


IST's Efforts on the Nexus of AI and National Security

We work with a diverse range of stakeholders across the AI ecosystem to produce:


risk mitigation strategies

Ongoing Monitoring and Maintenance	Types of risks mitigated
<p>Reporting mechanisms</p> <p>Establish clear and widely accessible reporting mechanisms for individuals to report suspected fraud schemes or malicious use of AI technologies.</p> <p>→ Accessible reporting mechanisms help quickly identify and address misuse.</p>	
<p>Public campaigns</p> <p>Ensure that reporting mechanisms are well-publicized and easily discoverable through various media outlets, public awareness campaigns, and educational materials.</p> <p>→ Publicizing reporting mechanisms ensures that more people are aware of how to report.</p>	


useful tools and frameworks



recommendations



CBMs that involve agreeing to, or communicating an intent to, renounce or limit the use of AI technologies in certain weapon and military systems.

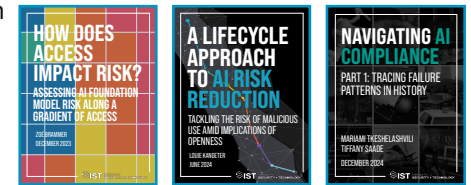


CBMs that encourage governments and industry players to agree on standards, guidelines, and norms related to AI trust and safety, as well as "responsible" use of AI technologies.

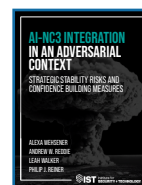
Implications of AI for national security and global stability: IST's efforts to understand the implications of AI began in 2017, with a roundtable featuring leading policymakers and developers on a workshop on AI's impact on national security issues. In 2018, IST undertook a joint initiative with Lawrence Livermore National Laboratory's Center for Global Security Research that aimed to articulate, understand and manage the long-term opportunities and risks posed by AI-related technologies for international security, global stability, and warfare.



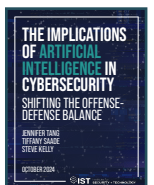
Accelerating multi-stakeholder coordination to mitigate the emerging risks of AI: As highly advanced artificial intelligence (AI) systems become increasingly integrated into critical aspects of society—from healthcare and finance to transportation and national security—policymakers and broader society are paying closer attention to the potential risks and opportunities associated with their development and deployment. With the support of the Patrick J. McGovern Foundation, IST engages with a diverse range of stakeholders across the AI ecosystem to better understand the emerging risks of AI foundation models and to develop technical and policy oriented risk reduction strategies, driving forward responsible innovation.



AI and ML integration into nuclear command, control and communications (NC3) systems: In January 2019, in collaboration with the Nautilus Institute for Security and Sustainability and Stanford University's Preventive Defense Project, IST hosted a multi-stakeholder discussion on the modernization of global NC3 systems. Building on this, IST convened scientists, engineers, policymakers, and academics to examine policy tools that could mitigate the risks posed by the integration of AI into NC3 systems. With the support of the State Department's Bureau of Arms Control, Verification, and Compliance and in partnership with Sandia National Laboratories, IST proposed confidence-building measures to limit the use of AI in weapon systems, encourage the creation of norms around the use of AI, increase lines of effective crisis communication, and bolster collaboration between private industry and government. As applications of AI tools continue to advance and its integration with NC3 systems and subsystems intensify, IST with the support of Longview Philanthropy has embarked on an initiative to raise awareness, foster dialogue, and establish frameworks for stable and predictable practices for AI-NC3 applications.



Implications of AI in cybersecurity: With the support of Google.org through its Digital Futures Project, IST is studying the applications of AI in cybersecurity and implications for the offense-defense balance. IST aims to provide a clear picture of current cybersecurity trends, cutting through marketing hype to offer a future outlook and actionable recommendations. This effort is part of a broader IST project to identify key cybersecurity areas needing focus, such as threat intelligence, automated defenses, and scalable security solutions. IST also co-leads a complementary effort with the World Economic Forum's Centre for Cybersecurity to understand the implications of AI in the cybercrime ecosystem.



Get in touch with us

WEBSITE
www.securityandtechnology.org

LINKEDIN
[@Institute for Security and Technology](https://www.linkedin.com/company/institute-for-security-and-technology)