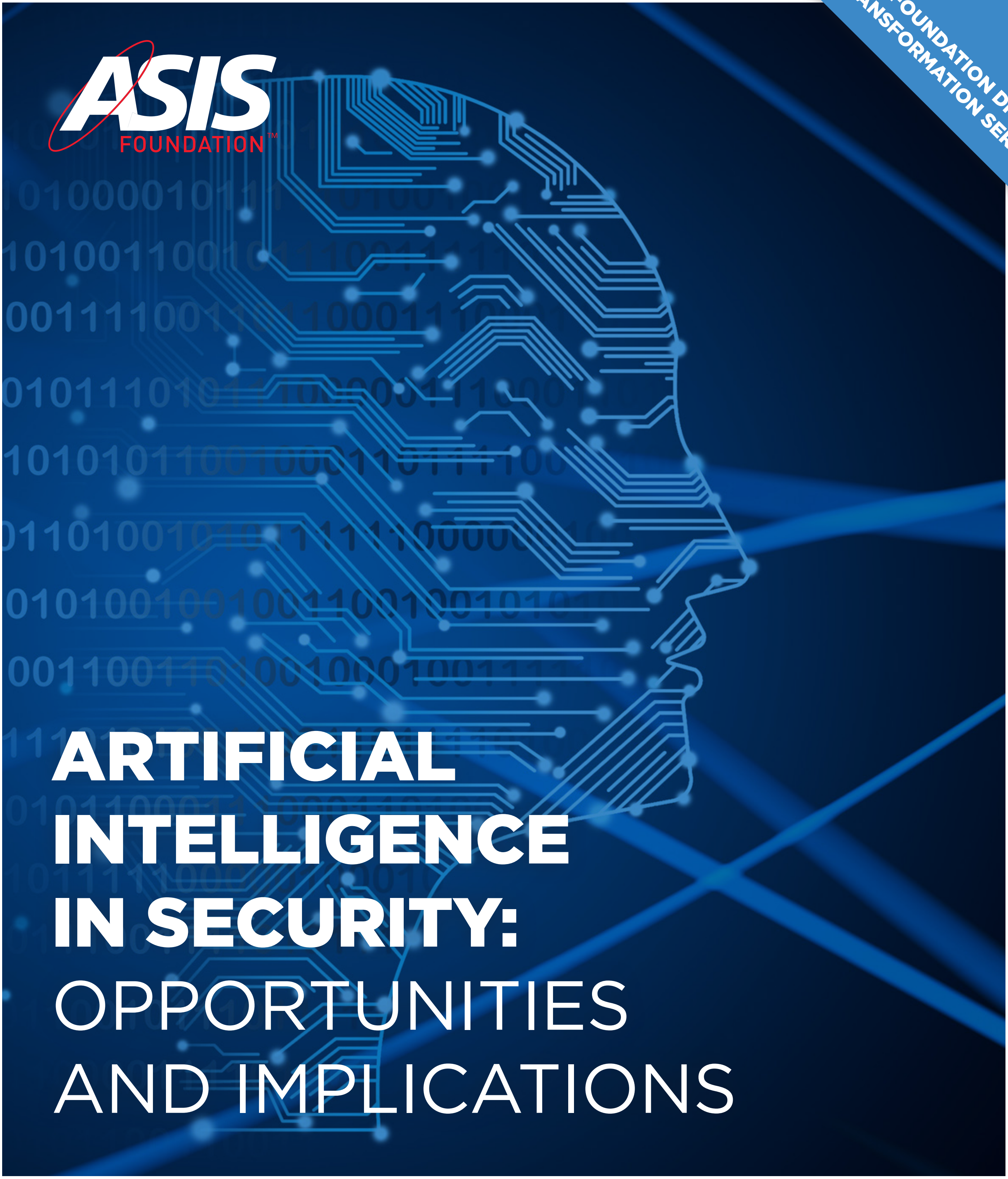
ASIS FOUNDATION DIGITAL TRANSFORMATION SERIES

ASIS FOUNDATION™



ARTIFICIAL INTELLIGENCE IN SECURITY: OPPORTUNITIES AND IMPLICATIONS

EXECUTIVE SUMMARY

Security as an occupational discipline has long used technological advancements to enhance its efforts in the protection of assets against malicious actions. In the contemporary era, such technological developments include advances in artificial intelligence (AI). However, many participants within the security sector describe AI beyond its technical capabilities, resulting in misunderstandings in functionality, opportunities for advancement, and knowledge gaps in risks associated with such development.

Types of AI Intelligence

Human intelligence has been a natural choice for benchmarking the evolution of AI, so AI capability is divided into the following broad categories:

- **Artificial Narrow Intelligence (ANI)**, also referred to as 'Weak AI' or 'Narrow AI,' is an approach that focuses on solving very specific tasks within the scope for which they have been designed. Narrow AI is very good at completing repetitive tasks and in many instances performs much better than humans. Examples include Siri, Google Translate and IBM's Watson.

