ASIS FOUNDATION DIGITAL TRANSFORMATION SERIES

ASIS FOUNDATION

ARTIFICIAL INTELLIGENCE IN SECURITY: OPPORTUNITIES AND IMPLICATIONS

FULL REPORT

Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments

Research Report: 2021

Authors:

Dr. Michael Coole

Associate Professor Peng Lam

Associate Professor Martin Masek

Mrs Deborah Evans

Mrs Nicola Lockhart

Mrs Jennifer Medbury

ASIS International
1625 Prince St.
Alexandria, VA 22311
www.asisonline.org

Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments

Artificial Intelligence Research Report

CONTENTS

Executive Summary

Security as an occupational discipline has long used technological advancements to enhance its efforts in the protection of assets against the malicious actions which pose a risk to objectives. In the contemporary era, such technological developments include advances in Artificial Intelligence (AI). However, many individuals 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.

The Project

The purpose of this project was three-fold. First to provide a thorough articulation of what is meant by the broadly used term 'Artificial Intelligence', including terminology, technical dissection, and progressive computational techniques along the spectrum of AI. Second, to gain understanding of how AI is embedded in current security technologies; and third, to provide articulation of the opportunities for technology enhancements using AI developments and the potential risks associated with such development and adoption within security technologies. The project applied a Three Phase qualitative study to support evidence-based findings and outcomes into the use of AI in the protection of assets.

Phase One employed a systematic literature review which provided a practical articulation of AI problem domains and spectrum of AI paradigms, and focused the study towards 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 (BACS) architecture across the range of Observe, Detect, Control and Respond technologies. Phase Two provided the basis for Phase Three, the articulation by experts of future opportunities for security technology developments using AI advances, their associated risks, as well as the broad risks likely to arise from the general development of AI in security technologies. Phase Three drew on the findings from Phases One and Two as a framework for establishing research supported progress in the security technologies body of knowledge from an AI perspective.

Project Findings

Phase One found that AI is characterised by the AI problem domains and spectrum of AI paradigms to perform computational techniques aligned to definitions of Narrow, Broad, General and Artificial Super Intelligence. Phase One also found that technology is used in physical security to achieve a degree of control over an environment, and such control is achieved through combinations of Observing, Detecting, Controlling and Responding technological measures. These technological measures 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 achieve predefined outcomes.

The outcomes of Phase One established that AI 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 the application of AI to security technology, AI is focused on predefined outputs using computational techniques and executing rules as they relate to security interest characteristics or patterns which are aligned to environmental threat stimuli changes.

Phase Two resulted in several key outcomes. 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 (Figure 20) 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 (Broad AI).

Phase Two also found that Machine Learning in Security Technologies is at an elementary level, with evidence of this found in Machine Learning within 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’ or apply context to gain human-level insight.

Phase Three sought to identify areas for potential advancements in security technologies using AI, and where evident, highlight risks in their uptake for security management. Phase Three developed *The Security Technology Intelligent Autonomy Scale* (Appendix E) 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. The scale facilitated expert identification of current levels of intelligent autonomy, as well as the potential levels of intelligent autonomy that security systems may be capable of reaching, from *Level 1: Manual* (an absence of AI) through to *Level 11: Post-autonomous* (General AI). Phase Three uncovered that the highest Mean rating of intelligent autonomy in security technologies was *Level 5: Automated* (Narrow AI), which suggests that despite the adoption of AI, security technologies have limited intelligent autonomy and are not currently capable of making complex decisions through the application of AI algorithms.

Phase Three found that 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 currently and is likely to remain this way until the age of quantum computing arrives.

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. Opportunities will also emerge for more extensive applications of observation and response technologies, though development will likely be constrained by *Political*, *Social* and *Legal* factors. Phase Three found 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.

Phase Three also uncovered that the development of AI in security technologies is not explicitly tied to the technical opportunities and capabilities of systems. There are *Political*, *Economic*, *Social*, *Environmental* and *Legal* (PESTEL) benefits and risks which may equally affect the development and deployment of AI. 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.

Phase Three uncovered that there are profound risks of developing AI in security technologies, the consequences of which may not currently be fully comprehendible. The quest for technological advancement may create political divides, upset balances of power, encourage exploitation of underdeveloped nations, and 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.

Summary of Findings

The overwhelming consensus from participants within the project is that AI in security is at an elementary stage, with a limited capacity for intelligent decision making and autonomy. The view to date is, 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, humans 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 ten 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.

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 Adoption framework.
- Development of an Artificial Intelligence Risk Decision Matrix for security mangers to evaluate the benefits and risks of AI technologies for their environment.
- An AI security vulnerability and criticality assessment should be developed to guide deployment of AI systems in cloud-based 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.
- 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.

Introduction

The ASIS Foundation serves the security industry by supporting the ongoing research and education of ASIS members around the world. Security as an occupational discipline has long used technological advancements to enhance its efforts in the protection of assets against the malicious actions which pose a risk to objectives. In the contemporary age, technological advancements include Artificial Intelligence (AI). However, to date automation across the security domain has been limited and narrow, with technology used in limited linked applications such as facial recognition, identification card analytics, and video analytics for alarm purposes. However, developments across the Artificial Intelligence space means there are opportunities for technology enhancements not yet recognised by security managers or technology companies, resulting in gaps of opportunities that could see security more effective, more efficient, safer and more integrated than current technological solutions.

Consequently, a thorough articulation of the current use of security technology along with artificial intelligence capability may provide for the articulation for opportunities for automation enhancement, increasing security effectiveness, capacity, efficiency, and safety.

Background

The formal recognition of security as a vital organisational undertaking is credited to Fayol (1916), who emphasised the existence of six activities associated with organisational management, of which security was one. Such recognition places security as a central pillar to organisational success. However, security's necessity extends beyond the corporate remit, with Smith & Robinson (1999) articulating that security as an industry has become not only an integral part of modern business, but also broader human activities capturing areas of social congregation for leisure or living, with the functional aspects of security necessary to ensure protection and safety of individuals and assets accordant with Hobbes social contract emphasis (Hobbes, 2018). This contextual sphere of activity places industrial and community level security undertakings within the non-traditional domain of security, that is within the criminological sphere.

Within this sphere Fisher, Halibozek and Green (2008, p.8) state that security implies a stable, relatively predictable environment in which individuals or groups may pursue their ends without disruption or harm, and without fear of disturbance. An exploration of security literature presents the theme that such a stable or predictable environment is achieved through means of control, real and perceived (Garland, 2001; Hobbes, 2018; Meerts, 2019). That is, such a stable relatively predictable environment is achieved through various active and passive process or mechanisms of control, where Coole argued that functionally, security relates to the degree of control or influence over an environment to achieve protection (Coole, Brooks, & Treagust, 2015). Consequently, across the organisational-corporate and societal spheres, such control is often achieved via a security or crime prevention function (Fischer et al., 2008).

Countless protection functions across the organisational-corporate or society spheres are achieved using physical security measures. Physical security is defined by ASIS International as 'that part of a security program concerned with physical measures designed to safeguard against a security incident using measures as a device, system, or practice of a tangible nature to protect people and prevent damage to, loss of, or unauthorized access to protected assets' (ASIS International, 2009, p. 3). According to ASIS International, the physical security realm traditionally consists of four overarching functions of physical security measures, including: Observing, Detecting, Controlling and Responding (ASIS International, 2015, p. 17). These functions are underpinned by the role elements of Detect, Delay and Deny (p. 17).

The achievement of many Physical Security functions is typically facilitated using electronic systems or devices (technologies) to enhance the protection of assets through means of automation. Automation is defined as "*the execution by a machine agent (usually a computer) of a function that was previously* *carried out by a human*" (Parasurman & Riley, 1997, p.231). Consequently, the functions of physical security are achieved using people, barriers and technology, with a strong focus toward technology in the contemporary security environment.

Linked to the definition and functions of physical security are the occupational roles of security practitioners which are of significance for the study. Smith and Robinson (1999) categorised the occupational functions into four defined practice areas of specific foci. These included Security Managers comprised of those who have the responsibility to protect, Security Technologist, comprised of individuals such as scientist or engineers who design and develop the physical security systems to protect, Security Technicians, as those who install and maintain physical security technical systems, and Security Guards, as the operational personnel or workers who have direct functions to protect. Smith & Robinson (1999) noted that each category of personnel require and hold different education, training and knowledge associated with their occupational context; nevertheless, arguably all must understand the functions of physical security to achieve their ends.

The introduction of automated electronic systems as a supplement to manpower and physical mechanisms traces back to Edwin Holmes, who offered the first use of automation in the pursuit of private security in 1858, with the first burglar alarm system developed by Augustus Pope. In 1874, American District Telegraph was founded, also offering alarm systems and security response (Fisher et al., 2008). Consequently, the historical use of technology for the physical protection of assets has resulted in contemporary security approaches relying on advances in technology to achieve desired states of awareness, influence or control. These advances such as enhanced connectivity, Big Data and the Internet of Things (IoT) often rely on the integration of many individual complex technologies, with the outcomes of such integration resulting in the achievement of protective security objectives. As Fennelly's (2012) work highlighted:

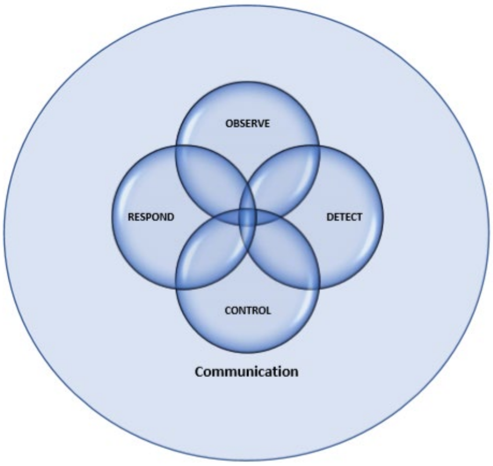
> "*protection through physical security has become increasingly complex, where often many complex, interrelated controls or systems are integrated to protect in a given security context and the function of Integration can be considered in its own right."*

Furthermore, much of the early focus on security technology was towards security at the facility or corporate level, within an often clearly bounded environment and defined threat thesis. In contrast, security in contemporary society is also prevalent at the broader community level, with community security sitting within the crime prevention lens of protection occurring in a broader, less bounded environment. Here the threat is less defined and more abstract, there exists less direct environmental and adversarial control, particularly as the adversary's intent and capability is not necessarily known.

Despite such a broad security purview, ASIS's overarching functions of physical security technological measures - *Observing, Detecting, Controlling and Responding* – are functionally best understood as a matrix of technological components (Figure 1). These components are integrated through technological communications networks, which provide limitless combination opportunities for exerting systemised control over a facility, space or urban environment. As Sun Tzu (1971, p. 64) pointed out, there are not more than five primary colours, yet in combination they produce more hues than can ever be seen; there are not more than five cardinal tastes, yet in combination they yield more flavours than can ever be tasted. Sun Tzu's mixing of principles analogy applies to physical security's technology functions for a facility or broader community well-being; where in combination, Observing, Detecting, Controlling and Responding technologies along with Communications technologies provide an endless series of integrated physical security functionality for the protection of assets, people, or information.

Figure 1

ASIS International's published overarching functions of physical security technology measures.



Contemporary technological advances have seen the integration of Physical Security technological measures onto what are perceived seamless system-based platforms, previously understood as automation; defined as "*the execution by a machine agent (usually a computer) of a function that was previously carried out by a human*" (Parasurman & Riley, 1997, p.231). Such automation is increasingly being captured under an umbrella term of Artificial Intelligence (AI), which more correctly is used to enable more complex automation. However, the term artificial intelligence has many variations in its definitions, with it often a misused term when considered against more precise definitions based on human cognition replication and computer science literature. Yet, to date there is a dearth of literature clearly aligning security technologies and their functions, features and technical processes within the AI problem domain and spectrum of AI paradigms.

Consequently, clearer articulation is required to facilitate a more precise understanding of where individual and integrated security technologies lie on the spectrum of AI paradigms, how the adoption of new technologies will facilitate enhanced decision making or increase efficacy in the protection of assets through its adoption. Such precise articulation will enable security practitioners to consider future directions for security technology based on accurate language aligned with AI experts. As such, it has become necessary to clearly articulate to security professionals how current and future advances in security technologies differ from the broad umbrella term Artificial Intelligence and align more precisely with the problem domain and spectrum of AI paradigms.

Artificial Intelligence has become an imprecise umbrella term - as such, it has become vital to articulate the AI problem domains and spectrum of AI paradigms for security practitioners.

Research Methodology and Objectives

Research Aim and Objectives

Prior to the undertaking of this project, limited research has investigated the current use of technology in the protection of assets and where opportunities exist for corporate entities or technology development industries to fuse Artificial Intelligence (AI) with the roles, functions and tasks of corporate security.

Consequently, the overarching aim of the project was;

> *To uncover opportunities and practical implications for corporate security managers within the context of security technology enhancement using Artificial Intelligence.*

The benefits of this include directing future technological efforts in intelligent integration, automation and autonomy; providing direction for necessary commercial research and development; increase corporate security efficiencies; reduce risks to personnel; and more effectively manage organisational risk through the aid of enhanced intelligence.

In alignment with the overarching research aim, there were two broad objectives of this research project;

a) *To understand the current state of Artificial Intelligence in security technologies*, and;

b) To investigate future developments for Artificial Intelligence in security technologies, and their associated risks.

Research Questions

The study Research Questions were aligned to the project aim and objectives. Phases One and Two were designed to develop an understanding of the current state of AI in security technologies, and Phase Three built upon this articulation by investigating the opportunities and risks which may be presented through the enhancement of AI in security technologies.

The research questions for each Phase of the research are as follows;

Phase 1

- How is control (real and perceived) currently achieved over physical environments using technology?
- What constitutes the Artificial Intelligence problem domains and spectrum of paradigms?
- How is artificial intelligence indicatively used in physical security?

Phase 2

- How do the capabilities of artificial intelligence techniques map to activities that can achieve control over a physical environment?

Phase 3

- What are the priority security technology capability areas in the protection of assets for enhancement through artificial intelligence?
- What vulnerabilities exist in the artificial intelligence techniques that may be used to achieve control in a physical environment?

Methodology

The study employed a Three Phase qualitative design, with each stage aligning to the research aims and objectives and building upon previous phase outcomes. The Phase One literature review was designed to provide an understanding of the current state of Artificial Intelligence (AI) in security technologies. Phase Two reflected on the literature review, drawing upon expert knowledge to clarify and confirm indicative findings of Phase One through an Expert Focus Group. Phase Three again built upon previous findings to explore the opportunities and risks of AI through in-depth interviews with international security and AI experts.

Phase One – Literature Review

Phase One employed a systematic literature review (Jesson, Matheson & Lacey, 2011), identifying existing literature with regard to the roles, functions, and tasks of security technologies in the protection of assets, along with the technical articulation of the AI problem domains and spectrum of artificial intelligence paradigms. Clarifying the current state of security technology roles, functions, tasks and analytical methodologies was necessary to develop an understanding of how artificial intelligence enhances security protection both now and in the future. Therefore, the systematic literature review in Phase One developed an overarching conceptualisation and articulation of security technology in the protection of assets.

The systematic literature review also identified existing themes within the artificial intelligence literature. This aimed to uncover the framework and terminology used across the AI domain for later alignment with security technologies in Phase Two, in addition to exploring the literature defined risks and limitations required for Phase Three. The outcome of the Phase One literature review was articulation of the current state of AI in security technologies, and how security technologies currently use AI analytical methodologies in the protection of assets.

Analysis Phase One

Phase One data was analysed using a discourse content and thematic analysis technique (Braun & Clarke, 2006; Boyatzis, 1998), to develop the presented frameworks of security technology use as described in the security literature. For the technical security literature included in the systematic review, thematic categories were assigned by technology function, i.e.; Observing, Detecting, Controlling and Response technologies. Qualitative themes for risks, vulnerabilities and limitations were generated with an inductive approach to data analysis, deriving overarching themes from identified initial themes (Braun & Clarke, 2006; Boyatzis, 1998; Creswell, 2009). The Phase One analysis also facilitated the identification and articulation of AI problem domains and spectrum of AI paradigms as a priori benchmark for Phase Two, through indicative mapping of the alignment between security technologies and, AI computational techniques and the spectrum of AI paradigms.

Phase Two – Expert Focus Group

Phase Two conducted an expert focus group to link the roles, functions and tasks of security technology uncovered in Phase One to the current state of artificial intelligence. Phase Two sought to determine how AI computational techniques map to activities that can achieve control over a physical environment using security technologies. Security and AI expert participants were consulted to confirm and correct the alignment of security technologies with the AI computational techniques and spectrum of paradigms, and to produce novel visual maps representing the architectural locations of both AI and security technology components.

Focus Group Design

As a type of group interview, a focus group brings together the selected participants for in-depth discussion of a particular theme or topic. The interaction of the participants enables the participants' views to emerge and predominate, resulting in the focus group outcome being a collective of the group rather than being individual or researcher driven (Cohen et al, 2007, p.376). To support the philosophy of a collective outcome, the Focus Group followed a semi-structured design with three overarching objectives and research questions (see 2.3 Focus Group Objectives). The focus group was facilitated to ensure the session was open and allowed participants to direct the discussion, whilst ensuring the Focus Group objectives were met.

Focus Group Population and Sample

Purposive sampling is recommended where a depth of information is required, and particularly where random sampling techniques would not provide participants with suitable knowledge of the subject matter (Cohen et al, 2007, p.115). Subsequently the Focus Group used non-probability purposive sampling to access expert participants from the populations of security and artificial intelligence professionals. Criteria for inclusion as an expert participant included a combination of security or AI qualifications, professional experience, academic and research experience.

Focus groups requiring high levels of engagement with the focus group topic generally require a smaller cohort of participants, with the recommended number of participants for these groups at approximately six to eight (Babbie, 2016, p.314, Morgan, 2019, p.53). Subsequently the sample size consisted of eight participants, and was heterogenous, comprising five security and three Artificial Intelligence (AI) experts (Morgan, 2019).

Analysis Phase Two

Following the Focus Group on the 5th of October 2020, two debriefing sessions were conducted on the 6th and 12th October 2020 for review and preliminary analysis of Phase Two visual diagrammatic and audio data. Visual data was then recorded, scanned and documented for further qualitative analysis and to produce novel visual maps representing the architectural locations of both AI and security technology components. Audio data was transcribed and used to generate content codes to confirm and correct the alignment of security technologies with the AI computational techniques and spectrum of paradigms (Braun & Clarke, 2006; Boyatzis, 1998). Through qualitative analysis, data was transformed and developed into *The Security Technology-Artificial Intelligence Linguistic Cycle*, facilitating the articulation and communication of the alignment between how security technology operates through sensing, processing, deciding and acting, and the AI cycle of data input, computational technique, rule assessment and defined output. Articulation of the sequence of operational stages for security technologies provided clarity as to where and how AI techniques may be applied in preparation for Phase Three.

Phase Three – Expert Interviews

Phase Three undertook expert interviews with researchers, professionals and technical experts within the fields of artificial intelligence and security technologies (Cohen et al, 2007). Expert interviews firstly generated an understanding of the current level of intelligent autonomy achieved by security technologies to assist in facilitating the identification of opportunities for AI in security technologies within and across a defined scale. Consequently, a research instrument was developed for this purpose – *The Security Technology Intelligent Autonomy Scale* – providing a quantitative measure of AI in

Observing, Detecting, Controlling, Response and Integrated Security Technologies. Expert interviews also used the intelligent autonomy scale to link opportunities for enhancement to the various stages in *The Security Technology-Artificial Intelligence Linguistic Cycle* developed from Phase Two, i.e. Input/Sensing, Computational Technique/Processing, Rules/Deciding, and Output/Acting. In addition to identifying the opportunities for enhancement, expert interviews also explored the benefits and risks of AI for security management.

Expert Interview Design

In compliance with national and state Covid-19 restrictions and recommendations, expert interviews were conducted online via Microsoft Teams. The interview format followed a structured design with interview questions being provided to participants prior to the interview (see Appendix M). While the structure and sequence of interview questions were pre-determined, open-ended questions provided flexibility for participants to explore salient topics throughout the duration of the interview. This method resulted in all interviews surpassing the allocated time of one hour, with the duration of most interviews exceeding two hours, however this allowed for high levels of engagement with each participant and subsequently increased depth of data (Cohen et al, 2007).

Population and Sample

Non-probability purposive sampling was again used for Phase Three due to the requirement for expert knowledge in security and AI (Cohen et al, 2007, p.115). Expert criteria for inclusion included a combination of security or AI qualifications, professional experience, academic and research experience, or referrals from identified experts. A snowballing technique was used, relying on peer recommendations for inclusion of additional participants until the interviews reached saturation (Cohen et al, 2007). In total, eleven participants were interviewed, comprising both security and AI experts, from Australia and international locations.

Analysis Phase Three

Expert interviews with participants were recorded and transcribed for quantitative, thematic and content analysis. Quantitative analysis used descriptive statistics to identify the level of intelligent autonomy for Observing, Detecting, Controlling, Response and Integrated Security Technologies. An inductive approach to data analysis was used to identify and extract qualitative themes, with overarching themes and opportunities for enhancements being generated from analysis of initial themes (Braun & Clarke, 2006; Boyatzis, 1998; Cohen et al, 2007).

The structured design of the interviews allowed for analysis of numerical ordinal level data and qualitative themes to locate the opportunities for enhancement of AI within each category of Observing, Detecting, Controlling, Response and Integrated Security Technologies. A PESTEL framework was used for thematic analysis of AI benefits and risks identified by participants to generate overarching themes associated with AI use in security technologies (Business-to-you, n.d.; Akman, 2020; Green, 2018).

PHASE ONE: LITERATURE REVIEW

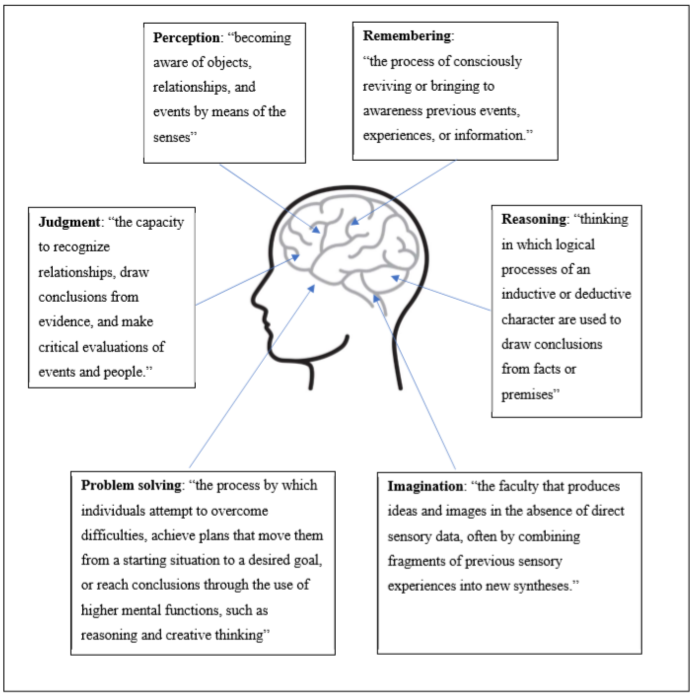
1.1 Cognition & Intelligence

Our understanding of the term 'Artificial Intelligence' originates from our analogate understanding of the human mind from cognitive psychology. Cognitive psychology is the scientific investigation of internal human processes involved in making sense of the environment and deciding what actions might be appropriate. The American Psychological Association (2020) defines cognition as "all forms of knowing and awareness, such as perceiving, conceiving, remembering, reasoning, judging, imagining, and problem solving. Along with affect [experiencing emotion] and conation [the proactive part of motivation], it is one of the three traditionally identified components of mind". Figure 2 outlines the different elements of the definition of cognition.

Pennington, Flanagan and Banyard (2000, p.1) argue that "one of the defining features that sets human beings apart from other animals is not only our ability to think, but also our ability to be aware of what we are thinking." Nonetheless, significant debate exists around exactly how the human brain processes information (Benton, 1980; Ashman & Conway, 1997; Ranti, Chatham & Badre, 2015). A key element of this debate is the relative importance of specific types of mental processes such as parallel processing, serial processing, top-down processing and bottom-up processing (Sloman, 1996; McClelland & Rogers, 2003; Dotan & Dehaene, 2020; Sackur & Dehaene, 2009). Top-down processing is also known as knowledge-based processing and involves using contextual or pre-existing knowledge to understand information while bottom-up processing, or text-based processing, involves decoding data and/or information (Tsui & Fullilove, 1998; Benton 1980). Disagreement on the importance of such processing is evident throughout the literature, for example, Tsui and Fullilove (1998) discuss significant and ongoing debate regarding the importance of top-down processing as compared with bottom-up processing for listening and reading comprehension.

Figure 2

Elements of Cognition (American Psychological Association, 2020)



Understanding how the human brain processes information is particularly complicated as there is a substantial body of knowledge that demonstrates significant differences in how people observe, encode, store, transfer, decode and understand stimuli, data and information (Baron-Cohen, 2017; Björklund & Bäckström, 2008; Miyake & Friedman, 2012). For example, Kanai and Rees (2011, p.235) assert there are individual differences in metacognitive ability – "the ability to appropriately link insight (confidence) to objective performance in a perceptual decision-making task" – which has a direct impact on judgments and decision-making. The concept of individual differences is also a feature in the growing body of knowledge on neurodiversity, which is defined as "the idea that people having a range of different types of brain, including those with and without autism, should be regarded as normal variations within the human population" (Cambridge Dictionary, 2020).

Although the 'mind as a computer' analogy has been widely used, this idea may oversimplify the complexity of how the human mind works (Sackur & Dehaene, 2009). For instance, computers are often described as logical, rules-based and emotionless. In contrast, the human brain can be impacted by emotions and cognitive biases but also uses short-cuts and heuristics, as well as logic and reasoning, when conducting problem-solving activities and making decisions (Howard, 1993; Andrade & Ariely, 2009; Maitland & Sammartino, 2014; Dietrich, 2010). For example, the concept of facial recognition demonstrates the differences between how the human mind and a computer processes information. Facial recognition software is based on comparing specific facial features (in terms of their size and proximity to other features) to images stored in the program's database (Sample, 2019). In contrast, research indicates there are individual differences in how people recognise faces and that the human brain uses different processes for identifying familiar faces (an almost instantaneous process) than when they are trying to determine whether multiple images are the same or different people (sequential processing) (Baron, 1981; Benton, 1980; Sandford & Burton, 2014). Consequently, there are differences in how humans and computers process information, solve problems and act, but analogies drawn between human cognition and computer processing may amplify misunderstandings of what is termed 'artificial intelligence'.

Despite the differences between the human mind and computing systems, a key aspect of artificial intelligence is the facilitation of automated processes to enhance decision-making. These automated decision-making processes exist in sectors such as communication, health, finance and law enforcement and may vary considerably between sectors (Araujo et al., 2020; Kitchin, 2017; Carlson, 2018). However, despite the diverse needs of each sector, and in order to make effective decisions, computing systems will need to replicate, or at least approximate, different elements of human cognition including perception, remembering, reasoning, judgment and problem-solving.

Artificial Intelligence 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.

1.2 Artificial Intelligence

Understanding the processes of the human mind are important to understanding current and future developments in Artificial Intelligence (AI). As Johnson (1986) explains, the notion of AI assumes that the mind can be described as some kind of formal system manipulating symbols that stand for things in the world. Consequently, AI is sometimes defined as the study of how to build and/or program computers to enable them to replicate the human mind. Current definitions add to the idea that observable performance in computer processing has features which would be attributable to mental processes in humans (Johnson, 1986).

Notwithstanding such descriptions, there is currently no single, precise or universally accepted definition for Artificial Intelligence (AI), as *intelligence* is a complex phenomenon that exists on a multi-dimensional spectrum with many facets, particularly in respect to the heterogeneity of the human mind. Thus, the attribution of 'Human-Level Intelligence' to a non-human system is fundamentally challenging (Goertzel & Pennachin, 2007, pp.VI-VII). Despite the absence of a universally accepted definition, AI in security technology may be understood as:

> *A branch of computer science that investigates and develops computational approaches and techniques that allow machines to perform tasks that would normally require some level of human intelligence. In other words, making machines intelligent.*

Regardless of variations in AI definitions, human intelligence has been a natural choice for benchmarking the evolution of AI and as such, AI definitions and descriptors have been associated with a machine's capability for mimicking human intelligence (Minsky, 1961; McCarthy, 2007). AI capability, relative to human intelligence, may therefore be divided into the following broad categories:

- Artificial Narrow Intelligence (ANI), also referred to as 'Weak AI' or 'Narrow AI' – approaches that focus 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 (Goertzel & Pennachin, 2007; O'Carroll, 2020).
- 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 (IBM Services, 2018, p.7). 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 in a banking system, or a software system supporting maintenance work on an oil rig.
- Artificial General Intelligence (AGI), also referred to as 'Strong AI' or 'Deep AI' – approaches that allow machines to perform intellectual tasks at the same level as a human. 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, be able to 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 (Goertzel & Pennachin, 2007; Maruyama, 2020, pp.242-251; O'Carroll, 2020).
- Artificial Super Intelligence (ASI) – approaches that hypothetically possesses ability and intelligence that far surpasses humans (O'Carroll, 2020; Pueyo, 2018).

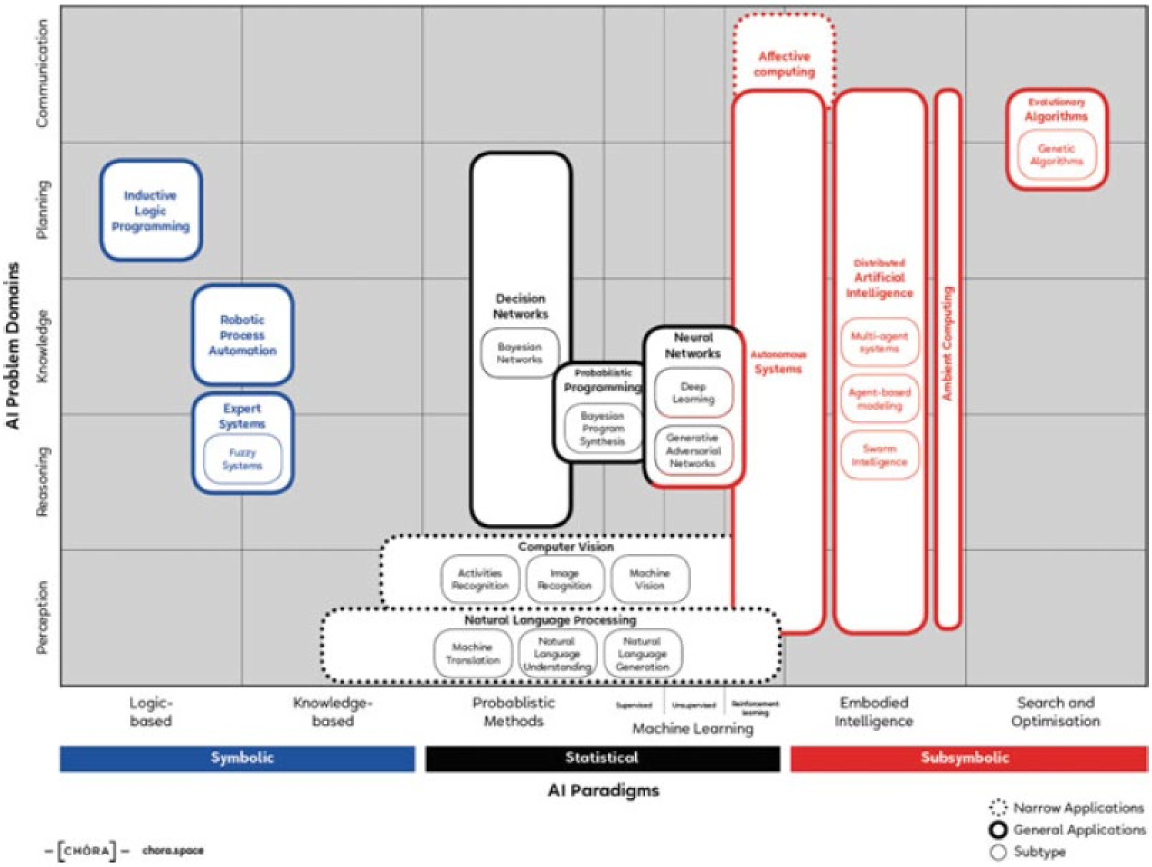
Drawing on these explanations, the current state of AI in security technology sits within the Narrow AI category (Maruyama, 2020, p.242; Pueyo, 2018, p.1731). Currently, there are a wide spectrum of AI paradigms and applications within this category. Corea (2019), presents a novel visual map of the current AI paradigms (Figure 3), depicting this wide spectrum of AI paradigms and their applications.

Corea's (2019) mapping represents an abstract range of problem domains on the y-axis, divided into the five categories of analogic human cognitive processes including: perception, reasoning, knowledge, planning and communication. AI paradigms used to address sub-sets of these problem domains are presented on the x-axis. While the AI paradigms presented in

Figure 3. are in a defined order, this does not imply any natural order or ranking of the paradigms in terms of superiority. They are simply different approaches to achieving some form of AI, each having unique benefits and challenges, noting some may be better suited to specific problems than others.

Figure 3

AI Knowledge Map (Corea, 2019).



In terms of functionality, an AI system is just like other system typologies, which receives a defined input, and based on that input and the functionality within the system produces a defined output (Ahmad, 2019; Brooks et al, 2017). Consequently, input/output considerations are necessary for understanding AI in the protection of assets.

The choice of input and output are governed by the application domain that the AI system will operate in (Kasabov, 2019). The inputs consist of data coming from various sensors and other systems that the AI operates in. The types of input can range from simple Boolean variables (e.g. whether a sensor is on or off), data on a continuous scale (e.g. output from a temperature sensor), multidimensional input (e.g. an image from a camera sensor consisting of a set of pixels, each having a red, green and blue component), and higher-level structured data (e.g. a record from a database). Output from an AI system can consist of the same type of data as input. Based on the defined inputs and outputs, an AI system can be integrated into a larger system without having to know the internal details of the AI. For example, a standalone narrow-AI voice recognition system (input is a sound sample, output is identity of speaker), a standalone narrow-AI face recognition system (input is an image, output is an identity of the person in the image), and a keyboard input sensor (user enters pin number) can have their output serve as the input to another AI system whose output may be to control access to a building. Over time, the AI technology may change (for example changing the face recognition system from a traditional computer vision system to a deep neural network), however if the input and output remain the same, the overall system can function (Kasabov, 2019).

Though the internal working details of an AI system do not necessarily need to be understood by a system integrator, some understanding is useful in order to keep in mind the benefits and limitations of each approach. Determining the structure and function internal to the AI system involves two considerations, deciding on the type of AI architecture to employ (as a high-level example, will it be a rule-based system, or a neural network, or something else), and second is how to configure and optimise the chosen system for the chosen task (for example, either manually, or automatically through a variety of learning or optimisation algorithms).

The breadth of articles discussing different types of AI can be confusing, mostly due to various authors mixing architectural paradigms and configuration paradigms in their explanations (for example attempting to compare a neural network, which is an architectural paradigm and evolutionary computation, which is a configuration/optimisation paradigm – indeed, evolutionary computation can be used to train a neural network). Consequently, to understand AI in the protection of assets it is necessary to understand the common AI architectural paradigms and common configuration/optimisation approaches used for each paradigm in understanding AI in the protection of assets.

1.2.1 AI Paradigms

1.2.1.1 Symbolic Paradigms

Symbolic paradigms from Corea's (2019) mapping (Figure 3), correspond to various AI methods that use symbols, which may be represented in rules, as their basis of manipulation. The internal working of such an AI process receives its input and based on this input satisfying a set of pre-defined rules produces an output. Rule-based systems were amongst the first AI systems, and under the name of Expert Systems promised to give computers the abilities of human experts. Feigenbaum (1982) has defined expert systems as intelligent computer programs which use knowledge and inference reasoning to solve complex problems normally requiring human expertise. The knowledge encoded in expert systems consist of facts and heuristics, commonly provided by human experts from relevant domains.

Early examples of such systems include MYCIN (Shortcliffe & Buchannan, 1975) for diagnosing infectious diseases. MYCIN's input were answers to a set of simple yes/no questions, used in conjunction with approximately 600 rules to produce a required output – a list of possible bacteria, a confidence level, a list of the rules that affected its diagnosis (to show system reasoning) and a recommended antibiotic and dosage. In a security context, alarm systems provide an example of symbolic operation. Where the input to the AI is a single *motion sensor*, the '*system activate*' switch is a single output is an audible *alarm* output is the rule, such a rule could take the form of: "**If** the *system activate* switch is on **and** the *motion sensor* senses motion **then** turn the *alarm* on." (Dutta & Bonissone, 1993; Kasabov, 2019). The actual syntax for writing a rule varies based on system implementation, but the common elements are:

1. Each rule consisting of a set of conditions (Inputs) and a set of resulting actions (Outputs) that occur if the conditions are met. The separator between conditions and actions in the example above is the '**then**' keyword (Dutta & Bonissone, 1993; Kasabov, 2019).
2. Checking or setting the state or value of the system inputs/outputs/variables (e.g. that the switch is on, or setting the alarm to on) – these usually involve testing for equality, greater than or less

than for Boolean logic-based (Boole, 1847), or the degree to which conditions are true for fuzzy logic-based systems (Zadeh, 1965).

3. The use of logic operators (and, or, not, etc.) to combine the conditions for a rule.

Although the example given above is trivial, and could be implemented just by wiring the sensor directly to the alarm without using any formal rule-based paradigm, for systems with a large array of sensors and actuators (alarms and other means of output), formal structures for the many resulting rules are necessary in order to manage total system complexity. The simplest structure, employed in traditional rule-based systems, is to order the set of rules by priority in a list. In order to produce the output from such a system, the rules are checked in order, and the first rule whose condition are all satisfied 'initiates', with its associated actions triggered, with all other rules in the set of rules ignored.

For complex systems where different rules apply in different circumstances, the rule-based system can be structured using graph-based formalisms – such as decision trees, behaviour trees or finite state machines (Zhang, Yao & Zha, 2018). Consequently, the design and selection of an appropriate set of rules to suit the intended purpose of a rule-based AI can range from manual (determined by humans) to fully automatic. In early expert systems, the focus was placed on interviewing and examining the work of human experts to manually construct a set of rules that captured their expertise (Tolun, Sahin & Oztoprak, 2016). Once an appropriate set of rules is captured, the execution of these rules can then be automated by machine implementation of the rule-based system.

Automated alternatives to manual rule design typically use either historical, or exemplar data to learn or derive the rules, where each data point consists of the set of inputs and the 'correct' output that the AI system should generate. Various algorithms then exist which take an initial set of rules and refine and optimise them towards the goal of the AI output matching the 'correct' output when presented with input from the exemplar dataset. For example, Learning Classifier Systems (LCS) use an evolutionary algorithm approach to evolve either individual rules or entire sets of rules through genetic operators such as mutation and crossover and the principle of survival of the fittest (Urbancowicz & Moore, 2009).

Corea (2019) listed logic-based and knowledge-based approaches under the LCS category, described as:

Logic-based AI approaches: comprised of representing knowledge of an agent's world, its goals and the current state via logical statements. By using inference or deduction involving these logical statements, an appropriate decision to achieve specific goals is obtained. Areas of logic-based approaches include knowledge representation, various form of reasoning (nonmonotonic, abductive and inductive) and computational logic.

Knowledge-based approaches: consisting of two main components, namely a knowledge-based component and an inference engine that acts as the control component for inferring new knowledge/decisions. The knowledge base contains information about the state of the world, and this could be represented in different forms: declarative; procedural; heuristic; structural or meta knowledge involving ontologies and huge databases. The second component, the inference engine includes techniques such a rule-based, model-based and case-based reasoning for inferring new knowledge/decisions.

1.2.1.2 Statistical Paradigms

The paradigms labelled by Corea (2019) (Figure 3) as statistical typically apply a series of mathematical operators to their inputs to produce an output. For example, a traditional computer vision AI uses the set of pixels belonging to an image as input and applies a set of operations based on the spatial location and colour of pixels to group sets of pixels belonging to distinct objects, and use measures taken from these groups to determine what the objects are (Voulodimos, Doulamis, Doulais & Protopapadakis, 2018). Another example of such AI is a neural network, taking a set number of numeric inputs and having the ability to transform them based on their value and order into an output by taking weighted sums of input subsets. This process of "summing" is typically performed numerous times, with the network arranged in layers, where the output of one layer, transformed using a set function, serves as the input into the next layer. Neural networks with 'many' layers are known as deep neural networks (Teschler, 2018; Zsigmond, 2020).

However, the design and configuration of neural network AI output to be effective at their given task typically cannot be done manually. For example, tuning a neural network to a particular task involves selecting the numeric weights applied to the input of each layer (Zsigmond, 2020). For most non-trivial tasks, it is impossible for a human to map the intended functionality of the AI to this large set of numbers. For this reason, determining the reasoning that architectures such as neural networks use to determine an output is challenging. Therefore, instead of manual configuration, training schemes are used, where during training the error - the difference between the output of the AI and what the actual output should have been - is used to modify the configuration of the system (Looney, 1997; Bataineh & Marler, 2017). Traditionally, the backpropagation algorithm (Rumelhart et al. 1986) is used in neural networks, where the error is fed back though the network and used to modify the weights in the direction that decrease the error. Neuroevolution is another alternative, using an evolutionary algorithm to evolve the neural network weights to minimise the output error (Heidenreich, 2019; Stanley & Miikkulainen, 2002). Corea (2019) identifies probabilistic approaches and machine learning techniques as being under this category and are described as follows:

Probabilistic approaches employ probabilistic representations, that capture uncertainty in complex relationships and knowledge of the world, in the form of probabilistic graphical models. The graphical models capture the distribution in the data and decisions can be obtained via statistical inference.

Machine Learning techniques automatically build models from data which can be used to make predictions or decisions. Typically there are three classes of machine learning techniques, namely unsupervised, supervised and reinforcement learning (Dunjko, Taylor & Briegel, 2016; Sathya & Abraham, 2013) though more recent research identifies four classes of machine learning which includes semi-supervised learning (Belaidouni & Miraoui, 2016; Rebala et al, 2019). Unsupervised learning techniques can find patterns from input data, without requiring these to be labelled. On the other hand, both supervised and reinforcement learning required labelled input data. Supervised learning includes classification techniques for determining which category an item belongs to and numerical regression which generate a function, which first captures the relationship between inputs and outputs, and subsequently used to predict how outputs will change as a function of the inputs. Lastly in reinforcement learning, the aim is to reward a learner agent for good responses and to penalise for poor responses, ultimately allowing the agent to learn an operational strategy for its problem domain (Belaidouni & Miraoui, 2016; Dunjko et al, 2016; Rebala et al, 2019; Sathya & Abraham, 2013).

1.2.1.3 Subsymbolic Paradigms

Sub-symbolic systems, also referred to as *Connectionist Systems*, are representative of neurons in the human brain. Sub-symbolic architectures can learn autonomously, following training and development of the neural network architecture (Kelley, 2003, p.849). Corea (2019) defines the sub symbolic paradigm as one where "no specific representations of knowledge should be provided ex-ante". Corea subsequently places the concepts of affective computing, autonomous systems, distributed artificial intelligence, ambient computing and evolutionary algorithms. If taking a systems-based view of AI architectural paradigms, only the distributed intelligence paradigm and some applications of evolutionary algorithms qualify, as the others are typically applications of AI approaches that are listed by Corea (2019) as either symbolic or statistical.

Affective computing focuses on recognising or simulating human emotions (Picard, 2000), which can be ill-defined, however the AI architectures employed in this area tend to be the standard symbolic or statistical paradigms. For example, a neural network can be used, with input coming from a variety of sensors measuring human attributes such as skin conductance, heart rate, and elements of speech of brain activity. Training data is usually obtained by exposing human subjects to stimuli that have been chosen and validated to induce a particular emotional response, where the subject indicates what their response is along a number of dimensions, typically including valence (good versus bad) and level of arousal (Picard, 2000).

Likewise, autonomous systems and ambient computing are defined by their types of input and output rather than the internal AI architecture. Ambient computing strives to make the computing experience transparent, through the use of unobtrusive sensors and program design rather than a specific novel architectural paradigm (Streitz & Privat, 2009). Similarly, autonomous systems, which is an AI application in the field of robotics are defined by their sensors and actuators, with aims such as navigation and obstacle avoidance (Pandey, Pandey & Parhi, 2017).

Distributed AI, with examples including multi-agent systems, agent-based modelling, and swarm intelligence, fits into the sub symbolic paradigm. An agent-based system is one where a set of actors (agents) work and interact autonomously (Vlassis, 2003). Typically each agent is given a simple set of behaviours (and can be modelled internally using either a symbolic or statistical paradigm) – but the key is that out of the interactions between the agents, a higher-level behaviour emerges that is of greater value than just the sum of its parts (Bertalanffy, 1968; Kelley, 2003). A good example of a non-artificial agent-based system is an ant colony. Each ant has a specific role with a simple set of behaviours, but combined the colony displays advanced behaviour, such as discovering the optimal path between a food source and the nest (Dorigo & Gambardella, 1997; Mohamad, Taylor & Dunnigan, 2006).

In AI Evolutionary Algorithms are typically used as a tool to optimize other AI architectural paradigms, such as neural networks and rule-based systems (Vikhar, 2016). However, they can also be used to allow the emergence of behaviour. For example, they can be used to evolve the individual agent controllers in agent-based systems or to evolve a set of simple rules into a rule-set displaying emergent behaviour (Dorigo & Gambardella, 1997). Within the AI literature, Corea (2019) identifies embodied intelligence approaches and search and optimization approaches as sub symbolic, and describes as follows:

Embodied intelligence approaches take into consideration *situatedness* and *embodiment* in the design of intelligent behaviour in embodied and situated agents. *Situatedness* is the coupling between the agent and its environment, and *embodiment* refers to the constraints associated an agent's body, perceptual and motor system. The study of embodied intelligence has been associated with early development in bio-inspired computational intelligence techniques in robotics, where the focus is on morphological computation and sensory-motor coordination in robotic models.

Search and Optimization approaches are techniques that can search a complex and ill-defined search space intelligently and efficiently.

In the modern threat environment, corporate entities and governments will continue to develop and use the latest technological advances for safety and security needs, and this includes in a broader public space threat environment. Contemporary advances in security technology include developments in artificial intelligence (AI) or machine intelligence. The spectrum of AI technologies enables machines to take inputs from the defined system environment, perform computational tasks, make decisions and implement actions as outputs to achieve objectives in the safest, most efficient and timely means possible with high degrees of accuracy, using advanced communication and integration means.

Developments across the Artificial Intelligence space means there are further opportunities for technology enhancements not yet recognised by governments and corporate entities including security managers, resulting in gaps of opportunities which could see security more effective, more efficient, safer, and more integrated than current technological solutions. To understand these gaps requires a thorough articulation of the current use of security technology, its functions, underpinning principles of automation, and integration with other technological systems based on, and aligned to the Artificial Intelligence (AI) problem domains and spectrum of AI paradigms.

AI provides machine support for decision making that can enable enhanced automation through the integration of complex outputs across a range of technological platforms that apply manual rules based on inputs from narrow AI technologies.

1.3 Understanding Physical Security Technology

The articulation and integration of multiple systems for the protection of assets for this study was based on ASIS International's published overarching functions of physical security measures (Figure 1) including Observing, Detecting, Controlling and Responding.

The study's background discussion highlighted that in the protection of assets security and crime prevention technology measures may consist of integrated combinations of observing technologies (Video Surveillance), detecting technologies (Threat Detection), controlling (Electronic Access Control) and response technologies (Active Measures) underpinned by Communications technologies (Networking Infrastructure) which support and facilitate security processes and operations in a systematic approach. Consequently, understanding both the current state, and future directions of AI in the protection of people, information and assets requires an understanding of where such technologies sit within the Spectrum of AI and how technological communication and integration facilitates advanced security outcomes.

1.3.1 Communication and Integration of Security Technologies

The basis for understanding the integration of detect, observe, control and respond technologies in the protection of assets stems from the early developments of automation. In their analysis of Building Automation and Control Systems (BACS) Brooks, Coole, Haskel-Dowland, Griffiths and Lockhart (2017) noted that automation across the built environment stems from drive for cost effective yet reliable efficiencies which remove the expensive and unreliable human element through digital or mechanical elements. Automation in the protection of assets is saliently understood by security professionals through the notion of the Intelligent Building System.

According to Brooks et al. (2017, p.18) an Intelligent Building is denoted as "An automated system where building services and corporate processes communicate with each other to exchange digital, analogue or other forms of information, to a central control point to manage the environment". Intelligent Building Systems may be referred to by various names, with Building Automation and Control Systems (BACS) supporting common linguistics. BACS require connectivity to function, facilitated through common language standardized logic code, via various communications networks that link and integrate the many discrete devices.

Understanding the Intelligent Building through BACS provides initial insights into the use of Artificial Intelligence in the protection of assets and reinforces the notion of integration; society now uses AI to achieve automation, as AI aims to be a problem-solving machine, integration underpins the embedding of automated process and is what many security professionals consider colloquially under the umbrella term, artificial intelligence. However, while automated processes may use AI, automation is not AI. Distinctions may be drawn between concepts and definitions of integration, automation and levels of autonomy and how AI performs within each classification or level. In systems engineering, differentiation between automation and autonomy is determined by the degree of human control during operation (Norris & Patterson, 2019). Consequently the following definitions are presented below;

Integration

'The act of combining or adding parts to make a unified whole' (Harper Collins Publishers, n.d.). The term *Integrated Security* implies a combination of security technologies, functions and components, or quite simply; 'an assimilation of different security services' (Dowell, 2015) which communicate to perform advanced functions in, as a minimum, an automated manner (Norman, 2014, p.263).

Automation

'The technique, method, or system of operating or controlling a process by highly automatic means, as by electronic devices, reducing human intervention to a minimum' (Harper Collins Publishers, n.d.). In automation, human decisions and logic are executed to 'accomplish a pre-set series of tasks within a known (or assumed) frame of reference', without decisions being made during operations (Norris & Patterson, 2019, p.1).

Semi-autonomy

Norris & Patterson (2019, pp.1-2) define semi-autonomy as 'anything that involves machine decision making (in response to external, unexpected events) during operation but that a human is involved in some of this process and provides some direct control. Therefore, semi-autonomous systems are those which are more independent and agile than automated systems and less self-contained than fully autonomous systems'.

Autonomy

Autonomy may be broadly defined as 'the condition of being autonomous; self-government, or he right of self-government; independence' (Harper Collins Publishers, n.d.). According to Norris & Patterson (2019, p.1), an autonomous system is one where decisions (in response to external inputs or signals of whatever complexity) are made within the system and do not involve human decision making'. Norris & Patterson further add that in theory, 'a fully autonomous system would be able to respond on its own to unknown and unexpected events via artificial intelligence without pre-programming by a human' (2019, p.1).

While these definitions provide a relatively solid foundation for differentiating between integration, automation and levels of autonomy, it is essential to understand definitions and interpretations may vary between academic fields, commercial sectors and industries. For instance, in the mining sector, Caterpillar defines 'autonomy' as referring to 'a state of equipment in which it can perform the programmed operations under defined conditions without human input or guidance' (Caterpillar, n.d.). However in disciplines such as robotics, space-based research or weapons engineering, systems or vehicles, operating within defined and programmed parameters may be considered automated rather than autonomous. In space-based research, Frost (2011, p.89) offers the following differentiation between automation and autonomy, and introduces the concept of intelligent autonomy;

> An *automated system* doesn't make choices for itself – it follows a script, albeit a potentially sophisticated script, in which all possible courses of action have already been made. If the system encounters an unplanned-for situation, it stops and waits for human help (e.g. it "phones home"). Thus, for an automated system, choices have either already been made and encoded, or they must be made externally. By contrast, an *autonomous system* does make choices on its

own. It tries to accomplish its objectives locally, without human intervention, even when encountering uncertainty or unanticipated events. An ***intelligent*** *autonomous system* makes choices using more sophisticated mechanisms than other systems. These mechanisms often resemble those used by humans. Ultimately, the level of intelligence of an autonomous system is judged by the quality of the choices it makes.

Frost (2011) clearly identifies *decision making* as the differentiating factor between automation and the various levels of autonomy. This notion is supported within the literature on Autonomous Weapons Systems (AWS) and Lethal Autonomous Weapons Systems (LAWS) in terms of decisions over critical functions such as the release of force (Heyns, 2016, p.4). According to Heyns (2016, p.6), AWS's may be differentiated by their level of autonomy in decision making and not by the weapons they use. Consequently, the higher the level of autonomy in an intelligent system, the lower the degree of human control and decision making.

However, though understanding conceptual and functional differences between levels of autonomy is essential for exploration of AI within intelligent systems, the majority of security technologies do not currently achieve semi-autonomous or autonomous operation. Integration and automated processes in BACS systems, as described by Brooks et al (2017), are essentially the current level at which security technologies achieve 'intelligent' operation. There are a vast array of built environment technological platforms integrating inputs, processes and outputs through BACS. BACS may integrate both facility services and utilities with security technologies, detection, observing, controlling and responding technologies with auxiliary technologies including lighting systems, intrusion/fire & life safety/smoke alarms, audio systems such as Emergency Warning and Intercommunication Systems (EWIS), incident management systems, vehicle control systems, situational and risk awareness tools and managed remote monitoring systems. Such technologies are integrated with their individual components, then collectively merged into the broader building automation architecture (Brooks et al, 2017, p. 21; Dowell, 2015; Norman, 2014, pp.263-279).

Brooks et al. (2017) body of work on BACS cited the European Committee for Standardization (CIBSE, 2000), specifically their International Standard for Building Automation (2004), to articulate the networked structure of BACS. The structure highlights that the building automation architecture and associated communications are divided into three distinct layers or levels comprising; Management, Automation and Field device level (Figure 4) (Brooks et al, 2017, pp.25-26). Such an architecture affords a clearer articulation of application of AI in security technologies, where future opportunities exist for capability enhancement, and facilitates the identification of risks associated with such adoption.

Management Level

The Management level is defined as an organisation's Information Technology and Communications (ITC) network. According to Brooks et al. (2017) the Management level comprises "operator stations, monitoring and operator units, programming units and other peripheral computer devices connected to a data processing device i.e., a server" (International Organization for Standardization, 2004, p. 53) to support the information exchange monitoring and management of the automation system (Brooks et al, 2017, p.25).

Automation Level

The Automation level is defined as a dedicated communications network for the sole purpose of building device connectivity, communication and control (automation). The Automation level comprises "control devices and monitoring and operator units, programming units, operator stations or panels and/or programming units connected to a data processing device i.e., a server" (International Organization for Standardization, 2004, p. 53). This level is associated with controllers that serve main plant, such as the air handling units, chillers and boiler units, etc. (Brooks et al, 2017, p.25).

Field Level

The Field level comprises of sensing or action devices that are generally self-contained physical units. Field level devices are connected to automation level controllers, either application specific or generic controllers. Application specific controllers operate using communications protocols such as M-bus or other proprietary protocols, referred to as Edge Devices (Brooks et al, 2017, pp.25-26; Fennelly, 2012, p.316).

Edge computing refers to ‘the enabling technologies allowing computation to be performed at the edge of the network, on downstream data on behalf of cloud services and upstream data on behalf of IoT services’ (Shi, Cao, Zhang, Li & Xu, 2016). Edge computing is ‘complimentary’ to cloud based computing, where the goal is to locate the cloud closer to the edge (Conde, 2018). By doing this, the data does not need to be sent to the cloud for processing thereby relieving pressure and improving the speed and performance of data (Anderson, 2019; Linthicum, n.d.). For physical and integrated security systems, this means edge devices with AI algorithms (such as IoT sensors, smartphones and security cameras) can process and extract selected information, such as an individual or a vehicle in a video image, thereby reducing bandwidth consumption and increasing efficiency, quality and accuracy (Anderson, 2019; Shaw, 2019).

Figures 4 (*Three-layer BACS Architecture*) and 5 (*Illustration of BACS Three-layer Architecture*) developed by Brooks et al (2017, p. 25) visually depict the integrated devices and components in each level.

Figure 4

Three-layer BACS Architecture (Brooks et al., 2017, p. 25)

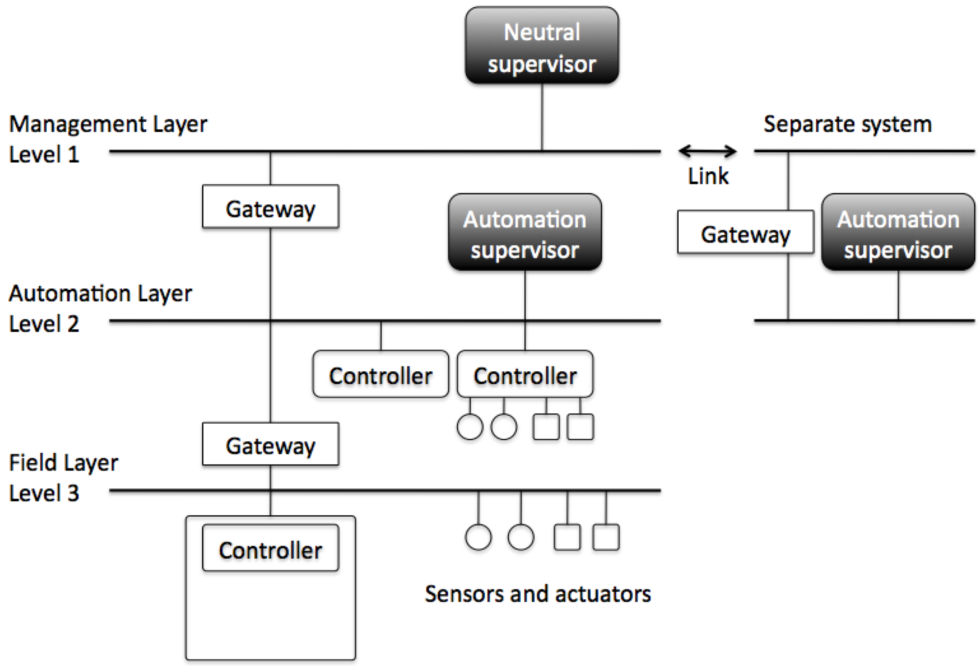


Figure 5

Illustration of BACS Three-layer Architecture (Brooks et al., 2017, p. 26)

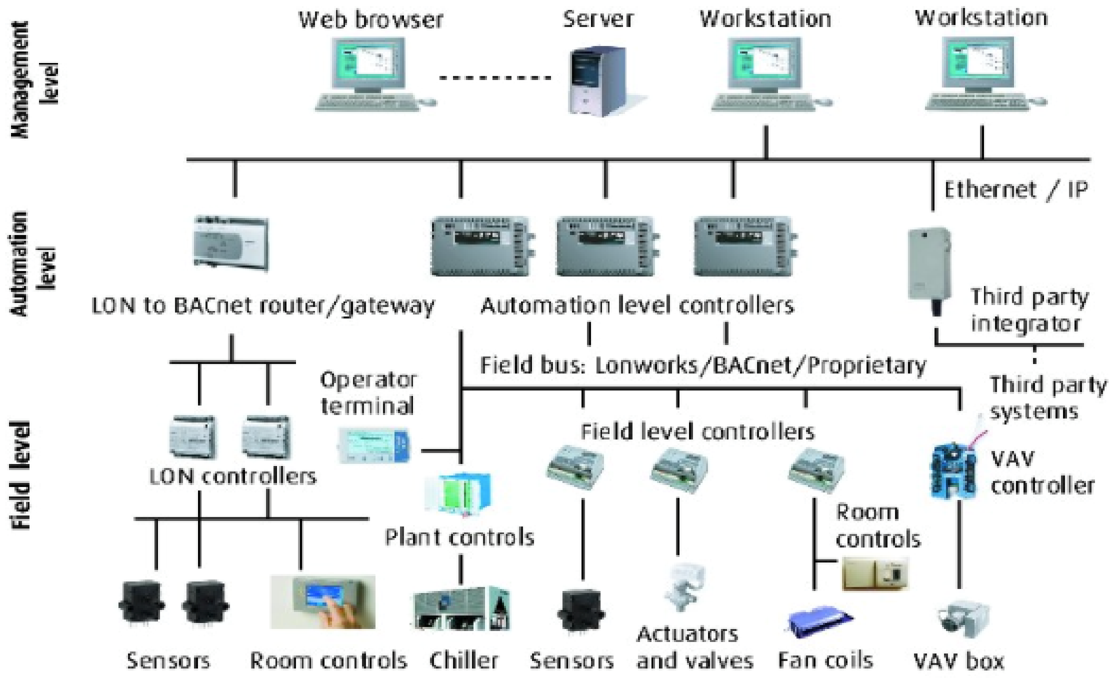


Figure 5 illustrates the connection between the various BACS technologies across the different network communication levels. Such connectivity is demonstrated in Table 1, which provides an example of how an individual BACS technology such as Lighting may be divided into field, automation and management level components.

