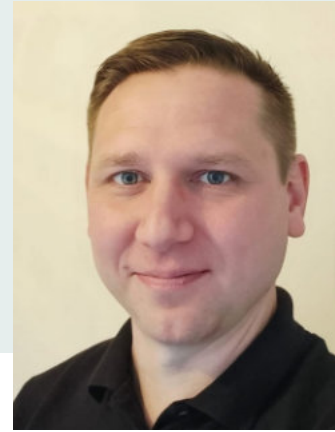


Beyond regulation:

Why the UK needs a national approach to AI assurance



Chris Dodd BEng IEng MIMechE RN
Managing Director and Co-Founder, R2AI

Today, Artificial Intelligence (AI) is moving faster than policy can comfortably track, and in many cases so fast that policies covering the subject don't even exist.

The UK is rightly proud of its global reputation for engineering excellence, safety assurance, and scientific integrity. All properly researched and planned. Yet in AI those same disciplines are being tested in real time.

So far we've seen local policy, strategy papers, and international summits. What we have not yet seen with AI is assurance: the practical, operational proof that today's AI systems can operate themselves safely, accountably, and within the boundaries of law and ethics.

At R2AI, we believe the next stage of national AI leadership will depend on building more trust than technology. That means building assurance into the development process. Our work focuses on turning responsible AI principles into measurable practice, through training, governance frameworks, and collaboration across sectors.

We are under no illusion about the scale of this challenge. In the history of modern technology, few advances have matched AI's pace, complexity, and impact. Meeting this pace will require partnership, shared standards, and the same spirit of engineering discipline that built the UK's reputation in the first place.

Regulation defines what is legal. Assurance defines what is safe.

The UK's AI Regulation White Paper¹ and the AI Safety Summit² at Bletchley Park signalled real intent to lead responsibly in this field. But policy and principle can only take us so far.

Regulation, by its nature, sets boundaries. It defines what to do and what not to do. Assurance is that element in the process which builds confidence that what is being done will work as intended.

In high-reliability sectors, from aviation to nuclear engineering, assurance is not an optional extra; it is the process that sits between innovation and implementation. It is how we move from theory to trusted reality. If we expect the public to trust AI in our hospitals, courts, classrooms, or command and control systems in Critical National Infrastructure, that same assurance mindset must now be hardwired into national AI governance.

I have spent most of my career as an engineer working on systems where failure simply was not an option, yet all too often a reality, serving within an ever-ageing Royal Navy fleet. But that is a story for another day. You do not certify safety by hope; you certify it through evidence, repeatability, and transparency. Those same principles now need to underpin how we govern AI.

Fragmented oversight, shared risk

Right now, the beginnings of the UK's AI oversight framework is spread across multiple regulators: the Information Commissioner's Office, the Competition and Markets Authority, the Medicines and Healthcare products Regulatory Agency, the Financial Conduct Authority, Ofcom, and others, each tasked with a different piece of the puzzle, none yet enshrined

in law. It is an understandable structure, AI cuts across every sector. But this fragmentation creates uncertainty for both innovators and operators.

If you are a company deploying AI in recruitment, finance, or healthcare, who do you go to for approval? Who verifies that your system is compliant, explainable, and safe to use? Each regulator interprets the risk through its own lens, but there is no single point of coherence. That lack of coherence is itself a risk.

Without an integrated assurance framework, we are asking industry to navigate a maze of overlapping responsibilities with no central assurance body to verify what 'safe AI' even means.

The missing link: assurance infrastructure

Assurance, therefore, is not a paperwork exercise; it is infrastructure. The structured process of verifying that a system can operate safely, ethically, and as intended before it reaches the wider public.

We already do this for aircraft, medical devices, and even offshore oil rigs. Yet when it comes to AI, a technology capable of shaping opinions, making hiring decisions, and influencing financial or judicial outcomes, we have no comparable mechanism.

An AI assurance infrastructure would mean that every high-impact system, from decision-support tools in the NHS



World leaders gathered for the inaugural Artificial Intelligence Safety Summit at Bletchley Park, UK.

to autonomous vehicles or generative AI models in education, goes through a repeatable auditable process of independent validation. It would confirm compliance with law, and compliance with its intent: that the system behaves safely, fairly, and transparently across its life cycle.

Digital responsibility

In engineering, you design for safety, then you prove it. You do not bolt it on later. The same logic applies to AI. Ethical and legal risks are not post-deployment issues, or just something to manage once the technology is already in service.

In our work supporting assurance development across sectors, we've seen, that assurance is most effective when it is treated as a design feature, not a defensive one. That means embedding governance and oversight from the very start of development, not waiting for a regulator to intervene.

This approach does not slow innovation; it strengthens it. Models that can demonstrate traceability, human oversight, and lawful data use, move faster and face fewer barriers when scrutiny inevitably arrives. Assurance becomes the enabler of responsible innovation, not the obstacle to it.

A national framework for AI assurance

It is time to treat AI assurance as a national capability, not a niche service. I believe the UK should establish a cross-sector National AI Assurance Framework, coordinated to bring coherence to the current patchwork arrangements.

Such a framework could establish shared assurance standards across sectors, ensuring that 'safe AI' means the same thing in healthcare as it does in finance or defence. It would provide certification routes for the systems that contain AI, enabling trusted deployment within the boundaries of UK law, and introduce consistent incident reporting, learning loops, and audit trails to maintain transparency when things go wrong. Crucially, it would also invest in people, building the skills and literacy programmes needed for professionals who understand both the technology and the governance behind it.

The goal is consistency. We already know how to do this. The same engineering mindset that built Britain's nuclear and aerospace safety regimes can be adapted to digital systems. The question is whether we are willing to treat AI with the same seriousness.

The government does not need to do this alone. In fact, it should not. Assurance thrives when expertise is distributed and validated through collaboration.

The UK's private sector already holds the practical experience: the engineers, ethicists, and governance specialists working with real-world AI deployments every day. Academia provides the research depth and ethical frameworks. Government sets the standards and expectations. Together, they can build a national assurance capability that reflects British values of safety, fairness, and transparency.

In my own work, I have seen this kind of partnership model deliver real results. Whether it is a small business trialling AI

or a complex, high-risk environment, the outcome is always the same: once people have clear governance structures, confidence replaces confusion, and innovation finds its footing.

Public confidence in AI does not come from press releases or political statements; it comes from visible, verifiable safety.

When an AI system produces a decision that affects a person's future, such as a mortgage, a job interview, or a parole hearing, the public expects the same level of scrutiny that we apply to any other safety-critical process. They expect oversight, documentation, and accountability.

That is why assurance is more than compliance. It is about being able to show, transparently, that humans remain in control, that there are guardrails, reviews, and real consequences if those controls fail.

Without that assurance layer, even the best-intentioned policy risks losing public trust. And once trust is lost, it is incredibly hard to regain.

The challenge before Parliament is not whether to regulate AI; that debate is already settled. The real question is whether we can build the necessary assurance systems that give regulation meaning.

References

¹ Department for Science, Innovation and Technology and Office for Artificial Intelligence (2023). *AI Regulation: A Pro-innovation Approach*. White Paper.

² UK Government (2023). *Chair's Summary of the AI Safety Summit 2023*, Bletchley Park.