- **Broad AI** is described as the integration of two or more narrow AI systems or techniques that make decisions to perform a task or process. Enterprises may use data specific to that business to train systems to address the specific business process, for example self-driving vehicles, analysis of investment strategies for corporate customers, or a software system supporting maintenance work on an oil rig.

- **Artificial General Intelligence (AGI)**, also referred to as 'Strong AI' or 'Deep AI,' is an approach that allows machines to perform intellectual tasks at the same level as humans. General AI is expected to possess theory of mind as well as being self-aware, able to understand belief, thoughts, emotions and expectations of people and able to interact socially. Like humans, general AI can reason, strategize, and make plans based on emotions and prior knowledge. Although general AI possesses self-awareness, it lacks emotion. Such advances are yet to be achieved in the current state of AI research and development.

- **Artificial Super Intelligence (ASI) –** approaches that hypothetically possesses ability and intelligence that far surpasses humans.

AI in Security Technology: Current State

Technology is used in physical security to achieve a degree of control over an environment through combinations of observing, detecting, controlling, and responding technological measures. These are integrated through various software, firmware, and hardware means at the field, automation, and management levels of built environment management technology architectures, underpinned by various computational techniques to achieved predefined outcomes.

Artificial intelligence techniques used in the protection of assets currently sit within the narrow or broad AI paradigms, with no evidence of general or artificial super intelligence in the protection of assets. In security technology, AI is focused on predefined outputs using computational techniques and executing rules as they relate to security interest characteristics or patterns aligned to environmental threat stimuli changes. Currently, there are a wide spectrum of AI paradigms and applications within these categories.

Artificial Intelligence (AI) is a subsection of computer science that investigates and develops computational approaches and techniques that enable machines to perform tasks that would normally require some level of human intelligence.

Current security technologies predominately use Narrow AI along the symbolic and statistical segments of the AI spectrum. For example, observation and detection technologies were found to be skewed towards threat or event diagnosis functionality in the protection of assets and use inputs (sensing) and computational techniques (sensing, perceiving, and knowledge) to achieve defined narrow AI outputs as alerts. In contrast, controlling and response technologies are skewed towards actions in controlling and responding to security events (knowledge and planning) using predefined rules, including broad AI outcomes to achieve security objectives. Furthermore, some security technologies such as drones and autonomous weapons systems produce what are perceived to be general AI outcomes, using combinations of narrow AI outcomes (or broad AI).

Machine learning in security technologies is at an elementary level, with evidence of it found in network video surveillance analytics; biometric system analysis and management; acoustic detection systems; and drone and robotics analysis and management. Currently these systems know only the data they have been provided and cannot yet interpret the 'unknown.'

AI in Security Technology: Future Developments

While AI in many security technologies may be considered narrow, or a series of narrow AI outputs (broad AI), and relatively unsophisticated, there are considerable opportunities for future developments. Currently those opportunities largely apply to security technologies in the statistical (probabilistic) AI paradigm, where complex decisions are not required to be made, but improvements in accuracy and reliability are highly desirable.

For significant and intelligent development to occur, AI must be able to interpret and contextualise dynamic environments, events, and situations, as well as understand and account for significant deviations or outliers from expected inputs, outputs, and norms. However, intelligent interpretation and contextualisation is currently beyond the capacity of machines and is likely to remain this way until the age of quantum computing arrives.

Opportunities will also emerge for more extensive applications of observation and response technologies, though development will likely be constrained by political, social, environmental, and legal factors. These factors are highly dynamic and therefore are likely to create a fluidity in how the development of AI transpires, with the adoption of AI fluctuating across global landscapes along with the perceived benefits and risks. At a more abstract level, opportunities for security technologies are likely to present where the risks of AI can be explicitly linked with lower consequences of AI failure.

In addition to potential benefits, there are profound risks of developing AI in security technologies, the consequences of which may not be fully comprehendible. The quest for technological advancement may create political divides, upset balances of power, encourage exploitation of underdeveloped nations, or promote the abuse of individual privacy and rights. Development of military and security response technologies with the capacity for autonomous use or release of force may eventually have the authority to determine life and death or inflict injury or harm onto humans. While this level of intelligent autonomy is not currently achievable in commercially available security technologies, the desire for military supremacy combined with the porous nature of military-commercial product exchange will likely see the autonomous use and release of force become a reality. The potential for harm to result from development and deployment of these technologies means there must be extensive legal, moral, ethical, and human rights considerations afforded to the discourse on intelligent autonomy, provided through enforceable international governance platforms.