Table 1

BACS Technologies and Levels

BACS Technologies	Devices and Components		
	Field Level	**Automation Level**	**Management Level**
Lighting	Active Infrared and PIR Movement Sensor, Light Sensor, Microwave Sensor.	Triggering of stimulus (active system) or absence of electrical current (passive system) at field level activates lighting or emergency lighting system according to pre-programmed rules set at management level	Functional aspects i.e. Control panel where pre-programmed responses to stimulus are input by user.

1.3.2 Roles of Electronic Security Technologies

1.3.2.1 Observing Technologies

ASIS International considers the core functions of observation technologies to be 'the detection of an approaching threat (such as an intruder or a storm), characterize the threat, aid in formulating a response, and assist in investigative efforts after an event occurrence' (ASIS International, 2015, p. 17). Network surveillance systems are perhaps the most prevalent of the observation technologies, and are typically multifunctional, incorporating an array of components including at the field, automation and management levels into a system of seamless integration.

1.3.2.1.1 Network Video Surveillance

Definition

Network Video Surveillance is primarily used to 'detect activities that call for a security response; collect images of an incident for later review or assist with incident assessment' and use as evidence if needed (ASIS International, 2015, p.315). Evolving from earlier Closed-Circuit Television (CCTV) systems, video surveillance systems have developed considerably, with modern and 'intelligent' surveillance systems capable of integrating detection elements such as sensors to provide additional functionality and minimise the monitoring requirements from operators (Garcia, 2001, pp.113-144; Kreugle, 1995).

Characteristics

Network Video Surveillance Systems consist of a combination of Field Level, Automation Level and Management Level devices and components installed at both internal and external facility locations (Garcia, 2001, pp.115-144). Network Surveillance System components and their indicative BACS architecture levels with AI alignment are presented in Table 2:

Table 2

Indicative Surveillance System BACS Architecture Levels and AI Analytical Alignment.

NETWORK SURVEILLANCE TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Network Video Surveillance	Cameras, Lenses, Hardware, Cabling, Transmission System, Video Recording Device & Hardware, Network Storage Devices, Controllers & Monitors, CPU, Power Unit	Signal transmission, processing & communication of alarm conditions, Video & Audio Streaming, Video Analytics (anomaly detection, behavioural analytics etc), sequence display, Object/ person tracking.	System Configuration & programming (schedules, zones, display sequence etc), cloud configuration, command signal (e.g. Pan, Tilt, Zoom (PTZ), user enrolment for integrated technologies (e.g. biometric sensors & devices), video assessment & instruction

Video Motion Detection	Cameras, Lenses, Sensors (PIR, Microwave), Cabling, Transmission System, Video Recording Device & Hardware, Network Storage Devices, Controllers & Monitors, CPU, Power Unit	Signal transmission, processing & communication of motion detection, Alarm Activation, Video & Audio Streaming, Video Analytics (motion history imaging, behavioural analytics). Object/ Person Tracking	System configuration & programming, (schedules, zones, display sequence etc), cloud configuration, command signal (e.g. PTZ)

Indicative Analytical Methodology

Network surveillance systems may employ both rules-based and statistical analytical methodologies, with the video surveillance functions using expert-rule setting, and the video analytics (e.g. object recognition or anomaly detection) components using statistical (probabilistic and machine learning) methodologies. Expert Rule-setting occurring at the management level for network surveillance systems follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. motion is detected), THEN the command signal initiates a defined output, such as an alarm and may automatically record streamed video (Dutta & Bonissone, 1993; Garcia, 2008.).

Functionality and Role in Protection of Assets

The function of video surveillance systems has historically been to observe and monitor people, objects and events. However, in addition to observation and monitoring, surveillance systems can generate and directing an appropriate response (Garcia, 2001, pp.113-144; Kreugle, 1995). In terms of outcomes, surveillance systems are also expected to both act as a deterrent and produce evidence such as identifiable images in the event a person violates policy, procedure or engages in criminal behaviour (Fischer & Green, 1998).

To achieve these outcomes, a variety of components are available to increase the overall functionality of surveillance systems. Depending on the application, environment and functional requirements, surveillance camera features may include fixed or 'PTZ' (Pan, Tilt, Zoom) installation, 360-degree rotation, high definition up to 30 Mega pixels, wireless/ internet protocol transmission (IP), weatherproof, night vision (thermal or infrared operation for night-time and/or poorly lit areas), multiple alarms, auto focus and tamper resistance (Active Online, 2019; Rai et al, 2018).

The most prevalent surveillance systems are network or 'Internet Protocol (IP)' and use Power over Ethernet (PoE) to supply power and data to the camera from a network video recorder (NVR). Functionality includes simple plug-and-play operation, high resolution video, remote access, fewer wires/cables, encrypted data, programmable recording (such as when motion or audio is detected) and integration of technologies such as biometrics as seen in sophisticated *Intelligent* Video Surveillance Systems (Spy Monkey Surveillance, n.d.; Videosurveillance.com, n.d.).

Integration

Network video and intelligent video surveillance systems are highly integrated with modern security and building systems, including security lighting, access control systems, and may also incorporate a range of security sensors and technology such as biometric facial recognition systems (Brooks et al, 2017). Advanced capabilities of integrated surveillance systems may include predefined rules-based analysis at the automation level, including motion detection, object (pattern) detection and tracking (including abandoned object detection), object classification, tripwire/line crossing detection, license plate recognition (LPR), vehicle make and model (Pattern) recognition (MMR) and colour recognition

(CR) as well as user-defined object search and image processing (Active Online, 2019; Rai et al, 2018; Robotech Vision, n.d.; Spy Monkey Surveillance, n.d.; Tawiah, 2010, pp.12-13).

Recent developments in both integration and automation have seen a trend toward robotics for automated security guarding and response, with these systems reliant on video surveillance for analytics and assessment. Examples include security robotic devices such as ROSA (Responsive Observation Security Agent) for automated security guarding of building perimeters and parking lots, in addition to SCOT (Security Control and Observation Tower), FRED (Facial Recognition Entry Device) and ROAMEO (Rugged Observation Assistance Mobile Electronic Officer) (Globe Newswire, 2019; Robotic Assistance Devices [RAD], n.d.).

Artificial Intelligence in Network Video Surveillance

Algorithmic developments have resulted in a high degree of AI in intelligent network video surveillance systems, particularly for security and broader law-enforcement purposes. ASIS International (2015, p. 473) defines *Intelligent Video* as 'a learning algorithm wherein the process evaluates the camera image for some period and develops a sense of normal behavioural patterns… after the learning algorithm has had some time to learn what is normal for the scene it can be used to alert on what is not normal.' Subsequently, the use of intelligent video surveillance systems to track and monitor humans, vehicles and events has significantly expanded in recent years, employing rule-based, statistical or analytical programmes to detect and respond to security threats (Denman et al, 2006; Tawiah, 2010). Intelligent video systems can incorporate algorithms to analyse video content and establish 'norms' such as patterns of behaviour. This technology, known as *Video Analytics* or *Video Content Analysis* includes algorithms for functions such as event detection, vehicle monitoring and traffic control, object detection and tracking, predictive analysis and behavioural analytics such as anomaly detection (Calavia et al, 2012; Sreenu & Durai, 2019; Tawiah, 2010).

The algorithmic approach may vary between functions, for instance algorithms for object detection and tracking are usually either feature-based, model-based, or motion-based (Tawiah, 2010, p.12), while algorithms for anomaly detection will rely on behavioural analysis to learn normal (pattern analysis) (nominal) and abnormal (anomalous) behaviours, and automatically detect anomalous behaviour and events in individuals, groups and crowds (Charara et al, 2012; Popoola & Wang, 2012). This type of learning, referred to as *machine learning*, uses observed data to develop a baseline of nominal behaviour and subsequent deviations from the developed baseline model are regarded as anomalous (Popoola & Wang, 2012, p.5; Saini et al, 2016; Wu, 2017, p.3).

Limitations

There are a numerous limitations of network video surveillance systems. Tawiah (2010, pp.7-9) identifies several challenges of visual scene analysis, including analysis complexity (such as difficulty in distinguishing object features from noise), accuracy and false positive/ negative rates (algorithmic parameters), reliability (particularly with dynamic scenes), response time (pending accuracy requirements), and cost effectiveness (Tawiah, 2010, pp.7-9). Privacy, ethics and the potential for discrimination may be particularly problematic and have the potential to result in reduced public acceptance of widespread implementation of these systems (Vincent, 2018; Wolf, 2012).

Significantly, one of the major limitations of network surveillance systems is security issues, particularly for intelligent systems which utilise cloud-based management, processing capacity and storage. While cloud-based AI has the capacity to analyse more data and run the powerful algorithms required by complex video analytics, cloud-based AI may have a higher risk to security issues such as access security, security of retained data and vulnerability to attack (Sawhney, 2020; Suryateja, 2018).

1.3.2.1.2 Observing Technologies Overview

Observing technology across the protection of assets varies, however understanding current and future AI optimisation stems from understanding where current technologies align to the domains and continuum of Artificial Intelligence (AI). Therefore, Table 3 presents a summary of current Observing technology and indicative AI analytical alignment.

Table 3

Observing Technology and Indicative AI Analytical Alignment.

Technology		**Function/ Performance Measure**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Network Video Surveillance	Digital Video Surveillance	Observe, Detect, Recognise, Identify	Reflected Scene Image	CCTV cameras used to monitor and record a defined space	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
	Video Motion Detection	Probability of Detection	Movement within a Frame	CCTV cameras detect movement of an individual into a defined space	Statistical Probability (Probabilistic Method)

1.3.2.2 Detecting Technologies

Described by ASIS (2015, p. 17) as a partnership function to observation and the function that makes possible notification and response, detection of an adversary is one of the key components of the conceptual basis of security, being Deter, Detect, Delay and Deny (ASIS International, 2015, p. 16). Garcia (2008) succinctly defines Detection as "the discovery of an adversary action…(by) sensing covert or overt actions" (Garcia, 2008, p. 8). Fennelly (2012, p. 22) further details the function of a Physical Protection System (PPS) as including "exterior and interior intrusion sensors, alarm assessment, entry control and the alarm communication and display subsystem all working together … (to gain) knowledge of a person or vehicle attempting to gain unauthorized entry into a protected area by someone who can initiate and appropriate response". Accordingly, detection must be achieved to allow other security functions to be activated and take effect.

Fennelly (2012) and Garcia (2008) highlight that detection also requires some type of assessment to determine the validity of the stimulus, adding that "detection without assessment is not detection" (Garcia, 2008, p. 59). This assessment is achieved through the activation of an alarm system sensor of a Physical Protection System (PPS) in response to stimulus, which in turn, through automation, initiates an alarm signal which is reported and then assessed through seamless integration. The assessment of the alarm to be valid or invalid then triggers the activation of a control or response. According to Garcia (2008, p. 60) a detection system must provide two types of information: information about whether the alarm is valid or a nuisance activation, and causation details – who, what, where, when and how many.

1.3.2.2.1 Intrusion Detection Systems

Definition

Detection is achieved through various Intrusion Detection Systems (IDS), also referred to as an 'Intruder Alarm System'. An IDS is described by ASIS as a system designed to detect an undesirable entity, be it person or vehicle attempting to gain access, using interior and exterior sensors, video alarm assessment, access control and communication systems working in combination (Integrated) (ASIS International, 2015, p. 276). Intrusion detection systems are further explained by Brooks and Coole (2019, p. 1) as a "protection norm" within the built environment, both domestic and commercial, providing the "detection" element in a security system is a fundamental role in the Defence in Depth strategy.

Characteristics

An IDS comprises numerous subsystems across the field, automation and management levels; detection sensors and output devices such as annunciators and dialling devices, a control panel (central processors, memory input and output, power supply, etc.) and a user interface allowing interaction with the IDS. The detection function of the IDS is initiated by the sensors, described by ASIS (2015, p.276) as "the basic building blocks" of the IDS. All subsequent data processing and transmission, display and recording are triggered from initial detection stimulus, and is based upon the sensor technology. IDS components and their indicative BACS architecture levels with AI alignment are presented in Table 4:

Table 4

Indicative Intrusion Detection System BACS Architecture Levels and AI Analytical Alignment.

INTRUSION DETECTION SYSTEMS			
	Field Level	**Automation Level**	**Management Level**
Intrusion Detection Systems	Contact/Reed Switch, Glass-Break Sensor, PIR Sensor, Active Infrared, Microwave Sensor, Ultrasonic Sensor, Audio Sensors, Photoelectric Beams, Fibre Optic Cable Sensor, Microbend Cable Sensor, Capacitance Proximity Sensor, Pressure Sensor, Light Sensor, Dual-Technology Sensor.	Triggering of the stimulus at field level, sets off automated response such as activation of audible alarm bell, notification of human response force (emergency services), activation of integrated access control system (locking down of area etc)	System Configuration & Programming, Setting of parameters through system interface (schedules, zones, user enrolment etc), command signal, alarm assessment & instruction
Annunciator Panels	CPU, Alarm Modules, Alarm Display Unit, Keypad, Buzzer & LED Lights, Keypad Inputs for Remote User Interface, Input/ Output Connectivity, Power Supply Unit	Signal Processing, Audio Alarm Activation, Visual Alarm Activation, Logic Functions, Output Relay Changeover	System Configuration & Programming, Setting of Parameters through System Interface (Operating modes e.g. attended/ unattended, Maintenance, Surveillance)

Duress Alarms (e.g. Panic Bar, Duress Watch or Device)	Sensor, Transmitter, Receiver, Hardware, Communications, Cabling/ Connectors, Sirens, Power Supply Unit, CPU for Smart Devices	Alarm Activation, Activation of Live Video & Audio Streaming, GPS Location Tracking, Automated Check-in Prompts, SMS Notifications to Emergency Contacts	System/ Device Programming (Man-down/ Incapacitation, Location Setting, Emergency Contacts), Command Signal for Panic Bars/Buttons

Indicative Analytical Methodology

Intrusion detection systems may employ a rules-based analytical methodology with Expert Rule-setting occurring at the automation and management levels where activation occurs due to detection of stimuli. Rule-setting for intrusion detection systems follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. sensor detects the conditions or requirements for activation), THEN the command signal activates the system to communicate the alarm condition (Dutta & Bonissone, 1993; Fischer & Green, 1998).

Functionality and Role in Protection of Assets

The automated functions of the IDS comprise of several stages including: 1). The stimulus is the presence of the object or the person within the field of view or range of the sensor. As Fay (2011, p. 164) states "detection can happen at the sensor point through either breaking of the circuit; interruption of a light beam, movement, sound, vibration, or a change in an energy field", 2). The sensor reacts to the stimulus, and 3). A pre-defined algorithm based on expert rules setting decides whether to act in producing an alarm in response to the presence of the stimulus. Subsequently such an alarm is assessed, and a response is activated.

The design of an IDS considers numerous factors including the asset or area being protected; the environment, terrain and geography in which the sensor must operate. According to Brooks and Coole (2019, p. 3) the protection level provided by the IDS is directly proportionate to the threat and risk level; as threat and/or risk increases, so too should the protection capability of the IDS. One method for increasing the protection level is by using dual-technology detectors with overlapping fields of views, with the aim of increasing probability of detection. Most dual-technology sensors combine passive infra-red with microwave sensors; microwave sensors are particularly effective at detecting motion toward or away from the sensors, while infra-red sensors are more effective at detecting motion moving across the field of view. Dual-technology sensors are typically configured using "OR" gate logic (detection using single sensor stimuli), or "AND" gate logic (detection using dual sensor stimuli). According to Stoltfuz (2020) the AND/OR logic gates are precursors to AI and the building blocks for neural networks, with the logic gate doing the work of a human neuron in the decision making process.

Integration

Alarm systems are highly integrated, particularly with other security systems such as access control and video surveillance systems, and Building Automation and Control Systems (BACS) such as communications, lighting, and fire detection and suppression systems (Brooks et al, 2017; Norman, 2014). Integration of alarm systems with critical functions may assist security and other personnel responding to an incident – for example surveillance cameras integrated with an intrusion detection system may assist in the identification and verification of an adversary action (Garcia, 2001). This type of information can help interface users manage unauthorized access, false alarms, nuisance alarms and work within the context of the building to provide more efficient solutions to identified threat and risk vulnerabilities.

Artificial Intelligence in Intrusion Detection Systems

Alarm systems process signals from a variety of integrated sensors and intelligent technologies such as intrusion detection systems, biometrics, and video surveillance. Currently, many of the sensors and devices which detect alarm conditions have been developed to incorporate deep learning and machine learning algorithms – for example, network surveillance systems with video analytics and anomaly detection (e.g. acoustic event detection and behaviour analysis). However, while there is currently limited evidence of the alarm processing unit (and not the sensors or integrated technologies) having AI capabilities, there may be scope for this in future applications. Security World Market (2019) predicts AI may be used by alarm monitoring centres to predict false alarms and assist operators in responding to valid alarms. By distinguishing and utilising patterns in data sets, such as coordinated alarm events across integrated systems, AI may be applied to IoT ecosystems (Bonoan & Ataev, 2020).

Limitations

Brooks highlights that intrusion detection systems are only as effective as the quality of design, installation and the environmental threat (2014, p. 695) and as such, systems are subject to legislation and standards in an attempt to regulate the industry. A further limitation of detection systems is the Nuisance Alarm Rate (NAR) which may be generated through inappropriate design, detector selection, poor maintenance or failure. External factors such as inclement weather conditions or wildlife or vegetation may also affect the NAR, resulting in response malaise, a strain on resources (wasted police response), perception of reliability and the security decay of system efficacy (Coole & Brooks, 2009).

Additionally, detectors can be defeated through methods such as bypass (circumventing the detectors field of view), technical defeat (modifying the input or output function of the IDS) or spoofing (defeating the detector by moving through the field of view without stimulating the sensor above its activation threshold) (Brooks & Coole, 2019, p. 3).

1.3.2.2.2 Active & Passive Lighting Systems

Definition

Honey (2001, p. 113) defines security lighting as lighting "intended to illuminate areas in order to deter crime" and further highlights that lighting systems improve the likelihood of detecting unauthorized personnel and intruders by supporting the efficiency of other security installations (such as the augmenting and increasing the visibility of CCTV camera systems).

Characteristics

Automatic security lighting system components and their indicative BACS architecture levels with AI alignment are presented in Table 5:

Table 5

Indicative Lighting System BACS Architecture Levels and AI Analytical Alignment.

LIGHTING SYSTEMS			
	Field Level	**Automation Level**	**Management Level**
Active Lighting Systems	Active Infrared and PIR Movement Sensor, Light Sensor, Microwave Sensor.	Triggering of stimulus (active system) or absence of electrical current (passive	Functional aspects i.e. Control panel where pre-programmed responses to stimulus are input by

		system) at field level activates lighting or emergency lighting system according to pre-programmed rules, set at management level.	user. Commonly integrated into Building Management System and linked to other systems such as access control and emergency lighting and CCTV.
Passive Lighting Systems	Electrical Current Capacitor Sensor	Upon detection of disturbance in electrical current, system automatically activates emergency lighting system or contacts emergency response services.	Control panel where pre-programmed responses to stimulus are input by user. Commonly integrated into Building Management System and linked to other systems such as access control and emergency lighting and CCTV.

Indicative Analytical Methodology

Active and passive lighting systems may employ a rules-based analytical methodology with Expert Rule-setting occurring at the automation and management levels, where activation occurs as a result of a pre-programmed response to detection of stimuli, or a change in state. Rule-setting for lighting systems follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. the detection point, line or volume or change of state is breached), THEN the command signal activates lights to turn on or off (Dutta & Bonissone, 1993).

Functionality and Role in Protection of Assets

Technical lighting systems facilitate detection and observation and therefore may sit across both technological functions of security. According to Purpura, Fennelly, Honey and Broder (2013, p. 181) the two major purposes of security lighting are to create a psychological deterrent to intrusion, and to enable detection. The importance of lighting systems in the prevention and detection of crime is supported by Farrington and Welsh (-2004) who conducted a systematic review of studies and concluded that lighting is an effective crime countermeasure. Lighting may be used alone or alongside other measures, and also for illuminating areas to augment the ability of CCTV systems in capturing images (Atlas, 2013, p. 747).

Integration

Sensors which integrate lighting systems into building management systems have developed considerably. Previously, photocells were required to be set up and controlled separately from occupancy sensors and occupancy sensor technologies – such as Passive Infrared (PIR), ultrasonic, and microphonic – were specified depending on the obstructions in the space and use of adjacent spaces. Additionally, auto-on/auto-off or manual-on/auto-off, and vacancy time-out/delay settings were programmed during installation via dipswitch or pushbutton. Smart sensors now incorporate a wide variety of sensing technologies, and most network control systems have the capability to be addressed remotely via desktop or phone apps avoiding the necessary physical adjustments causing disruption to occupants.

Artificial Intelligence in Lighting Systems

AI is used within all stages of the lighting life cycle, including design, installation, commissioning, configuration and maintenance (Juslen, Nasir, & Nouri, 2019). Juslen et al (2019) proposed that AI could be used across network structure, where self-learning networks of lighting components could communicate and set themselves up without requiring human intervention, akin to auto commissioning systems seen in the IT industry (p.62). Juslen et al (2019) further highlight that by using the observations of indoor environments, AI based lighting systems can optimise and tune light parameters accordingly. Data may also be used for other integrated Building Management Systems (BMS) such as Heating, Ventilation and Air Conditioning (HVAC), or access management. It can also be used to optimise use of efficiency measures in lighting systems; for example by highlighting areas of overuse or underuse, and analysing the movement of people between different areas (Juslen et al., 2019).

Limitations

Juslen et al (2019) highlight key limitations for intelligent lighting systems. These include observations that occupants may not be trained to use the system properly or the AI enhanced lighting systems may be overly complicated and not appropriate for building users. Juslen et al. (2019) further assert that research output is often limited to a niche group of scientists and beyond the comprehension of the end-user.

1.3.2.2.3 Contraband Detection: X-Ray Technology

Definition

Enhanced detection of specific materials is achieved using x-ray technologies. There are two types of x-ray technology typically used in security applications; backscatter x-rays and transmission x-rays (Scientific Committee on Emerging and Newly Identified Health Risks [SCENIHR], 2012).

Backscatter X-rays – these units operate by exposing the subject to low energy X-radiation which passes through clothing but is “scattered” by dense objects. The radiation is scattered back into a series of radiation detectors which create images of the items or bodies being scanned, including concealed contraband underneath clothing. Backscatter x-ray systems use a narrow, pencil shaped beam that scans the subject at high speed in a horizontal and vertical vector.

Transmission X-rays – these units use a vertical fan-shaped beam of x-ray radiation that passes directly through the object or person being examined into a detection system placed on the side of the person opposite to the x-ray source. Similar to the x-rays used for medical examinations, a radiographic image is produced highlighting any items with a differing density (Scientific Committee on Emerging and Newly Identified Health Risks [SCENIHR], 2012).

Characteristics

X Ray detection systems and their indicative BACS architecture levels with AI alignment are presented in Table 6:

Table 6

Indicative X Ray Detection System BACS Architecture Levels and AI Analytical Alignment.

X RAY DETECTION SYSTEMS			
	Field Level	**Automation Level**	**Management Level**
X-Ray Technology	Backscatter X-ray Emitter and Receiver (narrow beam), Transmission X-ray Emitter and Receiver	Upon detection of organic or metal material in scanned volume, it is automatically coloured and presented to	The user interface and control panel which displays contents of scanned items and highlights/colours metals/organic materials.

	(vertical fan shaped beam).	human operator for assessment and response. The system pauses until acknowledged by operator.	

Indicative Analytical Methodology

X-Ray detection systems may employ a statistical (probabilistic) analytical methodology occurring at the automation and management levels, where detection of threat objects occurs as a result of manual or Automated Target Recognition (ATR) to analyse x-ray images. ATR images are analysed by image comparison, load characterisation or direct target detection with reference material obtained from an x-ray reference library (Open Access Government, 2016; SCENIHR, 2012).

Functionality and Role in Protection of Assets

X-Ray technology is commonly used as a scanning device at access and egress points at airports, postal hubs and ports and high security zones such as court houses to scan clothed individuals, baggage and packages to identify metal objects such as weapons and organic matter such as contraband food or drugs (Min, 2015; Liang et al., 2018; Gaus, Bhowmik, & Breckon, 2019). The system calculates material properties from dual energy multi view scans, highlighting objects that might contain explosives. X-ray images are inspected by trained human inspectors however it is generally accepted in the literature that considerable challenges arise from this process (Liang et al., 2018; Gaus et al., 2019).

Integration

A review of the literature shows limited integration of x-ray Technology. Min (2015) recognises that the use of x-ray may be limited without integration with a system-wide decision support tool. For example, in maritime security, Min (2015) highlights the need for further integration with intelligence systems that detail shipment or freight profiles based on shipping origin or surveillance information. Min (2015) discusses previous studies (Bertozzi & Ledoux, 2005) where the use of automated nuclear resonance fluorescence imaging technology was integrated with x-ray technology and the holistic detection process was able to be automated providing a 99% detection efficiency rate (Min, 2015, p. 5).

Artificial Intelligence in X-Ray Technology

Existing scanners calculate material properties from dual energy multi view scans, highlighting objects that might contain explosives. Although Min (2015) suggests that x-ray technology is limited in its application unless it is paired with intelligence to allow the imaging to be interpreted in a specific threat context, the application of Deep Learning techniques to computer aided security screening has shown positive results (Liang et al., 2018). Deep learning algorithms and architectures, such as Convolutional Neural Networks (CNNs) are currently under development for object detection and classification in security baggage imagery (Akcay, Kundegorski, Willcocks & Breckon, 2018; Gaus et al, 2019; Liu, Leng & Liu, 2019).

Limitations

X-ray images are typically inspected by trained human inspectors however there are accepted limitations within the inspection process (Liang et al., 2018; Gaus et al., 2019). Gaus et al (2019, p.1) discuss the complexity of screening items such as baggage in public places, and assert "complexity stems from the nature of compact, cluttered and highly varying baggage contents being examined within a limited timeframe. It can also be challenging due to inherent monotony leading to varying concentration levels". Lim, Cho, So, Min and Park (2016, p.5) highlight that in the maritime environment data extracted from 'stand-alone' x-ray scanning systems can be misinterpreted, misplaced, or misused which may impede rather than improve the efficiency of cargo inspection.

Concerns have also been raised over the potential health implications of using radiation for both individuals and technical operators.

1.3.2.2.4 Explosives/Narcotics Trace Detection

Definition

Trace detection captures and analyses the microscopic particles and vapours naturally emitted by organic substances (GE Homeland Security, 2008). The most common form of explosive detection is by mass spectrometry. However, other methods include ion chemiluminescence, mobility spectrometry and gas chromatography. Detection of explosives (and narcotics) by mass spectrometry is based on an atmospheric pressure chemical ionization source using the "separation and analysis of substances according to the masses of atoms and molecules of which the substance is composed" (Yinon, 2007, p.43).

Characteristics

Explosive/ Narcotic Trace Detection Systems and their indicative BACS architecture levels with AI alignment are presented in Table 7:

Table 7

Indicative Trace Detection System BACS Architecture Levels and AI Analytical Alignment.

EXPLOSIVES/ NARCOTICS TRACE DETECTION SYSTEMS			
	Field Level	**Automation Level**	**Management Level**
Explosives/Narcotic Trace Detection	Air Dispenser and Vapour Particle Collector, Swab Reader.	Upon detection of chemical compound in swab or air vapour volume, detail is presented to human operator for assessment and response.	The control unit that allows pre-programmed chemical detail of substances being scanned for (explosives, narcotics)

Indicative Analytical Methodology

Trace Detection Systems may employ a statistical (probabilistic) analytical methodology occurring at the automation and management levels, where detection of targeted compounds occurs as a result of electronic comparison of spectral signatures with a spectral reference library (National Research Council, 2004, p.19; Yinon, 2007).

Functionality and Role in Protection of Assets

Explosive trace or narcotics detection systems are used for screening of personnel, packages, vehicles, baggage, vehicles, mail and cargo (GE Homeland Security, 2008). Yinon (2007, p. 51) explains the operational process of an explosive detection personnel portal, where air is passed over the target being screened, which removes and collects and particles (e.g. explosive materials or narcotics). The air is then 'filtered to collect the particles, residue and condensable vapours in a pre-concentrator' which is then analysed using mass spectrometry and the results are displayed, listing any narcotics or explosives detected on the target being screened. Assessment and response are actioned by human operators.

Integration

The trace detection systems function as a standalone system with limited scope for integration in current practices. To date, there is currently limited literature to support widespread commercial use of AI-based explosive trace detection systems. However, Monea (2018) highlights that statistical algorithms for signal detection and estimation, spectral domain processing, adaptive filtering, image processing algorithms or advanced analysis and pattern recognition techniques based on machine learning are examples of AI computational techniques in trace detection.

Artificial Intelligence in Explosive Trace Detection

Homeland Security Research Corporation (2019) attribute the growing market in trace detection technologies to the implementation of new and maturing technologies such as AI and machine based automated interpretation. Whilst there is limited literature to support widespread commercial use of AI-based explosive trace detection systems there is evidence that supervised learning has improved the detection algorithm to be able to disclose the minute differences or organic explosives on polymeric surfaces (Serrano, Moros, Sanchez, Macias & Laserna, 2014). Additionally, machine learning has been found to be effective in detecting and classifying volatile compounds with a 90% accuracy achieved (Gao, Zhang, Wang, Diaz, Zhao & Willis, 2020).

Limitations

Currently there are no automated assessment and response mechanisms for trace detection systems, resulting in system reliance on human investigation and assessment (Liang et al., 2018; Gaus et al., 2019). False positive signals (sensors activated by harmless substances such as sulphur in peroxide products or chemicals in perfumes) and false negative signals (no detection is made despite the presence of a dangerous substance) may be an issue due to ultra-sensitivity of the system in use (Musil, 2017).

1.3.2.2.5 Acoustic Threat Detection/ Intelligent Acoustic Threat Detection

Definition

Acoustic Threat Detection may be defined as a system designed to "recognise the acoustic events that may inform us of possible threats" (Lopatka, Kotus, & Czyzewski, 2016). Acoustic Threat Detection systems may recognise sounds such as gunshots, explosions or screams. The sound recognition algorithms are based on the extraction of acoustic features and statistical pattern recognition (Zhuang, Zhou, Hasegawa-Johnson, & Huang, 2010).

Characteristics

Acoustic Threat Detection Systems and their indicative BACS architecture levels with AI alignment are presented in Table 8:

Table 8

Indicative Acoustic Threat Detection System BACS Architecture Levels and AI Analytical Alignment.

ACOUSTIC THREAT DETECTION SYSTEMS			
	Field Level	**Automation Level**	**Management Level**
Acoustic Threat Detection	Acoustic Sensor, Underwater Hydrophone Sensor.	Internal processor based on expert rules setting that upon detecting pre-programmed acoustic signatures displays positive detection detail for assessment and action and response by human	Control panel or user interface that allows the programming detail of acoustic features being listened for, for comparison purposes. Often integrated with GPS for location information and with

		operative. Often linked with automated GPS	RADAR/SONAR for underwater detection.
Intelligent Acoustic Threat Detection	Acoustic Sensor, Vibration/Seismic Sensor.	Automated comparison with pre-programmed sounds and deep learning algorithms. Internal processor based on expert rules setting that upon detecting pre-programmed acoustic signatures displays positive detection detail for assessment and action and response by human operative. Often linked with automated GPS	User interface employing expert rules setting and subsequent deep learning algorithms to allow machine learning of sounds and differentiation with "harmless" crowd/animal noises etc.

Indicative Analytical Methodology
Acoustic Threat Detection systems may employ both rules-based and statistical analytical methodologies, with basic acoustic detection systems using expert-rule setting, and Intelligent Acoustic Threat Detection systems using a statistical (probabilistic and machine learning) methodology. Expert Rule-setting occurring at the management level follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. the acoustic signal feature extraction matches the acoustic signal of a gunshot), THEN the command signal activates an alarm (Dutta & Bonissone, 1993). Intelligent Acoustic Threat Detection systems may use a statistical methodology (probabilistic and machine learning) to classify detected acoustic signatures based on training sets (e.g. supervised, unsupervised or machine learning) (Zöhrer & Pernkopf, 2016).

Functionality and Role in Protection of Assets
Acoustic threat detection is increasingly used for event recognition and detection, and may alert security response personnel to specific events such as the approach of humans or vehicles to a protected area, screams, gunshots or explosions (Dibazar et al., 2011; Lopatka et al., 2016). Systems may also identify other emergency situations where sounds such as a crying baby, cough, dog bark, the word "help", a siren and glass breaking are discernible (Romanov, Kharkovchuk, Sinelnikov, Abrash, & Filinkov, 2020). Acoustic threat detection is also commonly used underwater with hydrophones to protect maritime interests and ports from approaching foreign ships, submarines and divers (Borowski, Sutin, Roh, & Bunin, 2008).

Integration
Acoustic threat detection systems are commonly integrated with video surveillance systems (Valenzise, Gerosa, Tagliasacchi, Antonacci, & Sarti, 2007) to identify and discriminate acoustic events such as gunshots, screams and explosions. Systems may differentiate acoustic events from non-threatening ambient noise, and automatically initiate a response such as activating surveillance and alarm and access control systems for assessment and response (Berce, 2019). These systems may also be linked with GPS systems allowing for accurate location of the threat event (Valenzise et al., 2007; Romanov et al., 2020).

Dibazar et al. (2011) highlights the use of integrated acoustic threat detection in perimeter protection and intruder detection systems, combining acoustic, vibration and seismic phenomenology sensors to alert approaching vehicles, including identification of the vehicle type and distance from the target, discriminating between intruder footsteps and animals, and distinguishing between leaning on a fence and intentional climbing. Additionally, Borowski et al. (2008, p.1) discusses the use of acoustic threat technology underwater in combination with surface SONAR/RADAR detection to protect ports and maritime interests from sea-based invasions and terrorist threats.

Artificial Intelligence in Acoustic Threat Detection

AI use in Acoustic Threat Detection systems has increased significantly with the technology being developed for use in a variety of applications. While AI-based Acoustic Threat Detection is predominantly described within the literature as being used for hazardous event recognition (gunshots, explosions etc), analysis and classification of acoustic signals are being applied to detect a greater range of acoustic events and signatures (Baggenstoss, Springer, Oispuu & Kurth, 2019; Dibazar et al, 2011; Lopatka et al, 2016). For instance, acoustic detection, tracking and classification of approaching unmanned aerial vehicles (UAVs), or drones, is generating significant interest. Acoustic detection is advantageous when compared with infrared, radar or electro optic systems as these alternative systems may present weak signatures (Baggenstoss et al, 2019; Busset, Perrodin, Wellig, Ott, Heutschi, Ruhl, & Nussbaumer, 2015; Case, Zelnio & Rigling, 2008).

Dibazar et al. (2011) discusses the “smart fence” using acoustic threat detection sensors that can differentiate characteristics of threat events. For instance, sensors may characterise threat level based on the proximity and actions of adversaries, where distant vehicles may be designated a lower threat level than intruders approaching a secure perimeter, and an intruder attempting to scale a fence may be designated as a higher threat level (Dibazar et al., 2011, p.1).

Limitations

Zu et al. (2017) identify one of the main limitations of acoustic threat detection as the propensity for acoustic sounds to be distorted by environmental factors such as wind, road conditions, terrain and multipath. Subsequently, sensor sensitivity requirements and pre-programmed sound accuracy including sample diversity is key to ensuring accurate detection (Zu et al., 2017).

1.3.2.2.6 RADAR Systems

Definition

Radio Detection And Ranging, referred to as ‘RADAR’, is defined as ‘a device or system consisting usually of a synchronised radio transmitter and receiver that emits radio waves and processes their reflections for display and is used especially for detecting and locating objects (such as aircraft) or surface features (as of a planet)’ (Merriam-Webster, n.d.).

RADAR systems emit various frequencies and wavelengths (electromagnetic waves) which are digitised and recorded as waveforms to measure both the range and velocity of objects in the monitored space. The range measurement is determined by the time required for the signal to be reflected back to the receiver, and velocity is measured by the phase shift between emitted and reflected signals (Earth Observing Laboratory, n.d.; Hughes, 2011; Isoardi, 2012). This phase shift, or change in frequency of the returned signal, is known as the ‘Doppler Effect’. The Doppler effect refers to ‘the apparent difference between the frequency at which sound or light waves leave a source and that at which they reach an observer, caused by relative motion of the observer and the wave source’ (Encyclopaedia Britannica, n.d.).

Though RADAR technology has traditionally used for military and government applications, including target detection and recognition, weather observation, and ground-based surveillance, the technology is currently being incorporated into commercial security systems for surveillance and intrusion detection. Security applications generally use RADAR technology for perimeter and site protection, and detection of persons entering (or within) a monitored space including underground (Ground Penetrating Radar) and behind/within barriers such as walls (Through-the-Wall Sensors) (Butcher & Fretwell, 2009; Elprocus, n.d.; Hayes & Ericson, 2014; Lombardi, Lualdi, Picetti, Bestagini, Janszen & Di Landro, 2020).

Characteristics

There are many types of commercially available RADAR systems, available with several types and combinations of radiated waveforms used in security systems. Radiated waveforms include Continuous Wave (CW), Frequency Modulated Continuous Wave (FMCW), Pulsed, Pulsed Doppler, and Pulse Compression, with each type being suited to different environments and applications (Federation of American Scientists [FAS], n.d.a.; Hughes, 2011). RADAR systems and their indicative BACS architecture levels with AI alignment are presented in Table 9:

Table 9

Indicative RADAR Detection System BACS Architecture Levels and AI Analytical Alignment.

RADAR TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
RADAR Video Surveillance (RVS)	RADOME casing; Transmitter; Waveform generator; Power amplifier, oscillator & wave carrier; Receiver; CPU; Power Supply; Synchroniser; Antenna, parabolic reflector and planar arrays; Display unit/ interface; Systems integrated with video include network video components.	Transmission, Detection and Processing of RADAR signal; Signal analysis (object detection/ classification); Regulation of rate at which pulses are sent; Antenna Orientation; Duplexer function (alternative connection between transmitter/ receiver); GPS; Target location & tracking; Systems integrated with video may include auto function e.g. slew-to-cue, application updates etc; Adaptive Radar	Command and control; Output selection of frequency and pulse; Range determination; Operational mode (e.g. 360° continuous); Display mode; Programming (start/stop, sweeping pattern, blanking zones, scheduling); Systems integrated with video may include auto function e.g. PTZ, alarm assessment, scene investigation etc.
Ground Penetrating Radar (GPR)	GPR unit hardware & casing; Waveform generator; Power amplifier, Transmitter & receiver; CPU; Power Supply; Synchroniser; Antennae, Display unit/ interface	Transmission, Detection and Processing of RADAR signal; Signal analysis; Alarm, display and profiling outputs; Terrain object detection, modelling and classification	Command and control; Output selection of frequency & waveform; Depth determination; Operational mode (e.g. reflection or tomographic); Display mode
Through-the-Wall Sensor (TTWS)	TTWS unit hardware & casing; Waveform generator; Power amplifier, oscillator & wave carrier;	Transmission, Detection and Processing of RADAR signal; Signal analysis; Alarm, display and	Command and control; Output selection of frequency and pulse; Range determination; Operational mode (e.g.

	Transmitter; Receiver; CPU; Power Supply; Synchroniser; Antenna, Wireless Communications; Display	profiling outputs; Remote wireless transmission	remote, stand-off mode); Display mode; Programming (start/stop, sweeping pattern, blanking zones, scheduling);

Indicative Analytical Methodology

RADAR systems may employ both rules-based and statistical analytical methodologies, with the basic RADAR system using expert-rule setting, and intelligent signal processing using a statistical (probabilistic and machine learning) methodology. Expert Rule-setting occurring at the management level follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. transmission returns a signal), THEN the signal is processed to depict and display the target range & location (Dutta & Bonissone, 1993). Intelligent RADAR systems may use a statistical methodology (probabilistic and machine learning) to detect, model and classify detected objects, and in some cases adjust the transmitted signal to optimise returned signals (Haykin, 2006). Cognitive RADAR systems are capable of up to semi-autonomous decision making, in some cases (i.e. military), may operate at a sub-symbolic level.

Functionality and Role in Protection of Assets

RADAR technology for asset protection is typically used for perimeter or area surveillance, detection, and tracking, although the technology may be incorporated into a range of commercial security system types, such as RADAR Video Surveillance (RVS), Ground Penetrating Radar (GPR) and Through-the-Wall Sensors (TTWS) (Butcher & Fretwell, 2009, p.9; Fielding & Torbett, 2017; Hayes & Ericson, 2014; Hughes, 2011). Commercial and industrial applications may vary, with systems being used for port and waterways, airport security, border security, oil platforms, among other sites and facilities.

Functionality of RADAR based security systems includes wide area and site perimeter monitoring with extension of the perimeter well beyond physical boundaries, day/ night operability, all-weather detection, multiple target location and tracking, fusion of monitored tracks, flexibility in monitoring corners and nonlinear perimeter sections, and alarm or event generation. RVS systems may use target location data to either direct the operator to use the closest surveillance camera for target verification or slew-to-cue (Butcher & Fretwell, 2009, p. 9; Fielding & Torbett, 2017). Systems may be custom designed with options including mobile or static, manned or unmanned, vehicle mounted, generator or solar powered, and options for communication of data by microwave or satellite links (Kelvin Hughes, n.d.).

TTWS sense the presence of persons through walls and obscure barriers (concrete, wood, brick, glass etc) and may be used for law enforcement, tactical, emergency and rescue operations. Accuracy of results depend on the barrier's material and thickness, though generally these systems provide either one-dimensional or two-dimensional information such as the shape, layout and presence of solid structures within a room and the range and status of multiple targets. TTWS may also be sufficiently sensitive to detect whether a person is in a front or rear room of a building, and detect the motion of a person's chest upon inhale/exhale enabling these systems to be used to locate sleeping, injured or hidden persons (Hayes & Ericson, 2014).

Integration

RADAR technology has been successfully integrated into numerous commercially available ground-based, perimeter and drone surveillance and detection systems. Ground movement RADAR for site security is particularly effective when used in conjunction with digital surveillance and intrusion detection systems (Butcher & Fretwell, 2009, p.9). However sophisticated modern systems may also integrate thermal imaging cameras, long-range cameras, acoustic hailing/deterrence devices and other instruments such as distance measuring tools (Kelvin Hughes, n.d.).

While there is limited evidence to suggest security-based RADAR systems are currently integrated with biometric systems, combinations of RADAR and optical sensors may allow biometric data to be extracted, particularly where high-resolution images are produced (Hägelen, Kulke, Jetten & Möllenbeck, 2016). This suggests RADAR technology may be successfully incorporated into integrated security systems, thus allowing for further development and consolidation of RADAR with biometrics.

Artificial Intelligence in RADAR Systems

Artificial intelligence is being embedded in commercial RADAR-based technologies and devices across a range of industries and applications including air-traffic control, perimeter surveillance, medical appliances, manufacturing, military, and security (Turner & Scholten, 2019; Zussman, 2018). Advances in signal processing such as those seen in Adaptive RADAR and 'Cognitive Radar' aim to produce a higher level of sensing adaptation (Wilson, 2017).

Deep learning and machine learning algorithms may be applied to the signals received and processed by RADAR systems. For instance, RADAR integrated with RVS & GPR may be guided with training sets and classification descriptors (Kobashigawa, Youn, Iskander, & Yun, 2011). Subsequently, commercial security technologies incorporating RADAR with what is referred to as AI algorithms are becoming increasingly available, with systems such as ground-based RADAR and counter-drone RADARs providing high-level security solutions for airports, prisons and other facilities and complexes (Coulombe, 2019; Martelli, Bongioanni, Colone, Lombardo, Testa & Meta, 2016; SpotterRF, n.d.).

Limitations

RADAR technology has several inherent limitations, including range and bearing discrimination, range and bearing accuracy, and anomalous propagation among other technical parameters and issues. Limitations such as sensitivity to jamming, processing times and complexity may also vary depending on the types and combinations of waveforms used within the RADAR system (Hughes, 2011). Thus, the security application and environment require consideration to evaluate the overall impact of system limitations based on the manufacturers specifications and system capabilities.

Specific RADAR-based security technologies and devices also have unique limitations impacting on security functions, which may reduce the efficacy of the technology's performance in certain environments. For instance, GPR signals may be blocked or confused by metal materials used for tensile reinforcement of concrete slabs, as well as clay or salt contaminated soils. Similarly, TTWS may fail to detect a person standing against a barrier wall or hiding behind metal objects such as refrigerators (Dynamic Excavations, 2019; Hayes & Ericson, 2014). Furthermore, as TTWS detect motion, these devices may also be triggered by the movement of inanimate objects (fans, curtains etc), which may impact on operational decisions and outcomes (Hayes & Ericson, 2014).

1.3.2.2.7 SONAR Systems

Definition

Sound Navigation And Ranging, referred to as SONAR, is a 'technique for detecting and determining the distance and direction of underwater objects by acoustic means' (Encyclopaedia Britannica, n.d.). The operation of SONAR closely resembles RADAR but emits and receives acoustic waves rather than electromagnetic waves (Federation of American Scientists [FAS], n.d.b.). SONAR systems transmit signals which propagate toward an object or target and are then reflected by the object or target and recorded by a receiver (FAS, n.d.b.; Hansen, 2011).

SONAR technology is being incorporated in modern security systems for monitoring coastal waterways and inlets, harbour and port security, or for underwater surveillance and detection in offshore locations such as oil rigs (DSIT Solutions, n.d.). SONAR technologies may be used for surveillance or to detect and classify objects for monitoring and assessment.

Characteristics

Underwater SONAR has two forms of architecture – 'Active' and 'Passive'. 'Passive' SONAR does not emit signals, it detects, and records SONAR waves converted to an electrical signal. Active SONAR operates by transmitting and receiving signals. SONAR waves are converted to and/or from electrical signals by a transducer (hydrophone array). The emitter transducer (T_{dx}) and receptor transducer (R_{dx}) emit and detect acoustic energy reflected from a sound/energy source. SONAR signals are emitted in pulses, with commonly used pulse forms including Gated Continuous Wave (CW), Linear Frequency Modulated (LFM), Hyperbolic Frequency Modulated (HFM) and Pseudo Random Noise (PRN) Binary Phase-Shift Keyed (BPSK) Coded CW (Collins & Atkins, 1998; Hansen, 2011; Jiang, Yan, Wu & Ma, 2016).

Active systems use the audible frequency range (50 Hz – 20 kHz) but may use higher frequencies in some instances resulting in the range being between 50 Hz and 50 kHz (Yusof & Kabir, 2012, pp. 307-309). By measuring the time between emission and reception of the acoustic pulse and the strength of the returned signal, the transducer can determine the range and orientation of the target (Yusof & Kabir, 2012). SONAR systems and their indicative BACS architecture levels with AI alignment are presented in Table 10:

Table 10

Indicative SONAR System BACS Architecture Levels and AI Analytical Alignment.

SONAR TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Active SONAR	SONAR casing; Transmitter; Transducer array; Beamforming processor; Duplexer, Synchroniser; Receiver and Display Unit; CPU; Power Supply	Transmission, Detection and processing of SONAR signal; Signal analysis (object detection/ classification); Pulse regulation; Antenna Orientation; Duplexer function (alternative connection between transmitter/ receiver); GPS; Target location & tracking; Integrated systems may include auto function e.g. slew-to-cue, application updates etc; Adaptive Sonar	Command and control; Output selection of frequency and pulse; Range determination; Operational mode (e.g. 360° continuous); Display mode; Programming (start/stop, sweeping pattern, blanking zones, scheduling); Systems integrated with video may include auto function e.g. PTZ, alarm assessment, scene investigation etc.
Passive SONAR	Hydrophone array; Beamforming processor; Broadband	Detection and processing of SONAR signal; Signal analysis	Command and control; Operational configuration (target data, exclusion

	display; Narrowband display; Frequency analyser	(object detection/ classification); GPS; Automated, detection, target location & tracking, classification and alarm; Display output (SONAR overlay on aerial/ satellite images, nautical charts)	zones, 120º, 240º, 360º coverage, scheduling); Command & control of display mode (SONAR overlay on aerial/ satellite images, nautical charts)

Indicative Analytical Methodology

SONAR systems may employ both rules-based and statistical analytical methodologies, with the basic SONAR system using expert-rule setting, and intelligent signal processing using a statistical (probabilistic and machine learning) methodology. Expert Rule-setting follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. transmission returns a signal), THEN the signal is processed to depict and display the target range & location (Dutta & Bonissone, 1993). Intelligent RADAR systems may use a statistical methodology (probabilistic and machine learning) to detect, model and classify detected objects, and in some cases adjust the transmitted signal to optimise returned signals. Adaptive SONAR systems are capable of up to semi-autonomous decision making and may operate at a sub-symbolic level in some instances (i.e. military applications) (Li, Li, Cui & Lui, 2011).

Functionality and Role in Protection of Assets

SONAR security systems for critical installations such as ports, harbours, coastal waterways or oil rigs are typically designed to function as surveillance and detection systems, although due to integration capabilities, may be multi-functional in some instances. Commercially available systems include temporary, permanent, and portable deployments for surveillance and detection of a range of targets including swimmers, closed circuit divers, open circuit divers, unmanned underwater vehicles or other submersibles, as well as small surface vessels (e.g. kayaks etc) (DSIT Solutions, n.d.; Sonardyne, n.d.). Automatic systems may include detection, location of target, tracking, classification and alert, with real time and continuous monitoring, recording and playback.

Commercial systems are reported to have a low false alarm rate (FAR) and may be used in shallow or deep water, day or night and in all weather conditions (DSIT Solutions, n.d.). Specific environmental and operational conditions will determine range requirements which may be anywhere from a few hundred metres to several kilometres (Sonardyne, n.d.). Currently available commercial systems are likely to accommodate or exceed user range requirements, as powerful SONAR waves can be detected over a distance of up to 500km using a high-power active wave system (Yusof & Kabir, 2012).

Integration

Developments in technology have assisted the crossover of military technologies into commercial markets, resulting in SONAR technology being integrated with diver detection, surveillance RADAR and electro-optical devices (DSIT Solutions, n.d.). Integration capabilities vary substantially between products and manufacturers, and although there is limited evidence to suggest any standardisation of commercial SONAR security systems, existing products are reportedly capable of being integrated into widely used and multi-layered security surveillance systems, (DSIT Solutions, n.d.; Fortem, n.d.; Sonardyne. n.d.).

Artificial Intelligence in SONAR Systems

While SONAR systems developed for military applications may have AI algorithms embedded into their signal processing capabilities, there is currently limited evidence of AI in commercially available SONAR security systems. Commercial systems may offer embedded algorithms which are able to provide classification of targets, such as discrimination between fauna, flora and real threats (Sonardyne, n.d.), though the analytical technique may be statistical (probabilistic) rather than true 'machine learning'.

Despite current limitations in commercial systems, digital signal processing is being explored for amalgamation with AI across a range of technology types, including SONAR. For instance, 'Adaptive SONAR' (and RADAR) use adaptive waveforms to optimise their capabilities in a changing environment. Adaptive SONAR systems 'capitalize not only on enabling technologies like AESA, back-end software control, and processing, but also on enabling technologies for multifunction systems that blend communications, electronic warfare, detection and tracking, and weather forecasting' (Wilson, 2017).

Limitations

There are several limitations of SONAR systems relating to both the technical/ operational capacity of the system to obtain an adequate signal, and the functional capacity for these systems to adequately identify and classify a target. Operational limitations such as 'quenching' and 'reverberation' are common SONAR issues, and pulse waves emitted by Active SONAR systems have a number of inherent technical limitations. Environmental conditions may also affect both Passive and Active systems. Reflected signals from underwater surfaces (rocks, fish etc) create 'clutter' which degrades the performance of SONAR systems. High levels of ambient noise (hydrodynamic, seismic, biological or man-made), poor contrast and image instability (created by alternating environmental conditions and sensitivity of acoustic waves) may result in poor detection and imaging in terms of target isolation and pattern recognition, referred to as 'classification' (Hansen, 2011).

1.3.2.2.8 Detecting Technologies Overview

Detection technologies are widely used for the protection of assets. Understanding current and future optimisation of AI in detection technologies can be achieved by reviewing how current technologies align to the spectrum of Artificial Intelligence (AI) Paradigms. Table 11 presents a summary of current detection technology and indicative AI analytical alignment:

Table 11

Detecting Technologies and Indicative AI alignment

Technology		Function	Stimuli	Example	Indicative Analytical Methodology
Intrusion Detection Systems	Contact Switch/ Reed Switch (Electromechanical Sensor)	Detection	Current Interference (loss) through breaking of connection between the two points	Reed switches on windows or doors to detect opening	Expert Systems (Rule Setting)

	Glass Break Sensor (Electro-mechanical Sensor)	Detection	Vibration or acoustic technology – elevations in vibration or audio levels across a surface	Glass break sensors used on windows to detect glass being broken or manipulated	Expert Systems (Rule Setting)
	Passive Audio Sensor	Detection	Elevations in audio levels	Sensors detecting noise caused by intruder movement through detection volume	Expert Systems (Rule Setting)
	Passive Microphonic Cables	Detection	Detection of vibrations along line	Microphonic cables used to detect fences being cut or climbed on	Expert Systems (Rule Setting)
	Photoelectric Beams (visible or infrared)	Detection	Interruption or change of amount of light crossing light axis and arriving at receiver unit	Photoelectric beams used to detect fences being cut or climbed on	Expert Systems (Rule Setting)
	Active Infrared Motion Sensor	Detection	Interference of infrared beam between two points	AIR activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Fibre Optic Cable Sensors	Detection	Loss of signal due to damage or break in fibre loop	Activated when cable is damaged or cut – used to protect CCTV and other systems cabling	Expert Systems (Rule Setting)
	Microbend Fibre Optic Sensor	Detection	Loss of signal due to bend or movement in fibre loop	Activated when cable is bent or moved - used to protect CCTV and other systems cabling	Expert Systems (Rule Setting)
	Passive Infrared Motion Sensor	Detection	Increased Change in thermal energy	PIR activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Active Ultrasonic Motion Sensor	Detection	Changes in reflected Ultrasonic energy	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)

	Active Monostatic Microwave Motion Sensor	Detection	Changes in reflective electromagnetic energy wave	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Passive Infrared Sensors (Thermopile or Pyroelectric Detector)	Detection	Changes in energy emitted by human intruder moving through background thermal energy	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Capacitance Proximity Sensor	Detection	Change in capacitance (electrical charge) between protection loop and ground	Activated when an intruder moves near to or touches metal object	Expert Systems (Rule Setting)
	Pressure Sensor	Detection	Electrical contact between ribbon switches due to pressure	Activated when intruder passes over sensor entering protected area	Expert Systems (Rule Setting)
	Light Sensors	Detection	Change in predetermined light level	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Dual Technology Sensors	Detection	AND gate logic requiring activation of both types of combine sensors	Combination sensors used to overcome vulnerabilities of above individual sensors	Expert Systems (Rule Setting)
	Alarm Annunciator Panel	Detect, Response	Inputs from field sensors communicate audio visual signals for faults any failures of the critical system or process	Monitoring of door status at guarding station	Expert Rule-setting
	Duress Alarms	Detect, Response	Command Signal; Emergency signal transmitted from duress mechanism to central monitoring station	Magnetic reed switch activated upon gun being removed from holster of armed guard	Expert Rule-setting
Lighting Systems	Active Lighting Systems	Detection	Activation of sensors (PIR, Infrared,	Used to automatically illuminate an area	Expert Systems (Rule Setting). Probability (Past Performance)

			Ultrasonic, Microphonic)	upon activation of sensor	
	Passive Lighting Systems (Failsafe/ Emergency Lighting)	Detection	Failure of electrical system	Used to automatically illuminate an area upon failure of electricity	Expert Systems (Rule Setting)
X Ray	X-Ray	Detection	X-ray electromagnetic wave interaction with material	Used to detect presence of non-organic matter on human body	Probabilistic – Image Recognition and Comparison
Trace Detection	Explosive/ Narcotics Trace Detection (Ion Mobility Spectrometry)	Detection	Change in relative concentration of explosive molecules in saturated air	Used to detect explosives in parcels and luggage	Expert Systems (Rule Setting). Probability (Past Performance)
Acoustic Detection Systems	Acoustic Anomaly Detection	Detection	Audio sensors detect acoustic signature of pre-defined actions in each detection volume	Used to identify gunshot or aggressive behaviour	Comparative analysis using deep learning algorithms and inference
	Intelligent Acoustic Surveillance (Acoustic Event Detection)	Detect	Nominal and Anomalous Acoustic Input	Deployed where there may be privacy issues with video surveillance (e.g. bathroom facilities) to 'learn' usual noise and detects abnormal noise (e.g. scream)	Statistical Probability (Probabilistic Method) AND Statistical Machine Learning (Supervised, Unsupervised or Reinforcement Learning)
RADAR Systems	Radar Video Surveillance (RVS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR), Reflected Scene Image (Video)	Radar video system observes and detects an individual intruding into a defined space (e.g. rural airport)	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
	Ground Penetrating Radar (GPR)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR),	GPR transmits electromagnetic waves to detect underground shelters, landmines or buried explosives	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
	Through-the-Wall Sensor (TTWS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR),	TTWS device transmits RF energy through a	Expert Systems (Rule Setting) and

				wall and monitors the reflected energy to detect person moving in a room	Statistical Probability (Probabilistic Method)
SONAR Systems	Active Sonar Systems	Detect, Recognise	Transducers (hydrophone array) emit and detect acoustic energy reflected/ emitted from sound/energy source	Sonar deployed in a port/harbour to detect underwater divers or vessels approaching docked vessels	Expert Systems (Rule Setting)
	Passive Sonar Systems	Detect, Recognise	Transducer (hydrophone array) detects acoustic energy emitted from sound/energy source	Passive Sonar deployed by an offshore oil rig to detect approaching vessels	Expert Systems (Rule Setting)

1.3.2.3 Controlling Technologies

The function of controlling is usually associated with access and egress control in the protection of assets. Access control is described by ASIS (2015, p. 17) as any technique that is employed to limit or otherwise control access to an area, facility, compound, system, person or asset. This limiting and controlling of an environment is achieved electronically through the broad category of "access control systems" using several distinct and often integrated technologies.

1.3.2.3.1 Access Control Systems

Definition

According to Norman (2011, p. 17) electronic access control systems "facilitate automated approval for authorized personnel to enter through a security portal without the need for a security officer to review and validate the authorization of the person entering the portal, typically by using a credential to present to the system to verify their authorization".

Characteristics

Typical electronic access control systems have two main components; a system that verifies the presence of the authorized personnel, and a second that controls the access to the physical location. Access control systems contain field level components such as credential readers (keypads, card readers, biometric scanners etc.) located at the point of entry, connected through communications cabling to processors and central databases and request-to-exit devices (ASIS International, 2015, p. 241). Access control system components and their indicative BACS architecture levels with AI alignment are presented in Table 12:

Table 12

Indicative Access Control System BACS Architecture Levels and AI Analytical Alignment.

