

Engineering safe AI at scale

How DSTL and the IET are helping accelerate the safe adoption of AI across defence

A partnership shaped by purpose and complexity

The Defence Science and Technology Laboratory (DSTL) operates at the intersection of science, technology and national security, where performance, reliability and safety are non-negotiable. As AI becomes increasingly integrated into safety-critical applications, ensuring its safe adoption has become a strategic priority

As AI becomes increasingly integrated into safety-critical applications, traditional assurance approaches must evolve to address new levels of complexity, uncertainty and autonomy.

Recognising this challenge, DSTL partnered with the IET to develop guidance that would help organisations understand, assess and manage the functional safety implications of AI-enabled systems.

“MOD does not operate in isolation – it is part of a wider UK ecosystem. The increasing adoption of AI in safety-related applications has driven collaboration across MOD... to maintain safety”

This belief sits at the heart of the collaboration. It reflects a shift from standalone innovation towards a more connected, system-wide approach to safe technology adoption.

The challenge: enabling innovation without compromising safety

DSTL has long played a central role in functional safety across the Ministry of Defence. But the rapid pace of AI development has introduced new challenges, and opportunities, that traditional approaches cannot fully address.

AI and machine learning systems can be difficult to predict, creating assurance challenges for engineers, safety practitioners and programme teams. As adoption increased, DSTL identified a need for a more consistent approach to AI safety assurance.

The objective was not only to address emerging risks but also to remove uncertainty that could slow innovation, procurement or deployment decisions. By establishing practical guidance, organisations could move forward with greater confidence while maintaining robust safety standards.

The solution: creating a shared framework for AI assurance

To address this challenge, DSTL commissioned the development of guidance focused on applying functional safety principles to AI-enabled systems. The initiative was designed to provide organisations with a structured approach to understanding and managing AI-related safety risks.

The guidance supports several strategic objectives:

- Strengthening trust and assurance in AI-enabled systems.
- Establishing greater consistency in safety assessments and decision-making.
- Reducing uncertainty across assurance, procurement and governance processes.
- Enabling responsible innovation while maintaining operational safety.

Perhaps most importantly, it provides a common language for practitioners working across engineering, safety, policy and operational domains, helping organisations align around what safe and responsible AI adoption looks like in practice.

Why collaboration matters

Developing effective guidance required expertise that extended beyond any single organisation.

Through its collaboration with the IET, DSTL brought together subject matter experts from government, industry and academia, creating a collaborative environment where knowledge, experience and emerging best practice could be shared.

This cross-sector approach ensured the guidance reflected both technical rigour and practical application.

"Publication by the IET provides both reach and recognised authority."

Working together also generated value beyond the publication itself, creating opportunities for dialogue between AI specialists, safety engineers and decision-makers facing similar challenges across different sectors.

"Arguably, it was the process of collaboration that was more valued than the output itself."

Crucially, it positioned the work within a wider ecosystem, rather than as a defence-specific solution.

Delivering measurable value

The guidance is already helping organisations make more informed decisions about the development, adoption and assurance of AI-enabled technologies.

Key benefits include:

- Greater clarity in safety assurance activities.
- Improved consistency across programmes and stakeholders.
- Increased confidence in the assessment of AI-related risks.

Stronger collaboration between developers, engineers and assurance teams. By providing practical guidance and shared benchmarks, organisations are better equipped to navigate complexity, reduce uncertainty and accelerate responsible adoption of emerging technologies.

Looking ahead: building future capability

As AI technologies continue to advance, the need for effective assurance frameworks will only increase. Emerging developments, including increasingly autonomous and agentic systems, will introduce new safety considerations that demand ongoing collaboration and innovation.

For DSTL, this work represents an important foundation for future capability development – and not just a one-off project. The insights gained through the project continue to inform thinking around AI assurance, while the partnerships created through the initiative provide a platform for future research and knowledge sharing.

“Particularly around state-of-the-art AI technologies... and their challenges in relation to functional safety.”

By working with the IET, DSTL is better positioned to stay ahead of that curve.

Supporting a safer, more responsible future

The safe adoption of AI requires more than technological innovation. It requires trusted frameworks, expert collaboration and a shared commitment to responsible implementation.

For DSTL, this work reflects a broader commitment.

"The MOD are leading proponents of ambitious, safe and responsible adoption of AI."

The guidance forms part of a much larger effort to ensure that technological progress delivers positive outcomes – not just for defence, but for society as a whole.

“We trust that its value is much broader than defence and that it contributes to a safer world.”

A partner in better business

This is what effective partnership looks like in practice.

- Not simply delivering outputs but enabling capability.
- Not working in isolation but strengthening the ecosystem.
- Not just responding to change but shaping it.

For organisations navigating the complexities of AI, safety and innovation, the message is clear: The IET is not just a professional body. It is a partner in building the clarity, confidence and capability needed to move forward.

Connect with to our team to discuss Codes of Practice and Guidance collaboration opportunities in your sector