However, as alarming as these risks may be, social factors may inhibit the deployment of AI in the commercial sector, particularly in countries and regions with individualistic cultures. These socio-technical environments will be where safe, legal, and ethical use of AI can be deployed in socially acceptable ways with public endorsement occurring because of transparency. The key findings and recommendations offered in this report may assist in guiding the transparent and socially acceptable development of AI in security technologies.

Key Recommendations for Artificial Intelligence in Security Technology

Key project recommendations for the use and oversight of AI in security technologies include:

- Development of a security industry artificial intelligence guidance document.
- Development of an artificial intelligence risk decision matrix for security managers to evaluate the benefits and risks of AI technologies for their environments.
- Development of jurisdictional legislative frameworks to afford protection to citizens and technology consumers, and to hold accountable those who breach the frameworks designed to protect individual's rights.
- An AI security vulnerability and criticality assessment should be developed to guide deployment of AI systems in cloud-based environments.
- World innovators and leaders must consider how assistance may be provided to at-risk nations to ensure technological disadvantage does not create a new era of hardship or enable misuse, abuse, or exploitation by external forces with AI developments.

Conclusion

The overwhelming consensus from participants within the project is that AI in security is at an elementary stage, with a limited capacity for inteligent decision making and autonomy. The view to date is that AI does not 'think,' rather it computes, processes, applies rules, and, in machine learning applications, may even generate rules based on data learning. But AI is fallible and inflexible, and subsequently operates in an environment of black and white. In contrast, humans think and create, and therefore can comprehend unusual environmental changes or disturbances with sufficient fluidity to analyse and contextualise human constructs such as intent and motivation. Such a depiction shapes our understanding of AI in security technologies, as security often requires context and understanding, not just computing and processing. Currently, AI does not have the capacity for human understanding—it cannot adapt as a human can and therefore cannot provide assurances under dynamic conditions. Security by its very nature is dynamic, therefore the implications of assurance deficits are profound.

Though the levels of intelligent autonomy for security technologies are unlikely to change considerably in the next 10 years, the age of quantum computing is likely to facilitate AI developments beyond any current expectations. Quantum computing will likely be the key driver allowing AI to reach the point of singularity and achieve the level of post-autonomy. Until such time, AI may produce reasonable economic benefits such as increased productivity and reduced costs from enhancement of narrow AI tasks. However, the overwhelming benefit to humanity in the developing future will be the use of response technologies such as drones and robotics to remove humans from harm.

Without governance structures, the socio-political, legal, and security risks of AI may eclipse any deliverable benefit. Safety, privacy, individual rights, and the potential impact on humanity should be fundamental considerations for the use of any AI, including those used in security technologies. The quest for technological and military supremacy may undermine those basic rights, and the consequences may be irrevocable and irreparable. Humanity must therefore produce a viable platform from which AI development can be managed in a socially desirable way for the benefit of all.

Research Methodology

This research consisted of three phases. Phase one employed a systematic literature review that provided usable articulation of artificial intelligence (AI) problem domains and spectrum of artificial intelligence paradigms. It focused the study on the technological category areas where technology is used in the protection of assets. Phase Two employed a focus group analysis using a purposive sample of AI and security technology experts to confirm indicative findings from Phase One, and to add insight as to where AI is located within the current building automation and control systems architecture across the range of observe, detect, control, and respond technologies.

Phase two provided the basis for phase three, the articulation of future opportunities for security technology developments using AI advances, and their associated risks, or risks associated with AI security technology advances in general. Phase three also developed *The Security Technology Intelligent Autonomy Scale* to contextualise the extent to which AI is currently able to be used by intelligent systems, and the degree of decision making and control that intelligent systems may possess during operation. It drew on the findings from phases one and two as a framework for establishing research-supported advances to the security technologies body of knowledge from an AI perspective.

About the Author

Chief investigator Dr. Michael Coole is a lecturer in the Edith Cowan University (ECU) School of Science in Australia and a member of the ECU Security Research Institute.

He holds a Doctor of Philosophy degree from Curtin University of Technology and a Master of Science (Security Science) degree from Edith Cowan University. His research team included Associate Professors Martin Masek and Peng Lam, researcher Jennifer Medbury, principal research assistant Deborah Evans, and research assistant Nicola Lockhart.

To Learn More

Download the complete 300-page report. Complimentary for members of ASIS International.

ASIS FOUNDATION™

About the ASIS Foundation

The ASIS Foundation, a 501(c)(3) nonprofit affiliate of ASIS International, supports global security professionals through research and education. The Foundation commissions actionable research to advance the security profession. It awards scholarships to help chapters and individuals—including those transitioning to careers in security management—achieve their professional and academic goals. Governed by a Board of Trustees, the Foundation is supported by generous donations from individuals, organizations, and ASIS chapters and communities worldwide. To learn more or make a donation, visit www.asisfoundation.org.