ACCESS CONTROL TECHNOLOGIES			
	Field Level	Automation Level	Management Level
Access Control Systems	Card readers (EAC, Wiegand, Smart, 2FA, RFID, NFC), IoT IP address readers Biometric Readers.	Upon presentation of approved credential, automated granting of access through door locking systems, gate systems, turnstiles, lift calling units, automatic barriers.	Control panel; expert rules input and domain for pre-programmed access granting and denial rules. With Biometrics, deep learning algorithms embedded.

Indicative Analytical Methodology

Access control systems may employ both rules-based and statistical analytical methodologies, with physical credentials using expert-rule setting, biometric access control systems using a statistical (probabilistic) methodology. Expert Rule-setting occurring at the management level follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. a pre-approved credential is presented), THEN access is granted (Dutta & Bonissone, 1993). Biometric access control systems may use a statistical methodology (probabilistic) where enrolled credentials (e.g. fingerprint) are assessed on matched score to produce a measure of similarity (Dunstone & Yager, 2009; Vijaya & Kumar, 2011; Purpura, 2019).

Functionality & Role in Protection of Assets

The authorisation mechanism for access control systems primarily works on the basis of the validation of one or more authorized credentials, described by ASIS (2015, p. 241) as something that you know (for example a PIN), something inherent to you (for example biometric data – fingerprints, retinal scan etc.) or something carried (for example a key card). Furthermore, combinations - or layering - of these credentials allow for effective access protection of a facility. For instance, having a valid PIN or possessing a key card indicates the user has acquired that information or object, however, to overlay that with inherent personal data such as a fingerprint or hand geometry ensures that the bearer is in fact the authorised user.

Early electronic access control systems used a combination of active and passive technologies, including PIN keypads and access card technologies such as magnetic strip cards, barium ferrite technology, Hollerith, rare-earth magnet, and early Proximity technology (Norman, 2011). Recent advances include more advanced 125KHz Proximity card readers, 13.56 MHz contactless smart cards, remote access and smart locks capable of being controlled by mobile phone or other edge devices, RFID wireless transmitter systems and biometric readers. Current physiological biometric readers include DNA, facial recognition, hand geometry, fingerprint and iris recognition, retinal scanning and blood vessel identification (Di Nardo, 2009). Behavioural biometrics such as voice print, voice recognition, walking gait and typing rhythm are now also widely available (Purpura, 2019) using deep learning algorithms.

Integration

According to IHS Markit's 2018 "Smart Buildings Report", electronic access control systems are the main type of security system being integrated into Building Management Systems, with nearly 30% of all access control equipment installed and connected to these new integration platforms (Montany, 2018). The integration of an access control system into a Building Management System (BMS) can enhance the authentication decision by taking into account the information from other connected building management or security systems; for example, using surveillance system inputs to verify the identity of people presenting access cards at door readers. Access control integration is widely used in the protection of assets as access control systems are used to initialise other BMS processes; providing the starting command that tasks the BMS to initiate other outputs associated with the wellbeing or security of occupants. For example, the activation of HVAC and automatic lights in designated areas, lifts only travelling to pre-authorised levels (Brooks et al., 2017).

Artificial Intelligence in Access Control Systems

Recent advances in 'deep learning' and artificial intelligence are facilitating advancements in automated access control technologies. For instance, facial recognition technology is leveraging 'deep learning' and AI to evolve without the need for external updates. Technology developers claim 'deep learning' systems may evolve, and update thirteen times faster than face recognition software updated externally. Consequently, deep learning is argued to be a pioneering technique used to optimise AI by creating an artificial neural network which is then trained to recognise faces by presenting the system with millions of facial examples. This allows the system to quickly become more experienced and accurate at facial recognition than any human or existing technology.

According to Security Magazine (Brown, 2019) the use of artificial intelligence is making the management of access control more efficient due to the ability of artificial intelligence to take into account all information across integrated systems simultaneously. Such a statement makes note of integration as an AI multiplier. According to Brown (2019) "an AI-powered security system can always "see and hear" everything from everywhere, correlating disparate data across the environment to identify when an employee might pose a risk with ill intent. This leaves no gaps or missed opportunities."

Furthermore, electronic access control systems have been transformed to be risk-adaptive with the use of AI. An AI-powered access control system can identify anomalous events, insider threats, hazardous situations and dynamically change access permissions. For example, the application of AI to access control allows the system to identify unusual activity such as off-hours access and/or abnormal location access and combine it with other threat indicators to quickly identify insider threats. Intelligent access control systems may prevent people from entering an area where a traditionally "non-obvious" danger exists, and can allow a first responder with the right credentials to access a closed area when threatening conditions do exist (Brown, 2019).

Limitations

According to Purpura (2019, p. 210) biometric matching accuracy can be affected by environmental factors such as sunlight, impacting on facial scans, or a dirty environment may affect fingerprint readings; a retinal scan may be mismatched due to ocular disease; drug use can affect gait and slur voice patterns, and the human aging of more than 6 years reduces facial recognition. Spoofing of the biometric technologies is also possible as fingerprints can be imitated by lifting a latent print of an authorized personnel using sticky tape, or even a fake finger moulded out of playdough; glass eyeballs, contact lenses or high resolution photographs may spoof retinal scanners (Di Nardo, 2009; Purpura, 2019, p. 210).

1.3.2.3.2 Controlling Technologies Overview

Understanding current and future optimisation of AI in controlling technologies can be achieved by reviewing how current technologies align to the spectrum of Artificial Intelligence (AI) Paradigms. Table 13 presents a summary of current controlling technologies and AI analytical alignment.

Table 13

Controlling Technology and Indicative AI Alignment

Technology	Function	Stimuli	Example	Indicative Analytical Methodology
EAC Reader	Physical Access Control Functions	Presentation or input (e.g. PIN) of authorised credential at reader	Input of a PIN into a PIN pad reader	Expert Systems (Rule Setting)
Smart Card	Authentication Access Control Technology	Presentation of authorised credential at reader	Use of access control card containing pre-programmed identity information (access rights)	Expert Systems (Rule Setting)
2FA (Two Factor Authentication) Access Control	Authentication Access Control Technology	Presentation of two pre-authorised credentials. Combination of one Knowledge factor (e.g. PIN), one Possession factor (e.g. Smart Card) or one Inherent factor (Biometric credential)	Access control point requiring a key card and a PIN	Expert Systems (Rule Setting)
Biometric	Authentication Access Control Technology	Presentation and conversion of biometric data into algorithm that measures probability of match with pre-enrolled authorised data	Retinal eye scan/ Fingerprint reader	Probabilistic Programming - Algorithm Probability
RFID (Radio Frequency Identification)	Communication Access Technology	Presentation of RFID embedded chip to RFID reader based on pre-set permissions	Use of RFID Access card at an RFID reader	Expert Systems (Rule Setting)
NFC (Near Field Communication)	Communication Access Technology	Generation of electromagnetic fields created when two devices are brought together (e.g. smart phone & reader)	Using Smart phone and device reader to gain access	Expert Systems (Rule Setting)
Internet of Things Access Control (Intelligent Door Locking Systems)	Access Technology Architecture	Presentation of authorised credential at any one of connected access control readers using unique IP addresses and connected through the internet		Expert Systems (Rule Setting)

Bluetooth Access Control	Communication Access Technology	Presentation of Bluetooth enabled and pre-authorised device at Bluetooth reader using Wi-Fi or cellular data network		Expert Systems (Rule Setting)

1.3.2.4 Response Technologies

According to ASIS (2015, p. 17) "responding is the effort to neutralize, contain, or mitigate an event. It may also include an assessment that the event does not require immediate action." In the protection of assets, response typically includes the human-guard force intervention underpinned by human decision making. However, 'Response' as a function within the security domain has evolved considerably beyond the human guard-force intervention to a response mechanism which includes a range of response technologies. As part of this evolution, and particularly with high-value assets, advances in technology are combining delay and response elements into automated response technologies in lethal and less-lethal forms. Examples of these technologies include remotely operated weapons systems (ROWS) and millimetre wave technologies which use high-frequency electromagnetic waves to stimulate the nerve endings in an adversary's skin (Garcia, 2008).

Current *Security Response Technologies* may be grouped into the following broad categories: Communications, Dispensable Barriers, Vehicle Barriers, and Weaponry Systems. Each of these individual technologies are comprised of numerous elements or components which may have varying technical outputs, as well as having varying levels of control applied.

1.3.2.4.1 Communications Technologies

Definition

Communications systems are integral to the response function and include a range of devices and components such as telephones (including pagers and emergency phones), radio systems, public address systems, intercom systems, speakers/microphones and mass notification systems (Craighead, 2009, pp. 317-320; Garcia, 2001, pp.223-238).

Characteristics (Automation Architecture of Communications Technologies)

Communications technologies and their indicative BACS architecture levels with AI alignment are presented in Table 14:

Table 14

Indicative Communication Technologies BACS Architecture Levels and AI Analytical Alignment.

COMMUNICATIONS TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Telephones, Mobile Phones, Pagers, Emergency Phones	Handset, battery, transmitter, receiver, speaker, microphone, flashing lights/ strobes for emergency phones	Alarm signalling for emergency phones, automated call recording/answering/ redirection/ messaging,	System configuration & programming (e.g. answering/ call recording, call redirection, Integration with IoT devices

		Integration with IoT devices	
Radio Systems (e.g. two-way radio)	Transmitter, receiver, antenna, speaker, microphone, battery	Encryption, Designation of Spread-Spectrum frequency assignment, Integration with IoT devices	System configuration & programming, Selection of Frequencies, Talk-Groups, Digital Software
Public Address / Mass Notification System	Microphone, amplifier, loud-speakers, IP controllers, cabling or wireless transmitters/ receivers, monitor for interface and handset/station	Fault monitoring and bypass, ambient noise sensing and Automatic Level Control (ALC) etc	System configuration & programming through PA system (e.g. tones & scheduled announcements), grouping of PA zones, etc.
Emergency Warning Intercom Systems (EWIS Systems)	Master and sub-station units, IP Controllers, video display, speakers, microphones, wireless transmitters, receivers or cabling	Integration and automation with control systems (such as CCTV & access control) to allow remote operation and control. Signal processing with automated evacuation instructions	System configuration & programming through system interface (e.g. remote notification and operation, language, evacuation zones etc)

Indicative Analytical Methodology

Communications technologies are predominantly logic-based, with analytical methodologies at the automation level being rules-based with Expert Rule-setting. Expert Rule-setting occurring at the management level follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. a Public Address (PA) system processor detects ambient noise has reached a predefined threshold), THEN automatic gain control is applied to optimise volume output (Alango Technologies & Solutions, n.d.; Dutta & Bonissone, 1993).

Functionality and Role in the Protection of Assets

Typically, security communications systems are used for the control and coordination of security personnel within a facility although their use may extend to and benefit other employees, contractors, visitors or personnel within the facility. The function of communications systems is primarily to facilitate response forces (i.e. security guards, security managers and officers, building maintenance personnel, devices and systems) to communicate, share and transfer information with speed and reliability (Garcia, 2001, p.228; Norman, 2014, p.110). Specific communications systems include a range of functions. For example, two-way radios may enable 'Talk groups' to be created with communications software for groups of selected individuals, such as departmental groups, to enable private conversations or increase productivity by including only relevant personnel (Norman, 2014).

Integration

Communications systems may integrate a range of devices and systems, such as two-way radios, telephones, mobile phones, public address systems, megaphones, intercoms and speakers into a Consolidated Communications System (CCS) (Craighead, 2009, p.236; Norman, 2014, p.99). Communications systems, together with other security systems (e.g. access control, lighting, monitoring etc), are typically monitored simultaneously from a central command centre, with all systems being interconnected and transmitted over Internet Protocols (IP) (Craighead, 2009, pp.235-236).

Artificial Intelligence in Communications Systems

Artificial intelligence and deep learning algorithms are increasingly being incorporated into communications systems and networks, with emerging technologies such as Cognitive Radio (CR). CR is a radio system which can 'dynamically and autonomously adapt its operational parameters and protocols, and to learn from its previous experience' (Nguyen, Villain & Le Guillou, 2012). For example, in wireless communications, CR may combine deep learning-based sensing with active radio waveforms to operate intelligently in a range of radio environments to optimise channel use and avoid common interference and congestion issues (IEEE Spectrum, n.d.).

Limitations

There are numerous limitations of security communications systems, though these may vary depending on the specific communication and the mode of transmission. Network interruption or lack of signal reception where mobile phones are used, may be an issue in certain locations, and at times such as maintenance periods or in the event of network service outages (Norman, 2014, p.97). However, enough redundancy can typically be built into communications systems to limit the impact of system outages and signal issues. Eavesdropping, deception and jamming are technical limitations for frequency-based communications systems. For instance, adversaries may monitor, or 'eavesdrop' on conventional radio systems using a receiver or frequency scanner or transmit deceptive messages to confuse or distract the response force (Garcia, 2001, p.229). Signals may also be affected by 'jamming', which refers to the 'insertion of unwanted signals of a communications system for the purpose of masking desired signals' (Garcia, 2001, p.230).

Digital and consolidated communications systems, particularly those using software and transmission across IP, also have inherent limitations. Integration may be problematic where different technologies require different supporting platforms (Wiser, 2017). Network security, data breaches and denial of service may be a significant risk and limitation, particularly for software components which are cloud-based (Ducklin, 2015; Lukan, 2014; Morrow, 2018; Suryateja, 2018; Walloschek et al, 2011).

1.3.2.4.2 Dispensable Barriers

Definition

Dispensable Barriers are an active response system 'that are deployed upon requirement, that is, during an adversary attack' (Garcia, 2008, p.236). Dispensable barriers are based on an active versus passive dichotomy. Active barriers, on initiation stop or delay an adversary from accomplishing their objective. The system includes the necessary processes for decision making to determine when the dispensable barrier is to be activated, command and control hardware to implement the decision, the material that is to be deployed, and the associated dispensing mechanism (Garcia, 2008). Active systems may be achieved in an automated process, guard force processes or combinations of automation and guard force. In contrast, passive barriers are activated by penetration attempts by an adversary, rather than a command-and-control system.

Characteristics

There are numerous types of active and passive dispensable barriers, with a variety of materials available to impede an adversary. In passive systems, the materials may be dispensed automatically through the activation of a sensor, and in active systems, by command-and-control hardware. In active systems, command and control hardware operates the dispensing hardware (Integration), which may consist of storage tanks, activation valves, pressure regulators, safety valves, filters, power sources and plumbing hardware (Sandia National Laboratories, n.d., p.24). Dispensing hardware is the mechanism by which the material - designed to stop or delay the adversary – is released. Materials dispensed may include polyurethane foam, smoke, fog, irritants such as tear gas, deployable concertina, razor wire strips and other entanglement devices (Garcia, 2008; Sandia National Laboratories, n.d.).

Dispensable barriers and their indicative BACS architecture levels with AI alignment are presented in Table 15:

Table 15

Indicative Dispensable Barrier Technologies BACS Architecture Levels and AI Analytical Alignment.

DISPENSABLE BARRIER TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Active Dispensable Barrier	Command and Control Hardware, Communications mechanisms, Dispensing Hardware, Deployable materials, devices and components	Command and control activation of dispensable hardware and materials	Integration of command-and-control hardware with dispensable barrier, Setting of parameters through system interface (e.g. sensors and alarms).
Passive Dispensable Barriers	Sensors, dispensing hardware, and deployable materials, devices and components, Activation by penetration attempts by the adversary	-	-

Indicative Analytical Methodology

Dispensable barriers typically employ a rules-based analytical methodology, with Expert Rule-setting occurring at the management level (for active barriers) and the automation level for passive barriers where activation occurs due to penetration attempts by the adversary. Rule setting for dispensable barriers follows the knowledge base of 'IF… THEN…' where IF a condition is met (sensor detects conditions or requirements for activation, or a guard assesses the conditions as requiring deployment of a dispensable barrier), THEN the command signal activates the dispensing hardware and the dispensable barrier is deployed (Dutta & Bonissone, 1993; Fischer & Green, 1998; Garcia, 2008, Sandia National Laboratories, n.d.).

Functionality and Role in the Protection of Assets

The main function of dispensable barriers is to delay or prevent an adversary gaining access to an asset, or reaching their objectives (Garcia, 2008). In some instances, barrier systems will not delay an adversary indefinitely and consequently may require an adequate response force as well as use in conjunction with other structural barriers or physical security measures (Sandia National Laboratories, n.d.). Despite this, dispensable barriers are generally compact, storable and offer rapid deployment of physical impediments. Unsuspecting adversaries may not have the personal protective equipment or tools required to defeat the materials dispensed, which may significantly reduce their capability and success. For example, irritants and obscurants such as smoke and tear gas dispensed by active barriers are non-lethal but may temporarily incapacitate the adversary by rendering them 'blind'. Passive barriers such as sticky or slippery materials may be sprayed onto assets or walkways to effectively impeded adversaries (Garcia, 2008; Office of Technology Assessment, 1993; Sandia National Laboratories, n.d.).

Integration

While acknowledging that active barriers can be integrated into BACS, these is a dearth of literature on the use of dispensable barriers across the AI continuum and their integration with other building technologies.

Artificial Intelligence in Dispensable Barriers

There currently is a dearth of literature on the use of dispensable barriers across the AI continuum.

Limitations

Dispensable barriers have numerous limitations, including safety of personnel, environmental limitations, network security issues in integrated systems and additionally the cost and labour involved in cleaning some types of materials dispensed. Safety issues may arise as a result of the dispensing mechanism and barrier, or as a result of the materials deployed. Environmental conditions such as lightning, earthquakes and power issues may also affect operation of the dispensing mechanisms, particularly where dispensable barriers are operated remotely. Command signalling may also be vulnerable to network-based attacks, particularly when integrated in smart buildings and automated systems (AT&T Cybersecurity, 2019; Garcia, 2008; Kaspersky, 2019; Office of Technology Assessment, 1993; Sandia National Laboratories, n.d.).

1.3.2.4.3 Vehicle Barriers

Definition

Vehicle barriers are 'designed to control access or prevent vehicles access to a space' (ASIS International, 2015, p.151). Vehicle Barriers, like dispensable barriers, may be passive or active in their operation, however passive barriers are typically static with no mechanical functioning, and therefore are not considered a security technology within the scope of this report. Examples of active vehicle barriers include rising or retractable bollards and wedges, beam barriers and automatic gates (Australia-New Zealand National Counter-Terrorism Committee [ANZCTC], 2017).

Characteristics

Active vehicle barriers are physical barriers that are deployed by command signal activation. Command and control hardware initiate the vehicle barrier to raise or drop, with some types containing security sensors to detect vehicle speeds and loads, and when the vehicle has cleared the barrier zone to prevent the barrier causing damage or injury (Cybersecurity & Infrastructure Security Agency [CISA], 2019; Garcia, 2008). Vehicle Barriers and their indicative BACS architecture levels with AI alignment are presented in Table 16:

Table 16

Indicative Vehicle Barrier Technologies BACS Architecture Levels and AI Analytical Alignment.

VEHICLE BARRIER TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Rising Bollards	Hydraulics, Hydroelectric Drive Unit, Cabling, Power	Alarm Activation (Intrusion, Barrier Compromise or	Configuration of operating modes and controls (e.g. normal,

	Source/ Power Unit, Operator Console, Remote Control Unit, Sensors for Integrated Bollards,	Maintenance Alert), System Alerts, Logging of Barrier Cycles, Interlock Control (Sally port)	emergency, power failure), Manual/ Auto Selection, Setting of Operating Speed/ Cycle Time, Integration of Video/ Sensors through Interface
Rising Wedge Barrier	Hydraulics, Hydroelectric Drive Unit, Cabling, Power Source/ Power Unit, Operator Console, Remote Control Unit,	Alarm Activation (Intrusion, Barrier Compromise or Maintenance Alert), System Alerts, Logging of Barrier Cycles, Interlock Control (Sally port)	Configuration of operating modes and controls (e.g. normal, emergency, power failure), Manual/ Auto Selection, Setting of Operating Speed/ Cycle Time,
Beam Barrier	Hydraulics/ Motorised Beam Hardware, Cabling, Power Source/ Power Unit, Operator Console, Remote Control Unit, Keypad, Reader & Credentials for Integrated Units	Alarm Activation (Intrusion, Barrier Compromise or Maintenance Alert), System Alerts, Logging of Barrier Cycles, Interlock Control (Sally port)	Configuration of operating modes and controls (e.g. normal, emergency, power failure), Manual/ Auto Selection, Setting of Operating Speed/ Cycle Time, Integration of Video/ Sensors through Interface
Automatic Gates	Hydraulics/ Motorised Hardware, Cabling, Power Source/ Power Unit, Keypad, Sensors, Reader & Credentials for Integrated Units, Operator Console, Remote Control Unit,	Alarm Activation (Intrusion, Barrier Compromise or Maintenance Alert), System Alerts, Logging of Barrier Cycles, Interlock Control (Sally port)	Configuration of operating modes and controls (e.g. normal, emergency, power failure), Manual/ Auto Selection, Setting of Operating Speed/ Cycle Time, Integration of Video/ Sensors through Interface

Indicative Analytical Methodology

Active vehicle barriers typically employ a rules-based analytical methodology, with Expert Rule-setting occurring at the management level (for active barriers) where a command signal activates the barrier to open or close at the automation level, based on the management level rules. Rule setting for vehicle barriers such as automated gates and rising wedges follows the knowledge base of 'IF… THEN…' where IF a condition is met (valid credential such as RFID or EAC card is presented), THEN the system (or command signal given by a guard) initiates a system response (i.e. open automatic gate or lower wedge barrier) to permit access to the credential holder or vehicle (Dutta & Bonissone, 1993; CISA, 2019; Garcia, 2008).

Functionality and Role in the Protection of Assets

Vehicle barriers are used to prevent or delay vehicle and adversary access to a building, facility or protected area. These barriers are designed to deter or prevent ramming by cars, trucks, motorcycles and all-terrain vehicles (ATVs), or in the event the adversary does attempt to ram the barrier, the vehicle and its occupants may be impaired or disabled on impact. Vehicle barriers may also be used to compel adversaries to revert to traveling by foot which may slow them down, particularly in the event they require tools, equipment or weapons to execute their attack (ANZCTC, 2017; Garcia, 2008, p.222).

Vehicle barriers may be used for a variety of applications, including protection against vehicles which may be used as a weapon, or at entry and exit points of high security facilities. Vehicle barriers may be installed in environments such public spaces, embassies, military installations or other buildings requiring hostile vehicle mitigation (ANZCTC, 2017; Garcia, 2008, p.222).

Integration

Vehicle barriers typically operate by electronic command signal which activates the barrier to open or close, though barriers may be automated with the use of access control readers and credentials. Multiple barriers may be integrated together with other sensors and devices, for instance CCTV, intrusion detection, automatic number plate recognition, and intercom systems (Australian Bollards, n.d.; CISA, 2019).

Artificial Intelligence in Vehicle Barriers

There currently is a dearth of literature on the use of vehicle barriers across the AI continuum.

Limitations

Vehicle barriers may be penetrated in several ways, for instance if the barrier is bridged, removed or rammed. A vehicle may ram through or over the barrier, or alternatively once a vehicle has rammed through the barrier and been disabled, a second functioning vehicle may pass through the penetration point (Garcia, 2008, p.222). Hazardous cargo or explosive materials being carried by a vehicle may be ejected into a facility following an impact, and while this is not a limitation of the technology, it may be limitation of the effectiveness of current hostile vehicle mitigation measures (Smith & Wesson, 2011). Technical and environmental limitations may include corrosion and drainage issues, in addition to ongoing maintenance and servicing requirements (ANZCTC, 2017, p.17).

1.3.2.4.4 Weaponry

Definition

Weaponry may be defined as “the science of designing and making weapons” which may be used to deter, delay or respond to adversaries (Merriam Webster, n.d.). A variety of weapons exist for the protection of assets– both lethal and non-lethal (or ‘less-lethal, as all weapons have the potential to be lethal). Weapons used in security response may include conventional firearms, acoustic weapons, high-voltage or ‘electroshock’ weapons, ‘directed energy’ weapons (such as electromagnetic or optical distractors), remotely operated weapons systems (ROWS), and autonomous/ semi-autonomous weapons systems. These systems may have varying levels of automation, particularly in military applications where autonomy is driven by objectives of increased capability and decreased operational risk (International Committee of the Red Cross [ICRC], 2014, p.7).

Characteristics (Automation Architecture of Weaponry Technologies)

Characteristics vary considerably between weapon types. Acoustic weapons can utilise sonic, infrasonic or ultrasonic frequencies, although typically they combine two high-volume ultrasonic frequencies to warn, disorient or disable adversaries (Brehm & Wheeler, 2018). High-voltage weaponry, also referred to as ‘Electroshock Weapons’ use transducers to emit a high-voltage, low amplitude current to incapacitate an adversary. Several types are available ranging from small taser and ‘sweeping stun-gun’ devices to Small Arms Pulsed Electronic Tetanisation (SPECTER) weapons which fire a projectile from up to 100m of the target. The SPECTER projectile releases electrode darts that penetrate an adversary’s clothes and skin, delivering an incapacitating electric shock (Hambling, 2004; Joint Intermediate Force Capabilities Office, n.d.; New Scientist, 2020; Pettit, 2020).

'Directed Energy Weapons' utilise electromagnetic energy in the form of high-energy lasers or high-power microwaves, which are spatially directed to incapacitate an adversary (Lexington Institute, 2003). There are three primary classes of directed energy weapons – electromagnetic which direct high frequency radiation (e.g. microwave), sound and light directed energy which direct high intensity laser beams (e.g. optical dazzlers), and particle beams, which disrupt or damage the target's atomic or molecular structure, though particle beams are not sufficiently developed to operational levels (Joint Intermediate Force Capabilities Office, 2018; Wheeler, 2017).

Autonomous, semi-autonomous and Remotely Operated Weapons Systems (ROWS) are capable of being operated remotely or deployed for autonomous operation. ROWS use a command-and-control station to allow a response by security or military officers using lead, rubber or fragmented bullets (Garcia, 2008, p.239). Autonomous and semi-autonomous weapons may select and engage targets with limited intervention by a human operator using robotic mechanisms and platforms. Development of autonomous weapons systems is progressing rapidly, though highly controversial, and subsequently systems are not currently available for commercial security applications (Dickinson, 2018; Felt, 2020; Wolfe, 2019).

Weapons systems components and their indicative BACS architecture levels with AI alignment are presented in Table 17:

Table 17

Indicative Weapons System BACS Architecture Levels and AI Analytical Alignment.

WEAPONRY TECHNOLOGIES			
	Field Level	**Automation Level**	**Management Level**
Acoustic Weaponry	System Casing & Hardware, Transducer (Piezoelectric Transducer Array), Microphone/ Phraselator Device, Cabling, Power Unit, Communications	Command and control activation, Modulation of frequency & output	Output Selection (e.g. verbal instruction/ warning tone, sonic output), Frequency Selection (sonic, infrasonic or ultrasonic), Output volume
High-Voltage (Electroshock) Weaponry	System Casing, Microprocessor, Projectiles/Electrodes, Transformer/ Transducer, Battery	Command and Control Activation, Intensity Modulation of Intensity, Duration & Shape of Current	Directional Guidance, Manual Control
Electromagnetic Weaponry (Directed Energy Weapon)	High Power Radio Frequency (HPRF) & High-Power Microwave (HPM) Components, Electron Gun, Vacuum Chamber, Mirrors, Video Camera/ Optical Apertures, RF Antenna, Thermal	Command and Control, Target Range Finding, Target Tracking (Automated Mode), Re-optimise System Configuration, Disconnect on Warning/ Safety Shutdown, Status Detection/ Alarm	Software Controls, Mission Planning, Rules of Engagement, Target Identification & Engagement, Zone Setting ('No-fire Areas), Circumvention/ Reset, Error Correcting, Directional Guidance,

	Imager, Dishes, Power Unit		Manual Control (Contingency Mode)
Optical Distractors (Directed Energy Weapon)	High Energy Laser (HEL) Components, Electro-optical and Electrical Sensors, GPS, Optical Antenna Communications,	Command and Control, Target Range Finding, Target Tracking (Automated Mode), Safety Shutdown, Status Detection/ Alarm	Mission Planning, Rules of Engagement, Target Identification & Engagement, Zone Setting ('No-fire Areas), Circumvention/ Reset, Error Correcting, Directional Guidance, Manual Control (Contingency Mode)
Remotely Operated Weapons Systems (ROWS)	System Casing & Hardware, CPU, Command Display Unit, Cameras, Laser Range Finder, Transmitters & Receivers, GPS Components, Sensors, Communications, Weapons hardware and munitions	Command & Control, Activation of Live Video & Audio Streaming, GPS Location Tracking, Automated Target Tracking, Target Recognition, Alarm Activation,	System Interface for Configuration & Operation, Systems with Object Recognition: Configuration & Programming, Introduction of training sets, image processing and classification
Semi-Autonomous & Autonomous Weapons Systems	Embedded Electronic Components & Sensors (Facial and Object Recognition Sensors), CPU, Power Unit, Transmitters & Receivers, GPS Device, Control Display/ Monitor	Signal Processing, Navigation, Surveillance, Target Detection, Automated Target Recognition (ATR) & Decision Making on Use of Lethal Force	System Interface for Configuration & Programming, Introduction of training sets, predefined target modelling, image processing and classification, trajectory planning, engagement criteria

Indicative Analytical Methodology

Weapons systems may employ a rules-based analytical methodology, with Expert Rule-setting occurring at the management level for acoustic weapons, electroshock weapons, 'directed energy' weapons (such as electromagnetic or optical distractors), and remotely operated weapons systems (ROWS). Rule setting follows the knowledge base of 'IF… THEN…' where IF a condition is met (e.g. the responder assesses conditions as requiring weapons deployment), THEN the command signal activates the weapons hardware and the ammunition is fired (Dutta & Bonissone, 1993; Felt, 2020; Garcia, 2008, Sandia National Laboratories, n.d.). Advanced autonomous weapons systems may employ a statistical (probabilistic and machine learning) or sub-symbolic (embodied intelligence) analytical methodology, or a hybrid methodology with statistical and sub-symbolic applied to individual components across the system (Abaimov & Martellini, 2020).

Functionality and Role in the Protection of Assets

Weaponry and weapons systems may provide a lethal, less lethal or non-lethal response for the protection of assets, the effects of which are immediate. For instance, acoustic weapons in the audible range may create intolerable noise when directed at adversaries though the effects are non-lethal and temporary. Consequently, Acoustic weapons have been used by law enforcement to disperse protesters, and passenger cruise ships have also reportedly used non-lethal Long-Range Acoustic Devices (LRAD) to repel pirates. Acoustic weapons in the infrasonic or ultrasonic frequencies may result in headaches and nausea, both may be capable of immediately incapacitating an adversary, though there is limited evidence to support operational use and effectiveness (Brehm & Wheeler, 2018; Chang, 2016).

While directed energy weapons are predominantly used by military forces, some types are currently being incorporated into law enforcement and commercial security applications. Vehicle stopping technologies for instance transmit a non-lethal pulsed microwave beam to interfere with a vehicle's electronics to disable the engine (Teledyne UK, 2020). High power microwave weapons such as the Active Denial System or "Heat Ray" delivers non-lethal radiation to an adversary, and Optical Distractors use Long-range lasers, such as optical 'dazzlers' to deflect adversaries such as pirates attempting to board merchant ships. Remotely operated weapons systems (ROWS) are predominantly used by military forces (i.e. mounted on vehicles, ships or stationary platforms), though technical and operational support of ROWS may be provided by private military and security contractors in hostile environments within some nations (Kreig, 2018).

Integration

Advanced weapons systems such as ROWS and directed energy systems are currently being integrated with existing kinetic weapons and combat systems in military installations. For example, directed energy weapons are currently being installed on full-functioning U.S. Naval surface vessels with existing systems being modified to ensure compatibility and interoperability with other sensors and systems such as RADAR (Booz Allen Hamilton, n.d.; Trevithick, 2020). As development of these systems are largely funded and directed by government and defence organisations, integration is currently limited to military systems and technologies (Abaimov & Martellini, 2020). However, due to the evolution and crossover of military technologies to the private security sector, there may be future opportunity for integration development in security technologies.

Artificial Intelligence in Weaponry

According to Abaimov and Martellini (2020, p.161), ambitions for military superiority have resulted in "multi-billion-dollar investments into weaponized AI research, continuous acquisition of cyber arms and vulnerabilities, development and production of the new generation AI-enhanced weapons". AI and machine learning are increasingly being developed and incorporated into numerous weapons systems, particularly directed energy weapons and autonomous systems (Modern Diplomacy, 2020). However autonomous systems currently operate at a higher evolution along the AI continuum. Autonomous systems are defined as "any weapon system with autonomy in its critical functions. That is, a weapon system that can select (i.e. search for or detect, identify, track, select) and attack (i.e. use force against, neutralize, damage or destroy) targets without human intervention" (Decide and Act) (ICRC, 2016, p.1) Unlike automatic systems which operate within the statistical AI paradigm (probabilistic), autonomous systems may operate between the statistical and sub-symbolic paradigms (machine learning/ embodied intelligence), and fully autonomous systems may operate sub-symbolically with Broad AI capabilities including navigation, surveillance, target detection and decision making on the use of lethal force (Abaimov & Martellini, 2020; ICRC, 2014).

Limitations

Weapons technologies systems with artificial intelligence capabilities, are more developed than the governance and regulatory systems currently in place to provide oversight. Legal, ethical, safety and humanitarian issues dominate discourse on the use of weapons systems as security and military technologies. Advanced weapons systems with any level of autonomy may be exposed to significant security and safety risks with AI-specific cyber vulnerabilities being one of the key limitations in autonomous technologies, particularly with dependence on civilian infrastructure and commercial equipment (Abaimov & Martellini, 2020; Baker, 2011; Wolfe, 2019).

1.3.2.4.5 Response Technologies Overview

Response technologies across the protection of assets vary with many sitting in the military paradigm. Understanding current and future optimisation of AI in Response Technologies can be achieved by reviewing how current technologies align to the spectrum of Artificial Intelligence (AI) Paradigms. Table 18 presents a summary of current responding technology and AI analytical alignment.

Table 18

Responding Technology and Indicative AI Alignment

Technology		**Function/ performance measure**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Communication Systems	Telephones, Mobile Phones, Pagers, Emergency Phones	Response	Command Signal; Transmitter converts electrical signal into radio waves which are radiated from the device antenna to the radio receiver (mobiles) or transmits electrical signal via wire/ fibre optics (landlines)	Emergency phone in car park used to communicate with security staff	Expert Rule-setting
	Radio Communication	Response	Command Signal; Radio transmitter transforms energy into radio waves radiated from the device antenna to the radio receiver	Spread spectrum, Two Way Radio Communication	Expert Rule-setting
	Emergency Warning Intercom Systems (EWIS) Public Address / Mass Notification System	Control, Response	Command Signal by operator/ user; Transmission of voice audio or automated instructions via IP network protocol (e.g. VoIP, RTSP, TCP/IP, SIP)	EWIS system used for emergency evacuation of a facility	Expert Rule-setting

	IP Network Connected Intercom	Control, Response	Command signal by operator/user; transmits audio and visual data via IP network protocol (e.g. TCP/IP, SIP, RTSP)	Intercom installed at remote/ isolated substation and monitored/ controlled from central site.	Expert Rule-setting
	Wireless Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via encrypted wireless signal (radio waves)	Intercom installed within secure facility between access control points	Expert Rule-setting
	Wired Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via low-voltage electrical wiring	Intercom installed at vehicle gates to allow entry to authorised personnel	Expert Rule-setting
	Door Release Intercom	Control, Response	Command signal by operator/user allowing remote unlocking of doors via wiring from the intercom to the electrified locking hardware	Intercom installed at entrance door to allow access to authorised personnel	Expert Rule-setting
Dispensable Barriers	Active Dispensable Barrier	Delay, Response	Command Signal	Fog or smoke dispensed to obscure an adversary's vision	Expert Rule-setting
	Passive Dispensable Barriers	Delay	Contact with Adversary	Asset covered in Sticky foam entangles attacker and tools during an attempt to remove asset	Expert Rule-setting
Vehicle Barriers (Hostile Vehicle Mitigation)	Rising Bollards	Delay, Response	Command signal activates bollard to deploy	High security entry/ exit points	Expert Rule-setting
	Rising Wedge Barrier	Delay, Response	Command signal activates wedge barrier to deploy	High security entry/ exit points	Expert Rule-setting

	Beam Barrier	Delay, Response	Command signal activates beam barrier to deploy	Security entry/ exit points	Expert Rule-setting
	Automatic Gates	Detect, Delay, Response	Command Signal, PLUS Presentation of Authorised Credential	Security entry/ exit points (includes sliding, swing gates, bi-fold gates etc.).	Expert Rule-setting
Weaponry	Acoustic Weaponry	Delay, Response	Sonic, infrasonic, or ultrasonic frequencies. High-amplitude infrasonic weapons combine two high-volume ultrasonic frequencies at a difference of 8Hz	Passenger cruise ship using non-lethal Long-Range Acoustic Device (LRAD) to create ear-piercing noise to repel pirates	Expert Rule-setting
	High-Voltage (Electroshock) Weaponry	Delay, Response	Transducers emit high-voltage, low amplitude current to incapacitate adversary	Crowd dispersal using non-lethal electroshock ‘sweeping stun gun’	Expert Rule-setting
	Electromagnetic Weaponry (Directed Energy Weapon)	Delay, Response	Electromagnetic energy beam created by a Gyrotron (Vacuum Electronic Device (VED) to create high power high frequency THz radiation)	High power microwave weapon such as the Active Denial System or “Heat Ray” delivering non-lethal radiation to adversary	Expert Rule-setting
	Optical Distractors (Directed Energy Weapon)	Delay, Response	High intensity beams of light directed at adversary to create nausea, confusion, disorientation, or temporary blindness	Long-range dazzling laser directed at pirates attempting to board merchant ship	Expert Rule-setting
	Remotely Operated Weapons Systems (ROWS)	Delay, Response	Weapons operated remotely in response to event or attack	Deployed at critical installations in volatile security environments	Expert Rule-setting

	Autonomous & Semi-Autonomous Weapons Systems	Delay, Response	Predefined Stimulus for which autonomous systems may execute self-direction, self-learning or emergent behaviour to select and attack targets	Autonomous drones used to track and engage targets	**Semi-autonomous**: Statistical (Probabilistic & Machine Learning) **Autonomous**: Sub symbolic (Embodied Intelligence)

1.4 Interpretation and Findings Phase One

Phase One of the study sought to develop a robust alignment between the current status of security technologies and the Spectrum of Artificial Intelligence. Specifically, Phase one sought to respond the research sub-questions:

i) How is control (real and perceived) currently achieved over physical environments using technology?
ii) What constitutes the Artificial Intelligence problem domains and spectrum of paradigms?
iii) How is artificial intelligence indicatively used in physical security?

Sub-Research Question One: Phase One found that control is achieved in security and crime prevention using technology through combinations of Observing, Detecting, Controlling and Responding technological measures. These technological measures are integrated through various software and hardware means at the Automation and Management levels of built environment management technology architectures, underpinned by various computational techniques. Such techniques include artificial intelligence methods from the AI problem domains and spectrum of AI paradigms to replicate, as best as practicable, human cognition processes across the perception, reasoning, knowledge, planning and communication computational problem domains, using logic based, knowledge based, probabilistic, machine learning, embodied intelligence and search optimisation techniques.

Sub-Research Question Two: The review of literature highlighted that the current state of artificial intelligence is characterised by the AI problem domains and spectrum of AI paradigms to perform computational techniques aligned to definitions of Narrow, General and Artificial Super Intelligence. However, to date, much of the functional use of AI in the protection of assets sits within what is considered Narrow AI, defined as approaches that focus on solving very specific tasks and cannot cope with any tasks outside the scope that they have been designed for. Narrow AI is very good at completing repetitive tasks and in many instances performs much better than humans.

Sub-Research Question Three, Phase One developed indicative alignments (Figures 6, 7, 8 and 9) between the BACS architecture levels and AI computational tasks aligned to the Artificial Intelligence Problem Domain and Spectrum of AI Paradigms.

Interpretation Phase One: Phase One identified that artificial Intelligence techniques used in the protection of assets currently sit within the Narrow AI paradigm, and that significant limitations exist in General AI processing, with no evidence of General or Artificial Super Intelligence in the protection of assets. In the application of AI to advances in security technology, AI is focused toward predefined outputs based on rules as they relate to security interest characteristics. These characteristics include; recognition (pattern-anomaly recognition/detection), classification, search and combinations of these, based on analysis of security elements of interest including; colour, shape (object), movement, feature or behaviour. Furthermore, these characteristics and the efficacy of AI are influenced by factors including speed, efficiency, accuracy and integration. Phase One identified that in the protection of assets control is achieved through combinations of security technologies that Observe, Detect, Control and Respond. In addition, such technologies draw on computational techniques from the Artificial Intelligence Problem Domain and Spectrum of AI Paradigms.

Figure 6

Indicative alignments between observe technologies BACS architecture levels and AI computational tasks aligned to the spectrum of AI paradigms.

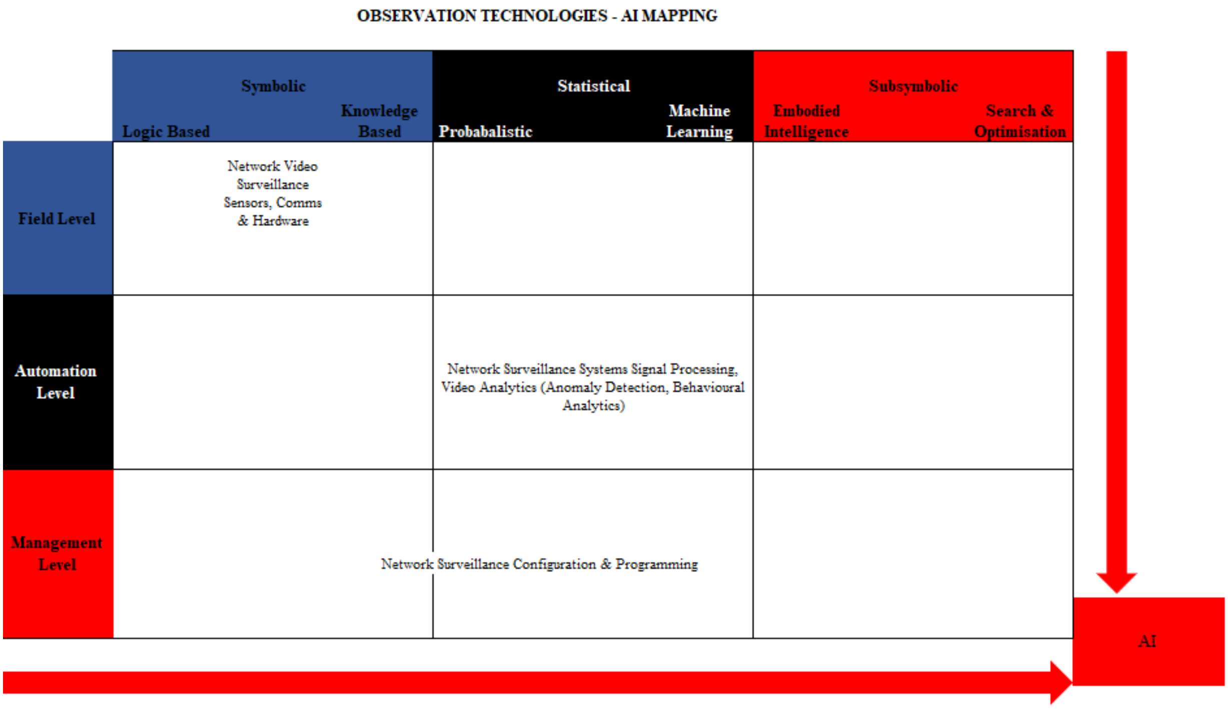


Figure 7

Indicative alignments between detect technologies BACS architecture levels and AI computational tasks aligned to the spectrum of AI paradigms.

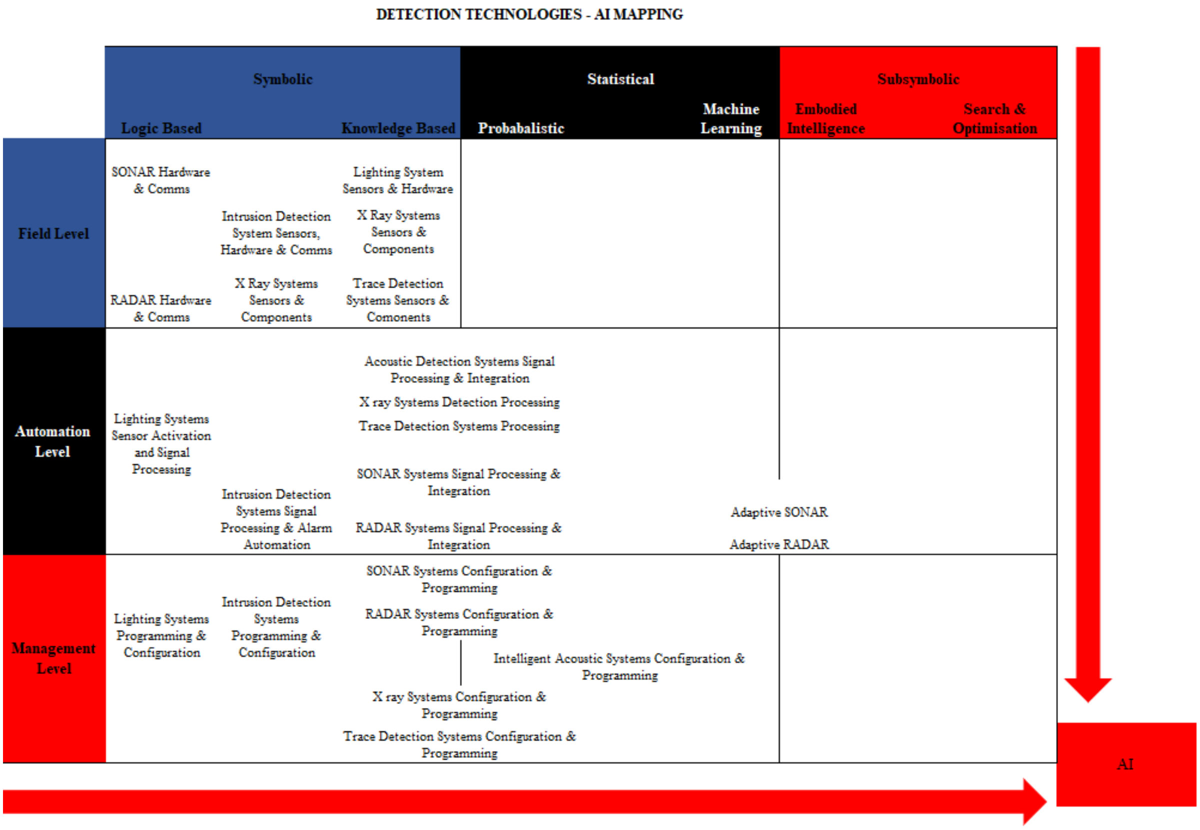


Figure 8

Indicative alignments between control technologies BACS architecture levels and AI computational tasks aligned to the spectrum of AI paradigms.

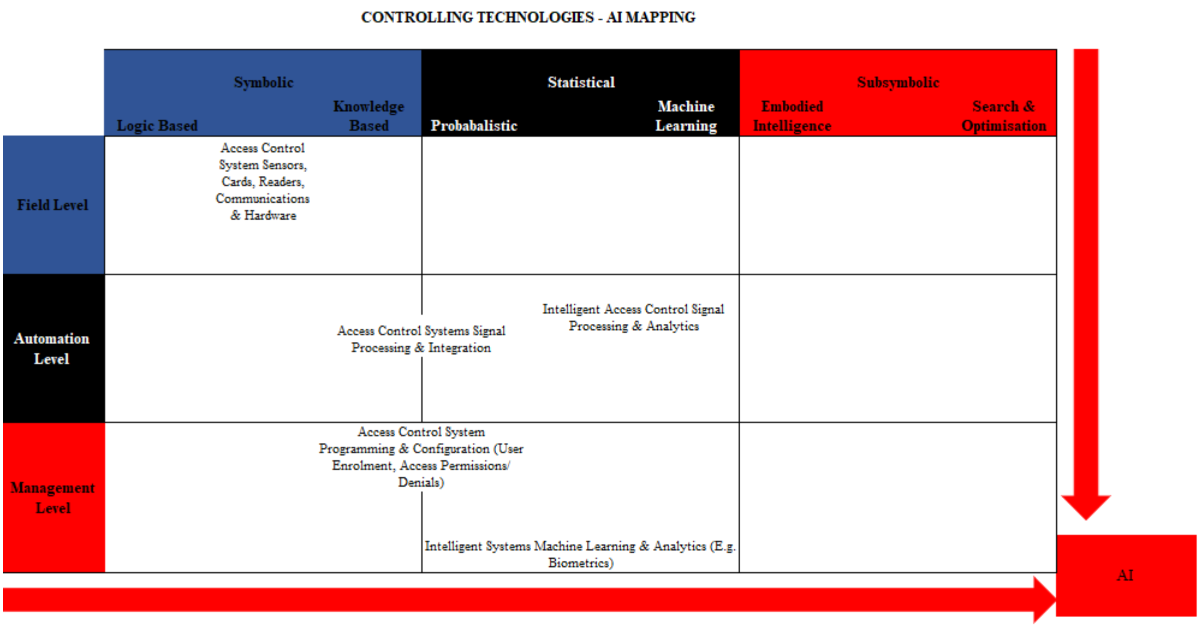
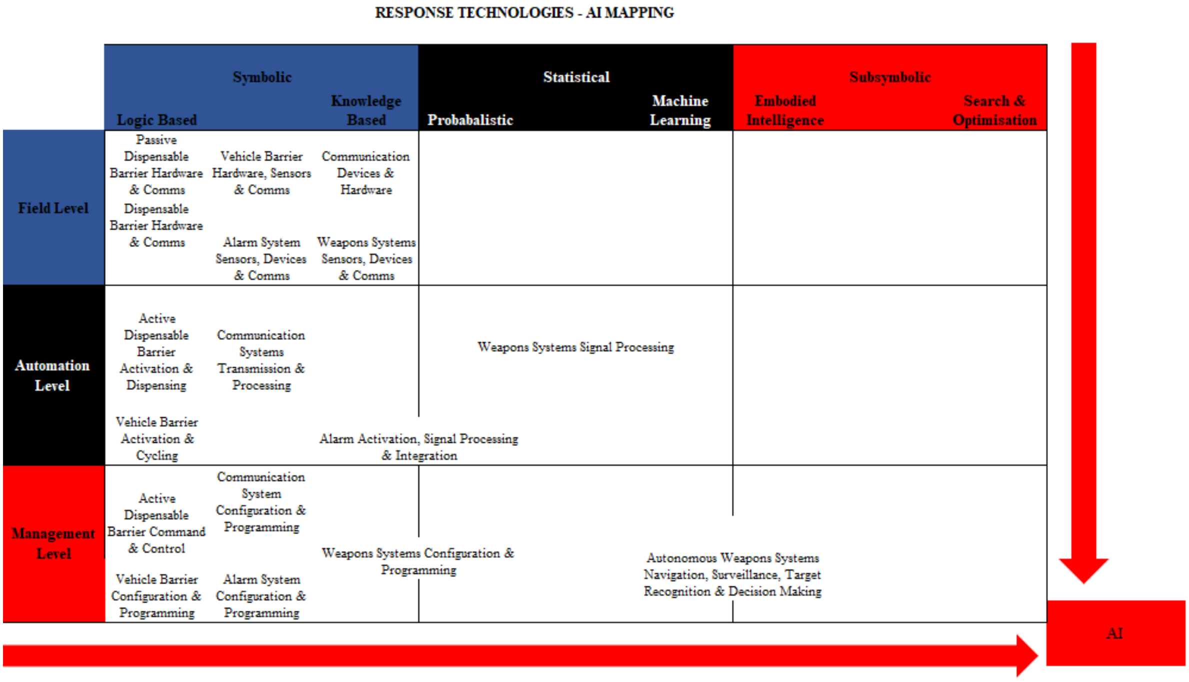


Figure 9

Indicative alignments between respond technologies BACS architecture levels and AI computational tasks aligned to the spectrum of AI paradigms.



1.5 Conclusion Phase One

Phase One employed a systematic literature review (Jesson, Matheson & Lacey, 2011), identifying existing literature with regard to the roles, functions and tasks of security technology in the protection of assets, along with the technical articulation of the Artificial Intelligence (AI) problem domains and spectrum of AI paradigms. Phase One found that control is achieved in security and crime prevention using technology through combinations of Observing, Detecting, Controlling and Responding technological measures. These technological measures are integrated through various software and hardware means at the Automation and Management levels of built environment management technology architectures, underpinned by various computational techniques.

The review of literature highlighted that the current state of AI is characterised by the AI problem domains and spectrum of AI paradigms to perform computational techniques aligned to definitions of Narrow, General and Artificial Super Intelligence. The review found that AI techniques used in the protection of assets currently sit within the Narrow AI paradigm, with no evidence of General or Artificial Super Intelligence in the protection of assets. In the application of AI to security technology, AI is focused toward predefined outputs using computational techniques and executing rules as they relate to security interest characteristics or patterns which are aligned to environmental threat stimuli changes.

PHASE TWO: EXPERT FOCUS GROUP

2.1 Introduction

This Chapter presents Phase Two of the research; the outcomes of the security and artificial intelligence (AI) Focus Group discussion. Phase Two sought to confirm findings from Phase One, reinforcing the current status of AI in the protection of assets, and to add additional outcomes that would aid in enhancing the security professionals' understanding of artificial intelligence in security technologies. Participants were invited to provide their expert insights into how the AI problem domains and spectrum of artificial intelligence paradigms align with the roles, functions and tasks of security technology through combinations of Observing, Detecting, Controlling and Responding technological measures. This alignment of security technologies with AI provides the foundation to establishing where further opportunities may exist for fusing security technology with developments in AI, as well as provide indicative risks associated with such advancements.

This report presents Phase Two of the sponsored study by ASIS International, providing an expert enriched benchmark for understanding future opportunities and risks in advancements in AI in the protection of assets.

2.2 Background

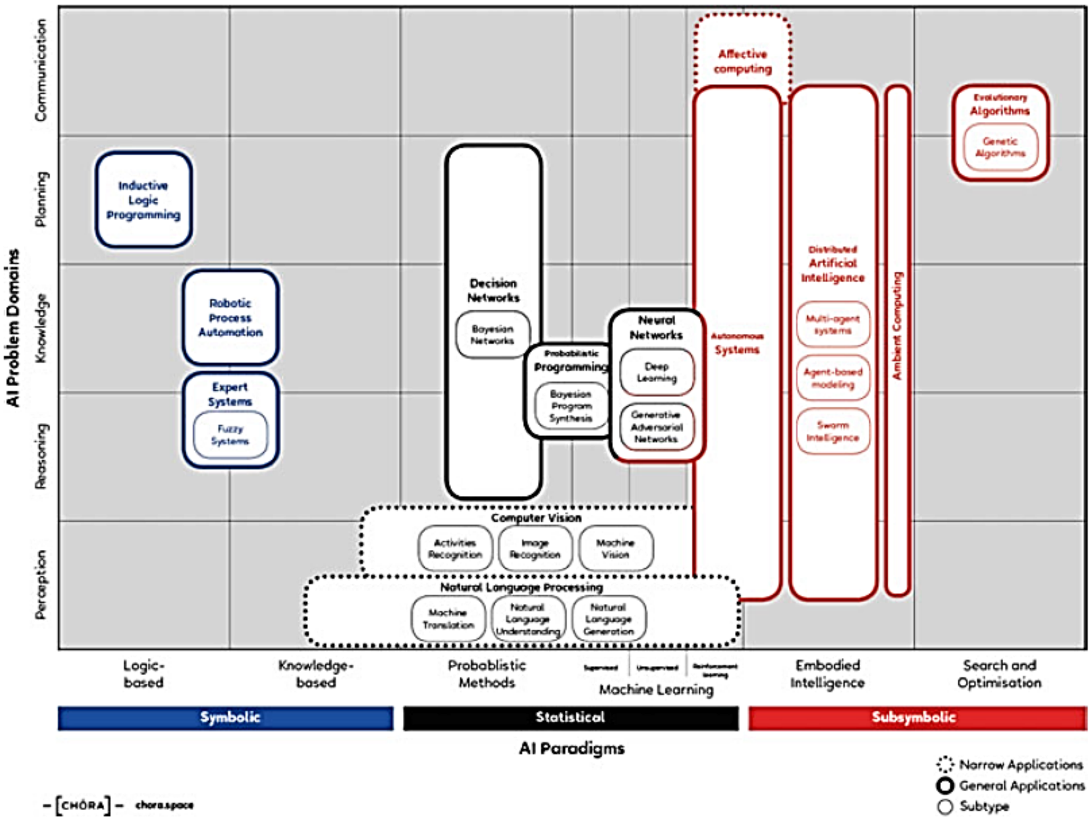
Phase One of the research employed a systematic literature review (Jesson, Matheson & Lacey, 2011), identifying existing literature with regard to the roles, functions and tasks of security technology in the protection of assets, along with the technical articulation of the Artificial Intelligence (AI) problem domains and spectrum of artificial intelligence paradigms. Phase One found that control is achieved in security and crime prevention using technology through combinations of Observing, Detecting, Controlling and Responding technological measures. These technological measures are integrated through various software and hardware means at the Automation and Management levels of built environment management technology architectures, underpinned by various computational techniques.

Phase One drew from Corea's (2019) *AI Knowledge Map* and highlighted that the current state of artificial intelligence is characterised by the AI problem domains and spectrum of AI paradigms to perform computational techniques aligned to definitions of Narrow, Broad, General and Artificial Super Intelligence. Corea's (2019) *AI Knowledge Map* presented a novel visual map of the current AI paradigms (Figure 10) for security professionals to understand the complex web of AI literature. This model represents an abstract range of problem domains on the y-axis, divided into the five categories of analogic human cognitive processes including: perception, reasoning, knowledge, planning and communication. AI paradigms used to address sub-sets of these problem domains are presented on the x-axis. This model formed the basis for the alignment of security technologies to the AI Spectrum (Paradigms and Problem Domains).

Such alignment showed that AI techniques used in the protection of assets currently sit within the Narrow AI paradigm from a single technology perspective, and Broad AI within the context of combinations or integration of specific technological outputs, with no evidence of General or Artificial Super Intelligence in the protection of assets. This review highlighted that in the application of AI to security technology, AI is focused towards achieving computational outputs for predefined tasks to execute rules as they relate to security interest characteristics, patterns or actions which are aligned to environmental threat stimuli changes and defined automated actions.

Figure 10

AI Knowledge Map (Corea, 2019).



Drawing from both the literature review and Corea's (2019) *AI Knowledge Map*, the application of AI to security technology was indicatively mapped using three formats; Security Technology Tables (see Appendix A), Alignment of Security Technologies with AI Paradigms (see Appendix B), and BACS/Intelligent Controller Architecture Levels (see Appendix C). These formats were used to indicatively present how computational techniques used in artificial intelligence map to the detecting, observing, controlling and responding technologies used to achieve control over a physical environment.

Consequently, Phase Two sought to respond to the Question: How do the capabilities of artificial intelligence techniques map to activities that can achieve control over a physical environment using security technologies? Phase two sought to respond the research question through a series of objectives and supporting sub-questions.

2.3 Focus Group Objectives

The objectives of the Focus Group were to:

1. Confirm and correct alignment of Security Technology Tables to AI techniques
2. Confirm and correct alignment of Security Technology to AI Spectrum
3. Locate AI (hardware, firmware, software) to network architectural understanding

2.3.1 Objective One

2.3.1.1 Alignment of Security Technology Tables and AI Techniques

Objective One was facilitated by the following Sub-research question:

Q1. Do you agree with the indicative alignment of security technology with AI techniques?

