CREST AI

Building trust in AI-enabled cybersecurity

Artificial intelligence is becoming part of cybersecurity service delivery.

Organisations are using AI to improve efficiency, enhance capabilities and support decision-making across a growing range of cybersecurity activities.

As adoption accelerates, trust, accountability and transparency become increasingly important. The CREST AI Principles provide a practical foundation for responsible AI use in cybersecurity, helping organisations, practitioners, buyers and regulators navigate this evolving landscape with confidence.

Introducing the CREST AI Charter and why it matters

- AI presents significant opportunities for the cybersecurity profession. It can help practitioners work more efficiently, improve service delivery and strengthen cyber resilience.
- At the same time, organisations need confidence that AI is being used responsibly, transparently and with appropriate oversight.
- The **CREST AI Principles** establish shared expectations for the use of AI across cybersecurity services and provide a foundation for trust, accountability and assurance.

The CREST AI Charter brings together organisations that support the responsible use of AI in cybersecurity; and recognise the importance of shared principles, professional standards and industry collaboration.

By becoming a CREST AI Charter Signatory, organisations agree to publicly support CREST's AI Principles by:

- **Recognising the Principles as a practical foundation for responsible AI use in cybersecurity, and;**
- **Helping shape trusted AI-enabled cybersecurity services**

By signing the Charter, your organisation is supporting the CREST AI Principles and contributing to a growing industry conversation around trust, transparency, accountability and assurance in AI-enabled cybersecurity services.

The CREST AI Charter is a voluntary commitment to support responsible AI use across the cyber profession. View it here

crest-approved.org/ai-charter

~60 FOUNDING SIGNATORIES

ACROSS 15 COUNTRIES

UNITED BY TRUST, TRANSPARENCY AND RESPONSIBLE AI USE

The CREST AI Principles

Nine principles for AI-enabled activities

1 Accountability and governance

We define the scope and purpose of AI-enabled activities and assess how they may affect service delivery, client outcomes, data handling, decision-making or operational risk. We apply oversight, testing and governance controls proportionate to the nature, scale and risk of the AI use.

2 Transparency of use

We inform clients of relevant AI use in our tools, technologies, methodologies and automations, including internal and third-party solutions where this may affect the service, data handling, decision-making, contractual commitments or client risk. We explain how AI is used where relevant, including the potential benefits, limitations and risks to the client.

3 Documentation and auditability

We document our AI use, including how services are delivered, conclusions reached, and standards met. We document validation, quality assurance and review processes to support reliable outcomes. Our AI use is traceable and reviewable, with records retained to support proportionate internal or external assurance where appropriate.

4 Boundaries and control

We ensure suitably competent personnel retain oversight of AI-enabled activities, including autonomous or semi-autonomous activities. They review outputs, challenge decisions and intervene where needed. AI-enabled activities operate within defined organisational controls and, where relevant, client-agreed scope and boundaries. We use technical and procedural controls to prevent AI from being used outside its authorised purpose.

5 Data, sovereignty and client control

We inform clients how AI-enabled activities may use their data, including whether data may be used to train models and whether data may be transferred outside their organisation, our organisation, or agreed jurisdictions. We handle client data in line with agreed legal, regulatory and contractual requirements. Client data is used and stored only within agreed purposes, controls and commitments.

6 Security and confidentiality

We protect client data, prompts, outputs and AI-generated artefacts through appropriate technical and organisational controls. We are transparent with clients about how their data is secured where AI-enabled activities are used.

7 Secure development of AI tooling

We use secure development, integration and assurance practices for AI tooling. We review and maintain AI tools throughout their lifecycles to ensure they remain reliable and properly governed.

8 Supply chain assurance

We identify material third-party AI technologies and providers used in AI-enabled activities, and assess the associated security, compliance, resilience and operational risks. Where third-party AI use may materially affect service delivery, data handling, client commitments or continuity of service, we apply appropriate supplier governance and risk management controls. We are transparent with clients about relevant third-party AI dependencies where they may affect the service, contractual commitments or the handling of client data.

9 Resilience and business continuity

We identify material AI dependencies in service delivery and assess the impact if those systems fail or become unavailable. We maintain proportionate fallback or degraded operating arrangements where practical. We are transparent with clients about how AI disruption may affect service delivery, service levels, data handling, decision-making, reporting, continuity arrangements and recovery expectations.

Want to learn more?

We see a future where AI strengthens trust, confidence and professionalism across the cybersecurity ecosystem.

View our Principles