Participants were provided with four categories of security technologies – Observing, Detecting, Controlling and Response Technologies – with each category presented in tables and indicatively aligned with the *AI Knowledge Map* (Corea, 2019). The AI Knowledge Map was introduced to participants as an appropriate model to communicate the AI problem domains and spectrum for security practitioners, and to align the analytical methods by AI researchers to security technology.
The technology tables were presented and discussed in the following order (see Appendix A: Indicative Technology Tables Presented to Participants);

- Observing Technologies and Indicative AI Analytical Alignment Table
- Detecting Technologies and Indicative AI Analytical Alignment Table
- Controlling Technologies and Indicative AI Analytical Alignment Table
- Response Technologies and Indicative AI Analytical Alignment Table

Participants were asked to evaluate the alignment of security technologies with the Indicative AI Methodology identified in each table. For example, Video Motion Detection in the category of Observing Technologies (see Table 19) was indicatively aligned with Statistical/ Probabilistic systems or methods from the *AI Knowledge Map* model (Corea, 2019). Video Motion Detection may use techniques such as background subtraction or estimation to establish or detect motion in video. Singh et al (2014) describe background subtraction as a comparison-based method that identifies motion by comparing the estimated background with the current frame. New frames are compared with the background model which is 'thresholded to segment the foreground objects' (Singh et al, 2014), in other words when a variation in pixels between frames reaches a set threshold, the system 'notes' this as sufficient change to imply movement. On this basis, Video Motion Detection was aligned with Statistical/ Probabilistic methods as motion detection is based on the probability of pixel change as movement (Table 19).

Table 19

Observing Technology and Indicative AI Analytical Alignment.

Technology		**Function/ Performance Measure**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Network Video Surveillance	Digital Video Surveillance	Observe, Detect, Recognise, Identify	Reflected Scene Image	CCTV cameras used to monitor and record a defined space	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)

	Video Motion Detection	Probability of Detection	Movement within a Frame	CCTV cameras detect movement of an individual into a defined space	Statistical Probability (Probabilistic Method)

Each of the Security Technology Tables were discussed, with the ***Function/Performance Measure***, the ***Stimuli*** and the ***Indicative Analytical Methodology*** evaluated to determine the accuracy of alignment between observational technologies and AI methodologies.

2.3.1.2 Objective One Findings

2.3.1.2.1 Disconnect between Security Terminology and AI Terminology

The issue of terminology disconnect between disciplines is well acknowledged in the academic literature. Cross-disciplinary studies often encounter challenges in communication due to a lack of common language or terminology among members or participants of cross-disciplinary research (Mennes, Pedersen & Lefever, 2019). Ambiguity of terms may arise from differences in interpretation of meaning between disciplines, particularly where each discipline has its own jargon or technical terminology, or if terms form part of multiple jargons among combined teams (Donovan, O'Rourke & Looney, 2015; Jeffrey, 2003; Mennes et al, 2019).

Although the challenges of cross-disciplinary communication were known prior to the Focus Group session, the Focus Group found a significant disconnect between security and artificial intelligence (AI) terminology. Current security literature was found to not effectively articulate language from the AI domain, and subsequently much of the technical terminology and jargon has not been embedded within security technology literature.

This resulted in difficulty in the interpretation of technological terms between security and AI participants. For instance, the meaning of terms contained in the descriptive headers and content used in the Security Technology Tables (Table 20) were interpreted differently by security and AI participants. Jeffrey's (2003) work highlighted the issue of collaboration between different discipline participants in the absence of pre-defined collective and cross-disciplinary reference terms, and this was experienced in Phase Two. However, discussion of cross-disciplinary reference terms during the focus group session allowed for specificity of the terms and meanings which had the most significant disparity in collective understanding.

Table 20

Descriptive Headers in Security Technology Tables Presented to Focus Group.

Technology		**Function/ Performance Measure**	**Stimuli**	**Example**	**Indicative AI Analytical Methodology**
Technology Category	Sub-category				
	Sub-category				

For instance, the column header ***Function/ Performance Measure*** was not explicitly defined as a security function within the context of security technology. The focus group participants questioned both the heading and content of data in this column, and agreed these terms were interpreted differently between security and AI domains. As one expert said:

> *"I think we need to actually as a group we need to actually have this common understanding of this terminology and what it means to you and how it maps across this domain" [Participant 7, AI Expert]*

The specific objectives, functions and performance measures of security technologies - ***Observe, Detect, Recognise, and Identify*** - were discussed at length by participants. Disparity in definitions and meanings of these terms between AI and security domains were noted, as were the application of the functions to security technologies which facilitated a deeper understanding of how AI is currently applied to security technologies.

In the security domain, the terms *Observe*, *Detect*, *Recognise,* and *Identify* are generally well understood as measurable functions of digital surveillance cameras, based on the DORI (Detection, Observation, Recognition, Identification) standard (European Committee for Standards, 2015; Infinity Electro-Optics, n.d.; US Department of Homeland Security, 2013, pp.27-28). In security surveillance literature, *Observation* refers to 'the function of detecting changes in a scene' (US Department of Homeland Security, 2013, p.48). However in AI, observation may refer to 'a piece of information received online from users, sensors, or other knowledge sources' and in the context of machine learning, an observation refers to a 'data point, row or sample in a dataset' (ML Glossary, n.d.).

> *"Digital video surveillance of observe, detect recognise and identify is only a methodology of applying the amount of pixels that are required to enable those functions to occur… that's done by people, it's not done by machines" [Participant 3, Security Expert].*
>
> *"The machine itself doesn't recognise observe, detect, recognise or identify [Participant 3, Security Expert]… They are the rules which we apply [Participant 1, Security Expert]… Which is knowledge based [Participant 2, Security Expert], so Observe Detect Recognise (DRI) equals knowledge based [Participant 1, Security Expert]"*
>
> *"This function of Observe, Detect, Recognise and Identify… I think it needs to be broken up… so is observe meaning that you just get that image? So, if that is the case it is input, right, for us (AI) its input. So basically you have the camera which is actually acting as some form of sensor that takes input, right. And then when you come to the detect, right, I mean we do have narrow AI techniques, I mean techniques right, that you can automatically nowadays do the detection that doesn't need humans to actually pick up the face… I am trying to actually work out this idea of detecting versus recognise – recognise is, are you saying do I recognise is a normal versus abnormal situation, right or do I recognise a face? Because the two are a bit different, right? Recognition is different" [Participant 7, AI Expert].*

"…if you want to recognise someone, there is enough pixels in the picture that I have seen that person before and can say I know who it is, or I saw the person in this other frame and they had a red jumper on and I'm recognising them. Identify is what's required to go before a court of law and have a, so that somebody says, I can identify that person from that frame, and unless there's enough pixels across the person's face, the law may say well we can't guarantee identification" [Participant 3, Security Expert].

Within the Focus Group, specific terms in AI, such as *Observe* and *Detect*, tended to represent attributes (e.g. input image) as opposed to implying a functional task such as monitoring for change. The column header ***Stimuli*** also created some discrepancies between security and AI terminology. The term Stimuli was noted to refer to the various inputs and characteristics such as colour (e.g. pixel colour/ intensity), shape (object), movement (e.g. difference in pixel intensity exceeds set threshold), feature or behaviour (e.g. anomalous behaviour), rate of change (e.g. amplitude of electromagnetic-wave). Subsequently *Stimuli* was understood to be the specific inputs (event characteristics or elements) which flag an event and begin the process of computing whether an input change is worthy of investigation.

Focus group participants further discussed the technical process involved in performance of security functions. For instance, '*Detection*' in AI terminology could be interpreted as the *Input* (or elements) which flag an event and begin the process of evaluating whether an input change is worthy of investigation. Subsequently the term *Detect* was understood as potentially being a process rather than a single stage of output. Participants agreed that the terms Observe Detect and Identify needed to be separated as security objectives, and from an AI perspective, the focus was on the event objectives.

"Detect means to kind of flag, you know, to raise some sort of flag that some change worthy of investigation or further analysis that's happened, with a high probability… so Detect means to flag something, some event, some characteristic has, you know, suggested this needs further analysis… it's a sifting process" [Participant 6, AI Expert].

"The next point I was going to make is that it's not just a single - in any moderately complex system- there's not just a single stage of detection. I mean, there might be multiple stages of detection. I'm not talking about repeated instances of someone triggering something, I'm saying that you detect this, you detect this, you detect this, and the AI level is going to do some analysis perhaps of each one individually or perhaps of a cooccurrence and all that sort of stuff. So when you say detect that's what I think you are meaning" [Participant 6, AI Expert].

"Really what we need to focus on is not necessarily the traditional security performance measures, but more about how we are using the technology and how that aligns to the problem domains of AI. So for example facial recognition is as a statistical system is a probabilistic system because it's based on a matched score that it will generate a matched score that says more likely than not this is MM or this is PL" [Participant 1, Security Expert].

The disconnect between security terminology and AI terminology was discussed further during the first focus group debriefing session. Consequently, to align security technology more accurately with the AI problem domains and spectrum, amendments were made to the descriptive headers within the Security Technology Tables (Table 21). The column header ***Function/Performance Measure*** was amended to

Security Function. Similarly, the column header ***Stimuli*** was amended to ***Event Characteristics or Elements***.

Table 21

Amendments to Descriptive Headers in Security Technology Tables.

Technology		**Security Function**	**Event Characteristics or Elements**	**Example**	**AI Analytical Methodology**
Technology Category	Sub-category	e.g. Detect, Observe, Recognise, Identify	e.g. difference in pixel intensity exceeds set threshold		
	Sub-category				

2.3.1.2.2 Requirement for Definition of Terms

To assist in reducing communication hurdles to cross-disciplinary collaboration, *Definitions of Terms* were developed for the project to reduce discrepancies in terminology and interpretation. Participants suggested a common language set be developed and introduced to facilitate the articulation of AI in security technology and supporting literature.

"See this is the disconnect because from the security literature, this is what we've always embedded down to, but as we move towards higher performance analytics in AI, we need to change the terminology"[Participant 1, Security Expert].

"What I was going to suggest is we call it 'Security Observation', 'Security Detect', 'Security Response' to move away from the AI Terms… so we have to use these terms" [Participant 2, Security Expert].

"I think we need to actually as a group we need to actually have this common understanding of this terminology and what it means to you and how it maps across this domain" [Participant 7, AI Expert].

"A common vocabulary" [Participant 3, Security Expert]… "Yeah, now that doesn't necessarily mean that someone in computer science in IBM will agree with our vocabulary, but it's about setting a vocabulary for security people" [Participant 1, Security Expert].

The focus group raised the need for a common vocabulary or language set for specific terms such as Observe, Detect, Recognise, and Identify. However, there were also AI systems and terms within the AI Knowledge Map that required discussion and clarification such as 'Rule-based' and 'Expert systems'. AI participants clarified the meaning of both 'rule-based' and 'expert' and established that Expert Systems are in fact Rule-based systems, though rules may be created by humans (including experts) or they may be machine-generated. This concept is well supported within the literature, with

Rule-based systems being identified as a simple form of AI and dependent upon expert systems which 'mimic the reasoning of human expert in solving a knowledge intensive problem' (Grosan & Abraham, 2011, p.149). Grosan and Abraham (2011, p.149) elaborate on the application of expert knowledge to Rule-based systems;

> A rule-based system consists of a set of IF-THEN rules, a set of facts and some interpreter controlling the application of the rules, given the facts. The idea of an expert system is to use the knowledge from an expert system and to encode it into a set of rules. When exposed to the same data, the expert system will perform (or is expected to perform) in a similar manner to the expert. Rule-based systems are very simple models and can be adapted and applied for a large array of problems. The requirement is that the knowledge on the problem area can be expressed in the form of if-then rules. The area should also not be that large because a high number of rules can make the problem solver (the expert system) inefficient.

I'm not an expert in expert systems, and very few people talk about expert systems these days, but to my limited, potentially biased knowledge, there's a very little difference if any between a rule-based system and an expert system [Participant 6, AI Expert].

"…we define the rules, as experts… so when we talk about an expert system, its usually a security expert or a technology expert that says this is the rule. So for example if I get a shift in frequency if I'm looking at an ultrasonic sensor, that frequency shift that is correlated statistically with an intrusion event becomes the rule. So it's set by the expert who has calculated that frequency shift… so what changes from an expert system, how do you differentiate the definition of expert rules-based system?" [Participant 1, Security Expert].

"An expert system is more of a marketing word I guess" [Participant 8, AI Expert].

Certainly one thing you can say is that a machine learning based system, in many cases, perhaps not all, in many cases you cannot clearly articulate why it decided this, right… whereas a rule-based system you can always do a trace through because it's really just a sophisticated IF THEN system [Participant 6, AI Expert].

So that's the rule, so the analytics is the rules put in place, so it's a rules-based system, not an expert system [Participant 1, Security Expert] Yep that's right [Participant 3, Security Expert] Yeah, yeah exactly that's right because its designed by a person [Participant 2, Security Expert].

AI participants explained rules-based and expert systems further, including a discussion of the earlier MYCIN system used for diagnosing blood infections in which the rules were built from extensive expert interviews (Russell & Norvig, 2020, p.23).

"I mean, to give you an analogy, which I think, you know, to my mind comes about because most of what I have read about, this dates back to the 1980's, one of the most successful systems at that time was a system that was supposed to replace medico's right, it was supposed to take IF, this blood pressure is above whatever, right, is called MYCIN… So you could now, in modern terms at least, and probably the concept was around in the 80's, you could have a system which just blindly looks at a whole lot of case histories and then comes up with a rule, you know, just of itself,

because every time this temperature was above a certain thing, AND you know the pulse was at this rate, the doctor prescribed X… you could have a system that does that, right… in a sense it's been informed by an expert because all of that data was coming from experts, but there was no expert actually in the process, you could imagine this as a way of doing it, whereas if you then brought in an expert and said, you know, this is what the systems come up with, does this make sense to you and would you alter any of these rules, would you add any other rules, or whatever – or you could bring the expert in from the start and have none of the machine learning – that to me warrants the title at least of an expert system because, you know, an experts actually been involved in assessing, generating and selecting whatever you want to call it, the rules [Participant 6, AI Expert]…

Although an expert system is understood to be a rules-based system, the consensus between participants in distinguishing between rules-based and expert systems is the 'source of the rules', and whether those rules are human generated, or machine generated.

"There is the notion of whether or not an expert has actually been brought in to either generate the rules or select the rules" [Participant 6, AI Expert]…

"An expert system is a rules-based system" [Participant 8, AI Expert]…

"It's still a rules-based system, but I think it's the source… I think it's the source"[Participant 7, AI Expert]…

"It's an expert informed, or semi-designed" [Participant 6, AI Expert]... "That's right, so basically, they are both rule-based, but one is really having – it is the source, right, one is from human giving you those decisions right, where as a rule-based system it is possible [Participant 7, AI Expert].

"So I guess… the rule can be human based, or machine based" [Participant 1, Security Expert]… "Generated" [Participant 7, AI Expert]… "Machine Generated" [Participant 1, Security Expert]...

Consequently, Table 22 presents terms explored, with the aim of establishing a common language set for the project. However, many of the terms are defined within either the security literature or AI literature, but few terms are defined in both domains. This observation supports the focus group finding of a terminology disconnect between security and AI domains, in that there is little evidence of any merging or common ground in language and understanding of terms. The lack of commonality between terms may signify a need for additional research, beyond this project, to develop a common language set for AI in security terminology.

Table 22

Definition of Terms.

Term	Security Definition	AI Definition
Active	Refers to a communication link that carries a continuous signal allowing immediate	

	detection of a break in the link (Garcia, 2008, p.326).	
Activity Recognition		The ability to recognise/detect current activity on the basis of information received from different sensors, including cameras, wearable sensors, sensors attached to objects of daily use or deployed in the environment (Hussain, Sheng & Zhang, 2019).
Affective Computing		The study and development of systems and devices that can recognise, interpret, process and simulate human affects (Banafa, 2016).
Agent		Anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators (Russell & Norvig, 2020, p.36).
Agent-based Modelling		A powerful simulation modelling technique, where a systems is modelled as a collection of autonomous decision-making entities called agents which individually assesses its situation and makes decisions on the basis of a set of rules (Bonabeau, 2002, p.7280).
Ambient Computing		The combination and evolution of voice and in-the-air gesture interfaces, speech recognition, the internet of things, cloud computing, wearable computing, the quantified self, augmented reality, haptics and, above all, artificial intelligence and machine learning (Elgan, 2018).
Analytics		Includes Descriptive Analytics, Diagnostic Analytics, Predictive Analytics and Prescriptive Analytics and uses Machine Learning techniques to find patterns and discover insights and relationships within data (Anodot, n.d.).
Artificial Intelligence		The study of agents that receive percepts from the environment and perform actions (Russell & Norvig, 2020, p.vii)
Artificial Narrow Intelligence		Also known as weak or applied intelligence, represents most of the current AI systems which usually focus on a specific task (Guo, Farhang-Razi & Algra, 2019, p.350).
Artificial Neural Networks		An Artificial Neural Network (ANN) is a computational model in machine learning, which is inspired by the biological structures and functions of the mammalian brain. Such a model consists of multiple units called artificial neurons which build connections between each other to pass information. The advantage of such a model is that it progressively "learns" the tasks from the given data without specific programming for a

		single task (Guo, Farhang-Razi & Algra, 2019, p.350).
Artificial Superintelligence		Intelligence that far surpasses human ability (Russell & Norvig, 2020, p.33)
Assessment	The determination of the cause of an alarm and information regarding the threat (Garcia, 2008, p.327)	
Automation Level	The Automation level is defined as a dedicated communications network for the sole purpose of building device connectivity, communication and control (automation) (Brooks et al, 2017, p.25).	
Autonomous Systems	One network or series of networks that are all under one administrative control. An autonomous system is also sometimes referred to as a routing domain. An autonomous system is assigned a globally unique number, sometimes called an Autonomous System Number (ASN).	A sub-field that lies at the intersection between robotics and intelligent systems (e.g. intelligent perception, dexterous object manipulation, plan-based robot control, etc) (Corea, 2019).
Bayesian Network		A Bayesian Network, also called Bayes Network, belief network, or probabilistic directed acyclic graphical model, is a probabilistic graphical model (a statistical model) that represents a set of variables and their conditional dependencies via a directed acyclic graph (Guo, Farhang-Razi & Algra, 2019, p.351).
Bayesian Program Synthesis		A form of probabilistic programming, where Bayesian programs write new Bayesian programs (instead of humans, as in the broader probabilistic programming approach) (Corea, 2019).
Communication	The function of transmitting or interchanging information, including both transmission of alarm signals to a central processing station and transmission of response information to security personnel (Garcia, 2008, p.328)	
Computer Vision	Computer vision is the ability for a computer to use an artificial intelligence algorithm to 'see' and interpret both still images and video (Moore, 2020).	The branch of AI concerned with the extraction of meaningful structures from images perceived by a system (Frankish & Ramsay, 2014, p.337).
Decision Networks		A decision network represents information about the agent's current state, its possible actions, the state that will result from the agent's action, and the utility of that state (Russell & Norvig, 2020, p.545).
Deep Learning		Machine Learning using multiple layers of simple, adjustable computing elements (Russell & Norvig, 2020, p.26)
Deep Neural Networks		A neural network architecture with many layers, typically 5–100. A network with only a few layers is called a shallow network (Guo, Farhang-Razi & Algra, 2019, p.353).

Detection	Determining that an unauthorised action has occurred or is occurring; detection includes sensing the action, communicating the alarm to a control centre and assessing the alarm (Garcia, 2008, p.330)	
Distributed Artificial Intelligence		A class of technologies that solve problems by distributing them to autonomous "agents" that interact with each other. Multi-agent systems (MAS), Agent-based modelling (ABM) and Swarm Intelligence are examples of this where collective behaviours emerge from the interaction of decentralised self-organised agents (Corea, 2019).
Embodied Intelligence		An approach to AI and cognitive science that largely renounces symbolic representations and formal reasoning, and emphasises context, physical embodiment, social interaction and sensorimotor behaviour over generally abstractness, individualism and logically rigorous thought (Frankish & Ramsay, 2014, p.337).
Evolutionary Programs/ Evolutionary Algorithms		Evolutionary programs can alter their own rules by using genetic algorithms (Frankish & Ramsay, 2014, p.337).
Expert System		An expert system is a computer system that simulates the ability or behaviour of a human expert on performing a task. An expert system incorporates the knowledge base that represents facts and rules, and the inference engine that uses the knowledge base to deduce new conclusions (Guo, Farhang-Razi & Algra, 2019, p.354).
Field Level	The Field level comprises of sensing or action devices that are generally self-contained physical units. Field level devices are connected to automation level controllers, either application specific or generic controllers (Brooks et al, 2017, pp.25-26).	
Fuzzy Systems		A specific example of rule-based systems that map variables into a continuum of values between 0 and 1, contrary to traditional digital logic which results in a 0/1 outcome (Corea, 2019).
Generative Adversarial Networks		Two networks that train each other (Corea, 2019).
Genetic Algorithms		Genetic algorithms are a machine learning method for finding solutions to certain kinds of problems, loosely analogous to the biological process of

		artificial selection (Frankish & Ramsay, 2014, p.338).
Identification	The positive assessment of a recognised object as a specific person, animal or thing (Garcia, 2008, p.333)	
Inductive Logic Programming		Inductive reasoning is a reasoning method which uses premises to supply evidence in order to support the conclusion. Opposed to deductive reasoning, inductive reasoning works as a down-top logic 356 Glossary which provides the conclusion by generalizing or extrapolating from special cases to general rules (Guo, Farhang-Razi & Algra, 2019, pp.355-356).
Machine Learning		Machine Learning is a field in computer science that builds computational models that have the ability of "learning" from the data and then provide predictions. Depending on whether there is a supervisory signal, machine learning can be divided into three categories: the supervised learning, unsupervised learning, and reinforcement learning (Guo, Farhang-Razi & Algra, 2019, p.357).
Machine Translation		Machine translation (MT) is the application of computers to the task of translating texts from one natural language to another (European Association for Machine Translation [EAMT], n.d.)
Machine Vision		Machine Vision is the technology used to provide image-based automatic analysis for applications in industry such as automatic inspection, process control, and robot guidance (Guo, Farhang-Razi & Algra, 2019, p.357).
Management Level	An organisation's Information Technology and Communications (ITC) network, comprising operator stations, monitoring and operator units, programming units and other peripheral computer devices connected to a data processing device i.e., a server to support the information exchange monitoring and management of the automation system (Brooks et al, 2017, p.25; International Organization for Standardization, 2004, p. 53).	
Multi-agent Systems		A collection of autonomous agents that need to coordinate their activities in order to achieve their individual goals. Coordination is achieved through negotiation or argumentation an, in most applications, requires that the agents learn to adapt to each other's strategies (Frankish & Ramsay, 2014, pp.339-340).

Natural Language Processing		Processing texts by computer for some practical or useful purpose (Frankish & Ramsay, 2014, p.340).
Natural Language Understanding		Natural Language Understanding (NLU) allows the chatbot to discern between entities, intent, timeframes and even locations, allowing understanding of an entire sentence or question (Adenin, n.d.).
Neural Network		Also known as artificial neural network, neural net, deep neural net; a computer system inspired by living brains (Guo, Farhang-Razi & Algra, 2019, p.358).
Observation	The function of detecting changes in a scene' (US Department of Homeland Security, 2013, p.48).	A data point, row, or sample in a dataset. Another term for instance (ML Glossary, n.d.).
Passive	Refers to a communication link that carries a signal only when an alarm occurs (Garcia, 2008, p.336).	
Probabilistic Programming		A framework that does not force you to hardcode specific variables, but rather works with probabilistic models (Corea, 2019).
Recognise		Object recognition refers to the detection in images of a known instance (a specific object, e.g. y mug) or a class of objects (e.g., the set of all mugs) (Frankish & Ramsay, 2014, p.340).
Reinforcement Learning		Reinforcement learning is a type of dynamic programming that trains algorithms using a system of reward and punishment. The algorithm is exposed to a total random and new dataset and it automatically finds patterns and relationships inside of that dataset. The system is rewarded when it finds a desired relationship inside of that dataset but it is also punished when finds an undesired relation. The algorithm learns from awards and punishments and updates itself continuously. This type of algorithm is always in production mode. It requires real time data to be able to update and present actions. The agent learns without intervention from a human by maximizing its reward and minimizing its penalty Guo, Farhang-Razi & Algra, 2019, pp.360-361).
Respond Response	The effort to neutralise, contain, or mitigate an event (ASIS, 2015, p.17), or The element of a physical protection system designed to counteract adversary activity and interrupt the threat (Garcia, 2008, p.338).	
Robot		A physical device capable of behaviour in the world involving interactions with its

		environment through sensors and actuators (Frankish & Ramsay, 2014, p.340).
Robotics		The field of study dedicated to the science and engineering of robots (Frankish & Ramsay, 2014, p.340).
Robotic Process Automation		Technology that extracts the list of rules and actions to perform by watching the user doing a certain task (Corea, 2019).
Search & Optimisation		Tools that allow intelligent search with many possible solutions (Corea, 2019).
Sensor	A device that responds to a stimulus associated with an unauthorised action, such as an intrusion into a protected area or an attempt to smuggle contraband through an entry (Garcia, 2008, p.339).	
Statistical System		The statistical approach is based on mathematical tools to solve specific sub-problems (Corea, 2019).
Sub-symbolic System		The sub-symbolic approach is one that no specific representations of knowledge is provided ex-ante (Corea, 2019).
Supervised Learning		Training a model from input data and its corresponding labels (Guo, Farhang-Razi & Algra, 2019, p.361).
Swarm Intelligence		See 'Distributed Artificial Intelligence'
Symbolic System		A computer program that performs computations with constants and variables according to the rules of algebra, calculus, and other branches of mathematics. Also known as algebraic computation system; computer algebra system; symbolic computation system (The Free Dictionary, n.d.).
Unsupervised Learning		Unsupervised learning is a type of machine learning algorithm used to draw inferences from sets of data consisting of input data without labelled responses, e.g., cluster analysis. This means that the system is exposed to a total random and new dataset and it automatically finds patterns and relationships inside of that dataset (Guo, Farhang-Razi & Algra, 2019, p.362).

Nevertheless, this discussion led to the development of Figure 11, the Security Technology-Artificial Intelligence Linguistic cycle. The cycle aims to communicate the alignment between how security technology operates through sensing, processing, deciding and acting, and the AI cycle of data input, computational technique, rule checking and defined output.

Figure 11

The Security Technology-Artificial Intelligence Linguistic cycle

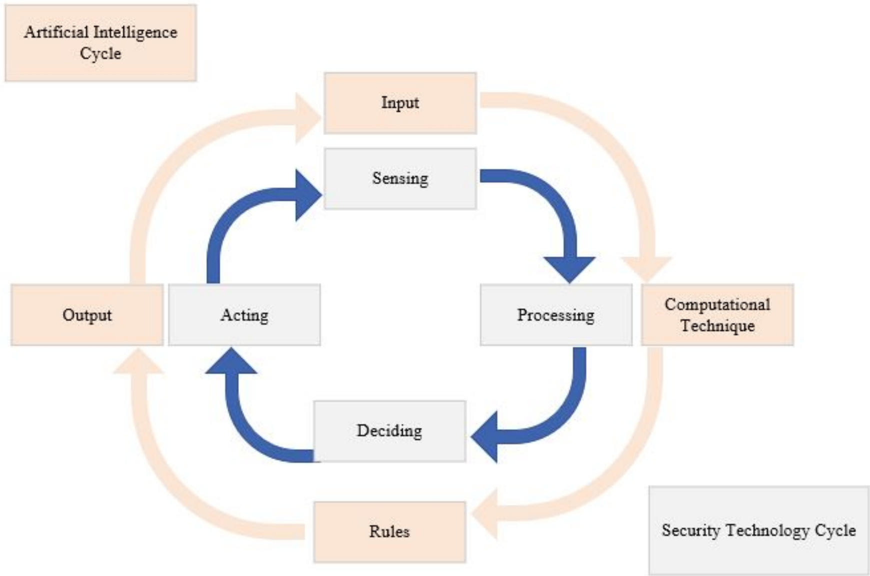


Figure 11 presents the analogous language used between security technology process functionality where all technologies sense, process, decide and act, and the AI computational process where input data are subjected computational techniques (algorithms), then computational outputs are compared to rules and where a set threshold is met, outputs are triggered, resulting in actions.

A common language set may facilitate the development of ***A Practical Guide to Artificial Intelligence in Security Technologies***. For instance, an 'AI Security Vocabulary' may define, explain, and detail a range of AI terms in relation to their function and role in security technologies. For example, denoting the term *Computer Vision*, how it is applied to Network Surveillance Systems, Robotics, Drones, and other technologies in which it is incorporated, and how it may be utilised by security professionals to assist in the development of security objectives and achievement of outcomes.

2.3.1.2.3 Alignment of Security Technologies with AI Computational Techniques

Following the Focus Group and debriefing sessions, further amendments were made to the descriptive headers and content within the Security Technology Tables to differentiate AI problem domains and methods (such as Computer Vision, Machine Learning, Neural Networks etc) with the specific analytical techniques and computational processes (such as Principal Component Analysis, Independent Component Analysis, etc) used by security technologies in processing and decision making.

"How important is it for the security people, to really truly know the exact technique that's being employed" [Participant 7, AI Expert]… "I guess what we want to do is get a ballpark understanding of the techniques for security people, so you know if you look at knowledge on a continuum, what we want to do is say look, you know, it's a rules-based system" [Participant 1, Security Expert].

To achieve this, two specific amendments were subsequently made to the headers in the Security Technology Tables (see Table 23). Firstly, the contents of the column 'AI Analytical Methodology' were transferred and merged into the 'AI Paradigm' column, using this column as a singular reference point in alignment with Corea's 'AI Knowledge Map' model (Corea, 2019). Secondly, the column header 'AI Analytical Methodology' was amended to 'AI Computational Technique' and populated with specific techniques and algorithms used by each security technology.

Table 23

Further Amendment to Descriptive Headers in Security Technology Tables.

Technology		**Security Function**	**Event Characteristics or Elements**	**AI Paradigm**	**AI Computational Technique**
Technology Category	Sub-category			e.g. Statistical (Probabilistic); Computer Vision (Activities Recognition)	e.g. Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Sub-category				

This amendment was applied to all four Security Technology Tables (Observing, Detecting, Controlling and Response Technology Tables) and populated accordingly as shown in Table 24 (Observing Technologies and AI Analytical Alignment), Table 25 (Detection Technologies and AI Analytical Alignment), Table 26 (Controlling Technologies and AI Analytical Alignment) and Table 27 (Responding Technologies and AI Analytical Alignment).

Table 24

Observing Technology and AI Analytical Alignment.

Technology		**Security Function**	**Event Characteristics or Elements**	**AI Paradigm**	**AI Computational Technique**
Network Video Surveillance	Digital Video Surveillance	Observe, Detect, Recognise, Identify	Colour, shape, contrast, silhouette, feature, movement	Computer vision (Probabilistic Programming) Statistical	Background Analysis, and Statistical Modelling (Subspace

					Learning, Kernel Density Estimation or Gaussian Mixture Method).
	Video Motion Detection	Detection	Contrast, Silhouette, as indicative of Movement (Statistical Change in Colour/shade)	Computer vision (Probabilistic Programming) Statistical	Pixel Matching, Foreground Analysis, Background Subtraction, or Gaussian Mixture Method (GMM)

Table 25

Detecting Technologies and AI Analytical Alignment

Technology		**Security Function**	**Event Characteristics or Elements (Stimuli)**	**AI Paradigm**	**AI Computational Technique**
Internal and external Intrusion Detection Systems	Contact Switch/ Reed Switch (Electromechanical Sensor)	Detection	Current Interference (loss) through breaking of connection between the two points	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Glass Break Sensor (Electro-mechanical Sensor)	Detection	Vibration or acoustic technology – elevations in vibration or audio levels across a surface	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Passive Audio Sensor	Detection	Threshold change in sound levels	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Passive Microphonic Cables	Detection	Vibration phenomenology changes - the system is sensing noise levels of differing types and amplitude	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Photoelectric Beams (visible or infrared)	Detection	Light Interruption/loss	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Active Infrared Motion Sensor	Detection	infrared beam Interruption/loss	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Fibre Optic Cable Sensors	Detection	Change in transmitted light signal	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output

	Microbend Fibre Optic Sensor	Detection	Change in transmitted light signal	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Passive Infrared Sensors (Thermopile or Pyroelectric Detector)	Detection	Thermal energy threshold change	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Active Ultrasonic Motion Sensor	Detection	Changes in reflected Ultrasonic energy	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Active Monostatic Microwave Motion Sensor	Detection	Electromagnetic energy wave threshold shift	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Capacitance Proximity Sensor	Detection	Threshold change in capacitance (electrical charge)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Pressure Sensor	Detection	Electrical signal generation	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Seismic detector	Detection	Change in levels of vibration	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Light Sensors	Detection	Change in predetermined light level	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Laser detectors	Detection	Change in reflected laser beams/patterns	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Dual Technology Sensors	Detection	AND gate logic requiring activation of either or both types of sensors	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Alarm Annunciator Panel	Detect, Response	Inputs from field sensors communicate audio visual signals for faults any failures of the critical system or process	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Duress Alarms	Detect, Response	Command Signal; Emergency signal transmitted from duress mechanism to local or central monitoring station	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output

	Electromagnetic Buried line sensor system	Detect response	Changes in electromagnetic signal between two underground cables	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Buried line sensor systems Hydrostatic	Detect response	Changes in hydrostatic pressure by deformation of containment vessel	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Taut wire detection systems	Detect response	Changes in the measured level of strain-by-strain gauges	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Taut wire detection systems	Detect response	Changes in the measured level of movement micro by sensor switches	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Microwave detection systems	Detect response	Change in the microwave patterns between bi static transmitter and receiver	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Guided radar systems	Detect response	Change in the radar pulses between transmitter and receiver cables	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Buried Geophone systems	Detect response	Changes in noise levels detected by sensors	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Network Video Surveillance	Anomaly Detection (Nearest Neighbour-based Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning); Computer Vision (Activity Recognition)	K-Nearest Neighbour; Relative Density
	Anomaly Detection (Classification-based Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning); Computer Vision (Activity Recognition)	Neural Network; Bayesian Network; Support Vector Machine; Rules
	Anomaly Detection (Clustering-based Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning); Computer Vision (Activity Recognition)	K-Means; Hierarchical
	Anomaly Detection (Information Theoretic Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning); Computer Vision (Activity Recognition)	Entropy; Relative Uncertainty; Kolomogorov Complexity
	Anomaly Detection (Spectral Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning);	Principal Component Analysis;

				Computer Vision (Activity Recognition)	Compact Matrix Decomposition
	Anomaly Detection (Statistical Technique)	Detect, Recognise, Notify	Nominal and Anomalous Human Behaviour (Action Recognition)	Statistical (Probabilistic and Machine Learning); Computer Vision (Activity Recognition)	Gaussian Mixture models; Histogram; Kernels
	Video Motion Detection	Detect, Recognise, Notify	Movement (Statistical Change in Colour)	Statistical (Probabilistic); Probabilistic Programming	Pixel Matching, Foreground Analysis, Background Subtraction, or Gaussian Mixture Method (GMM)
Biometric Detection Systems	Biometric Facial Recognition – Holistic Matching Method	Detect, Recognise, Identify	Eigenfaces/ Eigenvector & Pixel Intensity Array without the detection of facial features	Statistical (Probabilistic); Probabilistic Programming	Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Feature-based (Structural) Method	Detect, Recognise, Identify	Facial features create a geometric relationship between measurement points of the individual's unique features	Bayesian Program synthesis Probabilistic Programming (Statistical)	Structural Similarity Measure (SSIM), and Feature Similarity Measure (FSM), Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Hybrid Method (Combination of Holistic & Feature-based Methods)	Detect, Recognise, Identify	3D images to capture data (E.g. eye-socket curvature, chin shape).	Bayesian Program synthesis Probabilistic Programming (Statistical)	Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Thermal & Infrared (IR) Detection	Detect, Recognise, Identify	Heat signature of Reflected Infrared (IR) Radiation (Near-Infrared (NIR) & Short-wave Infrared	Bayesian Program synthesis Probabilistic Programming (Statistical)	Feature Vector analysis and classification from minutiae points in blood perfusion image

			(SWIR) ranges); and Thermal Infrared Radiation (Mid-wave Infrared (MWIR) and Long-wave Infrared (LWIR) ranges) generated thermograms from unique physiological inputs such as blood vessels or vascular network		(application of medial axis transform (MAT), morphological grey level erosion or Sobel operator on bit-plane imaging)
	Multimodal Biometric Systems for CCTV	Detect, Recognise, Identify	Biometric measurement from single or multiple sensors of multimodal characteristics (E.g. face, gait, height, iris, fingerprint etc)	Bayesian Program synthesis Probabilistic Programming (Statistical)	Structural Similarity Measure (SSIM), and Feature Similarity Measure (FSM), Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Multi-algorithmic Biometric Systems	Detect, Recognise, Identify	Biometric measurement from a single sensor analysed with two or more algorithms	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Multi-instance Biometric Systems	Detect, Recognise, Identify	Biometric measurement from single or multiple sensors of multiple instances of the same biometric characteristic(s)	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Multi-sensorial Biometric Systems	Detect, Recognise, Identify	Biometric measurement of a biometric characteristic using multiple different sensors, and	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or

			analysed using one or more algorithms		Independent Component Analysis
Lighting Systems	Active Lighting Systems	Detection	Output changes in environmental sensors	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Passive Lighting Systems (Failsafe/ Emergency Lighting)	Detection	Electrical Current interruption	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
X Ray	X-Ray	Detection	X-ray electromagnetic wave interaction with material	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Trace Detection	Explosive/ Narcotics Trace Detection (Ion Mobility Spectrometry)	Detection	Threshold change in concentration of explosive/drug molecules in saturated air or surfaces	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Acoustic Detection Systems	Acoustic Anomaly Detection	Detection	Nominal and Anomalous Acoustic Input	Bayesian Program synthesis Probabilistic Programming (Statistical)	Feature Vector & Principal Component Analysis (PCA); Mel-frequency cepstral coefficient (MFCC); Gaussian mixture model (GMM); Hidden Markov models (HMMs)
	Intelligent Acoustic Surveillance (Acoustic Event Detection)	Detect	Nominal and Anomalous Acoustic Input	Bayesian Program synthesis Probabilistic Programming (Statistical) or Machine Learning with Supervised, Unsupervised or Reinforcement Learning	Feature Vector & Principal Component Analysis (PCA); Mel-frequency cepstral coefficient (MFCC); Gaussian mixture model (GMM); Hidden Markov models (HMMs)
RADAR Systems	Radar Video Surveillance (RVS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR), Reflected Scene Image (Video)	e.g. Symbolic (Logic Based);	Pulse-doppler signal processing; Linear Frequency Modulation (LFM)

	Ground Penetrating Radar (GPR)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR)	e.g. Symbolic (Logic Based);	Pulse-doppler signal processing; Linear Frequency Modulation (LFM)
	Through-the-Wall Sensor (TTWS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR)	e.g. Symbolic (Logic Based);	Pulse-doppler signal processing; Linear Frequency Modulation (LFM)
SONAR Systems	Active Sonar Systems	Detect, Recognise	Transducers (hydrophone array) emit and detect acoustic energy reflected/ emitted from sound/energy source	e.g. Symbolic (Logic Based);	Sonar waves converted to/from electrical signal by hydrophone arrays; emitter transducer (T_{dx}) and receptor transducer (R_{dx})
	Passive Sonar Systems	Detect, Recognise	Transducer (hydrophone array) detects acoustic energy emitted from sound/energy source	e.g. Symbolic (Logic Based);	Sonar waves converted to electrical signal by transducer (hydrophone array)

Table 26

Controlling Technology and AI Analytical Alignment

Technology		**Security Function**	**Event Characteristics or Elements?**	**AI Paradigm**	**AI Computational Technique**
Access Control Systems	EAC Reader	Identify, Control	Presentation or input (e.g. PIN) of authorised credential at reader	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Smart Card	Identify, Control	Presentation of authorised codec	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	2FA or multifactor (Two or Three Factor Authentication) Access Control	Identify, Control	Two or three pre-authorised credentials. Combination of one Knowledge factor (e.g. PIN), one Possession factor (e.g. Smart Card)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output

			and/or one Inherent factor (Biometric credential)		
	RFID (Radio Frequency Identification)	Identify, Control	Presentation of RFID embedded chip to RFID reader based on pre-set permissions	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	NFC (Near Field Communication)	Identify, Control	Generation of electromagnetic fields created when two devices are brought together (e.g. smart phone & reader)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Internet of Things Access Control (Intelligent Door Locking Systems)	Identify, Control	Presentation of authorised credential at any one of connected access control readers using unique IP addresses and connected through the internet	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Bluetooth Access Control	Identify, Control	Presentation of Bluetooth enabled and pre-authorised device at Bluetooth reader using Wi-Fi or cellular data network	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Biometric Systems	Biometric Facial Recognition	Detect, Recognise, Identify	Biometric measurement (Signature)	Bayesian Program synthesis Probabilistic Programming (Statistical)	Statistical Probability Match Score (Probabilistic system)
	Biometric Facial Recognition – Holistic Matching Method	Detect, Recognise, Identify	Eigenfaces/ Eigenvector & Pixel Intensity Array without the detection of facial features	Bayesian Program synthesis Probabilistic Programming (Statistical)	Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Feature-based (Structural) Method	Detect, Recognise, Identify	Facial features create a geometric relationship between measurement	Bayesian Program synthesis Probabilistic Programming (Statistical)	Structural Similarity Measure (SSIM), and Feature Similarity Measure (FSM), Principal

			points of the individual's unique features		Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Hybrid Method (Combination of Holistic & Feature-based Methods)	Detect, Recognise, Identify	3D images to capture data (E.g. eye-socket curvature, chin shape).	Bayesian Program synthesis Probabilistic Programming (Statistical)	Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Biometric Facial Recognition – Thermal & Infrared (IR) Detection	Detect, Recognise, Identify	Heat signature of Reflected Infrared (IR) Radiation (Near-Infrared (NIR) & Short-wave Infrared (SWIR) ranges); and Thermal Infrared Radiation (Mid-wave Infrared (MWIR) and Long-wave Infrared (LWIR) ranges) generated thermograms from unique physiological inputs such as blood vessels or vascular network	Bayesian Program synthesis Probabilistic Programming (Statistical)	Feature Vector analysis and classification from minutiae points in blood perfusion image (application of medial axis transform (MAT), morphological grey level erosion or Sobel operator on bit-plane imaging)
	Multimodal Biometric Systems for CCTV	Detect, Recognise, Identify	Biometric measurement from single or multiple sensors of multimodal characteristics (E.g. face, gait, height, iris, fingerprint etc)	Bayesian Program synthesis Probabilistic Programming (Statistical)	Structural Similarity Measure (SSIM), and Feature Similarity Measure (FSM), Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Multi-algorithmic Biometric Systems	Detect, Recognise, Identify	Biometric measurement from a single sensor analysed with two or more algorithms	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or

					Independent Component Analysis
	Multi-instance Biometric Systems	Detect, Recognise, Identify	Biometric measurement from single or multiple sensors of multiple instances of the same biometric characteristic(s)	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis
	Multi-sensorial Biometric Systems	Detect, Recognise, Identify	Biometric measurement of a biometric characteristic using multiple different sensors, and analysed using one or more algorithms	Bayesian Program synthesis Probabilistic Programming (Statistical)	Combination of Principal Component Analysis, Linear Discriminant Analysis, or Independent Component Analysis

Table 27

Responding Technology and AI Analytical Alignment

Technology		**Security Function**	**Event Characteristics or Elements?**	**AI Paradigm**	**AI Computational Technique**
Communication Systems	Telephones, Mobile Phones, Pagers, Emergency Phones	Response	Command Signal; Transmitter converts electrical signal into radio waves which are radiated from the device antenna to the radio receiver (mobiles) or transmits electrical signal via wire/ fibre optics (landlines)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Radio Communicat ion	Response	Command Signal; Radio transmitter transforms energy into radio waves radiated from the device antenna to the radio receiver	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output

	Emergency Warning Intercom Systems (EWIS) Public Address / Mass Notification System	Control, Response	Command Signal by operator/ user; Transmission of voice audio or automated instructions via IP network protocol (e.g. VoIP, RTSP, TCP/IP, SIP)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	IP Network Connected Intercom	Control, Response	Command signal by operator/user; transmits audio and visual data via IP network protocol (e.g. TCP/IP, SIP, RTSP)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Wireless Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via encrypted wireless signal (radio waves)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Wired Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via low-voltage electrical wiring	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Secured monitoring solutions via hardwired or IP networks	Control, Response	Data gathering solutions as isolated copper and fibre systems transmitting to central control locations	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Door Release Intercom	Control, Response	Command signal by operator/user allowing remote unlocking of doors via wiring from the intercom to the electrified locking hardware	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Dispensable Barriers	Active Dispensable Barrier	Delay, Response	Command Signal	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Passive Dispensable Barriers	Delay	Contact with Adversary	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Vehicle Barriers (Hostile Vehicle Mitigation)	Rising Bollards	Delay, Response	Command signal activates bollard to deploy	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output

	Rising Wedge Barrier	Delay, Response	Command signal activates wedge barrier to deploy	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Beam Barrier	Delay, Response	Command signal activates beam barrier to deploy	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Automatic Gates	Detect, Delay, Response	Command Signal, PLUS Presentation of Authorised Credential	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
Weaponry	Acoustic Weaponry	Delay, Response	Sonic, infrasonic, or ultrasonic frequencies. High-amplitude infrasonic weapons combine two high-volume ultrasonic frequencies at a difference of 8Hz	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	High-Voltage (Electroshock) Weaponry	Delay, Response	Transducers emit high-voltage, low amplitude current to incapacitate adversary	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Electromagnetic Weaponry (Directed Energy Weapon)	Delay, Response	Electromagnetic energy beam created by a Gyrotron (Vacuum Electronic Device (VED) to create high power high frequency THz radiation)	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Optical Distractors (Directed Energy Weapon)	Delay, Response	High intensity beams of light directed at adversary to create nausea, confusion, disorientation, or temporary blindness	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Remotely Operated Weapons Systems (ROWS)	Delay, Response	Weapons operated remotely in response to event or attack	e.g. Symbolic (Logic Based);	Logic Programming or Rule based Process output
	Autonomous & Semi-Autonomous Weapons Systems	Delay, Response	Predefined Stimulus for which autonomous systems may execute self-direction, self-learning or emergent behaviour to select and attack targets	e.g. Symbolic (Logic Based); Bayesian Program synthesis Probabilistic Programming (Statistical)	Logic Programming or Rule based Process output; Multiple Algorithms for Classification, Regression or Clustering

2.3.2 Objective Two

2.3.2.1 Alignment of Security Technology to AI Spectrum

Objective Two was facilitated with the following Sub-research Questions:

Q2. Can you show where on the AI Domain continuum the following technologies, or elements of them are located:

Q3. Based on the response to question two, do you feel any technology categories and functions have been missed when employing technology in the protection of assets?

To achieve Objective Two, participants were provided with four AI Mapping Diagrams which indicatively mapped the Observing, Detecting, Controlling and Response Technologies with the BACS Architectural Levels (i.e. Field Level, Automation Level and Management Level) and the AI Paradigms (i.e. Symbolic, Statistical and Sub-symbolic paradigms) contained in the *AI Knowledge Map* (Corea, 2019). The AI Mapping Diagrams were used to visually depict where security technology system components may be located and to allow participants to validate or modify the placement of each technology and its components.

The AI Mapping Diagrams were presented and discussed in the following order (see Appendix A: Indicative Technology Tables Presented to Participants);

- Observing Technologies AI Mapping Diagram
- Detecting Technologies AI Mapping Diagram
- Controlling Technologies AI Mapping Diagram
- Response Technologies AI Mapping Diagram

This was a relatively complex process, as most security technologies consisted of elements or components across all three BACS Architecture Levels. For instance, in Network Video Surveillance, system components such as sensors, communications and hardware are located at the Field Level; while signal processing and analytics is generally located and performed at the Automation Level; and the interface for system management, configuration and programming is located at the Management Level. Further alignment of these components with the AI Paradigms was provided to indicate where on the AI spectrum these technologies and their technological components are currently located.

In addition, the placement of security technologies was complicated by the multifunctional and highly integrated nature of some technologies. As highlighted in Phase One, security is achieved through the mixing of technological measures across observe, detect, control and respond. Subsequently, some security technologies were included in more than one security category. For example, Network Video Surveillance as an environmental sensor system is primarily an observation technology, however when combined with Motion Detection or Anomaly Detection software, this technological treatment tool also achieves detection. The components of each system also sit across both the BACS Architecture and the AI Paradigms, for instance in Network Video Surveillance, the cameras, lenses and sensors are located at the Field Level (although technically it is possible some sensors may include some level of built-in automation and processing), while device connectivity, communication and control is located at the Automation Level and the programming and configuration of the system through the user interface is located at the Management Level (Brooks et al, 2017, p.25). In terms of alignment with the AI

Spectrum, Network Video Surveillance field-level devices may be located in the Symbolic Paradigm (Logic and Knowledge Based), while video analytics for anomaly detection are located in the Statistical (Probabilistic and/or Machine Learning) Paradigms. Therefore, capturing the location of security technologies and their components was the focus of Objective Two.

2.3.2.2 Objective Two Findings

The following results were found in response to Sub-research Question Two: *Can you show where on the AI Domain continuum the following technologies, or elements of them are located?*

2.3.2.2.1 Amendments to Observation Technologies Alignment with AI Spectrum

The indicative placement of Observing Technologies in alignment with the problem domains and spectrum of AI paradigms was generally accurate, with sensors and system components considered to be at the field level, signal processing and analytical techniques at the automation level and configuration and programming at the management level.

If we look at 'Observe' from taking an image in, that sits at the field device level, so you've got a lens and a camera arrangement in general [Participant 1, Security Expert]… That's video surveillance - observe - and then at the automation level we've got video analytics and at the management level we've got the human interface [Participant 2, Security Expert].

So at the automation level, when we talk about video analytics, really we're talking about a defined event, or an undefined event, depending on whether we're looking at an anomaly or a normal... so for example at an airport, a bag left behind is picked up in the analytical software as an alert, so it's an event. So it's an event based on a statistical analysis of the pixelization change that should occur… so when I walk and drop my bag, in the field of view of the CCTV system, then I walk away, I separate my pixelization from the pixelization of the bag. If that bags left for a period of time, the analytics send an alert, that says there's a bag left behind. So what we're looking at is a defined event. It's a rule that we set as humans in terms of predicting the scenarios that pose a risk [Participant 1, Security Expert].

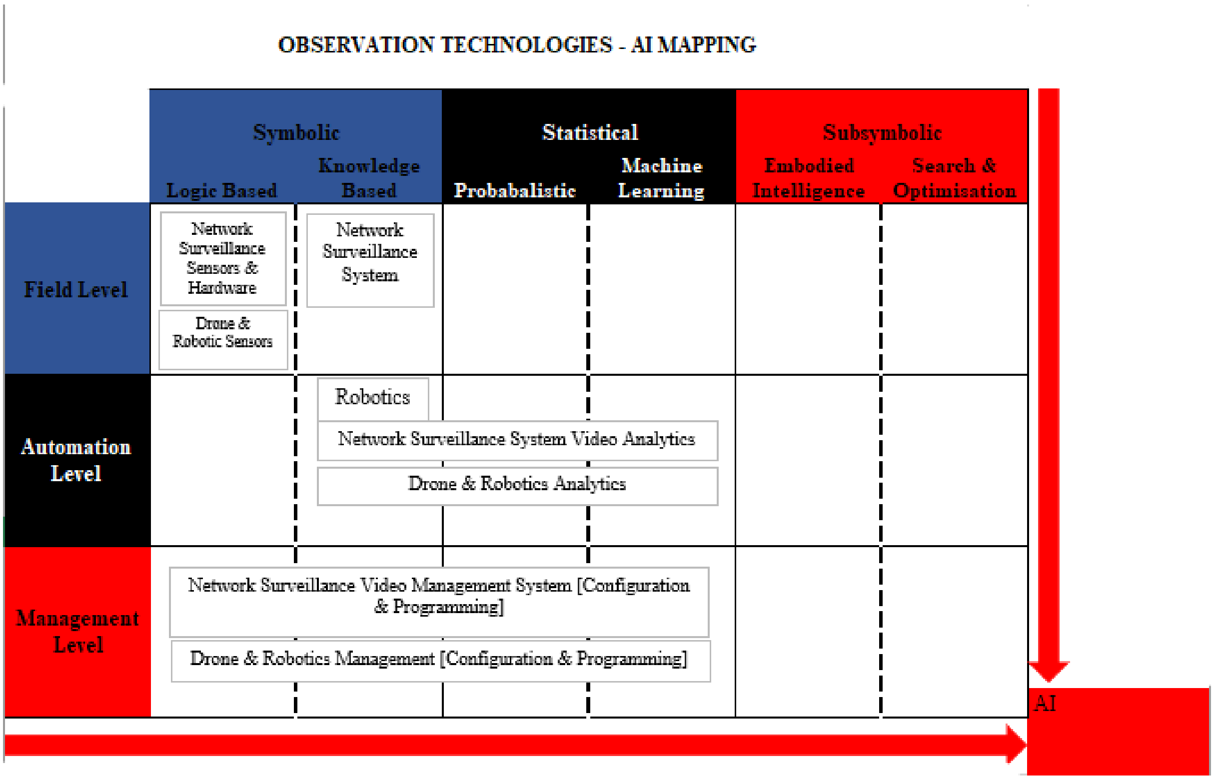
However based on participant discussions, there were some alterations made to the alignment and terminology of observation technologies. While the location of systems, sensors and hardware remained at the Field Level, amendments included further distinguishing between components as either logic based or knowledge-based components. At the Automation Level, signal processing and video analytics (including anomaly detection and behavioural analytics) were emphasised as being 'Analytics'; and configuration and programming at the management level was amended to 'Management' or 'Management System'. At the automation and management levels, observation technologies were realigned across both symbolic and statistical localities of the spectrum of AI paradigms as shown in Figure 12 (Phase One).

Findings showed that Observation Technologies were confirmed to be multifunctional systems and were aligned to more than one security category (i.e. Observe, Detect, Control, and Respond). Network Video Surveillance (Anomaly Detection and Video Motion Detection) were subsequently included in the category of Detection technologies. Drones and robotics were also added into the Observing Technologies AI mapping diagram in accordance with participant recommendations.

> *"The only other thing in Observe is whether we need to put Robotics in there" [Participant 2, Security Expert]*
>
> *"Any mobile robot is likely to use at least one of those technologies as well, and actually more likely to use those than vision" [Participant 6, AI Expert]*

Figure 12

Figure 12 presents the amended alignments between observation technologies BACS architecture levels, AI computational tasks and the spectrum of AI paradigms.



The Focus Group found all Observation Technologies sit in the Symbolic and Statistical AI Paradigms and included perception, knowledge and planning problem domains.

- Network Surveillance Systems are predominantly Symbolic (Logic-based and Knowledge-based) with Video Analytics and Management Systems (Configuration & Programming) spanning across Symbolic (Knowledge-based) and Statistical Paradigms (Probabilistic and Machine Learning).

- Drones and Robotic sensors are predominantly Symbolic (Logic-based). Drone and Robotics Analytics are aligned with Symbolic (Knowledge-based) and Statistical (Probabilistic and Machine Learning) Paradigms. Management Systems (Configuration & Programming) are aligned with across Symbolic and Statistical Paradigms.

- No Observation technologies currently sit within the Sub-symbolic AI Paradigm

2.3.2.2.2 Amendments to Detection Technologies Alignment with AI Spectrum

The indicative placement of Detection Technologies in alignment with the problem domains and spectrum of AI paradigms was again reasonably accurate (Phase One), with sensors, hardware and system components considered to be at the field level, signal processing and integration at the automation level and configuration and programming at the management level.

Detection technologies were discussed at length between participants, in particular Intrusion Detection Systems and biometric technologies and their alignment with the problem domains and spectrum of AI paradigms. At the field level, participants discussed field level sensors and agreed that sensors such as reed switches used in Intrusion Detection Systems were simple rule systems, based on simple threshold logic. By contrast, biometric systems were confirmed to be probabilistic in nature.

> *"In Intrusion Detection, we might use a standard electromagnetic reeds switch to say well someone opening the door and moving the magnet is indicative of an intrusion event… ok so in security technology it's a fairly low-level sort of switch technique. But really, in terms of the sort of AI problem domains..., we have sort of initially put it as an expert system or rule system, but is it a rule system?" [Participant 1, Security Expert]. "It's just a simple rule" [Participant 7, AI Expert].*
>
> *"No-one in their right mind would call that an expert system, it's a simple threshold logic" [Participant 6, AI Expert].*
>
> *"So for example facial recognition is as a statistical system is a probabilistic system because it's based on a matched score that it will generate a matched score that says more likely than not this is Participant 8 or this is Participant 7" [Participant 1, Security Expert].*

> *"a facial recognition system by and large is not just the software but it's an enrolled database. You can't divorce the two and the most important part about what I just said is that I don't think there is a system on the planet or more generally just outside even facial recognition that knows when its likely making a mistake right. And so if you give it a person that's not in the database it will typically match it to the nearest person, right. It will not say I don't know this person, right, I mean that's not to say people haven't produced systems that kind of claim they do that with some reliability etcetera, but again it's based on a kind of a, you know you've got some idea of a threshold…" [Participant 6, AI Expert].*

"If we look at facial recognition as sort of sitting within that sort of statistical probabilistic - it probably sits at the automation level indicatively … you have sensory input, so you have raw input and that sits at the field device level so a camera taking in the initial reflected scene is a sensor input system but the actual processing that gives you a matched score that says this is actually Participant 3 or this is Participant 1, that really sits at that automation/ management level…" [Participant 1, Security Expert].

"I was wondering whether we need to call it a network video surveillance system/ CCTV system at the field level but when we start getting into automation level does it become video analytics?" [Participant 2, Security Expert].

"So at the field level, we have a lens and a capture system. At the automation level and the management level we're really going to a video analytics system, and that's not captured very well in the literature at all" [Participant 1, Security Expert]… "that's right, and in fact all we're doing is capturing visual data and then various systems deal with it in different ways. And one of the ways they use Observe, Detect Recognise and Identify is just to enable people to build a system that a guy sitting in front of a screen can visually see things… that's where it comes from which is not associated with analytics at all" [Participant 3, Security Expert].

Terminology remained an issue throughout the participant discussion on Detection technologies, with participants querying and differentiating between sensors and detection, and the meaning of the term 'Detect'. The overlap of detection with other functions (Observe, Detect, Control & Respond) was explored by participants.

"Here again we are using the word detection. To me this is actually a sensor which has a simple rule, and the sensor gives an output" [Participant 7, AI Expert].

'You talk about a camera, but a camera is also a detector in itself… it has detected enough photons to show different than black. I mean, you really can't make this division in any powerful, meaningful way, I mean it is always going to descend down to essentially indistinguishable, right – there is no sensor on the planet that in some primitive sense doesn't detect… its detecting something, its doing a little bit more than detection because its typically giving you not just a yes/no – you might argue a reed switch is only giving you a yes/no – but most of the other sensors give you a magnitude of the effect or some other measurement associated" [Participant 6, AI Expert].

"I guess what we're saying here is there's a difference between the sensing and the detection – detection is the outcome of threshold change" [Participant 1, Security Expert]… "Yes, it is" [Participant 3, Security Expert].

Machine Learning was discussed with the general outcome being that machine learning in security technology and systems is elementary, or in the early stages of development. Systems which claim to be adaptive systems, are typically self-calibrating systems rather than true machine learning.

"Normally machine learning usually has no explicit rules whatsoever in it… its data, data, data, you know, human labelling of what that data means…" [Participant 6, AI Expert]… "So you don't initially define the rules" [Participant 1, Security Expert]… "That's right" [Participant 7, AI Expert]… And very few of them to my knowledge further evolve right on new data – they're trained, you know, don't get me wrong I mean there's obviously a prototype system and there's you know some lifecycle of the development I'm not saying that isn't the case but normally you know there's some notion that this has now been trained and can be deployed and very few of them collect additional examples to self-train further on, so I'm not aware of any but there might be some" [Participant 6, AI Expert]…

"So what you're saying if we have a system where we initially want to develop the rules and then add to those rules over time, it's really a hybrid system?" [Participant 1, Security Expert]… "A hybrid of what?" [Participant 6, AI Expert]… "A hybrid of Rules and then some Machine Learning, or does that not exist either?... so what would you define as something that starts initially with some rule and then develops the rules further over time?" [Participant 1, Security Expert]… "Are you talking about rules as in starts with training sets for classification of…" [Participant 4, Security Expert]… Even things like initiated alarm - so for example if you've got an intrusion detection technology, you've got the field device the sensor that comes in with a frequency measure and over time that algorithm adjusts, so as the wind increases for example some technology will adjust for that wind… [Participant 1, Security Expert]…"Yeah that adaptability has been around for a while" [Participant 6, AI Expert]… "But what do you call it in AI?" [Participant 1, Security Expert]… "I was thinking more sophisticated adaptability where you get something where it does something completely wrong, and then it you know says ok that's some new data and makes some sort of adjustment… um some sort of self-calibration is more like what you're talking about, some sort of continual calibration… now I know that could be said to also encompass the other but I'm talking about a very low level calibration right so, you know, adjusting to the ambient temperature you know, kind of thing, rather than adjusting to an unexpected event that triggered the wrong rule" [Participant 6, AI Expert].

Acoustic anomaly detection was discussed in terms of the rules, algorithms and making predictions, and how this aligns to Corea's (2019) AI Knowledge Map. However the consensus on acoustic detection systems is that they are Expert rule systems;

"With acoustic detection systems, so basically it is you have a database, so it's like intrusion detection system basically signature based… [Participant 7, AI Expert]… "Yes it is, it's really good because they're doing it with drones as well because they reckon that drone signatures from RADAR and other inputs are really weak so they are using – to detect approaching drones – they're using the acoustic signatures because its much stronger" [Participant 4, Security Expert]… So basically there's a database just like intrusion detection system there's a database of all known attacks? [Participant 7, AI Expert]… "Yes" [Participant 1, Security Expert]… "No I asked that question – I asked that question and that's first "do you record drones" as they come on the market as your database – No" [Participant 2, Security Expert].

"We're not going to put an acoustic detection system under a deep learning system – were not – it's a sensor, that's all it is. It may be an expert system as a detection but that's it" [Participant 2, Security Expert]… "Or there's some analytics when we go into automation level" [Participant 1, Security Expert]… "So to put it the terms that I used before, this is someone who is saying you know we've got um a basic physical sensor capability which is based on acoustics, um we have

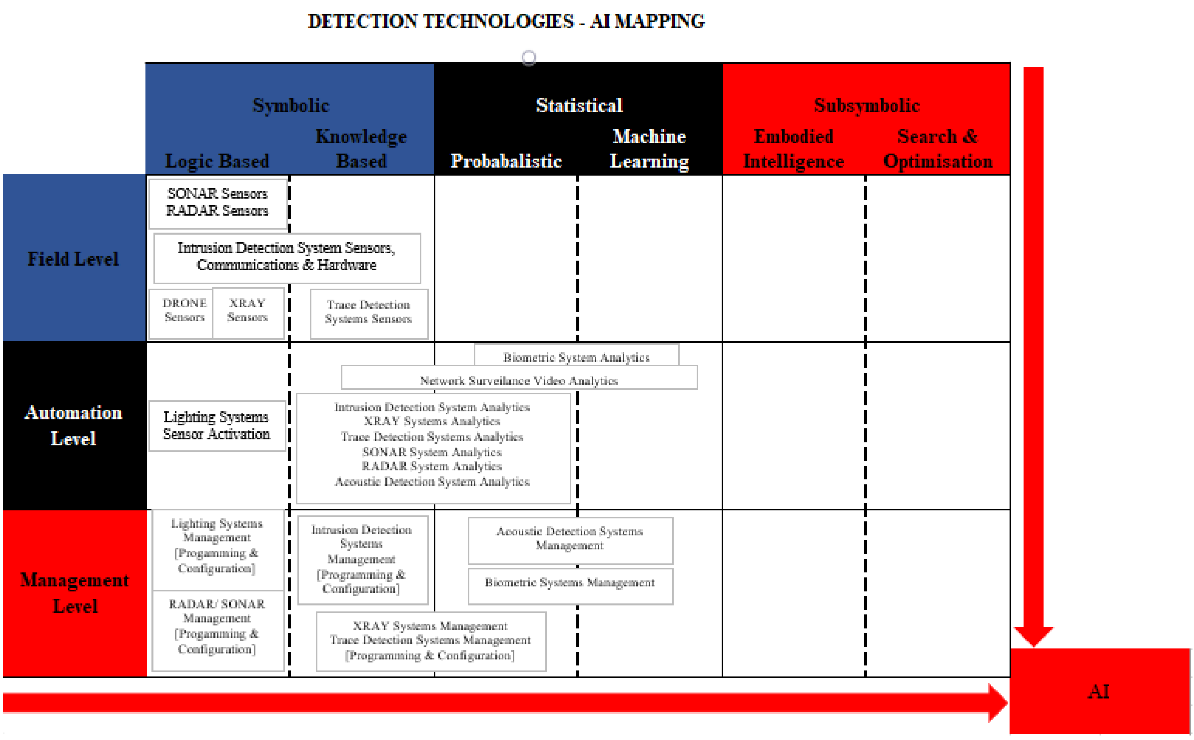
built this into an integrated system with you know deep learning and algorithms for inference or whatever that is meant to mean, and it can detect gunshots or aggressive behaviour, but all I could say is that, you know looking at, without even looking at the paper all I can say is that sounds plausible… if I look at the paper all I can say is yes it sounds plausible and I'm a little more confident because they've backed it up with that, but I could still not in any sense, get a feel for how market ready this is" [Participant 6, AI Expert]… "Well it's on the market now, so the technology's out there and I guess what we're trying to say is, is it really deep learning" [Participant 1, Security Expert]…

Participant discussions resulted in several alterations to the Detection Technologies AI mapping diagram including the addition of SONAR and RADAR sensors, specificity of biometric systems, and position adjustment of several technologies at the automation and management levels to include representation of Machine Learning. Drones and robotics were also added into the Detection Technologies AI mapping diagram in accordance with participant recommendations.

I was wondering whether on this [Detection AI Mapping Diagram], I don't think it captures the sort of 'sense' and 'detect' component… you've got SONAR and RADAR, but it needs a detection sensor" [Participant 2, Security Expert].

Figure 13

Figure 13 presents the amended alignments between Detection Technologies BACS architecture levels, AI computational tasks and the spectrum of AI paradigms.



The Focus Group found all Detection Technologies sit within the Symbolic and Statistical AI problem domains and spectrum of AI paradigms.

- Intrusion Detection System sensors, including Biometric sensors, communications and hardware are Symbolic with Management Systems (Configuration & Programming) located within the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms, and Intrusion Detection Analytics located in the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- Biometric System Analytics and Management Systems align to the Statistical (Probabilistic and Machine Learning) Paradigm.

- SONAR and RADAR sensors, communications and hardware are Symbolic with Analytics and Management Systems (Configuration & Programming) located across Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- Network Surveillance Video Analytics used for anomaly detection are aligned with Symbolic (Knowledge-based) and Statistical (Probabilistic and Machine Learning) Paradigms.

- Acoustic Detection System Analytics are aligned with both Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms, while the management of Acoustic Detection Systems is aligned with the Statistical Paradigm (Probabilistic and Machine Learning).

- Drone sensors as detection mechanisms are located with the Symbolic Paradigm (Logic-based).

- X-ray sensors are Symbolic (Logic-based) with both Analytics and Management Systems (Configuration & Programming) located within the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.
- Trace Detection Systems are Symbolic (Knowledge-based) with both Analytics and Management Systems (Configuration & Programming) located within the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- Lighting Systems including sensors, activation and management are aligned with the Symbolic Paradigm (Logic-based).

- No Detection technologies currently sit within the Sub-symbolic AI Paradigm.

2.3.2.2.3 Amendments to Controlling Technologies Alignment with AI Spectrum

The indicative placement of Controlling Technologies in alignment with the AI problem domains and spectrum of AI paradigms was supported by the group, with sensors, hardware and system components for raw data input considered to be at the field level, analytics the automation level and configuration and programming at the management level. Expert and rules-based systems were more dominant within these technologies, with the exception of biometric technology.

When we look at the indicative analytical technology, what we start to see again emerging except for biometrics – so in most of the smart cards, or different blue-tooth access cards etc, we're seeing a very rules-based system… you have a sensor or field device, that the system poles that field device when you present your credential, it takes that credential information to the database and it says yep Participant 1 has authorised access, and Participant 1 therefore meets that rule, that rule then initiates an outcome so as an input/output system it unlocks the door and allows access" [Participant 1, Security Expert].

When we go to biometric systems so example facial recognition, hand geometry etcetera, again we start to move into that sort of, more probabilistic system, so it's based on a statistical analysis that's if there's a matched score that's equivalent according to that rule, then that then allows the automation level rule to be executed and give me access, if it doesn't match, I don't get access [Participant 1, Security Expert].

Biometrics was strongly identified as aligning more closely with the Statistical AI Paradigm due to the probabilistic techniques employed by the various biometric systems, and therefore biometrics were separated from Access Control Systems and located as a separate technology (see Figure 14).

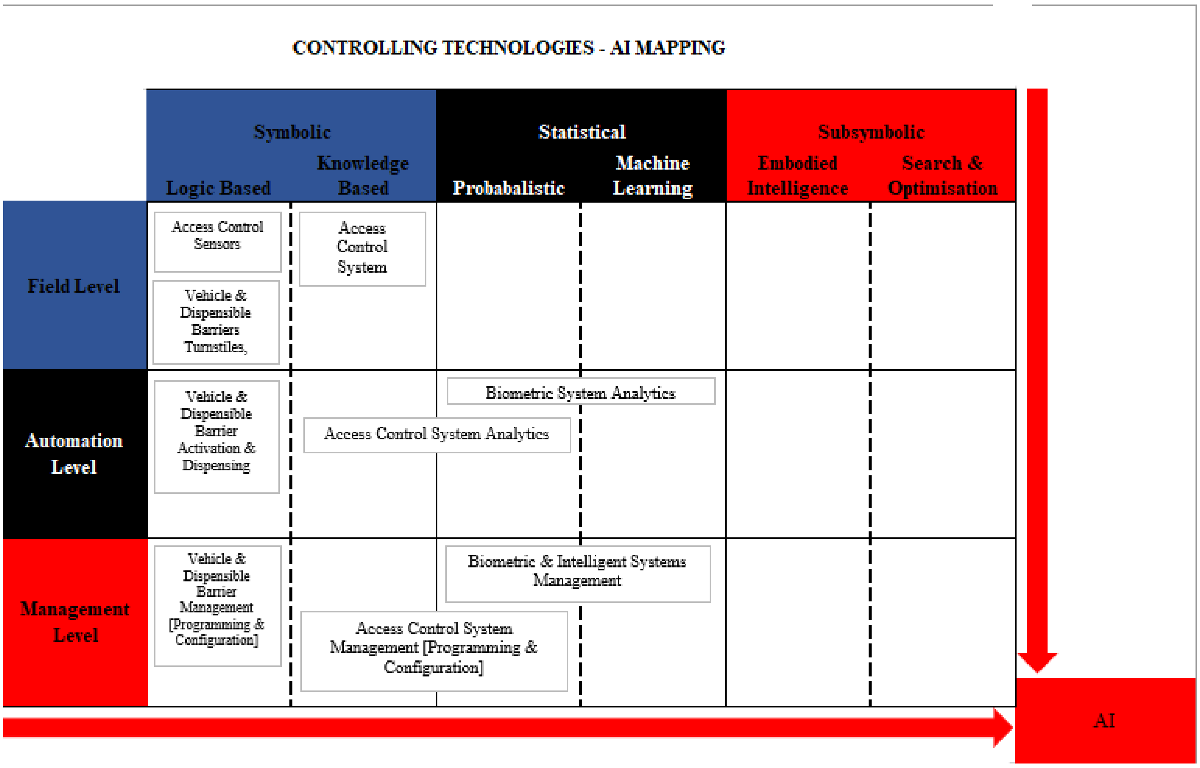
Participants also discussed in depth the interrelationship between the rule-based systems and the configuration and programming at the management level, and the security ramifications of management level access to system configuration.

"At management level is really where the human makes the decision, right?" [Participant 7, AI Expert]… "Yeah" [Participant 2, Security Expert]… "Yep" [Participant 1, Security Expert]… "Which decision? I mean if we follow the analogy that I thought we were using um, you know, there's rules in the system which are presumably based on some sort of statistics the developer has gathered, and it'll say you know this biometric is within or outside of tolerance from the enrolled um you know, database of people who have got access at various levels" [Participant 6, AI Expert]

The other thing that was being talked about is more about who is allowed to enrol people and disenroll people from that database, which is, you know, kind of a configuration thing right rather than in any deep sense, changing the rules underneath which the system's operating – it's just meagrely a configuration – I know it has important ramifications – but it's a configuration and it's an almost inevitable configuration you would have to have for the system to be useful in any realistic situation, right, you could have the ability to cancel someone's authorisation to add someone new etcetera" [Participant 6, AI Expert].

Figure 14

Figure 14 presents the amended alignments between Controlling technologies BACS architecture levels, AI computational tasks and the spectrum of AI paradigms.



The Focus Group found all Controlling Technologies sit in the Symbolic and Statistical AI Paradigms.

- Access Control System sensors, including Biometric sensors, align with the Symbolic Paradigm. Access Control Management Systems (Configuration & Programming) and Analytics are located within the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- Biometric System Analytics and Management Systems align to the Statistical Paradigm (Probabilistic and Machine Learning).

- Vehicle and Dispensable Barrier & Turnstile hardware and mechanisms are located in the Symbolic Paradigm (Logic Based).

- No Controlling technologies currently sit within the Sub-symbolic AI Paradigm

2.3.2.2.4 Amendments to Response Technologies Alignment with AI Spectrum

The indicative placement of Response Technologies in alignment with the AI problem domains and spectrum of AI paradigms placed sensory input at the field level, the application of defined rules at the automation level; and analytics, configuration and programming at the management level.

The Focus Group discussed the Response Technologies category of Autonomous Weapons Systems in relative depth, and found that although the literature aligns Autonomous Weapons Systems as starting to converge on the sub-symbolic and embodied intelligence AI spectrum, these systems still sit within the definition of Broad AI. Broad AI is the 'integration of two or more AI (Narrow AI) outputs within a specific process within an enterprise". The Focus Group found that Autonomous Weapons Systems integrate a series of Narrow AI tasks, and subsequently amendments were made to the alignment of Autonomous Weapons Systems with the AI Spectrum (Figure 15).

So if we're looking at the autonomous weapons systems, looking at the literature, they sit much higher on the AI spectrum that what the general security technologies do…apparently - that's within the literature… within the literature they're saying its very much AI and that um some of these systems are able to adjust automatically and make decisions, they're able to track, classify and do a whole range of things that would sit within – not within general AI – but closer to it" [Participant 4, Security Expert]… "So Broad AI" [Participant 1, Security Expert]…"Its Broad AI because it's got a capture, the drone capture devices they've got a radar linked into a capture drone… "An integration of [Participant 7, AI Expert]… "Yeah… drone, RADAR, surveillance" [Participant 2, Security Expert]… "So it's a series of like this is what Participant 7 was saying in her diagram, what you have is a whole series of input/outputs where you integrate different narrow AIs to get a broader AI outcome" [Participant 1, Security Expert]… "Broad AI means yes integrating Narrow AI technologies together to actually achieve a certain goal" [Participant 7, AI Expert].

"I'm well outside my expertise I guess because I don't know much about weapons systems but I would just be, and I notice you know that you didn't necessarily say that you agreed with that literature, but I'm just highly sceptical of it for several reasons – the first reason is just a very fundamental thing and I was going to say this earlier – we seem to be making a big meal out of a core problem that when someone says someone's adaptive you've got to say 'what do you mean'… What does it adapt to, and in what circumstances can you guarantee that it will do the right thing for that adaptation… and rather than accept adaptability is a general term that means it can adapt to everything" [Participant 6, AI Expert]… "Within the literature they tend to say adaptation for optimisation, so even with uh, they talk about cognitive RADAR and SONAR and that the incoming signals can be processed and adapted to optimise the environmental conditions, and they will change the frequency at which they project the outgoing signals based on the environment…" [Participant 4, Security Expert]… "Is that the system doing that?" [Participant 2, Security Expert]… "Yes, Cognitive RADAR and Cognitive SONAR" [Participant 4, Security Expert]… "So is that more a statistical" [Participant 1, Security Expert]… "It sounds probabilistic" [Participant 2, Security Expert]… "It sounds a sophisticated version of my temperature adaptation, frankly… I mean you know, taking in more variables and etcetera but it doesn't sound in any meaningfully AI sense/way anything more than you saying you broaden the range of adaptability, and I'll accept that – you can certainly broaden the range of adaptability, but nonetheless, nothing on the planet is as adaptable as a human, nor is it perfectly adaptable, nor can it actually ever be" [Participant 6, AI Expert].

Let me also articulate the second reason why I wouldn't accept that weapons based systems are actually the pinnacle of AI, um while I do accept the point that there are some very very clever people in this research – because it's very highly lucrative, um my observation is that firstly, you know, these people, you know, they're not particularly better than the people out in the commercial industry and so on – and so even then they might just get a little bit of a lead on it, there might be some wraps kept on it but eventually one of those persons are going to leave and they don't yet get all shot when they leave, you know, and so and one of them is going to either publish it or build a commercial product on it because it's usually adaptable to something else, right whatever it is they've discovered in the military area, and there's no perfect example of that than GPS in a way, you know… they tried to keep it a bit under wraps, until they got a better system and then they released the system, you know, that was less effective than they could turn on or turn off for their own military purposes, but essentially the notion that you can use GPS to know where you are, just, you know – there's such a commercial imperative that it's never going to stay in the military forever" [Participant 6, AI Expert]…

"And most things don't… anything in the military environment, be it the technology or the actual systems themselves, will always be leaked out" [Participant 1, Security Expert]… "I would suggest that anything which is in the weapons domain at best is a year or two in advance and mostly is probably slightly inferior because they've got you know, extra safety checks and all sorts of other things involved…" [Participant 6, AI Expert].

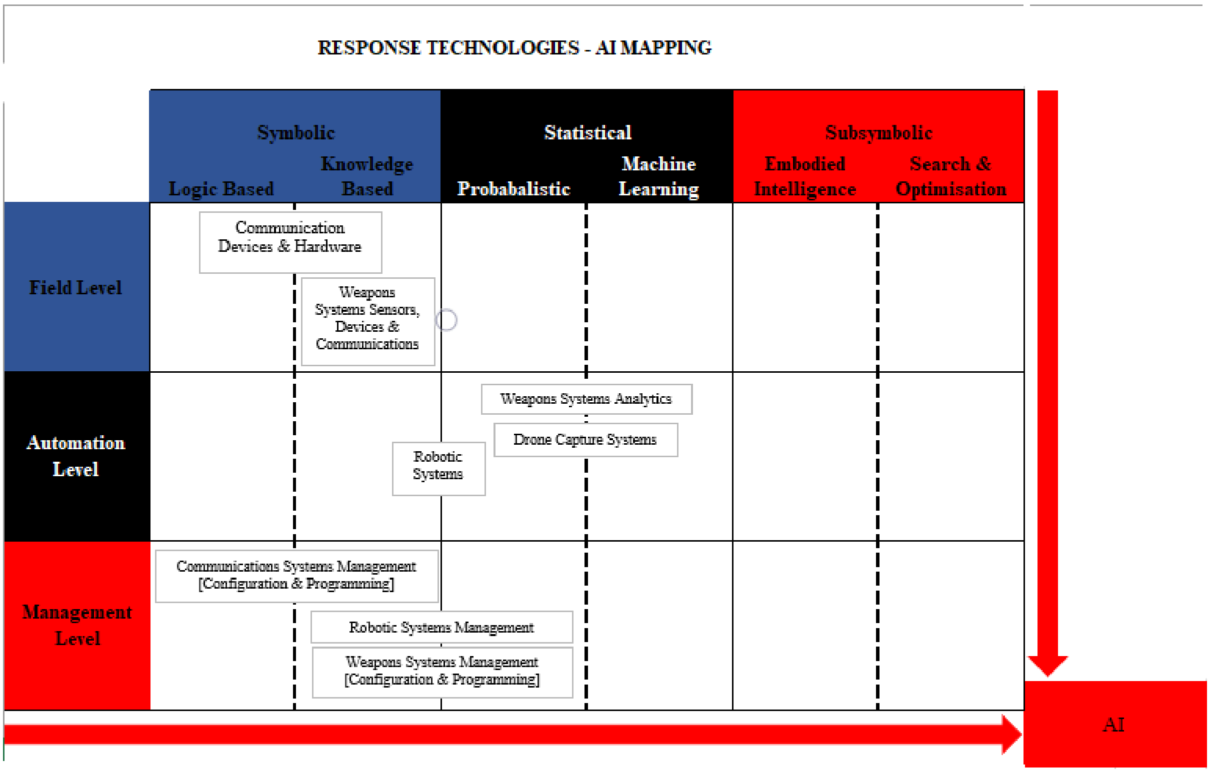
The focus group experts were of the view that Autonomous Weapons Systems are at best a year or two in advance, and are basically Broad AI, not close to General AI or the Sub-symbolic Paradigm of AI.

"So I guess where we're coming from that is really when we talk about automation in terms of response, what you're really doing is more moving into Broad AI where you've got the integration of Narrow AI input/output technologies that achieve broad AI that gives the perception of human intelligence but it's not really" [Participant 1, Security Expert]… "because when you break those things down to their sub-components, so detection, identification, classification based on training sets, and all of those components and then target tracking and all the rest of it, it's still those individual – it's just basically integration of those individual components" [Participant 4, Security Expert].

Drones and robotics were also added into the Response Technologies AI mapping diagram in accordance with participant recommendations. Consequently, Figure 15 presents the Amended alignments between Response technologies BACS architecture levels.

Figure 15

Figure 15 presents the Amended alignments between Response technologies BACS architecture levels, AI computational tasks and the spectrum of AI paradigms.



The Focus Group found all Response Technologies sit in the Symbolic and Statistical AI Paradigms.

- Communications System devices and hardware align with the Symbolic Paradigm (Logic-based and Knowledge-based).

- Robotic System automation and management functions are located within the Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- Drone Capture Systems used in response technologies are located within the Statistical Paradigm (Probabilistic).

- Weapons Systems sensors, communications and hardware are Symbolic with Analytics aligning with the Statistical Paradigm (Probabilistic and Machine Learning). Weapons Management Systems (Configuration & Programming) spanning align with both Symbolic (Knowledge-based) and Statistical (Probabilistic) Paradigms.

- No Response technologies currently sit within the Sub-symbolic AI Paradigm

2.3.2.2.5 Addition of Security Technologies for the Protection of Assets

In response to Sub-research Question Three; '*Do you feel any technology categories and functions have been missed when employing technology in the protection of assets?'* there were two technologies added to the Security Technology categories. Drones and Robotics were added to Observation, Detection and Response Technologies (see section 4.2.2.1). One access control technology (Wiegand Cards) was recommended to be deleted as this is an obsolete technology and no longer used.

2.3.3 Objective Three

2.3.3.1 Locate AI Hardware, Firmware, Software to Network

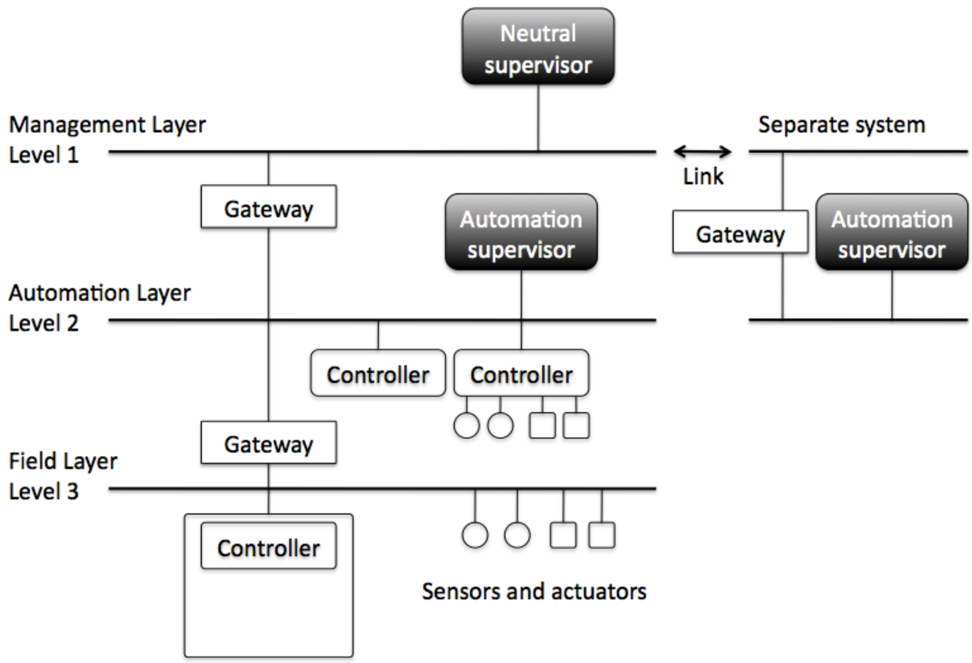
Objective Three was to locate where AI sits in relation to the network architecture of security technologies; that is AI networked integration in relation to the BACS/ Intelligent Controller Architecture levels. Objective Three was facilitated with the following Focus Group question:

Q4. Where does AI sit in relation to the BACS/ Intelligent Controller Architecture Levels?

Participants were provided with diagrams of the *BACS/ Intelligent Architecture Levels* and *Functional Architecture of an Intelligent Controller* (see Appendix C). Participants were asked to locate and draw where AI sits in relation to the BACS architecture levels (Figure 16).

Figure 16

Figure 16 presents that BACS/ Intelligent Architecture Levels from Phase One.



2.3.3.2 Objective Three Findings

The following results were found in response to the Focus Group Question *Where does AI sit in relation to the BACS/ Intelligent Controller Architecture Levels?*

2.3.3.2.1 Location of AI in BACS Intelligent Architecture Levels

Participants identified that Narrow AI could be at any level (Level 3 Field Layer, Level 2 Automation Layer, or Level 1 Management Layer) in the BACS architecture.

> *"I think Narrow AI will sit at all three levels" [Participant 2, Security Expert]… "You can use it at any level" [Participant 6, AI Expert]… "Yeah you can use it at any level" [Participant 7, AI Expert]… "I think it more naturally sits down lower because that's usually where the demands are clearer about what it has to do" [Participant 6, AI Expert]*
>
> *"…there's hardly a sensor around that doesn't have some implicit threshold of detection, noise rejection, whatever you want to call it, right, they're always packaged with something like this these days" [Participant 6, AI Expert]… "As field equipment has become more sophisticated, there's much more decision making made in the piece of field equipment" [Participant 3, Security Expert]*
>
> *"…you can more easily put AI the lower level you go in all of this because usually and I use that stress – usually, the demands are less, but I can imagine a system where the management is at a system where it's pretty flexible about [inaudible] tolerance and all sorts of things and you know, you might build in a bit of AI there that allegedly is at the management level" [Participant 6, AI Expert].*
>
> *"So basically narrow AI can be at any level" [Participant 1, Security Expert]… "You wouldn't often put it to the management level [Participant 6, AI Expert]… "Probably normally your Narrow AI would be much more at lower level at the execution level, right where basically you have all the sensors and things like that right" [Participant 7, AI Expert].*
>
> *"…Narrow AI can actually sit across the three levels, I agree to a certain degree but as we have discussed previously I also said that most likely that the narrow AI is more predominant at the lower level at the execution level, right, and it goes back to the thing we were talking about where basically with intelligent automation that you know, that you could actually have the broad and narrow AI at the coordination level, right, and then further up at management level, basically you have management basically being able to interface with the system that actually brings in the human decision as well as you know, to help with again the Broad AI and maybe you know, some narrow AI if its needed…" [Participant 7, AI Expert].*

Participants further discussed the preliminary '*Functional Architecture of an Intelligent Controller'* diagram and discussed how to achieve the mapping and alignment of the BACS architecture levels and Intelligent controller architectural levels, as shown in Figure 17 (see Table 28 for description of Levels):

Figure 17

Figure 17 presents the indicative Functional Architecture of an Intelligent Controller

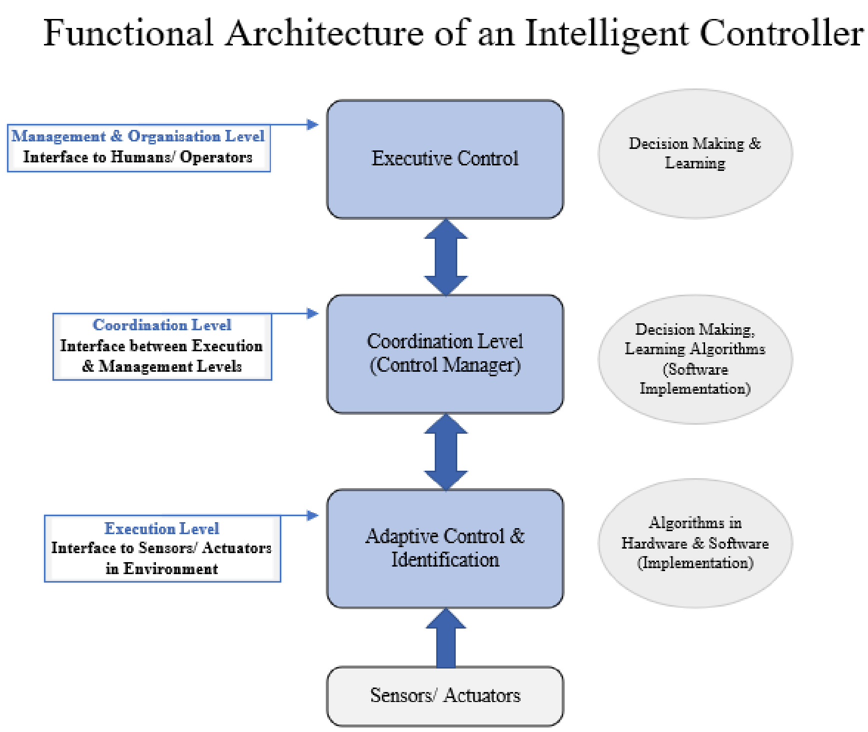


Table 28 presents a description of one participant's views on what denotes an Intelligent Controller for AI and its relationship to the BACS architecture levels.

Table 28

Description of Intelligent Controller Functional Architecture Levels.

Level	Description
Execution Level	Low level numeric signal processing and control algorithms.
Coordination Level	Provides for tuning, scheduling, supervision and redesign of the execution-level algorithms; crisis management; planning and learning capabilities for the coordination of execution-level tasks; and higher-level symbolic decision making for failure detection and identification (FDI) and control algorithms management.
Management Level	Provides for the supervision of lower-level functions and for managing the interface to the humans. The management level will interact with the users in generating goals for the controller and assessing capabilities of the system. This level also monitors the performance of the lower-level systems, plans activities at the highest level (in cooperation with the human) and performs high-level learning about the user and lower-level algorithms (excerpt).

Distinctions were made between the automation level in BACS Architecture Levels and the Coordination Level of the Functional Architecture of the Intelligent Controller:

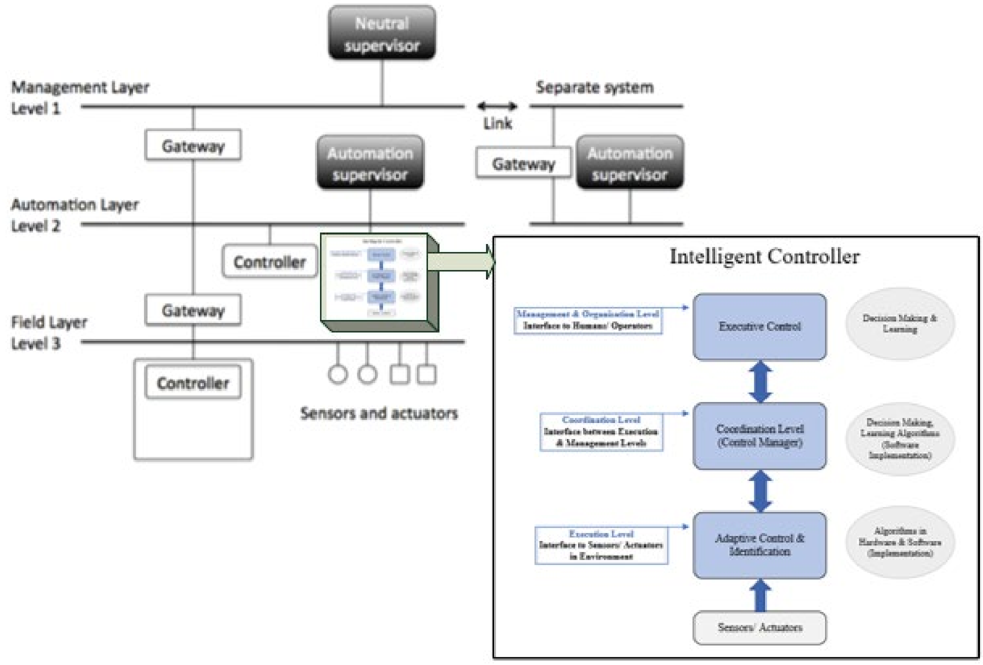
> *"How do we get both sides having a common understanding?" [Participant 7, AI Expert]… Your diagram, and it's quite a good diagram on the sort of smart controller, and certainly when you go to the Broad AI, you're really probably sitting more at…" [Participant 1, Security Expert]… "on the Coordination Level, more on that level" [Participant 7, AI Expert]… "So when you talk about the coordination level um, where would you align that according to the three levels that we're working off?" [Participant 1, Security Expert]… … the idea of what coordination level is which provides for tuning, scheduling, supervision, design of tasks and things, right" [Participant 7, AI Expert]…"It's almost a higher level up from automation [Participant 2, Security Expert]… "Yeah so it sort of sits between Management and Automation… I guess where I'm coming from is you know we haven't been able – ourselves yet - to map it directly into that three-level architecture, so for example this standard architecture according to the literature is that Field, that Automation, and that Management Level… so what you're sort of saying is that really its sitting in between the Management Level and the Automation Level" [Participant 1, Security Expert]…*
>
> *"This is the intelligent controller that does the automation work, and you have one layer that's actually automation, so for me then ok, that seems to sit into where the automation level is, but then we, so basically we take the sensors in, sensors and then there's automation and within your automation you have those levels in it and to me was a bit like that but then I had the, I wanted the discussion because with the intelligent controller, there's always the interface to the humans, right and on your diagram the interface to the human - the human having some input into it - seems to be at your management level, so I don't know whether this whole thing then becomes mapping to this because for me when you have your field level your basically saying there's the sensors, so it comes in, and then automation level is where basically you need your controller, so your controller is having this architecture in there… and my question really then became as part of the controller I do have the interface to the humans, but you seem to have it outside your controller" [Participant 7, AI Expert].*
>
> *"The expertise in setting up the systems and defining the parameters to take a system and say I'm going to put it in this environment, the expertise of the experts is actually done at the controller level and the person sets up those controllers to refine the characteristics of the way that system operates in that environment. The supervisor… is an operator, he's just a trained operator… all the smarts is actually down here, that bit there in the controllers, in the set-up of the controllers, how they configure the controllers, all of that smart work is done here – that is then passed up to this automation supervisor who makes decisions… this has actually come out of the BMS – BMS systems are different from ESS security management systems, they have a lot more intelligence in BMS systems than going to security management systems " [Participant 3, Security Expert].*

2.3.3.2.2 Development of 'AI Security Management Architecture' Levels

Participants reviewed the BACS/ Intelligent Architecture Level diagrams presented and collectively the AI security management architecture and intelligence controller diagram was conceived (Figure 18). Consequently, Figure 18 presents the AI Security Management Architecture.

Figure 18

The AI Security Management Architecture.



2.4 Interpretation and Findings Phase Two

Phase Two sought to respond to the Question: How do the capabilities of artificial intelligence techniques map to activities that can achieve control over a physical environment using security technologies?

Phase Two highlighted that the current capabilities of artificial intelligence techniques in security technologies facilitate the achievement of control over physical environments through the implementation of what is termed Broad AI, where two or more Narrow AI capabilities are integrated into a perceived single system output to achieve an automated security function.

Phase Two highlighted that commercial security systems predominantly use Narrow AI techniques to achieve defined individual outputs (Figure 19). However, through the integration of individual narrow AI outcomes, advanced technologies achieve what is perceived as human level cognition outputs using broad AI techniques. Figure 19 highlights the use of AI in security technologies where data inputs (sensory stimuli) are analysed into defined outputs using computer science computational techniques, and actions initiated as outputs according to defined rules to produce a predefined action which appears to be analogous with human cognition. However those outputs are not at the level of human cognition, that is, consistent with Pennington, Flanagan and Banyard (2000, p.1) (Phase One), and does not include the technology's ability to be aware of "what we are thinking", which may result in mistakes as outcomes are produced from the most likely data [Participant 6, AI Expert].

Figure 19

The use of AI in security technologies.

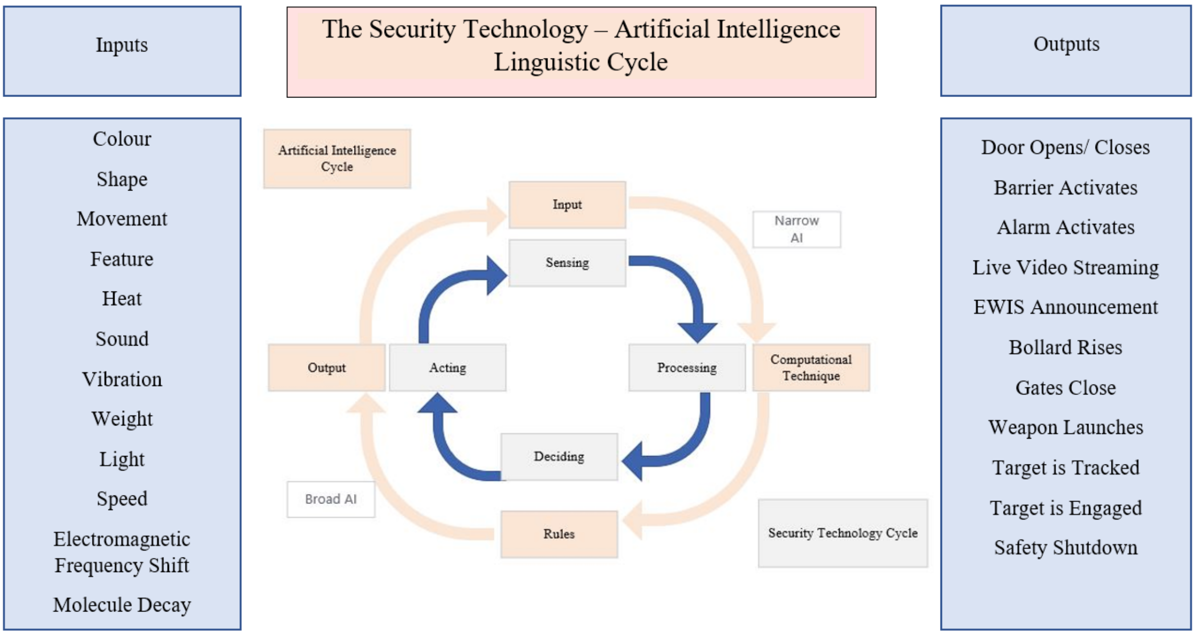


Figure 19 supports the finding in Tables: 24, 25, 26 and 27. For example, Table 24 highlights that security functionality or objectives are achieved using AI in Network Video Surveillance (NVS) based on combinations of Rules using Computational Techniques as Statistical Probabilities along the Statistical (X-axis; Figure 10) segment of the AI paradigms continuum. This included computational techniques of Background Analysis, Subspace learning, Kernel Density Estimation or Gaussian Mixture Method (GMM). For Detection, Table 25 highlighted that AI in Detection technology is predominately based on Logic Programming or Rules based programming to achieve predefined outputs, aligned to the Logic, Reasoning and Knowledge AI Problem Domains drawing on Logic-Knowledge based outputs along the Symbolic segment of the AI paradigms continuum.

Similar findings were found for Control technologies, with the exception of Biometric systems, whereas with the use of biometric systems in Detection, combinations of Logic Programming or Rules and Computational Techniques as Statistical Probabilities along the Statistical (X-axis; Figure 10) segment of the AI paradigms continuum were evident in achieving security function outcomes using AI. The findings in Figure 19 are supported through the narrative of focus group participants, for example:

> *"AI in security is… people use it as an umbrella term, everything's AI okay, but in actual fact it's not really, you know – at the best you've got some Narrow AI, they don't even know what Narrow AI is… they don't understand when you talk about Broad AI in terms of the integration of a series of inputs/outputs from Narrow AI, you've completely lost them, they've never heard of it… so how can they question what the vendors are selling them if they don't really understand what is and isn't AI?" [Participant 1, Security Expert].*

"In the literature, Narrow AI is an algorithm, a technique or software that incorporates algorithms that actually solve a single task so it does a specific thing… so it does, for example with facial recognition, so this is a piece of software that just does facial recognition. That is actually under narrow AI if you really look at that definition - it is techniques that actually address one very very specific task and is unable to generalise other tasks" [Participant 7, AI Expert]

"Broad AI… where basically the integration of two or more narrow AI techniques to do a specific thing, to perform a process... so it is a combination or an integration of narrow AI techniques into specific processes" [Participant 7, AI Expert]

"… security people go wow I've got all this AI but in actual fact they don't really have – what they have is a system that's defined by some sort of rules up front at the management level and the automation level executes those rules" [Participant 1, Security Expert].

"I don't know any security technologies at the moment that I know of across the world has got AI in it" [Participant 3, Security Expert].

"I don't know any system that's got true AI"… If you take Participant 7's [AI Expert] definition, there's nothing above Level 2… there's no AI even remotely approaching Level 3" [Participant 6, AI Expert].

"We will never on the level of General AI, at most if you integrate a number of Narrow AI technologies together to achieve what appears to be a slightly more complicated process then you have actually, you know, arrived at the level of Broad AI as defined by the IBM group" [Participant 7, AI Expert].

Such findings were reinforced through the narrative of Participant 6 [AI Expert] who drew on the example of self-driving cars:

"… many of the reinforcement learning approaches these days – by far the majority – are using deep learning within them, but the deep learning is basically a component that's taking some input and signalling some, you know, some characteristic of the data, and the reinforcement learning is bundling that up into a reward on the action you make on the basis of that – now this is sounding like human level intelligence of you know management whatever but it really isn't" [Participant 6, AI Expert]

"…people hope that it'll be the system that'll you know drive cars and they've certainly demonstrated with learning from these simulated car driving things that such a system does far better than any other system, right, if you put this in the loop, where it sees, you know, a computer graphics rendition of the scene and sees the person you know steering to avoid obstacles and whatever, this system trains itself overtime to do as good as your 16 year old kid who sits for hours and hours and hours and plays these games, so arguably its reaching close to human level performance, in that sort of sense, right, but clearly it has no knowledge of the world - and if there's something that comes up that was not represented in those simulations it will not do anything sensible, full stop [Participant 6, AI Expert].

Participant 6 [AI Expert] also expressed that no current system (such as biometric facial recognition systems or self-driving cars) has the capacity to recognise it has made a mistake. These systems are only able to work within the limited context which they have been provided. As they are not human, they are incapable of human-level insight and understanding.

"To take the auto driving car example, we kind of know that by and large, most of it is based on where we started, you know – the sensor's saying lane detection – I think you're likely to stray or have strayed across the line, sound alarm right, or if it's fully automated system turn the steering wheel… with the RADAR system or whatever- I think you're within 10 metres of a large object… doesn't even need to know what the object is, and it could be a flock of birds and that's, you know another situation but, and someone has made kind of a rule based decision, ok in that situation we better put on the brakes even though we can't be absolutely sure whether the flock of birds is going to dissipate, where a human would say oh it's a flock of birds, they're not dumb… and then someone could say ok well we've had too many automated cars put on the brakes when it's a flock of birds and drivers are complaining so we'll try and do a flock of birds detector, but we haven't got to the notion where this system has common sense real world knowledge that any 16 year old kid has, so you can achieve the level of driving capability of these 16 year old kids in the computer games and some of that transfers to the real world situation, and some of it doesn't…" [Participant 6, AI Expert]

In line with Phase Two findings, Figures 20 and 21 were developed. Figure 20 presents the compartmentalisation of Artificial Intelligence (AI) problem domains and spectrum of artificial intelligence paradigms in security technology. Figure 20 highlights that Observe and Detect technologies predominately use AI computational techniques to achieve event diagnosis inputs, whereas Control and Respond technologies predominately use AI computational techniques to achieve defined rule outputs as security actions.

Figure 20

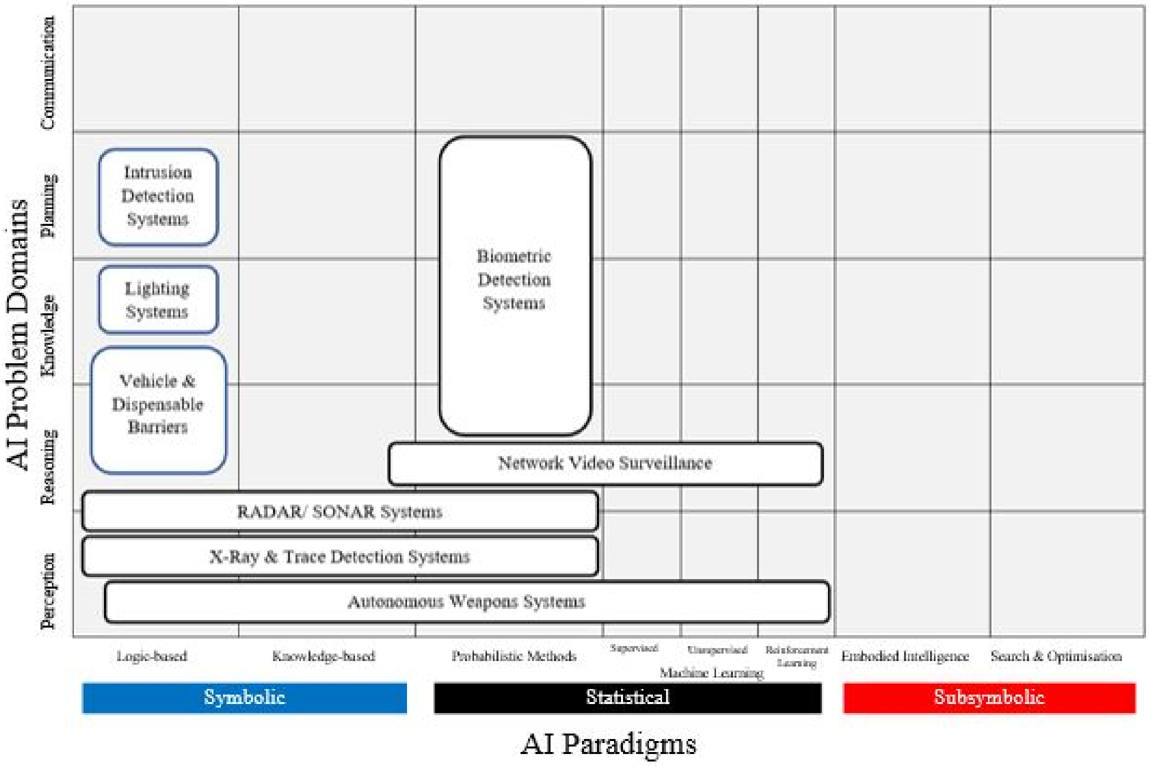
Compartmentalisation of AI Problem Domains and Spectrum of AI Paradigms in Security Technology

Narrow AI	**Broad AI**
OBSERVE/ DETECT	**CONTROL/ RESPOND**
Security Technology Cycle: ***Sensing*** ***Processing***	**Security Technology Cycle:** ***Deciding*** ***Acting***
AI Technology Cycle: ***Input*** ***Computational Technique***	**AI Technology Cycle:** ***Rules*** ***Output***

Figure 21 presents an adapted version of Figure 10, Corea's (2019) *AI Knowledge Map*, with superimposed security technologies, indicating the locality of various Observe, Detect, Control and Respond technologies along the visual map aligned to the Artificial Intelligence (AI) problem domains and spectrum of AI paradigms.

Figure 21

Security Technology alignment to the Artificial Intelligence (AI) problem domains and spectrum of AI paradigms.



2.4.1 Phase Two Focus Group Key Findings

Phase Two developed a list of key findings in understanding the use of AI in security technology in the protection of assets:

Objective 1

- Current Security Technologies predominately use Narrow AI along the Symbolic and Statistical segments of the AI spectrum.
- Observation and Detection Technologies are 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 (Figure 20) as alerts.

- 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.
- 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 (Broad AI).

Objective 2

- Machine Learning in Security Technologies is at an elementary level, with evidence of Machine Learning within Network Video Surveillance Analytics, Biometric System Analysis and Management, Acoustic Detection Systems, and Drone and Robotics Analysis and Management – where these systems know only the data they have been provided and cannot yet interpret the 'unknown'.
- AI can make mistakes, with no way of knowing a mistake has been made, impacting upon the accuracy and reliability of Security Technologies [Participant 6, AI Expert].
- There is no evidence of General AI or Artificial Super Intelligence (ASI) in security technologies.

Objective 3

- Narrow AI can be used across all three BACS Architecture Levels.
- BACS Architecture Levels differ from the Functional Architecture of Intelligent Controllers, as shown in 'AI Security Management Architecture' (Figure 18).

These key findings developed some key recommendations for communicating the opportunities and implications of using Artificial Intelligence (AI) in the establishment of secure physical environments.

Key Recommendations for Explaining AI in Security Technology:

- The development of a security technology and AI linguistics process diagram (Figure 19) demonstrating analogous technological processes for security professionals.
- Alignment of security technologies of Observe, Detect, Control and Respond to the AI problem domains and spectrum of AI paradigms Recommendation diagram (Figure 21).
- Develop a table of AI Computational Techniques, along with explanations and security technology examples for security managers to understand the use of AI in security technology.
- Development of ***A Practical Guide to Artificial Intelligence in Security Technologies***, including an 'AI Security Vocabulary' to define, explain, and detail a range of AI terms in relation to their function and role in security technologies.

2.5 Conclusion Phase Two

Phase Two employed a Focus Group technique to confirm findings from Phase One, reinforcing the current status of artificial intelligence in the protection of assets, and facilitate further outcomes that would aid in enhancing the security professionals' understanding of artificial intelligence in security technologies. Experts provided their insights into how the AI problem domains and spectrum of artificial intelligence paradigms align with the roles, functions and tasks of security technology through combinations of Observing, Detecting, Controlling and Responding technological measures. Phase Two established a robust foundation for understanding how security technology currently uses

Artificial Intelligence to achieve security objectives, and where further opportunities may exist for fusing security technology functions with developments in AI, as well as provide indicative risks associated with such advancements.

Informing Phase Three of the research, key findings from Phase Two included the understanding that Current Security Technologies predominately use Narrow AI along the Symbolic and Statistical AI spectrum. In addition, Observation and Detection Technologies are skewed towards threat or event diagnosis functionality in the protection of assets and use inputs (Sensing) and Computational Techniques (perceiving) to achieve defined Narrow AI outputs. In contrast, Controlling and Response Technologies are skewed towards actions in controlling and responding (Knowledge and Planning) to events using Rules and associated Outputs to achieve defined outcomes, including integrated Broad AI outcomes. In preparation for Phase Three a number of key recommendations were made, including; the development of a security technology and AI linguistics process diagram demonstrating analogous technological processes for security professionals along with the alignment of security technologies of Observe, Detect, Control and Respond to the AI problem domains and spectrum of AI paradigms diagram.

PHASE 3: EXPERT INTERVIEWS

3.1 Introduction

This Chapter presents Phase Three of the research; the outcomes of the security and artificial intelligence (AI) Expert Interviews. Phase Three sought to identify areas of opportunity for advancements in AI in security technologies, and where evident, highlight risks in their uptake for security management.

To achieve Phase Three objectives Security and Artificial Intelligence experts were invited to provide insights into the development of AI in security technologies. To draw out opportunities and risks, an evaluation tool was developed for Phase Three to assess the current level of intelligence for each security technology, and to establish the technical context in regard to where opportunities may emerge for future enhancements. Subsequently, Phase Three developed an AI technologies scale where each security technology category could be rated for their current level of intelligent autonomy, and used to indicate future enhancements (Appendix E). The *Security Technology Intelligent Autonomy Scale* was developed for this purpose, drawing from Phases One and Two outcomes.

The opportunities for AI in security technologies were extensive, with broad thematic areas for enhancement including *AI Testing and Standardisation*, *Technical Improvements*, *Analytics*, *Human-Machine Teaming*, *Innovative Development* and *Holistic Development in the Security Domain*. Opportunities were found to be located across all stages of the *Security Technology – Artificial Intelligence Cycle* (Figure 24): in Set Mission, Input/Sensing, Computational Technique/Processing, Rules/Deciding, Output/Acting and Review Mission. Articulation of the areas for enhancement and location of opportunities may provide security system designers, manufacturers and security users a greater understanding of both the capabilities and limitations of security technologies, as well as direct efforts for future development.

While the benefits of AI development were found to be moderate, and relating to mostly tangible gains such as increased productivity and reduced costs, the risks of AI development were profound. The benefits and risks identified by experts were analysed using a PESTEL framework, considering a number of *Political*, *Economic*, *Social*, *Technological*, *Environmental* and *Legal* factors. The PESTEL analysis indicated that the socio-political, technological, and legal risks of AI may eclipse any deliverable benefit. In essence, the overarching risks suggest deployment of AI technologies should be approached with extensive consideration and caution.

Despite these risks, AI developments in security technologies is expected to continue unabated, although experts do not anticipate significant changes to security technologies in the foreseeable future. Instead technologies are expected to become more accurate, reliable and generally more sophisticated until a paradigm shift in computing occurs. This paradigm shift will most likely be concordant with the arrival of quantum computing, however there is no consensus as to when this may occur. Until then, security technologies will continue to use and improve AI algorithms and computational techniques, while society monitors and engages with the opportunities and risks AI developments present.

3.2 Phase Three Aims & Objectives

The aims of Phase Three were to identify through *Expert Interviews* areas for potential advancements in security technologies using Artificial Intelligence, and where evident, to highlight risks in their uptake for security management. Phase Three responded to these aims through a series of supporting sub-questions which were applied to each category of security technologies across *Observe*, *Detect*, *Control*, *Respond* and *Integrated Security Technologies*.

The objectives of the Expert Interviews were to:

1. Identify the level at which each category of security technologies is currently located on the *Security Technology Intelligent Autonomy Scale*.

2. Identify opportunities for the development of artificial intelligence in each category of security technologies, including broad areas for enhancement and the specific locations for opportunities within *The Security Technology – Artificial Intelligence Cycle*.

3. Identify the benefits and risks associated with AI development in security technologies.

4. Indicate the expected development of AI in security technologies.

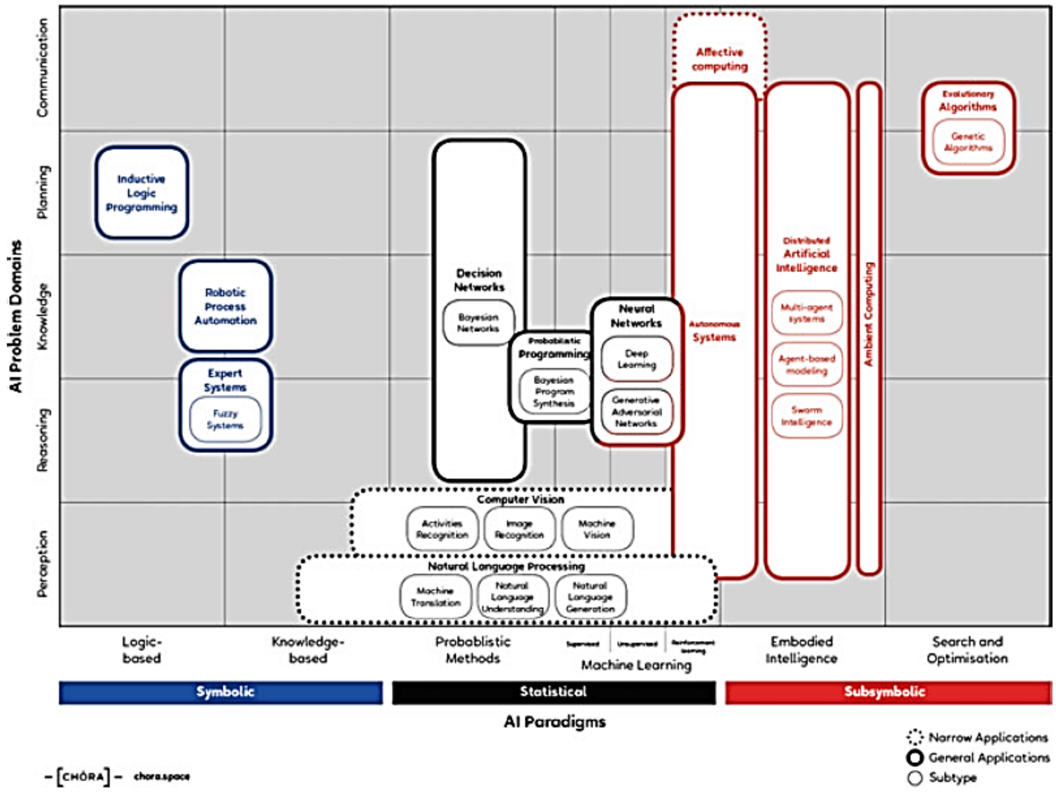
These objectives were addressed through a series of interview questions directed to security and artificial intelligence subject experts. Participants were provided with these questions prior to interviews as part of an information document outlining the interview procedure and aims of the research (refer Appendix M).

3.3 Background

Phases One and Two of this project provided the research foundation for understanding how security technologies currently use artificial intelligence techniques in the protection of assets. Phase One employed a systematic literature review (Jesson, Matheson & Lacey, 2011), identifying existing literature with regard to the roles, functions and tasks of security technology in the protection of assets. Phase One drew from Corea's (2019) *AI Knowledge Map* (Figure 22), to facilitate the technical articulation of security technologies within the Artificial Intelligence (AI) problem domains and spectrum of artificial intelligence paradigms.

Figure 22

AI Knowledge Map (Corea, 2019).

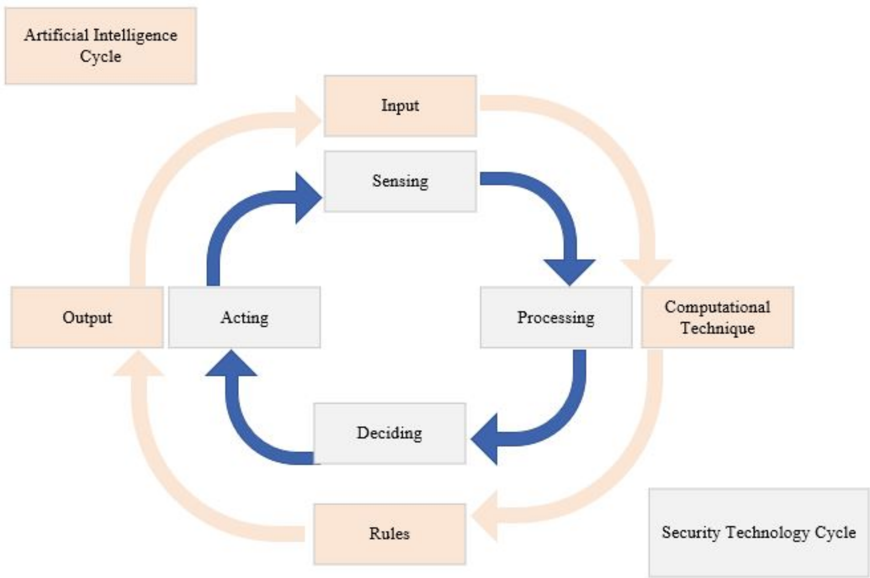


Phase Two conducted an *Expert Focus Group* to confirm and enhance findings from Phase One, reaffirming the current status of artificial intelligence in the protection of assets. Experts provided insights into how the AI problem domains and spectrum of artificial intelligence paradigms align with the roles, functions and tasks of security technology through combinations of Observing, Detecting, Controlling and Responding technological measures. Phase Two facilitated the development of AI maps for each of the security technology categories to enable security professionals to understand where specific security technologies may be located within both the AI spectrum and BACS architecture levels. Observing, Detecting, Controlling and Response technologies were subsequently represented as novel visual maps, indicating the alignment of hardware, firmware and software components with BACS architecture levels (Field Level, Automation Level and Management Level) and AI Paradigms (Symbolic, Statistical and Sub-symbolic).

Phase Two reinforced and enhanced the research foundation for understanding how security technology currently uses Artificial Intelligence to achieve security objectives, by articulating the sequence of stages by which AI and technologies concurrently achieve functional objectives. This sequence of stages, *The Security Technology – Artificial Intelligence Linguistic Cycle* (Figure 23), communicates the alignment between how security technology operates through sensing, processing, deciding and acting, and the AI cycle of data input, computational technique, rule assessment and defined output. Articulation of the sequence of operational stages for security technologies provided clarity as to where and how AI techniques may be applied.

Figure 23

The Security Technology-Artificial Intelligence Linguistic Cycle



Collectively, Phases One and Two articulated the AI paradigms and techniques used by each category of security technologies in addition to the concurrent security technology/artificial intelligence process in which this occurs. The sequence of stages in *The Security Technology – Artificial Intelligence Linguistic Cycle* was therefore introduced to participants in Phase Three of the research to assist in identifying the specific location of opportunities for AI development. Figure 24 depicts the stages in which security technologies process inputs (e.g. colour, shape, sound, movement, heat) and action outputs (e.g. alarm activation, EWIS announcement, target tracking, weapon launch) as presented to participants in the Expert Interviews.

Figure 24

Stages of the Security Technology – Artificial Intelligence Cycle

Input/ Sensing	**Computational Technique/ Processing**	**Rules/ Deciding**	**Output/ Acting**

The focus of Phases One and Two were relatively technical in terms of examining the various AI paradigms, the types of computational techniques used, and how and where they are employed by specific security technologies. In contrast, Phase Three explored beyond the technical and functional aspects of AI to capture conceptual evidence of how AI may impact on security as a domain within a multifaceted society. Subsequently, Phase Three incorporated expert knowledge to identify where opportunities for the development of artificial intelligence exist, the benefits and risks which may arise as a result of such development, in addition to the expected development of AI in security technologies.

3.4 Literature Defined Risks & Limitations of Security and Integrated Technologies

Integration of security, building, and IT technologies, with or without Artificial Intelligence (AI), present unique challenges, many of which are indeed known and accepted. However it is important to understand these underlying challenges as the use of AI in these technologies may change or amplify associated risks. Subsequently, the literature on security technologies and integrated technologies was explored to identify broad themes relating to the limitations, vulnerabilities and risks of security and integrated technologies to provide a foundation for where AI-based risks may emerge.

The predominant themes to emerge centre around technical issues such as computational capacity, latency and storage, as well as security-based vulnerabilities and risks which are specific to cloud-based computing. However, some associated risks and implications are perhaps more obscure, relating to broader management issues such as privacy and ethical considerations. Broad management and security risks may include data ownership, content disclosure, confidentiality, data location, control issues, regulatory and legislative compliance, forensic evidence issues, auditing issues, business continuity and disaster recovery, trust, privacy and ethical issues, and security policy issues (Onwubiko, 2010, pp.275-284). These management and security risks may vary considerably depending on factors specific to the individuals, groups and industries involved. While these broad management issues are central to integrated systems, and particularly integrated security & BACS systems utilising cloud computing, many are peripheral themes and subsequently are not discussed at length.

The major themes to emerge from the literature include; limitations of 'on-premise', cloud-based and hybrid systems; vulnerabilities of cloud-based systems; edge of network issues; data management and storage issues; commercial compatibility & standardisation, and safety, legal & ethical issues. Perhaps the most significant theme is security risks associated with 'connected' systems, spanning the entire spectrum of devices, components and infrastructure from on-premise to edge-of network, cloud-based and hybrid systems. Understanding these risks will help to shape how AI may be implemented and managed by security professionals to minimise the potential for adverse outcomes.

3.4.1 Overview of 'On-Premise', Cloud Based and Hybrid Systems

Integrated security systems and their components, including hardware, firmware and software may be categorised as either on-premise, cloud-based or hybrid systems. Each category may be differentiated as follows:

On-premise refers to an environment where 'resources are deployed in-house and within an enterprise's IT infrastructure' (Hughes, 2018). These resources and infrastructure generally consist of hardware, software and communications components connected to a central computer or server within a physical location. On-premise integrated security systems may include hardware (access doors, alarm points, alarm senders and receivers, intercom units, video cameras, card readers among other components), plus the network and communications infrastructure (in-house server hardware, computers, monitors etc) and integration software, software licences, integration capabilities and IT employees to manage the systems and issues which arise (Hughes, 2018; Pacom Systems Pty Ltd, 2019).

Cloud-based computing refers to 'the delivery of on-demand computing services – from applications to storage and processing power – typically over the internet and on a pay-as-you-go basis' (Ranger, 2018). Currently, cloud deployment is offered in various forms such as public cloud, private cloud, hosted private internet, and hybrid models (Burge, 2017; Hughes, 2018). In cloud computing, resources are hosted by a service provider and enterprises may access those resources at will (Hughes, 2018). Examples of cloud platforms include Microsoft's *Azure*, Amazon Web Services '*AWS*', and the Chinese Alibaba Group's *Alibaba Cloud*.

Hybrid systems currently offer some middle ground between on-premise, private cloud and public cloud systems. Hughes (2018) explains that 'a hybrid cloud infrastructure depends on the availability of a public cloud platform from a trusted third-party provider, and effective WAN connectivity between both of those environments.' Hybrid systems with private clouds are increasingly being used by larger enterprises and government entities that wish to retain custody of their own data. This type of model may allow for confidential data to be stored onsite, while third parties supply the processing power (Hodgson, 2018).

3.4.1.1 Limitations of On-Premise Systems

Until recently, integrated security systems have predominantly been on-premise, and although commercial vendors suggest hybrid and cloud-based systems may currently be overtaking on-premise systems, not all businesses have transitioned to the cloud (Leading Edge, 2020). The primary reason for this is on-premise systems are considered to be 'reliable, secure, and allow enterprises to maintain a level of control that the cloud often cannot' (Hughes, 2018). Although there is some disparity within the literature as to the security of on-premise systems, security concerns are frequently cited as a barrier to cloud-based deployment (Continuity Central, 2019; Leading Edge, 2020).

For those business who have not moved to cloud-based solutions, there are several limitations of on-premise systems. Limitations include storage, computational capacity and battery/ power issues, in addition to upfront costs, maintenance and fees for software updates and training (Kisi, n.d.; Sawhney, 2020). Limits in computational power of devices such as cameras, sensors or other devices highlight the need for more efficient processors and 'leaner algorithms' to reduce power and computational requirements (Sawhney, 2020). There are also significant technological vulnerabilities of integrated security systems, Building Automation and Control Systems (BACS), and their associated network and communications infrastructure. These vulnerabilities extend across the Field Level, Automation Level and Management Level (Brooks et al, 2017, p.134).

The trend toward hybrid and cloud-based systems may reduce the prevalence and availability of traditional systems. Organisations wanting to take advantage of analytics and other advanced algorithmic software will have to transition to cloud and Software as a Service (SaaS) systems (Hughes, 2018). The trend toward adoption of cloud-based systems is further supported within the literature with suggestions that perceptions of cloud-based security risk may also be shifting - a recent survey of security professionals found that 61 percent 'believe that the risk of a security breach is the same or lower in cloud environments compared to on-premise' (Continuity Central, 2019). This shift in risk perception, together with options for private & hybrid solutions to address security and data ownership issues, may eventually see on-premise systems become obsolete (Hodgson, 2018).

3.4.1.2 Limitations of Cloud-based Systems

Like many other business systems, integrated security systems are increasingly transitioning to cloud-based computing to take advantage of advanced analytics in artificial intelligence (AI) (Anderson, 2019). Although AI machines and algorithms can be implemented on devices or in the cloud (or a combination of both), AI has traditionally been deployed in the cloud due to the volume of data produced and computing resources required (IHS Markit, 2018; Sawhney, 2020). Sawhney (2020) highlights the enormity of AI processing requirements, explaining 'the cloud provides massively scaled computational power and very cheap memory and storage', adding 'it only makes sense that AI applications would be housed in the cloud, since the computation power of AI algorithms has increased 300,000 times between 2012 and 2019 – doubling every three-and-a-half months'. Cloud Service Providers (CSPs) have the capacity to provide computational power, memory and storage thorough enormous numbers of servers. For instance, Amazon Web Services (AWS) is estimated to have between 3 million and 5 million servers, with up to 80,000 servers in a single data centre (Anderson, 2019). However, while cloud AI has the capacity to analyse more data and run more powerful algorithms, cloud-based AI may be exposed to more issues around privacy and latency (IHS Markit, 2018).

Latency from cloud-based AI may become problematic when decisions need to be made in real time, and for security incidents or matters, delays in decision making due to computational latency may be a costly burden. Sawhney (2020) provides examples of this including a chemical plant predicting an imminent explosion, a security camera recognising an intruder at an airport, or an autonomous vehicle activating emergency braking upon predicting an imminent collision. To mitigate latency and bandwidth constraints, AI must be located at the 'edge' of the network (Kerner, 2019; Sawhney, 2020). Edge computing 'brings computation and data storage closer to the devices where it's being gathered, rather than relying on a central location that can be thousands of miles away' (Shaw, 2019). By reducing the distance between devices and remote services, real-time data does not suffer latency issues (Kerner, 2019; Sawhney, 2020; Shaw, 2019).

In addition to latency issues, ensuring the reliability, confidentiality, integrity and availability of cloud-based systems and services are critical for security enterprises. Tchernykh et al. (2016, p.2) defines integrity as 'the assurance that the information is trustworthy and accurate', and availability as 'a guarantee of reliable access to the information by authorized people'. Tchernykh et al. (2016) assert confidentiality, integrity and availability are the three most crucial components of cloud computing, and further discuss service interruptions as a limitation of cloud-based computing.

> Cloud-based services can crash just like any other type of technology. For example, access to information Amazon users has been limited for a long time due to distributed denial-of-service (DDoS) attacks in 2009. In 2013, a series of cloud outages are reported for Amazon, Microsoft and Google. Technical failures and data loss due to power outages are reported by Amazon, Dropbox, Microsoft, Google, and Yandex Disk. In the first quarter of 2014, Dropbox has experienced service outages twice. Bankruptcy is imposed for cloud storage company Nirvanix in 2013 (Tchernykh et al., 2016, p.2).

More significantly, the issue of availability extends beyond bankruptcy and temporary power outages to the service providers and the political and economic environments in which they operate. Governance and regulation of independent software vendors (ISV) and public cloud platforms where cloud-based power is leased may not be mitigated easily, if at all. With so many services, service providers, and products available, there are a myriad of potential issues ranging from privacy and confidentiality through to data ownership. Even with service contracts in place, questions may remain as to who owns the cloud and the data in it, and where might consumers stand should the ISV or their governing State decide to withdraw the service.

While these operational risks and constraints may be problematic for cloud-based systems, security risks dominate the literature. There are several significant security risks of using cloud-based systems. Network security is one of the major risks of cloud-based systems, in addition to vulnerability to attack, access security, and security of retained data. These risks are discussed in more detail in section 3.4.2 Vulnerabilities of Cloud-based Systems.

3.4.1.3 Limitations of Hybrid Cloud Systems

Hybrid cloud systems are 'a combination of public and private clouds implemented by different providers' (Balasubramanian & Aramudhan, 2012). As hybrid systems are able to store confidential data onsite, and use public clouds for processing power, they are able to take advantage of different levels of security to strengthen security and privacy for consumers (Balasubramanian & Aramudhan, 2012; Hodgson, 2018). However, there are some unique limitations of these complex systems, particularly in relation to managing different security platforms together (Balasubramanian & Aramudhan, 2012). Software and hardware incompatibilities may arise where cloud providers differ in their operating systems, programming languages or patch releases (Ward, 2018).

Much like public cloud systems, hybrid cloud systems also have the potential for service interruptions, therefore ensuring reliability, confidentiality, integrity and availability may prove challenging (Tchernykh et al., 2016). Funahashi & Yoshikawa (2011, p.292) explain: 'because on-premises and public clouds are connected via the internet, the quality of the network may be unstable'. Due to this internet component being present in hybrid systems, they may also be subject to some of the security risks which are present in public cloud computing systems. Additional security issues may also arise due to the complexity of integrating hybrid systems at a management level. Balasubramanian & Aramudhan (2012, p.39) identify five unique security issues of Hybrid cloud systems, shown in Table 29 (*Hybrid Cloud Security Issues*).

Table 29

Hybrid Cloud Security Issues

Hybrid Security Issue	Description
Absence of Data Redundancy	Redundancy is required across data centres to mitigate potential outages and may become problematic if redundant copies are not distributed across data centres.
Compliance	Coordinating, maintaining and demonstrating compliance between the public cloud and the private cloud are more difficult (e.g.: customer card data cannot be transferred from a compliant private cloud database to a less secure storage system in a public cloud).
Poorly Constructed SLAs	Service Level Agreements (SLAs) must be detailed sufficiently to ensure both public and private cloud providers are able to meet and maintain expectations. Integrating the capabilities of both may prove challenging.

Risk Management	Application Programming Interfaces (APIs) require complex network configurations, requiring additional knowledge & capabilities from administrators.
Security Management	Integration of authentication, authorisation and identity management between the private and public clouds is more complex.

These issues highlight that although hybrid cloud systems generally offer a higher level of security, the operation of public and private components in a single system may in fact create significant vulnerabilities through inadequate management or inconsistencies in management systems.

3.4.2 Vulnerabilities of Cloud-based Systems

Vulnerability to attack is a major challenge faced by Cloud Service Providers (CSPs) and has resulted in in security concerns being a significant obstacle to the widespread adoption of cloud-based systems (Suryateja, 2018). Vulnerabilities in cloud hardware such as Spectre and Meltdown, as well as malicious software such as Ransomware, may result in serious security breaches or render systems inaccessible to cloud consumers. The following examples highlight the extent of the damage which may occur due to these vulnerabilities.

> Code Spaces, a source code and project management services company which used Amazon Web Services (AWS) was compromised when an attacker gained access to the company's AWS control panel. The attacker tried to extort money in exchange for control of the AWS panel, however when Code Spaces refused, the attacker deleted resources, effectively destroying the company at that time (Venezia, 2014).
>
> Apple suffered what may be the largest high-profile cloud security breach due to the victims involved. Jennifer Lawrence and other celebrities had their private photos leaked online. Many of the victims initially thought that someone had hacked their individual phones. Instead, the iCloud service they used for personal storage had been compromised (Bradford, n.d.).

In the event cloud-based integrated security systems or Building Automation and Control Systems (BACS) are attacked, there may be significant ramifications for integrated systems and their users. For instance, in an attack on the Biostar 2 platform - a security system used by organisations to secure commercial buildings - a collection of biometric credentials and personal information was discovered.

> The breach, which was discovered by researchers Noam Rotem and Ran Locar alongside vpnMentor, included the fingerprint data of more than 1 million people, facial recognition information, unencrypted usernames and passwords, and other personal information of users of Suprema's Biostar 2 security platform. The information, which included a total of 27.8 million records totalling 23 gigabytes of data, was found in a publicly accessible database, although it's unclear whether any malicious actors accessed the data while it was unsecured (Porter, 2019).

This case raised significant concerns for the personal information of those users enrolled in the system, particularly as biometric credentials cannot be changed in the same way in which a password may be changed (Porter, 2019). However, it is not just the personal information of system users at risk from cloud-based security breaches, but also the operational aspects of integrated systems which may have wider safety and security implications. For instance, if a building's HVAC system is shut down, the data centre may melt down causing permanent damage to occur (Perelman, 2019), or if secure facility credentials are stolen, they may be sold on the black market and used to gain access to the facility (Kass, 2019).

Cyber-attacks on smart buildings have increased. Kaspersky (2019) reports that 'almost four in ten (37.8%) computers used to control smart building automation systems were affected by malicious attacks in the first half of 2019'. With 35,000 BAS systems connected to the public internet globally in 2019, the risk of attack may be high (AT&T Cybersecurity, 2019). An attack may result in the failure of one or several critically important smart building systems and may lead to potentially life-threatening situations (AT&T Cybersecurity, 2019; Kaspersky, 2019).

Subsequently, 'Siegeware', 'Jackware' and BAS attacks are emerging threats which may result in such life-threatening situations, particularly where an attacker is able to take control of a building and shut down any number of critical building operations. These critical systems include lighting control systems, fire detection and alarm systems, automated fire suppression systems, integrated security and access control systems, HVAC systems, power management and assurance systems, and command and control systems (AT&T Cybersecurity, 2019; Varghese, 2019). An attacker may take control of a building as a method for blackmail or extortion, or they may actively hack systems to aid in theft of assets among other exploits. Specific techniques may include generating false alarms to trigger a building evacuation, increasing the temperature in a server room to damage property, or stealing/ destroying sensitive laboratory research data. King (2016) explains how this type of scenario may occur.

> One of the challenges for enhancing cyber security within an intelligent building is the vast number of both logical and physical entry points that are subject to infiltration. As buildings begin to utilise myriad connected devices, various intelligent building technologies offer unique entry points for intruders, allowing them to exploit ineffective or non-existent security measures. In the case of physical security, intruders may exploit the technology to potentially steal physical assets from a building. In this scenario, an intruder might attempt to take control of a video surveillance or access control system to aid in a theft. In other instances, intruders may gain access to physical security systems to assess entry times into a building or view video footage to aid in theft preparation.

The rise in cloud-based cyber-attacks is not limited to BAS or building automation and control Systems (BACS). In all cloud-based systems, attacks and exploitation of hardware and software vulnerabilities are increasing in frequency, particularly due to the 'malicious use of cloud resources popularity and use of *Internet of Things* (IoT) devices' (Help Net Security, 2020). According to Raywood (2017), a Microsoft Security and Intelligent Report claimed that the number of attacks on cloud-based accounts increased by 300% in a 12-month period, with the majority of these attacks being targeted phishing attacks, breaches of third-party services, or the result of poor password management.

There are numerous methods by which cloud-based systems which may be exploited, with major vulnerabilities shown in Table 30.

Table 30

Vulnerabilities of Cloud-based Systems

Cloud Based Vulnerabilities	Description
Compromised Application Programming Interfaces (APIs)	Customers use APIs to manage and interact with cloud services, however as APIs are accessed by the internet, vulnerabilities in API software may be exposed to attackers.
Identity, Authentication, Authorization and Auditing (IAAA) Vulnerabilities	Cloud-specific IAAA vulnerabilities include denial of service by account lockout, weak credential-reset mechanisms, insufficient or faulty authorisation checks, coarse authorisation control, and insufficient logging and monitoring possibilities
Multi-tenancy Exploitation (Cross-tenant Access & Attack)	If the cloud service provider (CSP) hosts multiple tenants, an attacker may gain access to another user or organisation's assets or data by exploiting system and software vulnerabilities. Cloud storage is prone to cross-tenant storage access and cross-tenant network access.
Data Recovery Vulnerability	Recovery of a consumer's data may be possible where user resources are allocated to a different user at a later date. Additionally, consumers may not be able to verify the secure deletion of their data which may enable data remnants to be available to attackers.
Theft of Credentials	An attacker may be able to gain access to the consumer's systems and data by stealing or accessing the credentials of a user, a CSP administrator or other administrative user, or other organisations using the same CSP.
Unauthorised Access	Unauthorised access to networks, systems, management interfaces or data by consumer or CSP insiders may result in damage, theft, exploitation or extortion.
Back Door Access	Virtual machine template images may be used to provide back-door access for an attacker.
Loss of Stored Data	Stored data may be lost by malicious or accidental occurrences, or through physical events such as natural disasters. Examples include fire, earthquake, lost encryption keys, or CSP bankruptcy.
Security Auditing Controls	Standard security controls for auditing, certification and continuous security monitoring are difficult to implement in the cloud.
Compromised CSP Supply Chain	A consumer has little control over whether a CSP outsources its infrastructure, operations or maintenance and if so, how and when the CSP checks and enforces compliance to reduce risks.
Virtual Machine Escape	An attacker may exploit the CSP's operating system by tricking the host system into running code from a virtual machine or gain access to the hypervisor and all the virtual machines running on it.
Insecure or Obsolete Cryptography	Advances in cryptanalysis may defeat cryptographic mechanisms or algorithms. Cryptographic vulnerabilities may also arise from weak random number generation as a result of having several Virtual Machine Environments (VMEs) on the same host.
Internet Protocol Vulnerabilities	Cloud services are accessed through networks using the internet which may allow attackers to use man-in-the-middle and other forms of attack.
Hijacking/ Denial of Service (DoS) and Distributed Denial of Service (DDoS) Vulnerabilities	An attacker may insert malicious code, manipulate data or issue a DoS attack to render the cloud service inaccessible. DDoS attacks may be executed by an attacker using malware to take control of several victim systems, resulting in the cloud service being unavailable to legitimate users.
Injection Vulnerabilities	Injection vulnerabilities such as SQL injection, command injection and cross-site scripting may manipulate programs.

Authentication Vulnerabilities	Vulnerabilities may arise from user behaviour (weak passwords or reusing passwords) limitations of single factor authentication mechanisms, or credential interception and replay.

(Source: Ducklin, 2015; Lukan, 2014; Morrow, 2018; Suryateja, 2018; Walloschek et al, 2011)

Access Security for Cloud Applications

In addition to mitigating general and network-specific risks of cloud-based systems there are challenges in implementing and maintaining access controls for cloud-based systems. To reduce security risks for cloud applications, Garrity & Roberts (2018) recommend applying security controls in three main areas, as shown in Table 31.

Table 31

Access Security Controls to Reduce Risk for Cloud Applications

Access Type	Access Controls
Secure User Access	Elevate security policies for privileged applications and users. Enable single sign-on (SSO) with multi-factor authentication for all cloud apps. Audit user cloud app permissions using the concept of least privilege to provide access only to users who require access.
Secure Network Access	Deny traffic from known risky networks such as Tor, anonymous, and proxy networks. Only allow access to cloud apps based on the geolocation of the users. For high-risk cloud applications, require users to be connected to a corporate network.
Secure Device Access	Gain visibility into devices accessing cloud services. Block devices that are out-of-date and at risk of malware. Require newer mobile devices that have more advanced security controls such as a hardware security module (HSM), TouchID, FaceID and encryption. For high-risk cloud applications, require devices to be corporate-managed by an Enterprise Mobility Management (EMM) solution to ensure only trusted devices are able to access your cloud services.

While these advanced security controls offer high levels of security and encryption, they require access devices to be updated to the latest, most secure software. Failure to do so may result in outdated devices compromising or infecting the cloud software/network with malware (Garrity & Roberts, 2018). Consequently, maintaining adequate security may be problematic for organisations or individuals who do not make the required device upgrades. Social and equity issues may arise where individuals, employees or contractors are prevented from accessing cloud services due to outdated devices.

3.4.3 Edge of Network Issues

Edge computing refers to ‘the enabling technologies allowing computation to be performed at the edge of the network, on downstream data on behalf of cloud services and upstream data on behalf of IoT services’ (Shi et al, 2016). Edge computing is ‘complimentary’ to cloud based computing, where the goal is to locate the cloud closer to the edge (Conde, 2018). By doing this, the data does not need to be sent to the cloud for processing thereby relieving pressure and improving the speed and performance of data (Anderson, 2019; Linthicum, n.d.). For physical and integrated security systems, this means edge devices with AI algorithms can process and extract selected information, such as an individual or a vehicle in a video image, thereby reducing bandwidth consumption and increasing efficiency, quality and accuracy (Anderson, 2019).

However, edge of network computing has a number of technical and operational limitations, as well as specific security issues. Technical and operational limitations arise from the highly distributed nature of edge systems – each of the nodes are not independent, therefore many applications need to share security, customer and other information. However, coordinating the functions of each system result in problems with distribution, consistency and synchronization (Conde, 2018). Roman et al (2016, Table 3) consider challenges beyond devices and network infrastructure, to include the entire spectrum of edge paradigms, including 'fog computing', 'mobile edge computing' and 'mobile cloud computing' (see Table 32). This highlights the complex environment of edge computing and that a holistic approach to combining all edge paradigms is required to address technical, operational and security issues (Roman et al, 2016).

Table 32

Common challenges in Edge paradigms

Challenge	Description
Infrastructure	Interoperability, monitoring, accountability
Virtualisation	VM lifecycle, container and context awareness
Resources & Tasks	Resource location, task scheduling, offloading
Distribution	Cooperation, N-tier management, 'soft-state'
Mobility	Connectivity, seamless handoff
Programmability	Useability, session management

Source: Roman et al (2016, Table 3)

Edge of network systems and devices may be subject to many of the same security threats as cloud-based systems. Security risks are extensive, particularly if an adversary is able to access proprietary and personal information or control the services of the infrastructure. Linthicum (n.d.) highlights the criticality of identity and access management including encryption for device safety, urging readers to 'imagine if a jet engine could be hacked in flight'. For edge of network devices in integrated security systems, the shutting down of critical functions such as fire suppression systems may have fatal consequences (Perelman, 2019).

Roman et al (2016, pp.16-19) discuss numerous edge of network security issues (see Table 33) and emphasise security services and mechanisms are required for identity and authentication, access control, protocol and network security, trust management, intrusion detection systems, privacy, virtualisation, fault tolerance and resilience, and forensic evidence management. Overwhelmingly, security issues lead the discourse within edge computing literature, suggesting that edge paradigms are in their infancy (Conde, 2018; Linthicum, n.d.; Roman et al, 2016, p.25).

Table 33

Edge of Network Security Issues

Security Threat	Description
Denial of Service (DoS)	DoS attacks including distributed denial-of-service (DDoS) and wireless jamming, however these attacks may be limited to disrupt only those affected networks.
Man-in-the-middle Attack	Attackers may take control of part of the network to eavesdrop or inject traffic.
Rogue Gateway	Mobile devices and other edge of network participants may connect to a rogue gateway device deployed by an attacker, making eavesdropping or traffic injection possible.
Rogue Data Centre	An adversary may control an entire data centre, resulting in complete control over all services, information flows and interactions in the geographical location
Rogue Infrastructure	Elements of the core infrastructure are targeted by adversaries to take control of some services of upper layer infrastructure.
Physical Damage	Elements of the service infrastructure may be damaged physically by an attacker in the vicinity of the device.
Privacy Leakage	Internal and external adversaries may access information from the edge data centre, edge infrastructure or through APIs implemented.
Privilege Escalation	Adversaries may try to take control of services, particularly where edge data centres are managed by people with little training.
Service Manipulation	If an attacker gains control of the data centre and escalates privileges, they may manipulate services to enable the attacker to launch DoS attacks or tamper with information.
Virtual Machine (VM) Manipulation	A host system being controlled by an attacker may launch attacks on the VMs running in the system. VMs may be infected with malware or logic bombs, or the attacker may manipulate tasks or extract information.
Injection of Information	A device may be reprogrammed to fabricate or distribute fake data.
Misuse of Resources	Malicious programs to crack passwords or host botnet servers can be executed in the edge data centre to target local or remote entities or vulnerable IoT devices.

(Source: Roman et al, 2018, pp.11-25)

3.4.4 Data Management & Storage

Data management including data storage and preservation is a major challenge of integrated systems, particularly with security technologies such as surveillance cameras and body cameras worn by security and law enforcement personnel. Although cloud-based systems are better equipped to accommodate large quantities of data, the immense volume of data created over time may be problematic for storage, preservation and management. For instance, Security Informed (2016) estimates that a single day of video surveillance is estimated to be collecting more than 500 petabytes (PB) of data. In addition to the costs and resources required for storage, both organisations and governments are required to consider and accommodate legal issues such as long-term data storage of security events. Security Informed (2016) highlights surveillance storage issues in environments such as airports where a typical international airport may contain 20,000 CCTV cameras:

> Airport guidelines require any "event" to be stored for seven or more years! An "event" includes theft, reported injuries and any conflicts. These "events" happen anywhere between 20 to 40 times per day. For the purposes of estimating the amount of data these "events" produce, let's assume the following - if each "event" is ten minutes long and captured on at least ten cameras, that's about 18 GB of data that needs to be kept for seven plus years. With 20 to 40 of these occurrences happening daily, that's 360- 720 GB per day!

In addition to the capacity to store large volumes of digital data, further issues emerge concerning integration and storage of diverse formats, including forensic and evidentiary data. Lim & Lee (2013) raise the issue of the growing amount of digital evidence being encountered in criminal cases, investigations and military contexts, and argue the need for a unified digital evidence management system for forensic evidence from diverse sources. This is particularly important in respect to the trend of security integration & convergence, and how to manage vast quantities of digital data from multiple technologies, sources, departments & locations.

3.4.5 Commercial Compatibility & Standardisation

Industrial and commercial issues may present challenges for integrated systems, particularly with the abundance of products and services available and the absence of standardisation of digital services and devices (Ashford, 2014; Bandyopadhay & Sen, 2011). For instance, where a vendor exits the market or ceases to operate there may be issues with maintaining, sustaining and upgrading the system to meet the needs of the business or agency (Facilitiesnet, 2007). There may also be issues with standardising the technological components at an industry level as well as troubleshooting technical problems where individual components have been purchased through multiple manufacturers. Wiser (2017) discusses the issue of technical integration created by the abundance of manufacturers and products available:

> A system that involves intrusion and access control from one manufacturer, cameras from another, a video management system from a third, and a public address system from a fourth, puts the integrator in the middle. If part of the integration does not work properly, the integrator may have to troubleshoot with multiple manufacturers to fix the issue, which can require additional time and energy and add expense to the project. Or the integration may not include all of the available features of the individual technology components.

While the shift toward Internet Protocol (IP) technology has resulted in a higher level of product compatibility, the issue of standardisation remains due to the rapid progression of technologies, particularly between security sub-sectors and also within specific product groups such as biometrics (Ashford, 2014; Bandyopadhay & Sen, 2011). For example, in 2018 Delta Airlines opened their first biometric terminal in Atlanta but acknowledged that standardisation would be necessary to enable global implementation (Blue Swan Daily, 2019). Subsequently, further standardisation of integrated technologies across countries and regions may be required to facilitate widespread adoption and reduce compatibility costs.

3.4.6 Safety, Legal & Ethical Issues

Autonomous weapons systems, vehicles and robotics lead current discourse on safety, legal and ethical issues in AI and autonomy due to the high-consequence nature of these systems. The literature on autonomous weapons systems is littered with moral and ethical dilemmas and debate on whether machines should have the capacity – or authority – to make meaningful decisions on the use or release of force and whether these systems are compliant with International Humanitarian Law (IHL) (Heyns, 2016; Sharkey, 2016). Autonomous vehicles and robotics tend to lead discourse on safety with the increased drive for commercial development of driverless cars and robotics technologies.

While the majority of security technologies do not use or release force, the use of automated, autonomous or intelligent technologies may create safety, legal and ethical issues where technologies fail or infringe on individual privacy or rights. Safety issues may also create legal liability for manufacturers, vendors and security managers where harm occurs as a result of technical failures, through the bypassing or defeat of safety features, or as a result of security vulnerabilities being exploited. There are several examples provided within the literature which emphasise these safety and legal risks.

Security technologies may encounter safety issues if it is able to be overridden by either the system itself, its authorised users, or unauthorised adversaries (Köse, 2018). However firstly it is necessary to distinguish between overriding and defeating a system; overriding or 'bypassing' a system refers to the manual suspension of a system, and 'defeating' a system refers to an intentional and unauthorised act, which renders a function or device ineffective (Rockwell Automation, n.d.). In some instances, bypass operations are conducted on critical building systems such as emergency, fire and gas, and alarm systems where maintenance or other procedures are required (IBM, n.d.b). However, bypassing a system's safety functions is sometimes considered a 'taboo topic, almost frowned upon by safety professionals', as the existence of a bypass capability may create significant risks and liabilities if not managed properly (Rockwell Automation, n.d.). Bypassing or defeating the integrated system for malicious purposes may not only create security risks but also safety and legal issues for security management (Kont et al, 2015).

The literature contains numerous examples of how unauthorised adversaries may hack into and control or override a system through vulnerabilities created by 'connected' systems or devices Many of these safety and security risks have emerged from the IT domain and are widely recognised, even where solutions are yet to be determined and established. However, within the field of AI, which is anchored in the IT and engineering domains, there is limited discourse on the safety risks to integrated security systems from intelligent programs and their algorithms.

With the growth of AI, BACS and integrated security systems are able to take advantage of advanced algorithms such as behavioural analytics and other deep learning algorithms. AI is allowing integrated systems to be predictive, proactive, and adapt to environmental changes to increase life safety – for instance, an intelligent access control system may move from 'a gatekeeper to a life saver enabling safe passage for those who need to enter and assist', by moving people to a safer area of the building (Carter, 2019). However integrated systems utilising AI algorithms have their own unique safety issues, particularly where – and if – the system can override the operator's instructions. The Safety of Autonomous Systems Working Group [SASWG] (2018, p.23) provide an example of how an override may occur in automatic or autonomous weapons systems. This highlights how accidents or incidents may occur in the event that users do not fully understand the behaviour of the software system, or if all possible scenarios are not considered at the design stage.

> Operator's Choice Overridden by Software, pre-1999: During field practice exercises, a missile weapon system was carrying both practice and live missiles. Transit time was being used for slewing practice. Practice and live missiles were located on opposite sides of the vehicle. The operator acquired the willing target, tracked it through various manoeuvres, and pressed the

> weapons release button to simulate firing the practice missile. Without the knowledge of the operator, the software was programmed to override his missile selection in order to present the best target to the best weapon. The software optimized the problem ,de-selected the practice missile and selected the live missile. When the release command was sent, it went to the live missile. The "friendly" target had been observing the manoeuvres of the incident vehicle and noted the unexpected live launch. Fortunately, the target pilot was experienced and began evasive manoeuvres, but the missile tracked and still detonated in close proximity (SASWG, 2018, p.23).

While this case is 'pre-1999' and relates to autonomous weapons systems rather than integrated security systems, it raises questions as to how safe the assessments by AI, deep learning and predictive algorithms are, and to what degree are these systems able to execute instructions autonomously. Safety is of course one of the major themes of AI research, particularly in relation to specialised and high-risk areas such as autonomous vehicles and weapons systems (Levin & Woolf, 2016; SASWG, 2018; Yampolskiy & Spellchecker, 2016). However, in security applications, AI safety issues surface mostly in relation to individual security components such as automated robots.

For instance, in 2016 in the Silicon Valley's Stanford Shopping Centre, a robot security guard 'knocked down a toddler while on duty and then apparently just kept on driving' (Vincent, 2016). The Knightscope K5 robot – a five-foot, 300-pound machine – allegedly hit the 16-month-old's head causing the toddler to fall down and then ran over the child's right foot and scraped his leg (Vincent, 2016). Other incidents include a robot grabbing and crushing a worker to death at a German car factory (Phillip, 2015; Yampolskiy & Spellchecker, 2016). These incidents draw attention to the critical safety failures of algorithms in autonomous or 'intelligent' systems and the potential for legal liability arising from such incidents.

Yampolskiy and Spellchecker (2016, p.6) describe AI safety as a 'relatively young and underfunded field of study' and as such, lacking in legal and ethical development when compared to advances in technical abilities. They further assert that the 'problem of making a safe and capable machine is much greater than the problem of making just a capable machine' and the field of AI safety may need to draw from methods in the established scientific fields to navigate safety issues. Subsequently, safety issues and the legal ramifications of critical AI safety failures in integrated systems may be a major limitation of security technologies in the future.

3.5 The Security Technology Intelligent Autonomy Scale

To identify the opportunities and risks presented by AI development, contextualisation was required as to the extent to which AI is used by intelligent systems, and the degree of involvement, decision making and control that humans retain - or intelligent systems possess - while in operation. This concept refers to the *level of autonomy* a system may have, which may vary considerably between types and categories of security technologies.

Within the literature, attempts to define levels of autonomy have shifted from a focus on computer capabilities toward measurement of the interaction and collaboration between humans and machines to achieve outcomes (Sharkey, 2016, pp.23-38). Distinctions may be drawn between existing concepts and definitions of integration, automation and levels of autonomy, and how both humans and AI perform within each classification or level. Consequently, element definitions were required using a systems engineering approach, where differentiation between integration, automation and autonomy is determined by the degree of human control during various stages of operation (Norris & Patterson, 2019);

> ### Integration
> 'The act of combining or adding parts to make a unified whole' (Harper Collins Publishers, n.d.). The term *Integrated Security* implies a combination of security technologies, functions and components, or quite simply; 'an assimilation of different security services' (Dowell, 2015) which communicate to perform advanced functions in, as a minimum, an automated manner (Norman, 2014, p.263).
>
> ### Automation
> 'The technique, method, or system of operating or controlling a process by highly automatic means, as by electronic devices, reducing human intervention to a minimum' (Harper Collins Publishers, n.d.). In automation, human decisions and logic are executed to 'accomplish a pre-set series of tasks within a known (or assumed) frame of reference', without decisions being made during operations (Norris & Patterson, 2019, p.1).
>
> ### Semi-autonomy
> 'Anything that involves machine decision making (in response to external, unexpected events) during operation but that a human is involved in some of this process and provides some direct control. Therefore, semi-autonomous systems are those which are more independent and agile than automated systems and less self-contained than fully autonomous systems' (Norris & Patterson, 2019, pp.1-2).
>
> ### Autonomy
> Autonomy may be broadly defined as 'the condition of being autonomous; self-government, or he right of self-government; independence' (Harper Collins Publishers, n.d.). According to Norris & Patterson (2019, p.1), an autonomous system is one where decisions (in response to external inputs or signals of whatever complexity) are made within the system and do not involve human decision making'. Norris & Patterson further add that in theory, 'a fully autonomous system would be able to respond on its own to unknown and unexpected events via artificial intelligence without pre-programming by a human' (2019, p.1).

While these definitions provide a relatively solid foundation for differentiating between integration, automation and levels of autonomy, it is essential to understand definitions and interpretations may vary between academic disciplines, commercial sectors and industries. For instance, in the mining sector, Caterpillar defines 'autonomy' as referring to 'a state of equipment in which it can perform the programmed operations under defined conditions without human input or guidance' (Caterpillar, n.d.). However in disciplines such as robotics, space-based research or weapons engineering, systems or vehicles, operating within defined and programmed parameters may be considered automated rather than autonomous. In space-based research, Frost (2011, p.89) offers the following differentiation between automation and autonomy, and introduces the concept of intelligent autonomy;

> An *automated system* doesn't make choices for itself – it follows a script, albeit a potentially sophisticated script, in which all possible courses of action have already been made. If the system encounters an unplanned-for situation, it stops and waits for human help (e.g. it "phones home"). Thus, for an automated system, choices have either already been made and encoded, or they must be made externally. By contrast, an *autonomous system* does make choices on its own. It tries to accomplish its objectives locally, without human intervention, even when

> encountering uncertainty or unanticipated events. An ***intelligent** autonomous system* makes choices using more sophisticated mechanisms than other systems. These mechanisms often resemble those used by humans. Ultimately, the level of intelligence of an autonomous system is judged by the quality of the choices it makes.

Frost (2011) clearly identifies *decision making* as the differentiating factor between automation and the various levels of autonomy. This notion is supported within the literature on autonomous weapons systems (AWS) and Lethal Autonomous Weapons Systems (LAWS) in terms of decisions over critical functions such as the release of force (Heyns, 2016, p.4). According to Heyns (2016, p.6), AWS's may be differentiated by their level of autonomy in decision making and not by the weapons they use. Consequently, the higher the level of autonomy in an intelligent system, the lower the degree of human control and decision making.

However, while differentiation between automation and autonomy may become clear based on the degree of human involvement and control, the literature does not establish any robust characterisation of the extent to which intelligent systems are able to plan, direct and execute operations or missions. Furthermore, existing literature defined levels of autonomy do not explicitly link the concepts of *Narrow AI*, *Broad AI* and *General AI* with the characterisation of intelligent decision making and control at each ascending level.

Subsequently, the *Security Technologies Intelligent Autonomy Scale* (Appendix E, adapted from Chang, 2014; Proud & Hart, 2005) was developed to fully explore the concept of 'levels' of intelligent autonomy within the predefined AI categories of *Narrow*, *Broad* and *General AI*. This scale articulates the levels of intelligent autonomy which may be used by security technologies, and ranges from Level 1: *Manual* (absence of AI) through to Level 11: *Post-Autonomous* (General AI) and defines the technical AI characterisation at each increasing level of autonomy. At Level 11: *Post-Autonomous*, the stages of the *Security Technology – Artificial Intelligence Cycle* have been extended to align with the expectation that General AI will possess *Theory of Mind*, and be self-aware, understand belief, thoughts, emotions, expectations of people and interact socially (Goertzel & Pennachin, 2007; Maruyama, 2020, pp.242-251; O'Carroll, 2020). Such a theory suggests systems at the level of General AI would be capable of planning, directing, executing and reviewing security missions with no human intervention. Subsequently, the stages of 'Set Mission' and 'Review Mission' were added to the *Intelligent Stages of the Security Technology – Artificial Intelligence Cycle* (Figure 25) and presented to participants during the Expert Interviews as locations for opportunities and enhancement.

Figure 25

Intelligent Stages of the Security Technology – Artificial Intelligence Cycle

Set Mission	**Input/ Sensing**	**Computational Technique/ Processing**	**Rules/ Deciding**	**Output/ Acting**	**Review Mission**

Figure 26 captures a summarised depiction of the *Security Technology Intelligent Autonomy Scale*. For the full scale, see Appendix E.

Figure 26

Summarised Security Technology Intelligent Autonomy Scale

AI	LEVEL	CLASSIFICATION	DESCRIPTION
General AI	11	Post-Autonomous	System uses more sophisticated mechanisms to achieve autonomy. All decisions are made during operation without human involvement or human override. System writes/ modifies embedded code to respond to unexpected events. No human control.
Broad AI	9 - 10	Autonomous	System is scripted but decisions are made within the system during operation without human involvement. Responds to unknown or unexpected occurrences without human intervention and attempts to resolve. Low degree of human control.
	7 - 8	Semi-Autonomous	System is scripted with ability for some decisions to be made during operation with human involvement. System stops and waits for help for unknown or unexpected occurrences. Moderate degree of human control.
Narrow AI	5 - 6	Automation	System is scripted with no decisions made during operation. System stops and waits for help for unknown or unexpected occurrences. Full human control outside automated processes, including human override.
	3 - 4	Integration	System is scripted with all decisions encoded and embedded prior to operation. System stops unknown or unexpected occurrences. Full human control and override ability.
	1 - 2	Manual/ Unintelligent	All decisions made by humans. Complete human control.

3.5.1 Objective One

3.5.1.1 Location of Technologies on the Security Technologies Intelligent Autonomy Scale

Objective One sought to identify the level at which each category of security technologies is currently located on the *Security Technology Intelligent Autonomy Scale*. This objective was facilitated by the following sub-question presented to participants:

Q1. Based on the Security Technologies Intelligent Autonomy Scale, at which level do you believe Observation, Detection, Control, Response and Integrated Security Technologies currently sit?

Participants were provided with the Security Technologies Intelligent Autonomy Scale (Appendix E) and were asked to identify the autonomy level at which they believe each of the technology categories currently achieves. Participants were introduced to the concept of autonomy and how 'Intelligent Autonomy' may be achieved with Artificial Intelligence techniques and decision making.

Distinctions were drawn between *Machine Autonomy* and *Intelligent Autonomy*, with *Machine Autonomy* being described to participants as 'scripted' systems, with all decisions enacted by encoded or embedded script and unknown scenarios resulting in the machine stopping and asking for help. By contrast, *Intelligent Autonomy* was described to participants as sophisticated machine autonomy, with the system being capable of writing or modifying the script as well as being able to make decisions during operation. In *Intelligent Autonomy*, systems will respond to unknown situations or unexpected events and attempt to resolve the issue without human assistance or intervention.

3.5.1.2 Objective One Findings

Participants rated the level of autonomy for Observe, Detect, Control, Response and Integrated Technologies based on their expert knowledge of commercially available systems. Table 34 presents quantitative data for each of the technology categories, and while *Max.* participant scores reached as high as 8 or 9, *Mean* scores across the technology categories ranged from 3.90 (*Response*), to 4.44 (*Integrated Technologies*), 4.55 (*Control*), 4.70 (*Detect*) and 5.05 (*Observe*), indicating that security technologies currently achieve integration and automation, but not semi-autonomy, autonomy or post-autonomy levels.

Table 34

Participant Identified Levels of Intelligent Autonomy in Security Technologies

	Min.	Mean.	Mode	Median	Max.	Std. Dev.
Observe	3	5.05	4.0000	4.2500	9	1.9214
Detect	3	4.700	3.0000	3.5000	8	2.3118
Control	3	4.5500	3.0000	3.7500	9	2.0877
Respond	0	3.9000	3.0000	3.0000	8	2.7669
Integrated Technologies	0	4.4444	7.0000	5.0000	7	2.2283

Observe Technologies were rated as achieving the highest level of Intelligent Autonomy (M = 5.05, SD = 1.9214), driven by research and development in computer vision and analytics. Expert participants highlighted the level of autonomy may vary considerably between technology categories, based on the application and the country of deployment. Observation technologies in the form of mass surveillance systems, for example those used in countries such as China, were frequently cited by participants in reference to advanced development due to the size of the surveillance systems (e.g. 100,000 cameras or more in some provinces) and subsequent developments in the supporting systems required to process and manage the data and information generated.

Observe Technologies: Surveillance Systems

"In the marketplace at the moment, they fit within various levels – I'd say they start at about a level three and go up to level nine… so I guess you'd say the state of the art is up there at level nine and that's the sort of stuff that's happening in China at the moment with surveillance of mass populations and to a lesser extent in other countries… they've certainly been experimenting with these technologies for some time now, and having a large population allows them to do that and collect some really good data" [Participant 3, Security Expert].

"I would say that Observation technologies are probably a bit more sophisticated than detection technologies quite frankly" [Participant 6, Security Expert].

Expert participants noted inconsistency in development between industries, asserting that the IT domain leads AI development and is more advanced, for instance they compared the autonomous capabilities of Intrusion Detection Systems in the security and cyber domains, indicating cyber automation to be much higher in development. Participants further noted that the security industry is somewhat resistant to change which may be a contributing factor to disproportionate development across domains.

"Yeah, I think you'll find more in cyber – there'll be more of the Eights, getting into the Nines… you know there's a big difference between an IDS system and a cyber system" [Participant 2, Security Expert].

"I think from the beginning, it's been largely a belts and braces type of industry, and that lends itself to being quite resistant to change… so we're always following a number of years behind the IT industry and we don't incorporate everything that the IT industry does so they're far more sophisticated" [Participant 6, Security Expert].

Expert participants also highlighted discrepancies between how systems may perform under controlled conditions as opposed to dynamic, real world conditions. Manufacturers and product vendors may overestimate the capabilities of security systems which often do not perform optimally in a live operating environment. Such capability amplification inflates market perceptions of system capabilities, particularly in relation to AI. Subsequently accuracy in relation to the levels of autonomy achieved by security technologies may have limitations due to participants' perceptions of system capabilities.

"…salespeople would be sort of suggesting that surveillance technologies are probably semi-autonomous or maybe even more sophisticated than that, but in my experience that would be not representative of actual real-world conditions – that might be represented in expo or a laboratory conditions, but not in real world" [Participant 6, Security Expert].

"Vendors are great at sales pitches… when they test this stuff, they do it under a very, very controlled environment, with a defined set of parameters, and when you actually get it out in the field you never have an ideal or a controlled environment… the parameters vary wildly and stuff happens that just shouldn't happen because we still don't have the processing horsepower and the algorithms to do what a human does" [Participant 9, Security Expert].

Identification of the location of each category on the *Security Technologies Intelligent Autonomy Scale* also raised the issue of variance between systems under development compared with systems currently deployed. This is particularly evident in Response technologies which is reflected in the low *Mean* score (M = 3.90) and may primarily be due to social factors and a lack of trust of AI. A significant gap was found between the development and deployment of AI, particularly in the commercial sector as non-military applications may be subject to greater moral, ethical and legal critique. However while military applications of AI possibly avoid a degree of public scrutiny, participants generally felt that military development of AI is not significantly ahead of commercial development as a result of the porous nature of military-commercial exchange.

"Response technologies have the capability in the computational and processing techniques to be quite high up the scale, but then it drags itself back because of the trust factor – we won't allow it to do it on its own" [Participant 4, Security Expert].

I wouldn't even say it's holding the development of it back, I would say that it's holding the implementation of it back, the utilisation of it… I think we are probably all in agreement there is stuff out there, and then again we go to the military commercial split… so there's stuff out there that the military are legally entitled to use, if you want to call it that, that could be used in the commercial sector but the commercial sector morally, ethically, legally are not prepared to entertain" [Participant 4, Security Expert].

"… in selected areas, military are possibly a couple of years ahead, and other areas that are playing catchup" [Participant 11, AI Expert].

"… development can take part in the military environment that probably would not be permitted to take part in the commercial environment. But that firewall between the governmental military environment and the commercial military environment is so porous that you know it can be developed in the governmental space and then before you know it, it's just slipped through into the commercial space" [Participant 4, Security Expert].

Key Finding: There is a significant gap between the development and deployment of AI, particularly in the commercial sector as non-military applications may be subject to greater moral, ethical and legal critique.

The factors influencing how participants rate security technologies on *Security Technology Intelligent Autonomy Scale* may be indicative that what is evident in commercially available security applications is at the lower end of what is currently possible, due to resistance to change within the security industry, difficulty achieving optimal results in dynamic conditions, and caution exercised by the commercial sector in deploying AI systems.

3.6 Opportunities and Risks of Artificial Intelligence

3.6.1 Objective Two

3.6.1.1 Opportunities for the Development of Artificial Intelligence in Security Technologies

Objective Two sought to identify opportunities for the development of artificial intelligence in each category of security technologies, including broad areas for enhancement and the specific locations for technological opportunities within *The Security Technology – Artificial Intelligence Cycle*.
Objective Two was facilitated with the following sub questions presented to participants:

a) What enhancements could be made to Observe, Detect, Control and Response technologies to achieve a higher level on the *Security Technologies Intelligent Autonomy Scale*?

b) What Opportunities exist for Observe, Detect, Control and Response technologies using AI techniques, and at which stages of *The Security Technology – Artificial Intelligence Cycle* may these opportunities be located?

Expert participants were asked to refer to the *Security Technologies Intelligent Autonomy Scale* while considering the broad enhancements which could be made to each technology category. Participants were further presented with an unpopulated table of *The Security Technology – Artificial Intelligence Cycle* for each technology category to identify the specific locations of opportunities. Table 35 illustrates the unpopulated table presented to participants to locate opportunities for AI advancements in observation technologies.

Table 35

Location of Opportunities for AI Advancement in Observe Technologies

Technology	Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Observation Technologies						

3.6.1.2 Objective Two Findings

3.6.1.2.1 Opportunities for Observe Technologies

A number of technical and management opportunities in Observation technologies were identified by expert participants. Several overarching thematic opportunities for enhancement emerged from the data analysis, including *AI Testing and Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, *Human-Machine Teaming*, and *Holistic Development in the Security Domain* (see Table 36).

Table 36

Opportunities for Enhancement of Observe Technologies

AI Testing & Standardisation

- Design & development of common connectivity protocols across platforms, equipment & devices
- Uniformity of systems with development of international AI standards
- Development & standardisation of testing mechanisms

Technical Improvements

- Increased accuracy & reliability in live operating environments and under dynamic conditions
- Greater accuracy in object recognition and classification
- Data management systems for voluminous output
- Improvements in quality, resolution and processing capacity

Integration & Aggregation of Inputs

- Greater integration of security technologies to internal and external systems
- Expansion and aggregation of data and datasets for enhanced decision making
- Increased automation of Narrow AI tasks for enhanced decision making

Analytics

- Analytics & predictive analytics for pattern recognition and anomaly detection

Human-Machine Teaming

- Collaborative development between human and artificial intelligence for the optimisation of operational and organisational objectives

Holistic Development in the Security Domain

- Synchronicity in development across domains
- Establishment of protocols to achieve accountability and responsibility in the development and deployment of AI-enabled security technologies
- AI development linked with risk and consequences of failure
- Establishment of training policies and frameworks for machine learning applications
- Development and achievement of safety parameters to enhance commercial trust
- Research in consumer & market demands and drivers of development

Expert participants identified *AI Testing and Standardisation* as one of the fundamental areas requiring improvement to enhance the development of observation technologies. Standardisation and uniformity of connectivity protocols between systems and devices will reduce connectivity issues and increase the number and type of system inputs required for greater intelligent interpretation of visual data. Development and standardisation of testing mechanisms will also drive *Technical Improvements* in areas such as accuracy and reliability by providing assurances to security managers that security systems achieve the level(s) of proficiency declared by manufacturers and vendors under a variety of operational conditions.

Improvements in resolution and processing capacity are expected to enhance overall quality of inputs which may subsequently improve the application of advanced analytics. Integration and aggregation may enhance the identification of anomalies within seemingly innocuous environments by producing deeper insight from cumulative data analysis. For instance, aggregated analysis of video motion detection, network video surveillance, and biometric data may identify and locate several adversaries carrying out simultaneous attacks across facilities or buildings that may otherwise have not been detected by existing surveillance technologies. Aggregated analysis of data inputs gathered from multiple sources may equip security operators with more accurate information for enhanced decision making and security response.

The use of artificial intelligence to enhance decision making and response is a key concept in 'Human-machine Teaming' (HMT). This subdomain explores and promotes collaborative development through the merging of humans with complex machines to achieve specific goals and capabilities (Defense Innovation Marketplace, n.d.). McDermott et al (2018, p.1) describe HMT in practical terms while emphasising autonomy as the underpinning mechanism required to achieve collaborative outcomes:

> Autonomy should act seamlessly within a human operator's workflow, aiding performance by alerting them about behavior that deviates from normal, suggesting alternative solutions that they may not have considered, autonomously reorganizing priorities in response to their changing goals, or other collaborative activities (McDermott et al, 2018, p.1).

By this description, the goal of HMT is therefore to collaboratively assist humans rather than designate or redirect human activities to machines. This works successfully in a socio-technical system such as security, as the pairing of humans with intelligent machines retains an element of trust which may be deficient in machine only AI applications (Defense Innovation Marketplace, n.d.; Konaev & Chahal, 2021). Thus in observation technologies, HMT or collaborative development between human and artificial intelligence has emerged as a distinct opportunity for the optimisation of organisational and operational objectives.

Though there are an abundance of technical opportunities in observation technologies, technical development must be supported by enhancements across management, industry and internationally. Progress must occur at both micro and macro levels within the socio-technical system for *Holistic Development Across the Security Domain*. This means synchronistic development of AI must occur across all domains and disciplines which overlap the security function, including but not limited to safety, facility management, engineering, academia and commercial research and development. Social concepts such as trust, safety, responsibility, accountability and integrity as qualitative drivers of artificial intelligence must be explored fully to create the supporting macro foundations for micro evolution across all technology categories.

Key Finding: Development of AI in security technologies may be enhanced or constrained by synchronistic development in areas which overlap the security function.

Policies and frameworks must be established for the development and deployment of AI-enabled security technologies in socially acceptable ways. This presents a series of opportunities for the security industry to take the initiative and influence social development of AI technologies, particularly with observation technologies due to the significance of privacy as a fundamental human right. Opportunities exist not only for commercial, legal and political development of these policies and frameworks, but for the security industry as an exemplar of AI development.

Key Finding: Policies and frameworks must be established for the development and deployment of AI-enabled security technologies in socially acceptable ways.

In time, perhaps one of the most significant opportunities for Observation technologies is the exploration of how AI may set and review missions to develop and enhance the organisational security function. While the stages of 'Set Mission' and 'Review Mission' were located at Level 11 (*Post-autonomous*) on the *Security Technology Intelligent Autonomy Scale* - significantly higher than the Level 5 at which Observation technologies have currently been rated – opportunities for enhancement were identified within these stages.

Conceptually, for intelligent technologies to set a mission, they would need to interpret and import dynamic threat, risk and criticality factors – either present with the organisation or within the external micro and macro socio-technical and political environments - to determine security requirements based on organisational objectives and risk tolerance. For instance, a security system may import and interpret regional specific security data (e.g. country and regional security advice), local and national security data (e.g. crime reports and statistics), public and corporate security data and information (e.g. trends in security incidents and events), economic and political data (e.g. location and extent of economic conditions which may drive or undermine security conditions), integrate this with data generated from its own integrated security technologies (e.g. surveillance system, biometrics etc) to establish a dynamic baseline threat level. From here a system may align with organisational objectives to determine the overall level of security required *at that specific point in time*, and plan, direct and execute security missions (e.g. drone surveillance at certain times or locations, or in response to changing conditions or flagged visual input) to protect organisational assets.

Participants suggested 'Set Mission' may present opportunities in terms of the allocation of security resources to coordinate and oversee activities that would typically be carried out by a security manager. However based on current AI capabilities, those opportunities are not expected to be commercially deployed for some time.

"…if you have a truck that has a high value load coming in so the system – not the security system – the system or AI understands that truck is coming in, what time its coming in, so it diverts resources, drones or whatever, and controls the traffic to allow that truck to come straight in… it has more observation, surveillance put onto it until it pulls in, unloads and then pulls out again and then initiates changes for the AI system because then the value asset is in the building so then the resources are put onto that… so its dynamic" [Participant 1, Security Expert].

"That's what a human would do as a security manager – say 'I've got this value load in, the threats gone up a little bit, I need to put more resources into it" [Participant 1, Security Expert].

"I don't think that we are anywhere near a situation where, for instance, a set of parameters could be automatically generated from a TVRA for instance, and then those used as an algorithm to set the parameters for the observing. I think were miles and miles away from it… and the reason we are miles and miles away from that is because there are just so many differing ways of conducting that TVRA, there's limited standardisation in it, and whilst I accept that a TVRA needs to be subject specific, site specific, whatever we want to call that, I think that just creates so much complexity… I think it relies on so many other things happening first that it's as you said, way off there at Level 11" [Participant 4, Security Expert].

Similarly, an autonomous capacity to 'Review Mission' was identified by participants as an opportunity in both Observation and Detection technologies. Self-reflection, lessons learned, and adjustments or improvements based on those lessons were identified as specific opportunities in which AI could support decision making for security activities. For instance, reviewing specific security activities with dynamic feedback could enable threat level adjustments to direct the allocation of security resources. Data from near misses could also be incorporated into the 'Review Mission' stage for enhancement of machine learning algorithms. However like 'Set Mission' enhancements, those opportunities may not be commercially available for some time.

"…AI advancement in observation technology is really the system being able to learn lessons from itself and improve itself, so its lessons learnt. So how do we as humans reflect on what, how we did? Did we do it well as an individual and as a group? And how do we then put changes in the mission, input, processing, rules and output? Its reflection, the system reflection" [Participant 1, Security Expert].

"I'm also a great believer in using the safety approach of near misses in security. Safety are really good at recording and investigating near misses, security people are not very good at doing it… one of the opportunities is when we talk about review of the mission is AI can pick up those near misses and make notifications of near misses and lessons learned… my experience of security is that they're not very good at picking up on near misses, therefore they don't learn the lessons from them" [Participant 4, Security Expert].

For observation technologies, the areas for enhancement identified by participants were extensive and multi-faceted and included both technical and management opportunities. There may be considerable scope for security manufacturers and AI developers to explore the commercial prospects of such opportunities. Table 37 summarises the location of opportunities for Observation technologies, with details of specific opportunities contained in Appendix F.

Table 37

Location of Opportunities for Observe Technologies within the Intelligent Security-AI Cycle

Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission

Mission linked with organisational objectives	Technical improvements in quality, resolution, and noise reduction	Increased processing capabilities	Human-machine teaming	Systematic autonomous response	Autonomous review
Mission linked with criticality & risk	Improved testing	Technical improvements in dynamic environments	Analytics & predictive analytics for anomaly detection	Improved analytics and novel methodologies	Dynamic feedback
	Convergence & aggregation of data	Aggregation of data	Aggregation of data	Aggregation of data	
	Application of context	Application of context	Application of context	Application of context	

3.6.1.2.2 Opportunities for Detect Technologies

Opportunities for AI enhancement in Detection technologies were largely technical, with the overarching thematic opportunities consisting of *AI Testing and Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, and *Human-Machine Teaming* (see Table 38).

Table 38

Opportunities for Enhancement of Detect Technologies

AI Testing & Standardisation
- Design & development of common connectivity protocols across platforms, equipment & devices
- Uniformity of systems with development of international AI standards
- Certification and assurances in system capabilities

Technical Improvements
- Novel detection devices
- Reduced signal noise in devices for enhanced detection
- Enhancements for more dynamic environments
- Increased mobility in sensors/ detection devices
- Increased accuracy to for improved Probability of Detection

Integration & Aggregation of Inputs
- Greater integration and aggregation of multi sensor inputs
- Aggregation of multi-sensor data to provide context for enhanced decision making

Analytics
- Novel analytical methodologies
- Analytics to enhance decision making by providing ranked options for response
- Analytics to provide greater context in dynamic conditions

Human-Machine Teaming
- Collaborative development between human and AI to increase Probability of Detection

- Development of capable AI guardianship

As with all security technology categories, detection technologies may benefit from *AI Testing and Standardisation* and *Technical Improvements* such as increased quality, accuracy and operation under dynamic conditions. However the technical opportunities relating exclusively to detection technologies were development of *novel detection devices* and increased *mobility* in sensors/devices. Experts assert detection technologies have not fundamentally changed in many years, and that new and innovative detection technologies and devices may be required to drive AI development. Mobility of devices from connectivity, particularly with space and satellite research, will substantially increase data yields.

"I would expect to see more RADAR, SONAR, LIDAR technologies develop and probably new technologies that we don't fully understand yet… millimetre wave technology is another one that's becoming more common" [Participant 3, Security Expert].

"Mobility would be a key one… we will have 5G wherever you are on the planet, and so the ability to have sensors anywhere connected will amplify the amount of data you get" [Participant 2, Security Expert].

With improved connectivity, and subsequently an increase in the volume of data generated, experts identified aggregation of data and novel analytical methodologies as opportunities to enhance detection technologies to improve accuracy, reliability and the probability of detection. Complementary statistical analysis across multiple devices of the same type (e.g. Passive Infrared sensors), or across multiple devices (e.g. Passive Infrared, CCTV, LIDAR etc.) may improve the probability of detection without increasing the rate of false positives. However achieving this may be challenging without common connectivity protocols across platforms, equipment and devices, thus the development of international standards for connectivity protocols may drive the application of AI analytics in this application.

"Where the AI has the ability to come in is not in the technology itself so much, but it's in the data that's gathered from these multiple technologies, just as we were talking about multiple technologies being used in the [censored] and then analysing that, and from the different sensors to be able to say well this is or isn't valid – but at the moment that's just not occurring, not in any of the solutions… and you call them a solution if they've got multiple technologies in it but they just don't do it" [Participant 8, Security Expert].

"Inside each technology, there are generally the potential to use the same technology a number of times and look at that on a statistical basis… so you're increasing your whole detection metric by using the same technology in different ways. CCTV is the same, but nobody turns around and says well I want a visible camera and an infrared camera and I'm going to run complementary analytics on the two, because no one wants to spend the money [Participant 9, Security Expert].

However, above and beyond the collection and aggregation of data from devices, the concept of 'knowledge' and what knowledge consists of, means, or produces, must be revisited in an attempt to transform data into new and meaningful insights for security. While significant progress has been made in computer science fields such as Knowledge Representation, Automated Reasoning and Commonsense Reasoning, AI has a limited capacity to make human-level inferences about the world, situations or interactions between people (Russell & Norvig, 2020). The application of context is a pre-requisite for human-level understanding and must be applied to all stages across the *Security Technology – Artificial Intelligence Linguistic Cycle*, i.e. set mission, input/sensing, computational technique, rules/deciding, output/acting and review mission.

Key Finding: AI must be able to interpret and contextualise dynamic environments, events and situations. Development in this respect must also include the ability to understand and account for significant deviations or outliers from expected inputs, outputs and norms.

Take for instance a shopper who conceals an item under their arm – logic or rules would determine an attempt to conceal an item as shoplifting, however human understanding applies context and see that the shopper has placed the item under their arm to free their hands in order to move a child from harm's way. The importance of context in security cannot be emphasised enough, as context helps to understand intent which may inform and shape decision making on a case-by-case basis. This may be particularly useful in dynamic environments where security incidents can evolve or dissipate rapidly and decisions to engage or respond to a threat are often made quickly. While a capability for the application of context and understanding of intent certainly presents opportunities for detection technologies, AI contextualisation of knowledge is not currently anywhere near capable of this without human assistance, interjection, or override.

"We need to come to grips with this whole notion of what is knowledge. I mean, you stick a whole lot of text into a machine from Wikipedia… it is largely keyword search and clustering type of stuff that has no real knowledge in any significantly deep way, and it has no self-awareness in any way. It doesn't have any clue about what it's doing in any meaningful way" [Participant 11, AI Expert].

"You're getting the unconscious bias in as well, because the data sets themselves it is the human giving it the data sets or selecting those data sets as well as creating them… that's where the bias comes in to play… and even when it does create something, it's not necessarily creating something new like a human mind does, it's using what it has to create something so it's not necessarily new, its maybe just sort of transformed" [Participant 2, Security Expert].

"Knowing a rule is one thing, knowing something about the limits of the rule is another thing, and knowing something about the intention of the rule is a completely different thing… people decide whether the rule should be applied or not in a given situation and there isn't a machine on the planet that's even remotely approached this sort of level of intelligence" [Participant 11, AI Expert].

Key Finding: Algorithms must be able to understand and imply the underlying intention of programmed rules to add sufficient depth and meaning to decision making.

Current limitations in AI contextual knowledge may present opportunities for development of Security-AI Red Teams to reduce the likelihood of adversaries defeating the system (Eidson, 2019; Rizkallah, 2018). As systems increasingly adopt AI techniques such as machine learning, and progress to higher levels of intelligent autonomy, human-level input, assessment and control will decrease incrementally. Consequently both AI algorithms and security systems may be more vulnerable to attack, particularly with reduced capacity for human-level assessment of human factors such as intent. Expert participants assert that regardless of the number of adversarial moves an intelligent system can calculate, the adversary will continue to adapt to defeat the system.

Key Finding: Development of new domains such as Security-AI Red Teaming may arise and as a result of workforce opportunities created by AI.

> *"With a Red Team, I think it's the best way to know exactly what is your position, in evaluation in security" [Participant 7, Security Expert].*
>
> *"…we have people involved in the attacker side, it's not a machine… it's not machine against machine. Where people are involved in the attack, I can tell you as a Red Team leader you'll need a machine to think how I am as an attacker. I see the situation, I analyse it and say okay I will attack in this way. No machine can do that" [Participant 7, Security Expert].*

With a continued need for human involvement, Human-Machine Teaming (HMT) for enhanced decision making and response emerged once again as an opportunity for detection technologies. Experts identified collaborative HMT to increase the Probability of Detection, Development of Capable AI Guardianship as well as the provision of Ranked Options for Decision Making and Response. Ranked options may assist security operators to make compliant decisions based on the best probability of outcomes, though presentation of options may carry additional risks (Sharkey, 2016, p.35).

Areas for enhancement of Detection technologies are summarised in Table 39, with details of specific opportunities contained in Appendix G.

Table 39

Location of Opportunities for Detection Technologies within the Intelligent Security-AI Cycle

Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Mission linked with organisational objectives	Improvements in quality and capabilities of detection devices	Increased processing capabilities	Aggregation of data and data sets	Systematic autonomous response	Autonomous review
Mission linked with criticality & risk	Convergence & aggregation of data	Technical improvements in dynamic environments	Analytics & predictive analytics for anomaly detection	Improved analytics and novel methodologies	Dynamic feedback
	Application of context	Analytics in dynamic environments	Development of source data sets	Analytics & predictive analytics for	Application of context

				anomaly detection	
		Aggregation of devices and data	Red teaming capability	Convergence and aggregation of data	
		Application of context	Application of context	Application of context	

3.6.1.2.3 Opportunities for Control Technologies

Expert participants identified several opportunities for AI enhancement in Control technologies, which were largely technical and operational. Opportunities consisted of *AI Testing and Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, *Human-Machine Teaming,* and *Holistic Development in the Security Domain* (see Table 40). Though several of these opportunities have been discussed in preceding sections, and therefore will not be reintroduced in this section, there were some specific opportunities reported for Control technologies including development of first responder applications, and convergence of sensors and systems for multi-point identification and authentication.

Table 40

Opportunities for Enhancement of Control Technologies

AI Testing & Standardisation

- Design & development of common connectivity protocols across platforms, equipment & devices

Technical Improvements

- Increased accuracy & reliability to increase Probability of Detection
- Development of source datasets

Integration & Aggregation of Inputs

- Aggregation of data for distributed intelligence
- Increased automation
- Greater integration and aggregation of inputs between systems

Analytics

- Novel analytical methodologies
- Anomaly detection between distributed systems and devices
- Analytics to provide distributed intelligence for streamlined response

Human-Machine Teaming

- Collaborative development between human and AI to increase Probability of Detection
- Development of control technologies as a 'First Responder'

Holistic Development in the Security Domain

- AI development linked with risk and consequences of failure
- Regulatory frameworks for the acquisition of source data

According to the experts, there may be considerable scope to enhance control technologies for use as a first responder. For instance, the use of access control to apprehend or detain intruders has previously been implemented in some locations where it does not contravene legal or safety regulations; or in response to an active shooter incident to prevent further movement and casualties. Though there may be considerable scope for control technologies to facilitate a response, this may introduce additional risks in some instances, such as an active shooter being locked in with vulnerable persons, or adversarial knowledge of the response procedure being exploited to achieve criminal or terrorist objectives. As a prerequisite for control technologies to be used in this capacity, legal frameworks would need to be developed for specific uses (i.e. apprehension or detention would be illegal in many instances), moral and ethical considerations would need extensive contemplation, and social acceptance would need testing. Any instance where a security mechanism engages with humans or inhibits free movement of a person there may be extensive legal, ethical and social ramifications.

> *"the access control system could be integrated with the intrusion alarm system so if somebody got into an area of the building where they shouldn't be… we could actually automatically lock the fire escape doors from the lift lobby so that it was impossible for someone to leave the lift lobby… You can't do that in Australia because of our fire safety regulations… so that's not a particularly good example because in most countries you wouldn't be allowed to do that, but it just shows what you can do. So at least in theory you can arrest somebody just simply by locking them in an area in a room" [Participant 3, Security Expert].*
>
> *"Yes, but given there are a number of different active shooter scenarios, by relying on the technology to provide that response there is not an inconsiderable chance that the response gets it wrong and activates the wrong scenario and locks the active shooter in a room with a group of people" [Participant 8, Security Expert].*

Key Finding: AI adoption should be promoted in technologies, applications and environments where the risk of AI failure can be explicitly linked with lower consequences.

At a holistic management level, AI adoption should therefore be more explicitly linked with risk in terms of the consequences of failure. Essentially AI may be adopted more readily where there are less severe consequences of AI failure, and adoption should be reduced as the consequences of failure increase. Thus AI may be enhanced where the operational environment is less dynamic (i.e. reduced human factors, non-hostile environment), or where there may be more certainty in the outcomes. In a practical sense, this means promoting the use of AI in more logical applications such as presentation of user credentials (e.g. access card), or where statistical probability of detection is relatively high due to stability of materials or signatures (e.g. x-ray or trace detection of target compounds). There may also be increased opportunity in environments where there are several layers of defence in depth working simultaneously, so the consequences of failure are lessened due to the statistical probability of detection by another technical or procedural measure. Thus, defence in depth strategy should be actively applied to the implementation of AI techniques and consider the strategic placement of humans and machines to produce optimal security outcomes.

> *"… you put them in scenarios where it's not lethal… and I didn't mean just necessarily lethal in terms of human life, I mean it's not lethal to the mission, to the purpose if you make some mistakes" [Participant 11, Security Expert].*
>
> *"When I do red teaming or plan a defence security program, I always put into consideration three points that I mentioned as an attacker – where are the factors that rarely change, where are the factors I expect to be changing and where am I sure that there are factors which will change… so I can focus on where to put the human and where to put the machine" [Participant 7, Security Expert]*

Key Finding: Defence in depth strategy should be actively applied to the implementation of AI and consider the strategic placement of humans and machines to produce optimal security outcomes.

Despite these opportunities, Control technologies were somewhat limited in respect to opportunities for enhancement when compared with Observation, Detection and Response technologies. Participants indicated that although the technology has improved, the systems largely remain unchanged and are not currently taking advantage of AI opportunities. Many of the controlling technologies are logical in nature or are fixed structures (e.g. vehicle barriers) and therefore experts saw limited scope for applying AI to these technologies.

> *"Access control systems as a functional system haven't really changed much in the last 30 or 40 years in terms of what they do, you know the electronics inside the boxes have obviously changed but you know, when I look back to the 80's when I was working in I guess the state of the art in this area, they were doing exactly the same as they do now… but they were never intelligent then and even though we have the ability to make them intelligent now I'm not sure that intelligence is generally there" [Participant 3, Security Expert]*

Areas for enhancement of Control technologies are summarised in Table 41, with location of specific opportunities detailed in Appendix H.

Table 41

Location of Opportunities for Controlling Technologies within the Intelligent Security-AI Cycle

Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Mission linked with organisational objectives	Convergence & aggregation of data	Technical improvements in dynamic environments	Technical improvements in dynamic environments	Systematic autonomous response	
Mission linked with criticality & risk	Multi-point identification and authentication	Aggregation of data	Analytics & predictive analytics for	Establishment of policies and frameworks	

			anomaly detection		
	Application of context	Development of source data sets	Convergence and aggregation of data	Convergence & aggregation of data	
		Application of context	Application of context	Development of neutralisation capability	

3.6.1.2.4 Opportunities for Response Technologies

The results for Response technologies largely reflected those of Observe, Detect and Control technologies within the themes of *AI Testing and Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, and *Holistic Development in the Security Domain* (see Table 42). However the sub-themes of *Safety* and *Scale and Processing Capabilities* emerged under the overarching theme of *Technical Improvements*, and some significant and unique opportunities were presented in *Human-Machine Teaming* and *Innovative Development*.

Table 42

Opportunities for Enhancement of Response Technologies

AI Testing & Standardisation

- Design & development of common connectivity protocols across platforms, equipment & devices
- Development and standardisation of testing mechanisms

Technical Improvements

- Safety
- Increased accuracy & reliability
- Resilience and redundancy to respond to system failure and malicious threats
- Scale and processing capabilities

Integration & Aggregation of Inputs

- Increased automation and cross-platform control
- Greater integration and aggregation of response inputs
- Automation of response outputs

Analytics

- Dynamic & predictive analytics for anomaly detection
- Dynamic assessment for fast & distributed response

Human-Machine Teaming

- Collaborative development between human and artificial intelligence
- Development of response technologies as a 'First Responder' to gather intelligence for human decision making
- Novel Methodologies for dynamic human control & modification during operation
- Aggregated data to aid passive human response

Innovative Development

- Development of drone swarms & robotics for AI response forces
- Development of non-lethal or less-lethal neutralisation capability
- Military technology to shape development in AI-based response technologies
- Development of commercial counter drone technologies
- Development of automated multi-sensor crowd screening

Holistic Development in the Security Domain

- AI development linked with risk and consequences of failure
- Regulatory frameworks for the development of robotics

Safety was a frequently cited term in relation to response technologies, both as an opportunity and a risk of development. Technical improvements to achieve greater safety outcomes were presented as an opportunity, which may take the form of internal or external safety protocols, or through the development of supporting frameworks, policies and procedures. Guidelines for safety protocols in response technologies may be developed using standards for safety integrity (i.e. Safety Integrity Levels), and while there has been some progress toward the development of safety standards for response devices such as drones and robotics (e.g. *ISO 10218 Robots and Robotic Devices, ISO 21384-3 Unmanned Aircraft Systems*), further unification and assimilation of safety standards across countries is required (Pilz, 2018; Robotics Online, 2017). Where more than one response technology is used simultaneously (e.g. drones and robots), additional protocols and measures may be required to increase the safety integrity of interactions between these response devices, or address safety issues resulting from those interactions. Navigating this area may present unique opportunities for the security industry at national and international levels for the development of governance and regulatory frameworks.

There may also be scope for greater development of geospatial AI within response technologies, which for instance may include specific tools such as geofencing to define virtual perimeters (i.e. latitude, longitude, altitude, etc.) in which an object such as a drone or robot may operate. These types of tools may enhance security response through an increased spatial knowledge of security asset locations, in addition to achieving safety objectives for Human-Machine Teaming (HMT) by restricting the movement of response technologies in and around humans (Carter, 2019; Hahn, 2019; Smith, 2018).

While HMT appeared as a theme across all technology categories, it was more strongly represented within the category of Response technologies. As a result of current limitations of these technologies, coupled with immature regulatory frameworks, participants indicated that opportunities in the near future will be likely to excel within structures of collaborative development between human and AI technologies. The application of context to dynamic situations is currently only achievable by humans, which infers opportunities may be limited to applications where a human remains ‘in the loop’ as opposed to ‘on the loop’ (Barnett, 2020). This also signals that opportunities for development of response technologies at higher levels of intelligent autonomy, and in non-collaborative structures, may sit within the *Rules and Deciding* components of systems to be able to adequately assess evolving incidents and execute an appropriate response.

Key Finding: Human-Machine Teaming (HMT) may represent significant opportunities for enhancement of AI in security technologies.

"If I was the person threatening harm to myself or others, and a robot approached me, I don't know if that would necessarily calm me down… so is that actually going to be of any benefit? I think that in a lot of situations if you're just trying to interrupt somebody stealing or graffitiing – perfect. For someone who is going to do something quite drastic then the human intervention is going to be more effective" [Participant 5, Security Expert].

"There will be response technologies that have to assess situations as they develop, and decide on whether a crime is occurring or a security incident is occurring, and then take corrective action, or advise a human operator to take corrective action… That's where I think the big opportunities would be" [Participant 6, Security Expert].

"I think there'll always, at least from a civilian perspective, be a human in the loop to some point because it does come down to security situations, particularly if it's a response and you're responding to somebody, it becomes very dynamic very quickly. You need somebody that'll react to those circumstances and make their best call about whether someone that's presenting as being very angry and agitated is actually going to do something or not, and whether that means you're taking the person out of the response so they're not physically present so they still remain safe, or if the initial response is automated and then it could trigger an in-person response" [Participant 5, Security Expert].

"I think in the civilian market, you know a lot of that stuff is coming in terms of how we can automate UAV response for example to track intruders on site…" [Participant 5, Security Expert].

Many of these examples relate to commercial use of response technologies such as passive surveillance drones and robotics within a very public sphere. Technologies which sit outside of the public sphere, such as weaponry, including Acoustic, Electroshock, Electromagnetic, Optical, Remotely Operated Weapons Systems (ROWS) and Semi-Autonomous and Autonomous Weapons Systems (AWS) were discussed with participants in less detail as these technologies were marginally beyond the specific expertise of participants interviewed. Despite limited emphasis being placed on weapons systems in the interviews, the development and use of such systems outside of commercial security sector was certainly acknowledged by participants. Discussions of weaponry were generally indicative that the porous nature of military-commercial exchange may indeed see weapons systems permeate commercial enterprise at an increased rate.

Military technology is expected to shape development across a range of AI applications, and as a result of increased military-commercial exchange, this may see opportunities for innovative development in commercial response technologies and supporting frameworks. For instance, drone swarms and robotic response forces are predominantly being developed by military or for military and law enforcement applications, but there may be an urgent requirement for counter-systems to become available for commercial and private security applications. For instance, opportunities for counter-drone systems may also evolve in response to enormous growth in the use of commercial drones, particularly in countries where existing regulatory frameworks are immature or, for a myriad of reasons, difficult to enforce. Drone technology is increasingly being equipped with mechanisms such as collision avoidance, allowing drones to rapidly evade objects or hazards (e.g. projectiles launched by protestors), and research and development in sophisticated bio-inspired collision avoidance algorithms may elevate the need for counter-systems in the event these technologies are used by adversaries (Falanga et al, 2020). Thus, opportunities for counter-systems to deactivate or incapacitate a range of AI-enabled response technologies and devices, may significantly increase with progression toward autonomous response forces.

Key Finding: Military technology is expected to shape development across a range of AI applications as a result of the porous nature of military-commercial exchange.

> *"So the military have got drone response capability in various forms of cannons that shoot nets and all of that you know, but it came to light quite quickly that the law in the UK doesn't permit anyone except for the military to be able to shoot something out of the sky… especially because that thing you're shooting out of the sky is owned by somebody else. So there's a huge dilemma in the UK at the moment because of various other issues like Brexit and COVID and whatever, legislation has not made it through parliament yet that allows anyone to aggressively protect themselves from intrusion" [Participant 4, Security Expert].*
>
> *"One of the techniques that is used to sort out the drone problem is jamming, but the minute that you jam so that the signal can't travel from the control box to the drone, you also knock out the mobile phone signal as well… In Australia it is a real issue, you know… the only ones who can actually block out the electromagnetic spectrum is the military" [Participant 4, Security Expert].*

Development of less-lethal neutralisation capabilities also emerged as an opportunity for innovative development in response technologies, though the complexity of achieving this within current legal structures was noted by participants as being particularly challenging. 'Less than lethal' weapons are defined as 'weapons that are explicitly designed to and primarily employed so as to incapacitate personnel or materiel while minimizing fatalities, permanent injury to personnel and undesired damage to property and the environment' (United Nations, 2015, p.6). Currently there are a range of less-lethal weapons available for security and law enforcement response (e.g. taser/ electroshock, tear gas, sticky foam etc.) however currently the majority of these are manually operated. Opportunities may therefore exist for intelligent deployment capabilities for various classes of less-lethal weapons, and development of novel classes of weapons with intelligent autonomy may have a similar commercial viability in specific applications or environments.

> *"From a UN (United Nations) or a humanitarian aspect – non-lethal response… a non-lethal response to probably a lethal attacker" [Participant 1, Security Expert]*

Key Finding: Unique opportunities for innovative development of drone swarms and robotics are emerging for Response technologies.

Innovative development of both military and traditional security systems therefore present significant opportunities for response technologies, however there must be substantial supporting frameworks to guide and shape the design, development and deployment of such systems within the commercial sector. Once again, AI adoption should be more explicitly linked with risk in terms of the consequences of failure, with AI being implemented where there are less severe consequences of failure.

Specific areas for enhancement of Response technologies are summarised in Table 43, with location of opportunities detailed in Appendix I.

Table 43

Location of Opportunities for Response Technologies within the Intelligent Security-AI Cycle

Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Mission linked with organisational objectives	Convergence & connectivity across systems	Increased processing capabilities	Human-machine teaming	Systematic autonomous response	
Mission linked with criticality & risk	Aggregation of data	Technical improvements in dynamic environments	Analytics & predictive analytics	Convergence and connectivity across systems	
		Analytics & predictive analytics for anomaly detection	Fluidity in rules under dynamic conditions	Development of novel responses and neutralisation capacity	
		Aggregation of data; data management systems	System resilience and redundancy	Aggregation of data for distributed intelligence	
		Application of context	Application of context	Application of context	

3.6.1.2.5 Opportunities for Integrated Security Technologies

Integrated security technologies are a cumulation of Observe, Detect Control and Response technologies, and therefore many of the themes and opportunities for integrated technologies reflect those across all categories of security technologies. Themes which emerged for Integrated Security Technologies were *AI Testing and Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, *Human-Machine Teaming*, *Innovative Development* and *Holistic Development in the Security Domain* (see Table 44).

Table 44.

Opportunities for Integrated Security Technologies

AI Testing & Standardisation

- Design & development of common connectivity protocols across platforms, equipment & devices
- Development of testing, trialling and piloting of systems

Technical Improvements

- Increased accuracy & reliability
- Enhancements for more dynamic environments
- Resilience and redundancy in physical and management systems
- Improved processing capabilities
- Improved quality of systems and components

Integration & Aggregation of Inputs
- Greater integration and aggregation of multiple sensors & systems
- Automated decision making based on integration of multi-sensor inputs
- Aggregation of multi sensor data to provide context for enhanced decision making

Analytics
- Analytics to provide greater context in dynamic conditions
- Analytics to enhance decision making by providing ranked options for response

Human-Machine Teaming
- Collaborative development between human and AI to increase Probability of Detection
- Novel methodologies for dynamic human interaction & control
- Aggregated data to aid human response
- Integrated technologies to enhance human decision making

Innovative Development
- Development of automated multi-sensor crowd screening

Holistic Development in the Security Domain
- Synchronicity in development across political, economic, social, legal and cultural frameworks
- AI development linked with risk and consequences of failure
- Regulatory frameworks to assist development of integrated technologies
- Cultural change in security management to promote quality over cost

Participants identified greater integration and aggregation of data from multi-system and multi-sensor inputs as one of the primary areas requiring development, in addition to the application of context from analytical algorithms. Greater integration and aggregation of data with applied analytics has the potential to enhance security decision making and provide distributed intelligence between systems, facilities and security teams for improved security response. To achieve this, integrated systems essentially require transformation from a logical input/output type architecture to a more sophisticated architecture with a coordinated approach to decision making between devices and throughout the field, automation and management levels.

Key Finding: Integrated systems essentially require transformation from a logical input/output type architecture to a more sophisticated architecture with a coordinated approach to decision making.

"Most of the intrusion detection systems that hang off security management systems are very, very basic – they're usually switches, on and off, plus maybe five to ten percent of them being sensors. And those sensors are the cheapest ones that people can buy to put in because that what the clients are prepared to pay for… and so first of all the sensors and the technology hanging off security management systems is not there in the ability to be manipulated by AI to any large degree because they're all just rules based on and off" [Participant 9, Security Expert].

"at that [security management] level, that's where you would have the more general AI I guess that would be able to translate all of these different field device data sets into something that's a bit

more context driven… so you know the CCTV system might say this looks strange, Ill raise an alarm, but it doesn't go to an operator, it goes to the PSIM – the PSIM also picks up through the perimeter intrusion detection system that this event is happening in a similar location" [Participant 5, Security Expert].

"So the sensor will sense something that makes a decision, that decision gets fed up to the next level that's a bit smarter and has a few more components, and if those components say nothing's happening, then ok nuisance alarm – or we've got these other inputs, that's also a bit weird, let's keep escalating, and then eventually it gets to the point where you need to bring an operator in" [Participant 5, Security Expert].

However even with increased capabilities in machine decision making through aggregation of data and the application of analytics, emphasis was still placed on Human-Machine Teaming (HMT) for assessment and response to security events. Participants noted opportunities may exist for the presentation of tailored information to human operators, including ranked options with probable outcomes to enhance decision making. With the goal of HMT being to collaboratively assist humans, rather than redirect human activities to machines, HMT within integrated security management systems could offer substantial improvements and benefits for security operators without compromising critical decisions and potentially impacting on the security and safety of individuals.

Key Finding: Human-Machine Teaming and ranked options for response could enhance critical decision making for human operators.

"I do think that it will be difficult to expect a security management system to manage itself… I think it has to be able to present a level of automated information to a human operator or capable guardian somewhere, somehow, and then that person still make the final executive decisions… I think that's the only way you could possibly refer them and guarantee the appropriate provision of security and safety of individuals" [Participant 6, Security Expert].

"… at the end of the day we're not trying to develop these technologies to replace the organisations. We might be replacing certain jobs in the workforce, but the machine just is an operator, is just an actor within an organisation in the same way that a person is… So I just see these machines as becoming another actor within the organisation and the organisation will probably continue to be a combination of collaborative machines and people, and I think that's the future organisations we're going to have are going to look like that. And so I think that's where we're sort of heading rather than making them fully autonomous, they'll be fully autonomous for a particular task, but not to do things totally independent of their human masters" [Participant 3, Security Expert].

"… we're asking for the security technology to be the most effective aid to the human behind the desk, we're not asking for the security technology to make any critical response decisions that have the potential to affect people" [Participant 8, Security Expert].

Specific areas for enhancement of Integrated Security Technologies are summarised in Table 45, with location of opportunities detailed in Appendix J.

Table 45

Location of Opportunities for Integrated Security Technologies

Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Novel methodologies	Convergence & aggregation of data	Novel detection devices and analytical methodologies	Fluidity in rules under dynamic conditions	Systematic autonomous response	
	Increased use of detection devices in stable conditions and environments	Increased processing capabilities and analytics	Technical improvements in dynamic environments	Improved analytics and novel methodologies	
	Application of context	Technical improvements in dynamic environments	Application of context	Analytics & predictive analytics for anomaly detection	
		Application of context		Development of neutralisation capacity	

3.6.2 Objective Three

3.6.2.1 Benefits & Risks of AI Development in Security Technologies

Objective Three sought to identify the benefits and risks associated with AI development in security technologies, including integrated security systems. Objective Three was facilitated with the following sub-questions presented to participants:

a) What benefits would result from identified opportunities (across all categories of *Observe*, *Detect*, *Control*, *Response* and *Integrated Security Technologies*), and what risks do you foresee from future advancements?

b) From a security management perspective, which aspects of integrated security systems should retain human assessment and control.

c) Which of the *Literature Defined Risks of AI* apply to security technologies and integrated security systems?

d) Do you foresee any additional risks which require consideration?

Benefits and risks were raised with participants in all parts of the interview with (Parts 1-5, see Appendix M) to identify risks which may be specific to each technology category. Expert participants were provided with a comprehensive list of risks (Table 46) derived from the literature (see Section 3.4 *Literature Defined Risks and Limitations of Integrated Security Technologies*) which was used to guide responses to sub-questions and provide an opportunity for participants to add identify additional risks which may be specific to the security industry.

Table 46

Literature Defined Risks of Artificial Intelligence

Risk	Description
AI Failure	Insufficient redundancy built into technologies – technical or operational AI issues result in system breakdown or failures.
Latency Risks	Cloud-based delays in processing and communication create latency risks impacting on safety and operations.
Distrust of AI	Corporate or public noncompliance due to distrust of AI technologies or those responsible for system operation.
AI Mistakes	AI does not know how to respond to novel objects or situations without extensive training sets or data resulting in errors or consequences – for example system incorrectly classifies a civilian cruise ship as a naval warship and launches attack
Misuse	AI used inappropriately or for unintended purposes.
Limited Threat Management	Pattern recognition can be used to identify behaviour, however AI cannot contextualise the behaviour to determine whether it is malicious - e.g. if a sensitive file is accessed after office hours by an authorised person, the event may not be flagged.
Insider Threat	Insiders are able to access, modify, disable or override systems.
Cyber Security Risks	Numerous security risks - adversaries may hack, hijack or attack, compromise or disable the technology (e.g. data breaches, Denial of Service (DoS), Man-in-the-middle Attack, malware injection/ use of malicious software (e.g. ransomware, siege ware, jackware), autonomous cyber-attacks etc).
Cloud-based Security Risks	Attack or exploitation of cloud-based vulnerabilities (or use of malicious software such as Spectre & Meltdown) to disrupt, damage, gain access to, or take control of security system.
Edge of Network Security Issues	Attack or exploitation of edge of network vulnerabilities including Denial of Service (DoS), Man-in-the-middle Attack, Rogue Gateway, Rogue Data Centre, Rogue Infrastructure, Physical Damage, Privacy Leakage, Privilege Escalation, Service Manipulation, Virtual Machine (VM) Manipulation, Injection of Information, etc.
Safety Issues	Autonomous execution of instructions by system (system override) preventing human intervention and control.
AI Arms Race	Development of AI technologies creates an AI arms race in developing technologies such as lethal weapons systems.
Illegal AI trade	Technologies or systems traded through black or grey markets. State-sponsored acquisition of AI-tech by adversarial groups.
Lack of Transparency and/or Accountability	Auditing issues or inability to create transparency in system processing and decision making. Inability to determine responsibility for system shortcomings or failures (e.g. manufacturer, system or operator error ?)
Legal Risks	Legal damages or ramifications resulting from AI actions, omissions or failures. Forensic evidence issues (chain of custody, compatibility of various types of digital evidence) resulting in evidence or prosecution failures.
Economic Risks	Unemployment e.g. AI causes some security jobs to become obsolete.
Political Risks	E.g. market monopolies or reliance on imports of international technology.
Ethical Risks	Fairness and equity issues (e.g. aggregated data may result in vilification of certain groups of people). Ethical risks of the development and use of lethal AI systems (e.g. autonomous weapons systems).

Inadequate Governance and Regulation	Responsible supervision, regulation and governance of the design and deployment of AI systems, including governance and regulation of independent software vendors (ISV).
Data Ownership and Management Issues	Lack of regulation of AI software, hardware, and models, including; data ownership, data location and control, content disclosure, confidentiality, Regulatory and legislative compliance.

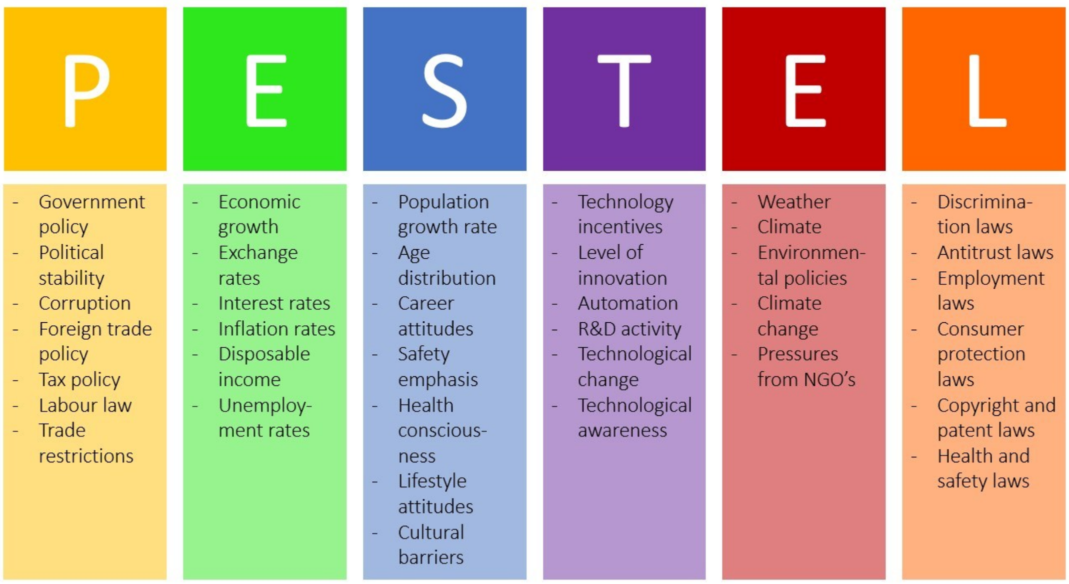
A 'PESTEL Analysis' model was used to present the benefits and risks of using Artificial Intelligence in security technologies. PESTEL is an acronym referring to the *Political*, *Economic*, *Social*, *Technological*, *Environmental* and *Legal* factors which may affect an organisation. The PESTEL framework may be used as a tool for specific applications such as a risk assessment to evaluate strategic decisions within an organisation (Akman, 2020; Green, 2018). The PESTEL framework was used to establish the areas which may be impacted by the design, development and deployment of Artificial Intelligence in security technologies.

Benefits and risks identified by participants were assigned to the relevant PESTEL classifications as illustrated in Figure 27 (Business-to-you, n.d.) for each of the security technology categories. Though the risks and benefits are clearly identified, there has been no application of criteria to determine the weight of each benefit and risk in relation to one another, or in a strategic management sense. A comprehensive analysis including the overall value of each benefit, and criticality of each risk is necessary for comprehensive assessment of AI risks within specific business contexts. Development of an AI risk decision matrix, aligning with the *Security Technology Intelligent Autonomy Scale* is therefore recommended as an outcome of Phase Three of this research.

Key Recommendation: Development of an Artificial Intelligence Risk Decision Matrix for security mangers to evaluate the benefits and risks of AI technologies.

Figure 27

PESTEL Analysis Framework.



3.6.2.2 Objective Three Findings

3.6.2.2.1 Observe Technologies

The use of AI in observation technologies was identified as having substantial benefits and significant risks in all PESTEL classification areas with the exception of *Environmental* factors as detailed in Table 47.

Table 47

Benefits and Risks of Artificial Intelligence in Observe Technologies

	OBSERVE TECHNOLOGIES	BENEFITS	RISKS
P	POLITICAL	Enhanced aggregation of disparate data for law enforcement	Political, cultural, social bias caused by machine learning algorithms. Difficulty in attributing accountability & responsibility for AI systems & decisions Lack of transparency in AI decisions Industry disruption from AI use, misuse or failures Privacy & human rights issues Political implications from the use or release of force
E	ECONOMIC	Improved allocation of resources Cost reduction Reduced disruption from nuisance alarms Reduced operator workload and fatigue Increased operational capabilities	Costs to industry/jobs from redirection of labour Economic costs of disruption/ downtime from system maintenance or failure Impact of system reliance (i.e. reduction in manual operational capabilities)
S	SOCIAL	Human-machine teaming to reduce safety risks Removal of humans from the front line	Absence of legal and political frameworks to govern the development and deployment of AI in socially acceptable ways Safety risks & issues Privacy, ethical & human rights issues Social risks of inability to revoke AI development Breaches of capable guardianship and duty of care Social implications from the use or release or force Misuse of data
T	TECHNOLOGICAL	Human-machine teaming to defeat human adversaries More extensive and directed surveillance Increased real-time processing Faster, more focussed decision making and response Diversity of applications	AI & technology failures Risks associated with machine learning databases Practical issues associated with generation or acquisition of datasets AI implementation & management issues Autonomous identification & classification failures

			Autonomous failure of human override Misuse of data Vulnerability from low accuracy/reliability of technology AI vulnerability to defeat by human adversaries Security and technology decay Data compromise/ breach
E	**ENVIRONMENTAL**		
L	**LEGAL**	Increased traceability of incidents	Absence of legal and political frameworks to govern the development and deployment of AI in socially acceptable ways Non-compliance Legal liability (e.g. from system errors) Duty of Care Privacy & human rights issues

Benefits

Technological and *Economic* factors produced favourable and tangible benefits, with AI enabled observation technologies enhancing operational capabilities to reduce costs, workloads and issues such as operator fatigue. *Social* benefits of increased safety and the removal of humans from the front line were frequently identified by participants as providing significant benefits, and are expected to weigh heavily in decisions to substitute AI-enabled machines and devices (e.g. drones) for humans operating in difficult or dangerous environments. Only marginal benefits emerged under the classifications of *Legal* and *Political*, with increased surveillance creating greater traceability of incidents and aggregation of data from multiple systems and sensors providing advantages for law enforcement.

Risks

Though the benefits may be considerable gains from a corporate security management perspective, the risks of AI in observation technologies were extensive and may have significant ramifications for corporate use.

Political Risks

Several political risks may arise from the use of AI in Observation technologies. Difficulty in attributing accountability and responsibility for AI systems, in addition to a lack of transparency in the way algorithms make decisions, may influence political regulation and limit the extent to which security organisations can implement and use AI. Consequently, governments may move to restrict the use of intelligent observation technologies, particularly where there is the potential for biases to be created (e.g. machine learning systems create biases directed toward ethnic, racial, social or minority groups), or where surveillance is perceived to infringe on privacy and human rights. This may be more evident in individualist populations where individuals have greater rights and freedom of speech.

> *"So Facebook developed a chatbot that could learn, and what it learned was to become racist and misogynistic and abusive… when they looked at the input provided to it, the input wasn't overly any of that… so when we say listen were going to create a rules based AI and then it'll migrate across into an AI that learns for itself, the biggest issue that we've got is how you train it because once it*

has the autonomy to learn for itself we then have the potential to lose control of it" [Participant 8, Security Expert].

"… there is no answer to that one… I'd want to see a computer up in front of the Human Rights Commission' [Participant 9, Security Expert].

"You could build a case for and against appropriate application of technology in a court of law and that's largely underpinned by personal bias" [Participant 6, Security Expert].

Key Finding: Observation technologies may carry significant risks resulting from privacy and human rights issues, particularly for individualist populations where citizens have greater rights and freedom of speech.

Economic Risks

The primary economic risks for Observation technologies were a redirection of labour, potentially impacting on security employment and reducing available jobs as a result of increased automation. For corporate entities, there may be economic costs of disruption and downtime in the event of technical difficulties or mishaps, and overreliance on AI systems may not only exacerbate downtime but also present additional risks of not being able to fulfil contractual obligations.

Social Risks

Privacy and human rights issues arising from AI also fall under the classification of *Social* risks, which again may be amplified in individualist populations that place greater emphasis on individual rights. This is particularly relevant for Observation technologies, as visual surveillance can create significant social opposition due to the potential for misuse and the social ramifications of personal visual data being exploited or exposed in public spheres.

"If you have a surveillance system that is designed to monitor the population as a whole, then you can use it to I guess gather information about people, and this is what China is doing. So you know they have a social credit system over there so they can use surveillance technologies to contribute to their data collection… the systems have a learning capability, so they start to learn who people are, how they go about their business, and whether they are complying with the social requirements, or if they are part of a racial group that is specifically being targeted in policies" [Participant 3, Security Expert].

"CCTV is far more intrusive because you record it, you can visually see it, you can put it on the internet, you can use it to look into people's windows – you can do all kinds of stuff with it, right… and it's just a far more visual medium which seems to lend itself to more commonly being abused" [Participant 6, Security Expert].

Privacy and ethical issues coupled with the absence of legal and political frameworks to govern the development and deployment of AI in socially acceptable ways carries significant social risks. This is of particular concern where there may be difficulty in retaining control of personal information generated by AI enabled technologies. The inability to contain development of technologies or revoke personal information (such as facial biometric data) that has been captured by observation technologies is a serious social concern, particularly when sensitive personal information is stored online, or where it may create safety risks for individuals such as victims of domestic violence.

"We keep telling people you've got an access control card and you've got a pin – both of those if you lose control of them can be replaced. You cannot replace your biometric… and once you've lost control of that biometric there's no getting it back" [Participant 6, Security Expert].

"When [shopping centre] put their car park systems online, where when you drove in it automatically registered your plate and followed you around the car park so if you forgot where you were, you could put your license plate information in, and it would tell you where your car was. And that seemed like a fantastic idea to [shopping centre] marketing, until a young lady who was a subject of a domestic violence order had her husband put the license plate details in and it told him where to find the car and assaulted her…" [Participant 6, Security Expert].

Key Finding: AI must be developed and deployed in socially acceptable ways, to minimise the risks associated with violation of personal data and information.

Technological Risks

There are considerable security risks generated by the technological limitations of AI-enabled observation technologies, particularly where the security function may overlap safety or other mission critical functions within an organisation. AI and technology failures, including low accuracy and reliability issues, may increase a security system's vulnerability to defeat. Limitations in computer vision are reported to have resulted in incorrect identification and classification of objects, which may allow an adversary to evade detection. There are also issues with generating or acquiring the necessary data sets required to facilitate greater accuracy in object recognition – for instance, acquisition of visual images for data sets my breach data privacy laws or duty of care obligations.

"Number plates are a perfect example, because a human can tell the difference between a real number plate and a fake number plate. So when you approach the gate, the guard looks at it and goes 'mate you've got a photocopy cello taped over your number plate there', the number plate recognition system just reads the letters and goes oh yeah sure I'll let you in' [Participant 8, Security Expert].

"The biggest flaw is still in the computer vision stage, and it is in the ability of a system to differentiate different objects. Even the top-notch systems that we've got occasionally classify a lamppost as a human… one of the things when you're looking at technology for automation, and

particularly from a security perspective, you want the system to be able to tell the difference between a human and a dog. So when a human walks along, and then as he walks behind the lamppost, gets down on all fours, and the computer vision says 'well now there's a dog in the frame and the lamppost is a human' it doesn't matter what rules you put behind it, nothing works… looking at things from an object classification, we're still at a very basic level" [Participant 8, Security Expert].

"My biggest fear is the misuse of the database information that's being gleaned on the general population" [Participant 8, Security Expert].

The management and storage of extensive amounts of data generated by observation technologies also presents risks for AI use. Data must be stored correctly to ensure it remains functional and accessible for legitimate purposes, in addition to ensuring data does not become corrupted, misused or hacked by adversaries. This presents significant challenges for observation technologies due to the volume of data generated by such systems. Implementation and management of AI technologies must have oversight to ensure integrity of systems and reduce vulnerability to risks such as insider threat. However this may be difficult to achieve where security or executive knowledge of AI systems is limited or where there is limited transparency of AI techniques in commercial security products resulting from intellectual property rights or trade secrets.

"If you have very large systems with you know say 100,000 cameras or more, you need technologies to manage the data because it's impossible for people to manage that information" [Participant 3, Security Expert].

"Wherever you still do have a human in the loop, you potentially have a situation where that human has a degree of access and insight into the system in an uncontrolled manner – and we do see that even with security guards, or other security practitioners that have field roles in the industry… and that's what you get, you get humans that have no management oversight, so they have no capable guardian, so who is the capable guardian of the capable guardian? And that's really where security governance comes into play, so where you don't have proper oversight of those humans in the loop, that's where misuse can occur, and I guess that misuse is strongly related to insider threat" [Participant 6, Security Expert].

Legal Risks

Observation technologies may create legal liability for security organisations as a result of privacy breaches or other infringements on human rights. System failures or errors may result in legal damages, though current legal frameworks for attribution of AI responsibility and accountability are significantly underdeveloped. Considerable progress must be made in determining legal responsibility and liability for AI programs and products, particularly for machine learning applications with limited transparency and traceability in decision making.

"I also think that the structure of the Western world, which provides people with the route of compensation via a legal system will always push us to a situation where we have to be able to say

that person made the decision, that person was responsible, that person was accountable, and I see a long, long way before you get away from that" [Participant 4, Security Expert].

"How do you make these systems accountable? How do you build the legal and regulatory frameworks around that to ensure that there is traceability… they talk about the black box of AI, and then when you start getting into machine learning, how do you know how the machine has arrived at those decisions that it's made, and then who is liable?" [Participant 8, Security Expert].

Key Finding: Governance and regulatory frameworks must establish mechanisms for oversight to prevent the misuse or exploitation of AI technologies.

3.6.2.2.2 Detect Technologies

The use of AI in detection technologies was identified as having moderate benefits and considerable risks in all PESTEL classification areas with the exception of *Environmental* factors as displayed in Table 48.

Table 48

Benefits and Risks of Artificial Intelligence in Detect Technologies

	DETECT TECHNOLOGIES	BENEFITS	RISKS
P	POLITICAL	Additional data for law enforcement & intelligence	Sovereign risk resulting on reliance on foreign technologies & entities Managing public perceptions of risk Privacy & human rights issues Political implications from the use or release of force
E	ECONOMIC	Improved allocation of resources Cost reduction Increased operational capabilities	Cost to industry/jobs from redirection of labour Costs associated with data ownership, storage and location Cost burdens of purchase, installation & maintenance
S	SOCIAL	Early intervention opportunities	Privacy & human rights issues Social implications from the use or release or force Misuse of data
T	TECHNOLOGICAL	More extensive and directed surveillance Faster, more focussed decision making and response Improved control of/ response to critical incidents Early intervention opportunities Threat mitigation opportunities	AI & technology failures Vulnerability to tampering & defeat Manipulation of data via weak system integrity Misuse of data Autonomous response may create additional risks/ issues Exploitation of security vulnerabilities of IOT and Edge of Network devices Security and technology decay Cyber & cloud-based risks

E	**ENVIRONMENTAL**		
L	**LEGAL**		Issues concerning data ownership, storage & location Legal liability (e.g. from system errors) Privacy & human rights issues

Benefits

Technological and *Economic* factors once again produced favourable and tangible benefits for AI enabled detection technologies. Increased operational capabilities and improved allocation of resources resulting from advances in detection technologies are expected to reduce security costs. AI in detection systems and devices may also facilitate more directed surveillance, enabling faster decision making and response to critical incidents. Early detection may also increase opportunities for threat mitigation as well as *Social* benefits from implementation of early intervention strategies.

> *"One of the benefits then is the opportunity for early intervention, and being able to mitigate a range of threats based on that early detection" [Participant 2, Security Expert].*

Law enforcement agencies may see *Political* benefits resulting from increased use of detection devices, particularly for intelligence purposes. For example, number plate recognition could be integrated with traffic surveillance and GPS systems, which may also be integrated more extensively with law enforcement databases and notification systems to assist in identifying and locating persons of interest. Development of these types of systems may improve policing outcomes and subsequently promote good governance practices from which the commercial sector can model.

Key Finding: AI enhancements in detection technologies may offer early intervention opportunities to respond to, and mitigate threats, sooner.

Risks

Though there may be moderate benefits from enhanced AI in detection technologies, there are considerable *Political*, *Economic*, *Social*, *Technological* and *Legal* risks associated with AI use.

<u>Political Risks</u>

Sovereign risk was identified as an issue for detection technologies, particularly with the adoption of new devices and systems (e.g. biometrics, acoustic detection etc.) which have emerged from global markets. Space technologies and satellites were used as an example to emphasise the potential outcomes from a loss of sovereignty over technologies and data. Political and economic forces affecting the availability of technologies and their components have the potential to disable industries or sectors, for instance any interruption to satellite launch capabilities may restrict or disable national communications. Therefore countries or regions without strong manufacturing industries may also experience higher risks in relation to sovereignty as reliance on international markets to supply AI enabled technologies increases.

> *"If a system is compromised by AI itself, then sovereignty over its own data… 5G is an example… such new technologies come through that we adopt, we may lose sovereignty over that. We're seeing now in space without saying that's why, you know, there's a bit of a space race on and Australia started its own space industry, realising that we rely on foreign satellites that if they get taken down, we have no way of launching our own satellites. So that's the first mission right now is to be able to have sovereignty over our space communications because we rely on it" [Participant 2, Security Expert].*
>
> *"Australia is not immune to this… Australia would be a lot better placed in many ways if we had certainly a better high-tech manufacturing. I think it's quite arguable that we don't lack so much the scientific and technical knowhow but certainly in some sense we're deficient… there's things that we probably could work out how to do in five years, but we couldn't do tomorrow which other countries on the planet can do" [Participant 11, AI Expert].*
>
> *"The US isn't exactly the paragon either… they have outsourced a lot of their manufacturing capability to Asia, to China and Taiwan and various other places. It's probably not the case that they would be particularly that hard pressed to restart it up in the US… they are probably in at least some danger of going down that track where if they were denied access to the silicon foundries offshore, you know Samsung and various companies produce a huge fraction of the worlds semiconductors and they're not based in the US" [Participant 11, AI Expert].*

Key Finding: Failure to establish local manufacturing industries may increase reliance on international supply markets and consequently increase sovereign risk.

Managing public perceptions of detection technologies may also be considered a political risk. Where new detection technologies are introduced, or where their use may be significantly increased, public perceptions of safety may require consideration. Management of health and safety concerns may extend to security personnel operating detection systems and devices in addition to the public or individual system users.

> *"We just need to be careful that we do understand any hazards that the technologies might present, or even perceptions of hazards that the public might be concerned about. For example if we start using millimetre wave technology, and even x-ray was mentioned there, if the public become concerned that they are going to catch some nasty disease from it then it will become a problem… I know that a lot of people are concerned about electromagnetic causes of health problems, so we just need to be aware that some of these technologies may be of concern to the public and so the risk is we don't manage those concerns or perceptions" [Participant 3, Security Expert].*

Economic Risks

Economic risks for detection technologies were similar to those of observation technologies. Redirection of labour and cost burdens of purchasing, installing and maintaining systems were identified as economic risks in addition to potential data ownership and storage issues. The storage and location of data may affect data availability for consumers, particularly in cloud-based applications. Political unrest, natural disasters, or other events may impact on the capacity of the hosting platform to provide services, and commercial disagreements or disputes may result in data or cloud services being withdrawn. This could have serious ramifications for corporate entities if user access to data is prevented, or corporate entities lose control of building systems and services as a result.

> *"After a disagreement, they turned them off, so their whole building had no air-conditioning, no power, standby generation – all the systems that were hosted in the cloud, they lost the lot" [Participant 9, Security Expert].*

Key Finding: Cloud-based systems increase economic risks resulting from loss of access to data and services.

Social Risks

Privacy, human rights issues, and the misuse of data again surfaced as social risks arising from the development of AI in detection technologies. However, minimal emphasis was placed on these risks in respect to detection technologies as participants generally indicated detection systems and devices are less socially intrusive than surveillance technologies, and therefore carry less social risk. Misuse of data primarily referred to biometric technologies and the loss of control of biometric data (see 5.2.2.1 Observe Technologies).

> *"I don't think there are too many risks at all as far as the detection devices go" [Participant 3, Security Expert].*
>
> *"The risks with detection technologies would definitely be lower, and that's because the detection technologies that we've talked about are not so easily used against people, against their will. So you've got passive infrared detectors, or you know some sort of radar system or microwave detection system, there's only so much you can use that for in a way that could cause discomfort or harm to an innocent member of society" [Participant 6, Security Expert].*

Key Finding: Detection technologies are perceived to carry less risk than observation technologies.

Technological Risks

The most significant risks associated with detection technologies centre around *Technological* factors more so than other PESTEL factors. Detection technologies – including IOT and Edge of Network Devices - have a range of security vulnerabilities, thus sensors and devices may be vulnerable to tampering, exploitation or defeat. Participants noted that current limitations of detection technologies, such as false positives/ negatives and susceptibility to defeat may be exacerbated by weak system integrity. System integrity issues may be further exploited in AI-enabled detection technologies, resulting in hacking, hijacking, manipulation of data or algorithms, or facilitate AI attacks such as the insertion of malicious code which could be triggered by external signals or inputs (Eidson, 2019).

> *"No one technology is capable… you can defeat every technology in isolation" [Participant 8, Security Expert].*
>
> *"The risks all, I believe, sit around integrity of that whole rules/deciding domain and see integrity of those rules and ensuring that they can't be tampered with… so it can't be manipulated" [Participant 4, Security Expert].*

Cloud-based applications raise issues of not only data ownership and storage, but broader cloud and cyber risks such as data control, confidentiality and content disclosure. Availability and accessibility of data and services are significant risks of cloud-based systems, particularly for AI systems as AI algorithms often require the increased processing power of cloud-based platforms to operate effectively. Denial or withdrawal of service or technological failures resulting from inability to access these services and data could have profound consequences for organisations using AI devices.

> *"The whole Internet of Things and the problem that you've got all these thousands of sensors that are attached to the internet that invariably are not properly secured… so firstly, how do you detect that a sensor has been attacked and secondly, what do you do about it? And this is where you get this whole grey area about where does the responsibility lie? Unfortunately over the years there's been a determination that logical security will sit in the hands of the IT world and physical protection will sit in the security world" [Participant 4, Security Expert].*
>
> *"When you run these computational algorithms on Amazon's cloud, you've got to remember that 'cloud' is just a marketing term for somebody else's computer… so you install this plug-in on your CCTV system, your camera detects a motion event of some kind, takes that footage and uploads it to the cloud. They run their GPU on it, send it back and say that meets your rules, it's an alarm or it's not. So what happens to your video motion when you can access the world? What happens to the footage you're uploading to their cloud? [Participant 4, Security Expert].*

Key Finding: Detection technologies – including IOT and Edge of Network Devices - are significantly vulnerable to defeat.

Legal Risks

Detection technologies were perceived to have less legal risks than Observation technologies, as privacy risks are less prominent due to the absence of visual surveillance. However detection technologies may still create legal liabilities for organisations as a result of perceived injustices, infringements on individual privacy (e.g. text analytics applied to organisational emails), or adverse outcomes resulting from the implementation of detection algorithms or processes (Krouse, 2019; Partnoy, 2018). System failures or errors may also result in legal damages or compensation, particularly where manufacturers and vendors purport that a system achieves detection proficiencies beyond its actual capabilities. However, developing AI testing mechanisms and standards may reduce these risks by holding manufacturers and vendors to greater account in regard to system performance.

Determining legal responsibility and liability for AI programs and products may help to mitigate some of the legal risks associated with AI use, however development of legal frameworks in this respect are still in their infancy.

Key Finding: Development of AI testing mechanisms and standards may reduce legal liability resulting from inadequate system performance.

3.6.2.2.3 Control Technologies

The use of AI in control technologies was identified as having minor benefits and considerable risks in all PESTEL classification areas with the exception of *Environmental* factors as displayed in Table 49.

Table 49

Benefits and Risks of Artificial Intelligence in Control Technologies

	CONTROL TECHNOLOGIES	BENEFITS	RISKS
P	POLITICAL		Difficulty in attributing accountability & responsibility for AI systems & decisions Lack of transparency in AI decision making Privacy & human rights issues Political implications from the use or release of force
E	ECONOMIC	Increased productivity	Costs associated with data ownership, storage & location
S	SOCIAL	Improved safety	Privacy & human rights issues Social implications from the use or release or force Breaches of capable guardianship and duty of care Misuse of data Exploitation of biometric data is irreversible
T	TECHNOLOGICAL	Reduced latency issues	AI & technology failures Autonomous failure of human override

		Increased adaptability from dynamic threat assessment	Misuse of data Vulnerability from low accuracy/reliability of technology Exploitation of biometric data is irreversible Manipulation of data via weak system integrity Cyber & cloud-based risks
E	**ENVIRONMENTAL**		
L	**LEGAL**	Reduction or elimination of safety risks	Legal liability (e.g. from system errors) Privacy & human rights issues Issues concerning data ownership, storage & location

Benefits

Marginal benefits were identified from AI enhancement in Control technologies. *Social*, *Legal* and *Technological* benefits were based on improved safety, adaptability and reduced latency issues. Use of AI in access control systems was discussed with Covid-19 safe applications provided as an example, whereby technologies could be used to provide automated control of entry to buildings based on fluctuating Covid-19 rules and restrictions. Adaptability based on dynamic conditions could assist security operators manage the movement of people within buildings based on the latest threat advice.

> *"… having smart access control – and it has actually started moving into that space – but more probably CCTV through people counting where the CCTV nominates ok you've got 50 people just moved into this space – now you've reached your two square metre limit and then we stop access… we went from two square metres to four square metres, now we're back to two square metres and next week we're back to four square metres, so you need to be able to have adaptable access control for that" [Participant 2, Security Expert].*

Convergence between buildings and departments, including physical and logical systems, may enhance productivity by automating functions. For instance employees could have physical or logical access restricted out of hours or in the event they are physically absent from the workplace. Convergence of systems and departments can assist in delivering security objectives by creating greater control and limiting opportunities for procedural vulnerabilities to be exploited.

> *"…Joe Bloe booked in sick today, okay well Joe Bloe's access is automatically blocked " [Participant 2, Security Expert].*

Risks

Risks associated with AI use in Control technologies were identified across *Political*, *Economic*, *Social*, *Technological* and *Legal* PESTEL categories.

Political Risks

Political risks may arise as a result of difficulty in attributing accountability and responsibility for AI systems, in addition to a lack of transparency in the way algorithms learn or make decisions. Issues in accountability, responsibility and transparency may result in increased regulation or restrictions on the use of AI in certain applications. Political risks may also arise where control technologies are perceived to infringe on individual privacy or human rights, for instance the use of biometrics in access control systems may be met with discontent, particularly in countries where corruption may be problematic and enrolled users fear authoritative misuse, or even for counties that simply have less developed data security frameworks. Users may be concerned with security of their personal information retained in databases, in terms of who is authorised to access that information, and what the consequences of data breaches may mean for them personally. Similarly, governments may be specifically concerned with foreign creation of - or access to - databases containing the personal profiles and details of national citizens and how this may affect both individual and national security.

"It's the misuse of the information… the problem is, nobody collects this information on an ad hoc basis. Someone collects the information, they have a purpose for it, and therefore they curate the database, which means the database is well looked after and accurate… and then they lose control of the database because their IT department got hacked by a foreign government… and then you've got a million people that are pre-enrolled in the system because they've stolen someone else's database" [Participant 8, Security Expert].

Social Risks

Privacy and human rights issues, in addition to misuse and the potential for exploitation of biometric data are considered social risks due to potential impacts on individuals, as well as broader implications for society. Analytics applied to control technologies to assess or predict individual or group movements may be met with opposition, especially if these systems have the potential to be misused or are able to facilitate any form of discrimination against individuals based on their gender, race, ethnicity, disabilities, or other unique attributes. Although many countries have legislative instruments to prevent such discrimination, this is not reflected world-wide and there may be significant social risks associated with AI-enabled control systems knowing and recording the movements of individuals.

Biometric based access control systems may also create social risks, particularly in applications such as airport security where users must oblige to enter or exit a country or province. Inaccuracies in biometric technologies are well known, for instance where Amazon's facial recognition falsely matched twenty-eight members of US Congress with people who had been arrested for a crime (Snow, 2018) – the implications for an individual could be immense if a false positive was to match their identity to someone who had committed a serious offence, particularly if an incident like this were to occur while travelling overseas. For this reason, acceptance of biometric technology is considerably low, therefore social distrust is likely to continue to limit widespread adoption of biometrics.

“Twenty or thirty years ago, biometrics was going to change the world, and you still don’t see a fingerprint reader anywhere… its very rarely used outside of high security applications. The technology will progress and well see it in some applications, so were seeing as I said biometrics in phones and consumer equipment, but whether that actually flows through to what the public would perceive as a more authoritarian use for security and policing applications, for example, is different question” [Participant 5, Security Expert].

Key Finding: AI development in Control technologies may encounter social and ethical challenges arising from potential misuse of personal data such as movement patterns.

Technological Risks

There are several technological risks associated with control systems, including misuse of data, manipulation of data made possible through weak system integrity and security vulnerabilities created by low accuracy and reliability of technologies. Control systems incorporating biometric technologies may have an increased risk of defeat or failure due to the high rate of false negatives and positives, in addition to the relative ease at which biometric information (e.g. fingerprints) can be copied or imitated to deceive access control systems.

Where access control and other systems are operated in the cloud, they may be subject to significant cyber security risks, and technical failures may occur as a result of internet or broadband interruptions, or though denial or withdrawal of cloud-based services. Autonomous failure of human override may be particularly problematic for control technologies, especially in the event functions such as safe egress are impacted, or movement of individuals is restricted.

AI failures may also be an issue for technologies where algorithms are unable to contextualise sufficient detail required to either trigger or reject an action or response – for instance, a system may reject access to in individual with an altered gait or crutches resulting from a sports injury, or fail to detect and prevent access to an individual wearing an explosives belt as the system is unable to contextualise the size of the person’s torso in relation to the rest of their body. These types of scenarios present significant operational limitations for control systems as contextualisation is imperative for the development of system intelligence. However for contextualisation to improve in this areas, significant development in datasets must occur first to provide systems with baseline data from which anomalies can be distinguished.

“Well I don’t think at this stage they’re able to say ‘well that person’s gait doesn’t look right, and if you take the basic dimensions of the body, it looks like they’ve got something unusual about their torso’ it’s not able to do that yet” [Participant 4, Security Expert].

Key Finding: Greater contextualisation is required to enhance control technologies, however for this to occur, significant development in datasets is required.

Legal Risks

Control technologies were found to carry the same legal risks as Detection technologies, with legal liability potentially resulting from perceived injustices, infringements on individual privacy (e.g. analytics applied to the movement of individuals and groups or biometric information), or adverse outcomes resulting from the use AI in control systems. Failures or errors may also result in legal damages or compensation, where manufacturers and vendors may purport that a system achieves control proficiencies beyond its actual capabilities. However, developing AI testing mechanisms and standards may reduce these risks by holding manufacturers and vendors to greater account in regard to system performance.

Consideration should also be provided to whether enrolment of vulnerable persons in public control systems should occur, and the potential privacy implications of retaining such information, as well as potential safety risks and legal ramifications associated with breaches of access to that data.

Key Finding: Consideration must be afforded to ensure sufficient protections are in place to ensure the privacy and security of vulnerable persons enrolled in control systems.

3.6.2.2.4 Response Technologies

The use of AI in response technologies was identified as having moderate benefits and extensive risks across all PESTEL classification areas as displayed in Table 50.

Table 50

Benefits and Risks of Artificial Intelligence in Response Technologies

	RESPONSE TECHNOLOGIES	BENEFITS	RISKS
P	**POLITICAL**		Political, cultural, social bias in machine learning Uncontrolled UAVs may ground commercial & emergency aircraft Civil liberty breaches from AI systems or devices physically engaging with humans Difficulty in attributing accountability & responsibility for AI systems & decisions Lack of transparency in AI decision making Reputational damage and loss of trust Privacy & human rights issues Political implications from the use or release of force
E	**ECONOMIC**	Improved allocation of resources Cost reduction Reduced operator workload and fatigue Increased operational capabilities	Low resilience and redundancy across systems and devices Cost burdens of purchase, installation & maintenance Reputational risk/ loss of consumer confidence

S	**SOCIAL**	Human-machine teaming to reduce safety risks Removal of humans from the front line Convenience	Safety risks & issues Uncontrolled UAVs may ground commercial & emergency aircraft Civil liberty breaches from AI systems or devices physically engaging with humans Privacy & human rights issues Fatal consequences of AI failures Social implications from the use or release of force Breaches of capable guardianship and duty of care Low public user/acceptance of AI Misuse of data Reputational risk (personal, corporate or government) Inappropriate training of personnel
T	**TECHNOLOGICAL**	Increased accuracy Faster, more focussed decision making and response Increased mobility & connectivity Breaking the Kill Chain	AI & technology failures Fatal consequences of AI failures Misguided AI systems not following programming Misuse of data Vulnerability from low accuracy/reliability of technology Complacency in operation of unmanned systems Limited counter drone technologies available commercially Exploitation of security vulnerabilities of IOT and Edge of Network devices Security/ IT divide Black Box of AI issues & risks Criminal exploitation of AI based on known vulnerabilities or responses Adversarial exploitation of AI Difficulty in reverse-engineering machine learning algorithms for transparency Low resilience and redundancy across systems and devices Inappropriate training of personnel
E	**ENVIRONMENTAL**		Uncontrolled UAVs may ground commercial & emergency aircraft
L	**LEGAL**	Reduction or elimination of safety risks	Absence of legal frameworks for counter drone applications/ response Non-compliance Legal liability (e.g. from system errors) Privacy & human rights issues Safety risks & issues Fatal consequences of AI failures Black Box of AI issues & risks Adversarial exploitation of AI Civil liberty breaches from AI systems or devices physically engaging with humans

Benefits

Moderate benefits were identified from AI enhancement in Response technologies. *Technological* benefits were based on increased accuracy, mobility and connectivity, as well as more focussed decision making and response capabilities. The capacity to interrupt and stop an attack vector by an automated or autonomous response could produce immense benefits in high-risk environments, provided the response is appropriate and measured. *Social*, *Economic* and *Legal* benefits primarily related to reduced costs and improved safety - participants frequently identified the removal of humans from the front line, or from dangerous situations and environments as one of the major benefits of AI in response technologies.

Key Finding: Response technologies may provide a considerable benefit where humans can be removed from the front line.

"The major one for me, is it has the capability to remove humans from positions or situations of danger. It also has the ability to break the kill chain… so if you have an automated response that's working well, you can stop an attack vector via an automatic response" [Participant 4, Security Expert].

"I think the real benefits are around removing the danger for a person responding, but giving them some type of semi or fully automated avatar of some kind" [Participant 5, Security Expert].

Risks

Risks were identified across all PESTEL categories for Response Technologies, with the most extensive risks relating to *Political*, *Social*, *Technological* and *Legal* factors.

Political Risks

There are significant political risks associated with the design, development and deployment of AI-enabled response technologies, principally in respect to drones, robotics and weapons systems. The capacity for the use or release of force from Unmanned Aerial Vehicles (UAVs), drone swarms, robotics or Autonomous Weapons Systems (AWS) is a contentious political issue, and while global regulatory frameworks to restrict their development and deployment is limited, a number of individuals and organisations have instigated action to ban or control the use of Lethal Autonomous Weapons Systems (LAWS) (Campaign to Stop Killer Robots, 2019; Lethal AWS, n.d.) which has the potential to result in profound limitations for the development or use of such technologies.

Autonomous use or release of force by machines is subject to extensive moral, ethical and human rights considerations, guided by questions as to whether machines should have the power to determine life and death, or inflict injury or harm onto humans. Legality of autonomous release of force is also a contested issue, highlighting the limitations of machines to exercise judgement, apply meaningful control and comply with rules of engagement, International Humanitarian Law (IHL), as well as adhere to principles of war such as proportionality (Heyns, 2016; Sharkey, 2016). The capacity for humans to remain 'in the loop' so as to override the system is also incredibly important – both conceptually and operationally – and has a profound impact on the autonomous development of response technologies. Though these issues are certainly more prominent in the military domain, the basic tenets apply to all response technologies and devices that may have the capacity to apply or use force against humans. The development of machine or robotic response forces to engage with humans on any level is both a problematic and complex endeavour.

Key Finding: Autonomous use or release of force presents extensive moral, ethical and human rights considerations.

> *"… and so now there are systems in place to make sure that the presence of the human always overrides the other directions that the robot may have. And I think in security, we have to have similar rules where it has to be clear as to who the robot or the system is working for, and the people that it's working for really should always have that ability to override when necessary… so I personally think it's always risky if machines are going to be totally autonomous" [Participant 3, Security Expert].*

In addition to those risks associated with autonomous use or release of force, there may be further political risks arising from black or grey market acquisition and trade of response devices. For example, commercial markets flooded with unregistered drones or drone equipment and parts, may inadvertently drive black market development and acquisition of unregistered UAVs. This may escalate illegal trade and impact on how drone technologies are able to be controlled, essentially hampering current efforts to regulate drone use by mandatory registration of drones and pilots. Black or grey market trade of any response devices, or device components, may also have the potential to facilitate adversarial or terrorist activity or obstruct law enforcement efforts to neutralise resulting threats. These are critical risks for which governance and regulatory frameworks must be developed to address as a matter of priority.

Key Finding: Urgent attention must be afforded to the development of regulatory frameworks to minimise the risks of back or grey market trade of response technologies.

> *"… it's interesting that in the States now this whole issue of drones is becoming a real big worry for people because you've got this situation where the use of drones is almost uncontrolled… there are pieces of legislation around but it doesn't help and I'll go back to the Gatwick Airport incident of a couple of years ago… they still don't know for sure if there was ever actually a drone, and if there was, was it one or was it more than one?" [Participant 4, Security Expert].*

<u>Economic Risks</u>

Economic risks to corporate entities primarily consisted of cost burdens relating to purchase, installation and maintenance. Other risks related to economic consequences arising from reputational damage, or legislative changes restricting use of technologies after corporations may have already invested extensive sums of money in purchasing response devices to build autonomous response forces.

At national and international levels, economic risks may arise from uncontrolled use of technologies such as drones in the event civil aviation airlines are grounded. Participants highlighted the potential for economic disruption which may result from uncontrolled drone use grounding airlines, providing the example of Gatwick Airport in the United Kingdom being closed due to drones endangering the safe operation of aircraft. The economic flow on effects of this type of disruption could be extensive, affecting numerous industries beyond aviation.

"There's the reputational risk aspect, and the actual causing injury or harm risk if the product doesn't work as intended" [Participant 6, Security Expert].

"They shut down Gatwick Airport for about 48 hours – Gatwick Airport being the second busiest airport in the whole of the UK – over Christmas" [Participant 4, Security Expert].

<u>Social Risks</u>

Participants identified considerable social risks from the use of AI in response technologies. As previously identified for observation technologies, privacy issues arising from surveillance may be problematic for drones or UAVs which either incorporate surveillance devices or use computer vision to assess and respond to threats. The development and use of machines such as drones may also create additional social risks, such as facilitating the perpetration of crimes which may weigh heavily as a social risk. Criminal use of AI-enabled response technologies is already occurring, for instance with drug cartels in Mexico using drones with pre-programmed GPS coordinates to send their goods into the United States (Ovchinsky, 2018 as cited in Sukhodolov et al, 2020). Social risks arising from criminal or adversarial exploitation of technologies must therefore be afforded consideration as to how society may wish to deal with negative social outcomes, or more importantly whether it will be possible to control the technologies responsible for negative social outcomes once they have been widely distributed.

The use of AI to create 'smarter' and more autonomous response technologies or devices may encounter significant social resistance, and this may be particularly challenging for the security industry. According to Fischer & Green (1998, p.3) the term *Security* 'implies a stable, relatively predictable environment in which an individual or group may pursue its ends without disruption or harm, or without fear of disturbance or injury'. To achieve a secure, stable environment, security also implies some form of response to a threat, and with that response, some form of interaction with people. However, a security response must be able to interact responsibly and in socially acceptable ways with the humans it protects, as well as the humans which are perceived to threaten security. Therefore the design, development and deployment of response technologies must prevent any infringement on the rights of individuals or groups. At this point in time, no fully autonomous machine has been developed with a capability to provide a measured, socially acceptable response in which it can irrefutably guarantee the human rights of the individuals in which the response is directed. The essence of this notion is that implicit and explicit compliance with societal constructs is currently beyond the capabilities of any machine.

Key Finding: Implicit and explicit compliance with societal constructs is currently beyond the capabilities of any machine.

Technological Risks

There were numerous technological risks of response technologies identified by participants, which may broadly be categorised as security vulnerabilities; technological failures; transparency issues; and failures of human operation and management.

Security vulnerabilities largely refer to cyber risks of connected systems and devices, the misuse of data, or exploitation of AI resulting from criminals or adversaries being able to predict or anticipate a machine's responses. There may be instances where adversaries are able to evade or defeat response systems by analysing the responses generated by various algorithms, especially where algorithm codes are open source and available publicly, or where systems are able to be purchased or acquired by adversaries who may then determine the techniques or methods required to defeat such systems.

"… remember, we're not the only people that know about these things, because someone will put the whole lot on the internet, and the guy who's going to be the active shooter knows that if he does a, b, and c, the building's going to be locked down and he's inside with a gun' [Participant 9, Security Expert].

"… when they first started screening people on planes, not in Australia but in the Middle East, the first of the new x-ray systems were always the best and were always bought by the PLA (Palestinian Liberation Army)… they'd buy the first x-ray and then they'd work out how to defeat it' [Participant 9, Security Expert].

Additional security risks may arise from systems and technologies which have not previously been anticipated or encountered and therefore may expose individuals, corporate entities, or governments to novel security risks. For instance, in an *Image Reconstruction Attack*, data from mobile robots may be captured and reconstructed to provide adversaries with relatively accurate images of what the robot was observing (Chelani et al, 2021; Oh & Lee, 2019). This type of technique could be used by adversaries for intelligence or reconnaissance missions, and may have considerable implications for military or high security environments.

Key Finding: Technology users must be aware that novel security vulnerabilities may arise through the development and deployment of innovative AI technologies.

"… there's a whole lot of data and privacy issues, because these things learn on massive amounts of data… so you develop a robot with some abilities to do something, but when you bring it into your house, and you want it to do a job that you're interested in, you might have to teach it various things, you might have to show it things it hasn't seen before and so on… but as some researchers have kind of demonstrated at least in principle, there's a shocking amount of information that you can get out of that, you know – people can recreate what your bedroom looks like" [Participant 11, AI Expert].

Technological failures are a considerable risk of AI and may include tangible or measurable outcomes such as disproportionate response, or may be more difficult to quantify such as machine learning algorithms failing to 'learn' in the intended way. Lack of transparency in these systems, often referred to as the 'Black Box of AI' increase technical complexity, and may result in more difficulty in resolving technical issues. Where AI algorithms are become rogue or misguided, difficulty in reverse-engineering these systems may further decrease transparency which may be problematic where AI failures have caused safety hazards or injuries.

> *"Disproportionate response – that's a pretty big risk. You know if the system inadequately addresses the threat, and then provides a disproportionate response, and someone gets hurt or killed, then that's a major risk" [Participant 6, Security Expert].*

Incorrect or inappropriate training of personnel may have serious ramifications for effective management and use of response technologies. Complacency in operation of unmanned systems may become an issue, particularly where response technologies may have the capacity to inflict harm, such as autonomous weapons systems or even robotic systems which may unintentionally injure someone as a result of complacent operation. Similarly, complacency where human life is *not* at risk may also contribute to a disproportionate response and potentially risk non-human assets. Protocols for good management practices should explicitly define organisational responsibilities to prevent issues emerging and contributing to technological failures.

Key Finding: Protocols for good management practices should explicitly define organisational responsibilities in relation to AI.

> *"if you don't have human life at risk, you might miscalculate the response… or a potential complacency of response, because you don't really mind if that machine gets ruined, but you don't know what else does" [Participant 2, Security Expert].*

Environmental Risks

Identification of environmental risks were limited, with participants recognising the impact of UAV and drone use on commercial aviation and emergency response. The grounding of emergency or humanitarian response aircraft as a result of drone use (legal or illegal) may inhibit firefighting or other response efforts which may be catastrophic in terms of environmental consequences as well as pose unacceptable risks to human life.

Legal Risks

There are several novel legal issues and risks to emerge within the category of security response technologies. Considerable progress is required in the development of legal and regulatory frameworks to address criminal prosecution and legal liability resulting from AI crimes. Mohsin (2020) provides insight as to the current status and future challenges which may be expected as development of AI continues;

> There is no regulatory framework at the moment which targets AI-crimes, majority of the stakeholder's policy recommendations are targeted towards fairness, accountability, and transparency in machine learning or they target the R&D of robots, none of the legal texts cover the much-debated topic of criminal liability of an AI as the law is still undecided. The traditional method of regulation of AI has been by, product licensing, research and development oversight and tort liability. However, the current AI principles are not enough to determine the threat posed by AI acting autonomously in the future (Mohsin, 2020, p.1).

These concerns were echoed by participants raising issues of how criminal risks may be mitigated, and how criminal acts may be investigated and prosecuted. For instance, how may the legal construct of 'intent' be established in the event a system is hacked and the algorithms are modified, resulting in the machine learning process being manipulated or deviating from the original programming. At what point will the legal system be prepared for that level of forensic detail, particularly when there is currently limited understanding of how machine learning algorithms make decisions. Criminal exploitation of AI may further complicate our ability to 'deconstruct' an AI crime scene, as well as follow established judicial processes to achieve legitimate and just outcomes for victims of crime.

"Crime related risk… and that's sort of an encapsulation of both the system being weaponised without you knowing or compromised in some sort of way with ill intent" [Participant 2, Security Expert].

"How do you prove criminal involvement? Let's say a drone responds to a burglar and accidently flies into him and kills him, but the system responded – who wrote the rules? How did that AI make that decision? You don't know how it came to make a decision because you don't know how it learned what it did" [Participant 2, Security Expert].

The legal ramifications of machine-based response from drones or robotics for instance raised issues of legal liability for the use of force, or the use of lethal force in the event of a fatality. For instance, if a robot uses excessive force to apprehend an intruder and causes a fatal injury, then who is liable for the injury and death? Should liability be attributed to the programmer, the software developer, or the entity who is physically responsible for the robot? Again, and as Mohsin (2020) suggests, we have not established the necessary legal frameworks required to sufficiently administer or apportion responsibility and liability for AI response technologies. Furthermore, as progress in made in this area, there may be disparity between countries in how legislative frameworks are approached - some countries may lack of governance and regulation, while others may overregulate and essentially suppress technological development. Both scenarios may be problematic in their own right.

Key Finding: We have not yet established the necessary legal frameworks required to sufficiently administer or apportion responsibility and liability for AI response technologies.

"Civil liberty issues associated with programmed objects or response technologies – apprehending or physically engaging with humans will be an issue of hot debate, because you know someone will potentially cry poor and say that robot assaulted me and there was a misuse of force there and I've now sprained my ankle or what have you and then someone got to be responsible for that… so who's the primary responsible entity in that sociotechnical system?" [Participant 6, Security Expert].

"I think some States would actually legislate to control these systems – the types and degrees – things like detention by robots, or holding people through autonomous systems… so you're not actually allowed to use a robotic system to actually detain someone, and then you've put all these robotics out in the area and they're not allowed to do anything bar follow someone, and they're only allowed to follow on your own land, they're not allowed to go into public lands or other areas" [Participant 1, Security Expert].

Perhaps one of the more understated risks identified for response technologies is duty of care issues, which relate very specifically to the security industry. Security response implies that a capable guardian will provide a duty of care to victims, perpetrators, employees, bystanders or other individuals who may be present. Arguably, an autonomous response technology must be able to provide capable guardianship and ensure legal duty of care obligations are met. There are no clear methodologies or precedents in legal discourse to govern how the establishment of capable guardianship may be tested or applied to non-human autonomous responders. Until response technologies endure this potentially extensive process, there may be a very real possibility that autonomous response technologies do not receive the legal approval necessary for corporate entities to invest in.

Key Finding: Autonomous response systems and devices may breach legal duty of care obligations.

"… the nature of the security industry means that you must still have a capable guardian factored into the security equation, otherwise by definition, you largely don't achieve a secure environment" [Participant 6, Security Expert].

"There are compliance risks as well, because if you look at the way legislation is written, in some cases for the people responding with a duty of care – there is a reason security officers have very sophisticated standard operating procedures, and they have a way to make sure that they and the customer don't expose themselves to liability. So if you have an autonomous system that would have to go through the same process" [Participant 5, Security Expert].

"Crime will exist if you don't have a capable guardian… I'm not sure whether the security industry or our clients would ever accept than an advanced sentient AI is the capable guardian, I think it would always have to be a human operator" [Participant 6, Security Expert].

3.6.2.2.5 Overarching Risks for Integrated Security Technologies

The use of AI in integrated security technologies was identified as having extensive risks across all PESTEL classification areas as displayed in Table 51. These risks build upon all previously identified risks across Observe, Detect, Control and Response Technologies, and are subsequently considered overarching risks for the use of AI in security technologies.

Table 51

Overarching Risks of Artificial Intelligence in Integrated Security Technologies

	INTEGRATED SECURITY TECHNOLOGIES	**OVERARCHING RISKS**
P	**POLITICAL**	Political, cultural, social bias in machine learning Difficulty in attributing accountability & responsibility for AI systems & decisions Lack of transparency in AI decision making Reputational damage and loss of trust Privacy, ethical & human rights issues Foreign espionage The weaponization of AI Privatisation of military AI Industrial relations issues resulting from use, misuse or failures of AI Regulatory risks & implications (i.e. overregulation, industry self-regulation) Criminal exploitation of AI Unauthorised use of data from social media platforms Disproportionate development between countries Political divide between developing and developed nations Exploitation of underdeveloped countries (experimental AI)
E	**ECONOMIC**	Low resilience and redundancy across systems & devices Cost burdens of purchase, installation & maintenance Reputational risk/ loss of consumer confidence Industrial relations issues resulting from use, misuse or failures of AI Economic costs of disruption/ downtime Economic costs of security and AI decay (e.g. upgrades, security patches, etc.) Reliance on the cyber domain for design & development of security technologies
S	**SOCIAL**	Absence of legal and political frameworks to govern the development and deployment of AI in socially acceptable ways Safety risks & issues Privacy, ethical & human rights issues Lack of moral and emotional intelligence in AI Public discontent concerning government use of AI Low public/ user acceptance of AI Social fear/distrust of AI Misuse of data Reputational risk (personal, corporate or government) Low consumer tolerance of retrospective remedies Criminal exploitation of AI Unauthorised use of data from social media platforms Exponential development may inhibit social responsibility Social risks of inability to revoke AI development AI vulnerability to defeat by human adversaries Lack of red teaming capability to reduce adversarial threats
T	**TECHNOLOGICAL**	AI & technology failures Misuse of data

		Security/ IT divide Black Box of AI issues & risks Criminal & adversarial exploitation of AI Difficulty in reverse-engineering machine learning algorithms for transparency Low resilience and redundancy across systems and devices Privatisation of military AI Dangers of uncontrolled AI Cyber & cloud-based risks Hacking/ hijacking or modification of script Forensic auditing & assessment Complexity of systems Reliance on the cyber domain for design & development of security technologies Reduced physical response force for critical incidents Neglect of mission critical functions Potentially limited capacity for 'untraining' and 'retraining' machine learning algorithms Data integrity, control and management issues Regulatory risks & implications (i.e. overregulation, industry self-regulation) Limited source data
E	**ENVIRONMENTAL**	Environmental consequences of AI use, misuse or failure
L	**LEGAL**	Legal liability (e.g. from system errors) Regulatory risks & implications (i.e. overregulation, industry self-regulation) Privacy & human rights issues Safety risks & issues Security/ IT divide Industrial relations issues resulting from use, misuse or failures of AI Lack of security governance/ oversight Forensic auditing and assessment Inadequate human assessment resulting from reliance on AI systems Reduced physical response force for critical incidents Limited regulatory frameworks for source data (e.g. databases with image classifications) Legal limitations on acquisition of source data Retrospective legal action

Overarching Risks

There were a number of significant overarching risks of AI which emerged across all categories of Political, Economic, Social, Technological, Environmental and Legal factors in the PESTEL framework.

Overarching Political Risks

The politicisation and weaponization of AI emerged as a significant overarching risk, with the use and exploitation of AI technologies and algorithms by militaries or by foreign governments raising serious concerns. The lack of governance and regulatory controls in addition to the lack of transparency in processes and protocols are currently fuelling an 'AI Arms Race' which has created considerable controversy, particularly in relation to the design and development of Autonomous Weapons Systems (AWS) and Lethal Autonomous Weapons Systems (LAWS). Military advantages created by AI not only affect the political balance of power, but may create significant ethical and human rights issues in regions and nations which are not able to accelerate development at the same rate as global superpowers. Consequently, disparity in AI development between nations may facilitate the political, military or economic exploitation of underdeveloped regions or nations.

Key Finding: Disparity in AI development between nations may disrupt the balance of power, and facilitate the political, military or economic exploitation of underdeveloped regions or nations.

"Technology is power" [Participant 11, AI Expert].

"Those who don't advance at the same rate will be disadvantaged…it could be in two ways, the first one just could be the fact that they don't benefit from the opportunities or efficiency advantage that this sort of technology would bring… the other disadvantage is where an outside power who has that sort of technology, who then bring that technology into the poor or not as advanced and use that technology to their advantage… that's already happening in Africa" [Participant 4, Security Expert].

"This sort of technology will become very powerful because of the manner in which the data is not only used but harvested as well… if you use Zambia as an example, so the main airport in Zambia, the main power generated in Zambia and the television station in Zambia are all now owned by the Chinese. Because they gave the Zambian government loans, knowing full well that they would never be able to repay them, and when it came time to call in the loan, they took the collateral which they'd been given. The security was yeah you can have the airport, you can have this, and they now own it all. If you look at Ghana the Chinese own the two major ports in Ghana on exactly the same principal" [Participant 4, Security Expert].

"…and yeah, we get back to the whole Cambridge Analytica discussion that we had previously, and you think about those sorts of concepts that were present there… you then add that to what we've just discussed with a more advanced entity bringing this technology to a less advanced entity on the promise of improved efficiencies for example, however there is another side to the equation" [Participant 4, Security Expert].

The political risks of AI are potentially so detrimental to developing nations that development of global governance frameworks must be given urgent and absolute priority. As acceleration of intelligent autonomy will be heavily influenced by government structures, it is likely to create a political divide between the developed and developing worlds. Therefore, world innovators and leaders may need to consider how such a divide will impact on the developing world, and 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 and exploitation to be driven by external forces. The provision for technological aid should not be considered a burden, but rather a moral responsibility which rests on the shoulders of all who endorse and advance the AI and technological revolution.

Key Recommendation: 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.

However, care must be taken to ensure technological aid for developing nations is provided within such a way that sensitive technologies, such as weapons systems, are not corrupted to facilitate criminal or terrorist objectives. Determining ways in which such obligations may be met is another matter well beyond the scope of this research.

Authoritarian control through the use of AI in security technologies is another key political risk. As these technologies become more prevalent, and their use more widespread, there may be an increased risk of governments, law enforcement or other agencies using technologies to exert control over their citizens. While the risk of this may be seemingly lower in countries such as the USA, Australia and Europe, complacency in relation to political transparency should not be ignored. Ethical checks and balances in both the public and private sectors may therefore require systematic review to ensure the use of AI to monitor, detect, control, or respond to security events remains aligned with public notions of what is socially and ethically acceptable.

Key Finding: Ethical checks and balances may require systematic review to ensure the use of AI remains aligned with public notions of what is socially and ethically acceptable.

"I have spoken to a number of people who live in and around that city… They're absolutely terrified by the whole thing, but if you speak up against the government you disappear. So it's not 'Yes the people are accepting of it', they don't really have a choice" [Participant 8, Security Expert].

"… they have one set of rules for citizens, and they have another set of rules for foreigners. So foreign visitors don't see a lot of this, but when you get to know the citizens, you realise how paranoid they are about doing something wrong which the authorities won't be happy with… They did that when I was there once where they had a group of teenage girls who were sitting at a fast-food restaurant just doing what teenage girls do and just chatting to each other. And the manager of the fast-food restaurant had complained to the police that these girls had been sitting there for too long and the police arrested them and took them to prison and sat them in a cell for a couple of hours and told them they hadn't committed a crime yet, but this is where they'll be living if they do" [Participant 3, Security Expert].

Overarching Economic Risks

The overarching Economic Risks of AI largely reflect the themes which emerged for Observe, Detect, Control and Response technologies, with various cost burdens associated with purchase, installation and management of AI systems or disruption and downtime resulting from their use. Loss of consumer confidence may have broader consequences for corporations and product manufacturers as reputations may be more challenging to restore. Reliance on AI systems may create significant operational consequences for organisations in the event AI systems are inoperable or suspended from use.

Furthermore, regulatory changes affecting the use of AI, for instance legislative changes resulting from safety issues or social discontent, may have significant economic impacts on organisations which have implemented AI technologies. Industrial relations issues may also surface in response to the adoption of automated or autonomous systems, particularly if the use of smart systems result in employment functions becoming obsolete.

Overarching Social Risks

One of the most prominent overarching social risks is the fear or distrust of AI and the potential for loss of control over technologies which may cause, or be perceived to cause, harm to the individual. The inability to revoke technologies once they have been deployed in the public sphere was a frequently cited concern raised by participants. For some technologies such as surveillance systems, development of AI is likely to be constrained in those technologies which are perceived to infringe on privacy and human rights, or where their use is not considered to provide an overall benefit to society. This may be particularly evident in individualist countries such as the USA and Europe where individual rights are emphasised and both government and corporate conduct is able to be widely debated, publicly protested and where necessary, overturned.

Key Finding: Deployment of AI is likely to be constrained in those technologies which are perceived to infringe on privacy and human rights, or where their use is not considered to provide an overall benefit to society.

"The developed nations and their paranoia about privacy may well become an issue for them because the whole power of the use of data will be overridden by this paranoia about privacy… it will make it more difficult for developers in those nations to move it forward at the same speed as the less constrained regions of the world like China, Russia, North Korea or some of the Asian countries" [Participant 4, Security Expert].

"We want to develop and implement technology and AI sentience in a socially responsible way, that is transparent, that has been debated in public, it has been endorsed by our politicians, and it has an appropriate legal and governance frameworks wrapped around it so that it doesn't backfire and start picking up guns and shooting us" [Participant 6, Security Expert].

The absence of legal and political frameworks to govern the design, development and deployment of AI in socially acceptable ways currently drives social concerns. Retrospective action is not a viable social solution for risks that arise from AI technologies, therefore manufacturers may need to ensure technologies are safe and socially acceptable prior to commercial development and deployment. Legislative frameworks must also be developed to afford protection to citizens and technology consumers, and to hold accountable those who breach the frameworks designed to protect the rights of individuals. This extends to behavioural and personal profiling from information collected from social media platforms for the purposes of consumer marketing, political campaigning, or psychological operations, as seen in the Cambridge Analytica scandal in which the personal data of Facebook users was allegedly harvested to influence the 2016 US presidential election and the UK Brexit campaign (Cadwalladr & Graham-Harrison, 2018).

Key Recommendation: Legislative frameworks must be developed to afford protection to citizens and technology consumers, and to hold accountable those who breach the frameworks designed to protect individual's rights.

> *"... now we're moving into the Wild West of AI and autonomous systems... we're investing in this technology, we want to get it out to the market and sell it, and well look at the safety and the security risks afterwards... let us know if it hurts you and we'll fix it" [Participant 2, Security Expert].*

Overarching Technological Risks

There were two predominant overarching technological risks for integrated security technologies, *security vulnerabilities* and *transparency issues*. The types of security vulnerabilities such as hacking, hijacking, manipulation or defeat of AI systems has largely been covered for each of the security technology categories of Observe, Detect, Control and Respond. In essence, these systems are insecure and the risk of the vulnerabilities of connected devices being exploited means that consideration may need to be afforded to all AI-technologies prior to implementation and use, even more so for systems being deployed in the cloud. Subsequently AI use should be assessed at a security vulnerability level and linked to criticality and consequences of risk. In a practical sense, this may mean development of ways in which integrated technologies can be segregated by their vulnerability level and subsequent deployment into less secure environments (such as the cloud) may be applied to non-critical security functions only. The security industry must use a needs-driven approach to shaping the way in which security technologies are developed and deployed, rather than trying to adapt security needs to solutions which were never designed with security in mind.

Key Finding: The security industry must take the lead in shaping the design of AI technologies to meet security needs.

In terms of transparency issues - often referred to as the 'Black Box of AI' – this is a major overarching technological risk for AI use in all technology categories including integrated systems. The security industry must consider that if transparency is an issue for Observe, Detect, Control and Respond technologies individually, transparency issues will certainly be compounded for integrated and aggregated technologies. The more AI layers which are overlayed to integrated systems, the more difficulty there may be in unravelling what has occurred within the black box. In essence, we are developing AI systems for deployment which will exert at least some level of control over critical security functions, and we do not fully understand how they operate – we do not know with any certainty how decisions are being made, and as we are unable to forensically audit such systems to evaluate and assess the algorithms, we cannot guarantee the safety, accuracy, and reliability of those systems.

Key Recommendation: An AI security vulnerability and criticality assessment should be developed to guide deployment of AI systems in cloud-based environments.

Overarching Environmental Risks

Environmental risks were not highly regarded by expert participants, though the potential for environmental consequences through 'miscalculation' of AI algorithms was identified as a potential risk. The possibility that this may occur generally relates to AI and technical failures of security technologies in certain environments such as chemical plants, nuclear plants etc, rather than being a risk directly resulting from the use of the technology. While there may in fact be environmental consequences resulting from the manufacture and use of computers or other equipment required to house AI algorithms (e.g. environmental waste issues), these risks were not explicitly identified by participants. Similarly, environmental impacts from the use of AI-enabled technologies such as weapons systems, robotics or drones (e.g. destruction of vegetation, impact on wildlife etc.), these risks were not explicitly identified.

"In an operational environment, might be a chemical plant… and if that response is wrong, " [Participant 2, Security Expert].

Overarching Legal Risks

Legal liability was the most significant legal risk identified by participants. Liability was perceived to arise in response to privacy and human rights issues, as a result of injury or harm from inadequate safety protocols, or as a result of AI or technological failures. However liability may be further compounded by the issue of underdeveloped legal and regulatory frameworks, resulting in the attribution of responsibility and accountability for AI outcomes being fraught with complexity. We are yet to establish methods to forensically audit machine learning algorithms in a way that meets the requirements of established legal procedures. Therefore the overarching legal risk to users of AI in the security domain is simply *uncertainty*. From a legal perspective, *uncertainty* as to how, when and why legal liability may arise and the extent to which they may be liable. This may be an unacceptable level of risk and one which may mean that adoption of AI is met with extreme caution.

Key Finding: *Uncertainty* is the overarching legal risk to users of AI in security.

"It's all very well for us to talk about AI, but the first time that somebody comes up with an AI system and somebody gets killed by it – and remember that's already happened with Tesla a number of times – no care or no responsibility… so people will say well I'm not going to put AI into our systems unless I know it's going to work… So somebody else has to go and do a load of research, and pay a load of money to get it built, then you have to get the computational power to be able to make these complex decisions, and then if it gets it wrong and someone goes down the gurgler then we're back to where we started from" [Participant 9, Security Expert].

3.7 The Future of Artificial Intelligence in Security Technologies

3.7.1 Objective Four

3.7.1.1 Expected Development of AI in Security Technologies

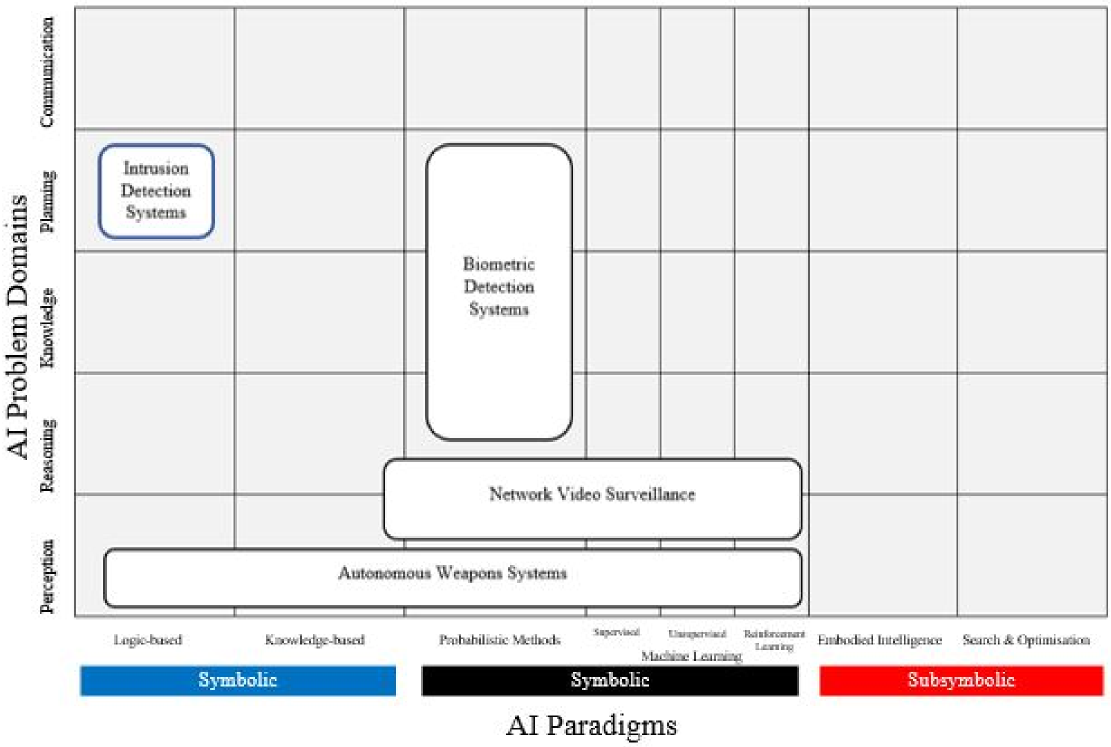
Objective Four explored the expected development of AI in security technologies, seeking participants' insights as to how they believe security technologies may change in the next ten years. Objective Four was facilitated with the following sub-question presented to participants:

a) Do you feel the location of security technologies on the *Security Technologies AI Knowledge Map* will change considerably in the next ten years?

Expert participants were presented with the *Security Technologies AI Knowledge Map* (Figure 28) with the expectation that they would indicate, either verbally or diagrammatically, where the location of the technologies depicted on the map (i.e. *Intrusion Detection Systems*, *Biometric Detection Systems*, *Network Video Surveillance* and *Autonomous Weapons Systems*) may be located in ten years' time.

Figure 28

The Security Technologies AI Knowledge Map.



Despite the *Security Technologies AI Knowledge Map* (Figure 28) being included in the information document provided to participants prior to the interview, the majority of participants did not provide insights as to how specific technologies may align with the AI Knowledge Map (Corea, 2019). Instead, participants provided broad insights in respect to the extent to which they believe security technologies will change as a result of AI, and the factors which will both drive and inhibit such development.

3.7.1.2 Objective Four Findings

Expert participants provided broad insights as to whether they believe security technologies will change significantly over the next ten years with the incorporation and expansion of AI, how security technologies may change as a result, and the factors which may drive or inhibit this development. Analysis of participant insights are discussed in three main areas; *Expected Areas of AI Development in Security Technologies*, *Expected Drivers of AI Development in Security Technologies* and *PESTEL Factors Expected to Influence AI Development*.

3.7.1.2.1 Expected Areas of AI Development in Security Technologies

In broad terms, expert participants do not anticipate a significant degree of change in how AI will be deployed in security technologies over the next ten years. Participants indicated they expect security technologies – particularly surveillance technologies – and integration of those technologies to improve in their sophistication, but do not anticipate any significant change in fundamental operation of those technologies as a result of AI development.

Key Finding: Experts expect security technologies will become more sophisticated, but do not anticipate those technologies will significantly change in the next ten years.

"No I don't think so. I think we'll see more sophisticated applications of the same technology. I don't think there's going to be a substantive shift... I think the capabilities will be more reliable, but I don't think we're suddenly going to see fully autonomous response teams or anything like that. I think it's going to be just a continual integration of the technologies into the current security workflows" [Participant 5, Security Expert].

"There's going to be no real shift... I think it will be pretty standard for the next thirty years" [Participant 1, Security Expert].

"I think surveillance technologies will change a lot… in the next 10 years for the rest of them, no not much" [Participant 3, Security Expert].

“Yeah, but I don’t think it’ll be a quantum leap. I think it will be a largely linear progression over time, and I don’t think that linear curve is very steep, I think it’s quite flat” [Participant 6, Security Expert].

“Yeah, massive changes – we’re about to see the disruption of the alarm monitoring sector. We’re about to see the advent of robotic and drone response. We’re about to see satellite integration to other technologies that we’re doing” [Participant 2, Security Expert].

However, participants did indicate there may be significant changes to the security industry as technologies develop. Connectivity of systems and devices are expected to be enhanced through the use of satellites, which will significantly expand the number of connected sensors and devices. Virtual assistants such as Alexa may see an increase in self-monitored security solutions and consequently a decreased need for alarm monitoring centres. Participants also indicated the use of drones and robotics are likely to increase, though as these systems are do not currently possess any capacity for advanced decision making, their use in the next ten years is likely be limited to surveillance missions or as an extension of human instruction.

“So the central system will say ‘that’s abnormal behaviour from that person, robot go out and intercept that person or detain that person’ and it’s the sum of those systems working together… I don’t believe the robot will have the ability necessary to make that sort of higher-level decision to detain or not detain… certainly not in the next 10 years anyway” [Participant 1, Security Expert].

Key Finding: Experts predict transformation within the security industry, such as an increase in the use of drones and robotics, and a decrease in the need for alarm monitoring centres.

3.7.1.2.2 Expected Drivers of AI Development in Security Technologies

Three technical factors are expected to directly and explicitly drive the development of AI in security technologies; *Quantum Computing*, *Crossover of Military Technology*, and *Enhancement of Narrow AI Techniques*.

Quantum computing is expected to be an explicit driver of AI, though there was no unanimous agreement among participants as to when they believe quantum computing will be developed sufficiently for widespread commercial use. Built from theories of quantum phenomena, quantum computers employ quantum bits – ‘qubits’ – which can be in a 0 *and* 1 quantum state simultaneously as opposed to current computers which store information as either a binary 0 or 1 state (Cho, 2020; IBM, n.d.a). Katwala (2020) explains ‘if you ask a normal computer to figure its way out of a maze, it will try every single branch in turn, ruling them all out individually until it finds the right one. A quantum computer can go down every path of the maze at once’. Quantum computing therefore has the potential to revolutionise modern technology, not only through speed and efficiency, but achieving far beyond what is currently possible with current computers. Participants indicated that quantum computing will be the key driver to facilitate machines reaching the point of singularity, and subsequently, achievement of *Artificial General Intelligence* will largely depend on the arrival of quantum computing.

> *"I don't think we'll have a significant change in technologies until we have a different technology for computers… So quantum computers are still a few years away… I think it'll be 50 years before quantum computing is really running everything… in 10 years' time you might start to see the first quantum computer, but it'll probably be the size of a 10-storey building or something" [Participant 3, Security Expert].*
>
> *"I don't think we're going to get to that stage in quantum computing in my lifetime… quantum computing over the last 20 years has gone from a theory to being able to entangle four bits between two locations… I don't think we're going to get there in my lifetime, but you know who knows" [Participant 8, Security Expert].*
>
> *"Once we start getting into quantum computing, all bets are off" [Participant 8, Security Expert].*
>
> *"So, if you can get a quantum computer to start applying machine learning, you can't even describe that computational power… this is where you'll get to the point of singularity" [Participant 2, Security Expert].*

Crossover of military technology to the commercial sector is also expected to drive the development of AI in security technologies. In respect to AI development, military and security should not be viewed as completely separate domains due to the porous nature of military-commercial exchange, in addition to the global increase of military privatisation. The majority of experts did not view the military domain as more advanced, and in fact identified some areas of AI such as autonomous vehicles as being behind private sector development. However, as the private sector may face greater limitations in deploying AI technologies, the testing or piloting of certain systems may occur more quickly within military research, and subsequently drive certain applications of AI development in the commercial sector. Already, the military and security domains are relatively enmeshed, and AI exchange is likely to create a finer, more fluid line between the technologies in these domains.

Enhancement of Narrow AI techniques is also likely to drive development in security technologies. The systems that operate on the basis of performing specific tasks (i.e. Narrow AI) are likely to dominate commercial progress due to the higher rates of accuracy and reliability that will result from AI enhancement. Where the decision is logical or probabilistic, AI will drive development of these specific technologies more quickly than applications where more complex decision making is required. For instance, probabilistic technologies such as biometrics, acoustic detection, or trace detection systems assess the probability that the features in an image are likely to be a specific person, an acoustic signal is likely to be a specific gunshot, or the chemical compounds in a trace detection sample are likely to be those of a particular explosive. AI is likely to enhance accuracy of these types of technologies, particularly because the decision is essentially a 'yes' or 'no', i.e. within a certain probability, either it matches, or it doesn't. Therefore enhanced accuracy and reliability will drive development of technologies where there is greater certainty in decisions, enabling technologies to reduce or replace the need for humans.

> *"… the detection of explosive and things like that, the machines are good, and they really can eliminate the human factor or replace him… its 1 or 0, I detect or not, and if there's any suspicion, I send it to a sniffer - another machine - and then if you get 'yes' 'yes', you go to the human factor" [Participant 7, Security Expert].*

Key Finding: Quantum Computing, Military Technology and Narrow AI Techniques are expected to drive AI development in security technologies.

3.7.1.2.3 PESTEL Factors Expected to Influence AI Development

Several PESTEL factors are expected to influence and *shape* – rather than explicitly drive - the adoption of AI. How the adoption or repudiation of AI will be shaped depends largely on the *Political*, *Economic*, *Social* and *Legal* structures of the countries, regions or cultures in which organisations or entities reside. Such factors are expected to either promote or inhibit development of AI, though this process may be fluid and highly responsive to dynamic conditions.

In countries and regions considered to be more individualistic in culture, and in those which are inclined to encourage free and fair debate may see AI development restricted through demands for governance and regulation to ensure the safety, privacy and rights of individuals. The development of legal frameworks to determine how responsibility and liability for AI failures may be attributed is likely to inhibit adoption of some AI technologies, particularly those that are capable of making decisions in any type of autonomous manner or where there may be ambiguity as to 'right' or 'wrong' outcomes. However, until the decisions made by machine learning algorithms can be understood and reverse engineered in such a way that forensic evaluation and assessment is possible within established legal procedures, then attribution of responsibility and liability may be largely unattainable.

> *"I think it'll be held back for a whole host of reasons. There's going to be privacy reasons, it's going to be legal liability reasons, it's going to be the inability to get insurance reasons, and then just waiting for the government of any persuasion to provide a framework for it to be utilised on their buildings in their technology solutions… I can think of legislation that's been waiting for 10 years… and unless it elicits winning votes it's not going to get there. So just to provide the regulatory framework it's going to be years away" [Participant 9, Security Expert].*

Key Finding: Demand for legal and regulatory frameworks are likely to inhibit AI development in individualistic countries and cultures.

In some instances, development of such legal frameworks may be accelerated to facilitate technological and economic development if this is deemed to be in a country's best national interests. Political and economic drivers such as progress in AI-based defence and weapons technology may promote AI development in some nations, particularly in the military sector, even without establishment of the necessary legal frameworks. Participants were acutely aware of the potential for the weaponisation and the politicisation of AI to disrupt the balance of power, create political divides and adversely impact on developing nations. Thus, the fusion of economic and political factors will most certainly influence the development of AI across all sectors and within all technologies, including security. In societies that are less transparent or loosely governed, AI may progress rapidly regardless of safety or legality, and with or without endorsement from the masses. Whether or not that development will be moral, ethical or acceptable is entirely subjective and dependent upon not only the application of the AI, but the moral compass of those who make those judgements.

Key Finding: Political factors in some regions will accelerate AI development, regardless of safety, legality or public endorsement.

"They are accelerating really fast, but they've got this communist regime that's really loose around how the rules are sort of implemented... And basically, the only rule is, if it's good for the government then that's what we're going to do. It's just not the right rule to have, in my opinion, if you're developing a highly sophisticated AI intelligence, it needs to be done in a more socially respectful way" [Participant 6, Security Expert].

"it's the governments, the legal, the political constructs across the world...there needs to be a unification of all those things that would then define the appropriate and socially acceptable development and deployment of AI" [Participant 6, Security Expert].

Political factors may also influence social contracts with citizens, with the use of AI in military and security technologies having the potential to influence the degree to which governments are able to facilitate AI development, and subsequently how government use will influence AI adoption in the private sector. The social contracts in countries or regions with collectivist or nationalist cultures may also influence a greater uptake of AI technology for use in security environments. Nationalistic countries such as Israel that tend to have a high security and military presence may accept and collaborate with government, law enforcement and private sector for the implementation of AI more readily than perhaps in countries where citizens are more suspicious of government intentions and motives. Despite a potentially higher acceptance, public opinion and subsequent social contracts may be heavily influenced by both internal and external factors such as political change, economic conditions, or international conflicts.

Key Finding: Social contracts in collectivist or nationalistic cultures may influence a greater uptake of AI, especially in countries that have a high security or military presence.

The social factors expected to influence the use of AI are diverse, and may change depending on individual, cultural and even organisational influences on different groups or individuals. For instance, some individuals or organisations may consider AI and high levels of security as an indicator of wealth, privilege, status, or simply used as a tool to enhance perceptions of how secure an individual or facility may be.

> *"…and so I said to them, given the threat level is so low, why do you want to have incredibly sophisticated security technologies in your building? And their response basically was because many of our clients come from overseas… and when they come to our building, we want them to feel as though we understand security and that their business is going to be protected here… and so it was all about image because the threat level was incredibly low" [Participant 3, Security Expert].*

However, social fears could essentially derail the implementation of AI at any level of society, and in any context whether that be for corporate, law enforcement or national security purposes. Experts warn that devices and methodologies must be socially acceptable and not infringe on the privacy or rights of individuals. For the corporate security industry, and on the basis that security technologies are essentially in the public sphere, social acceptance will most likely be the most influential factor shaping the development of AI.

Key Finding: Social fear of AI is capable of derailing the implementation of AI.

> *"…just a simple thing like facial recognition – socially it's acceptable in China, socially it's not acceptable in the UK. There was actually an outcry in the UK about six months ago because the Metropolitan Police ran a trial on facial recognition software" [Participant 4, Security Expert].*
>
> *"Twenty or thirty years ago, biometrics was going to change the world, and you still don't see a fingerprint reader anywhere… its very rarely used outside of high security applications. The technology will progress and well see it in some applications, so were seeing as I said biometrics in phones and consumer equipment, but whether that actually flows through to what the public would perceive as a more authoritarian use for security and policing applications, for example, is different question" [Participant 5, Security Expert].*
>
> *"…I think that the development of the technology is progressing unabated, you know I think it's growing and developing all the time… I think the controlling factor here is the social acceptance" [Participant 4, Security Expert].*

3.8 Interpretation and Findings Phase Three

Phase Three sought to identify areas for potential advancements in security technologies using Artificial Intelligence (AI), and where evident, highlight risks in their uptake for security management. Phase three responded to the following overarching research questions:

a) What are the priority security technology capability areas in the protection of assets for enhancement through artificial intelligence?

b) What vulnerabilities exist in the artificial intelligence techniques that may be used to achieve control in a physical environment?

Phase Three developed *The Security Technology Intelligent Autonomy Scale* (Appendix E) 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. The scale facilitated expert identification of current levels of intelligent autonomy, as well as the potential levels of intelligent autonomy that security systems may be capable of reaching, from *Level 1: Manual* (an absence of AI) through to *Level 11: Post-autonomous* (General AI). The highest Mean rating of intelligent autonomy in security technologies was *Level 5: Automated* (Narrow AI), which suggests that despite the adoption of AI, security technologies have limited intelligent autonomy and are not currently capable of making complex decisions through the application of AI algorithms.

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 at this point in time, and is likely to remain this way until the age of quantum computing arrives.

While AI in many security technologies may be considered *narrow, or a series of narrow AI outputs (broad AI),* and relatively unsophisticated currently, there are considerable opportunities for future development. 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. Opportunities will also emerge for more extensive applications of observation and response technologies, though development will likely be constrained by *Political*, *Social* and *Legal* factors. 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.

However development of AI in security technologies is not explicitly tied to the technical opportunities and capabilities of systems. There are *Political*, *Economic*, *Social*, *Technological*, *Environmental* and *Legal* (PESTEL) benefits and risks which may equally affect the development and deployment AI. 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.

There are profound risks of developing AI in security technologies, the consequences of which may not be fully comprehendible at this point in time. The quest for technological advancement may create political divides, upset balances of power, encourage exploitation of underdeveloped nations, and 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 power 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 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 discourse on intelligent autonomy, provided through enforceable international governance platforms.

However alarming 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 as a result 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.

3.8.1 Phase Three Expert Interview Key Findings

Phase Three developed a list of key findings in understanding the opportunities and risks for AI in security technologies:

3.8.1.1 Key Findings - Objective One

Objective One identified the level at which each category of security technologies is currently located on the *Security Technology Intelligent Autonomy Scale*.

- Security technologies currently achieve no higher than Level 5 on the Security Technologies Intelligent Autonomy Scale
- Observe Technologies were rated as achieving the highest level of Intelligent Autonomy (M = 5.05, SD = 1.9214), driven by research and development in computer vision and analytics.
- Response Technologies were rated as achieving the lowest level of Intelligent Autonomy (M = 3.90, SD = 2.7669), limited by concerns such as safety, liability, and social acceptance.
- There is a significant gap between the development and deployment of AI, particularly in the commercial sector as non-military applications may be subject to greater moral, ethical and legal critique.

3.8.1.2 Key Findings - Objective Two

Objective Two identified opportunities for the development of artificial intelligence in each category of security technologies, including broad areas for enhancement and the specific locations for opportunities within *The Security Technology – Artificial Intelligence Cycle*. The broad thematic areas for enhancement included *AI Testing & Standardisation*, *Technical Improvements*, *Integration & Aggregation of Inputs*, *Analytics*, *Human-Machine Teaming*, *Innovative Development*, and *Holistic Development in the Security Domain*. Specific locations of opportunities occurred across all stages of the intelligent cycle.

Key findings across the categories of Observe, Detect, Control, Response and Integrated Security Technologies were:

Observe

- Development of AI in security technologies may be enhanced or constrained by synchronistic development in areas which overlap the security function.
- Policies and frameworks must be established for the development and deployment of AI-enabled security technologies in socially acceptable ways.

Detect

- AI must be able to interpret and contextualise dynamic environments, events and situations. Development in this respect must also include the ability to understand and account for significant deviations or outliers from expected inputs, outputs and norms.
- Algorithms must be able to understand and imply the underlying intention of programmed rules to add sufficient depth and meaning to decision making.
- Development of new domains such as Security-AI Red Teaming may arise and as a result of workforce opportunities created by AI.

Control

- AI adoption should be promoted in technologies, applications and environments where the risk of AI failure can be explicitly linked with lower consequences.
- Defence in depth strategy should be actively applied to the implementation of AI and consider the strategic placement of humans and machines to produce optimal security outcomes.

Respond

- Human-Machine Teaming (HMT) may represent significant opportunities for enhancement of AI in security technologies.
- Military technology is expected to shape development across a range of AI applications as a result of the porous nature of military-commercial exchange.
- Unique opportunities for innovative development of drone swarms and robotics are emerging for Response technologies.

Integrated Security Technologies

- Integrated systems essentially require transformation from a logical input/output type architecture to a more sophisticated architecture with a coordinated approach to decision making.

- Human-Machine Teaming and ranked options for response could enhance critical decision making for human operators.

3.8.1.3 Key Findings – Objective Three

Objective Three identified the benefits and risks associated with AI development in security technologies. A PESTEL analysis determined benefits and risks according to *Political*, *Economic*, *Social*, *Technological*, *Environmental* and *Legal* factors across each of the security technology categories. The key findings for Observe, Detect, Control, Response and Integrated Security Technologies are as follows:

Observe

- Observation technologies may carry significant risks resulting from privacy and human rights issues, particularly for individualist cultures where citizens have greater rights and freedom of speech.
- AI must be developed and deployed in socially acceptable ways, to minimise the risks associated with violation of personal data and information.
- Governance and regulatory frameworks must establish mechanisms for oversight to prevent the misuse or exploitation of AI technologies.

Detect

- AI enhancements in detection technologies may offer early intervention opportunities to respond to, and mitigate threats, sooner.
- Detection technologies are perceived to carry less risk than observation technologies.
- Failure to establish local manufacturing industries may increase reliance on international supply markets and consequently increase sovereign risk.
- Cloud-based systems increase economic risks resulting from loss of access to data and services.
- Detection technologies, including IOT and Edge of Network Devices, are significantly vulnerable to defeat.
- Development of AI testing mechanisms and standards may reduce legal liability resulting from inadequate system performance.

Control

- AI development in Control technologies may encounter social and ethical challenges arising from potential misuse of personal data.
- Greater contextualisation is required to enhance control technologies, however for this to occur, significant development in datasets is required.
- Consideration must be afforded to ensure sufficient protections are in place to ensure the privacy and security of vulnerable persons enrolled in control systems.

Respond

- Response technologies may provide a considerable benefit where humans can be removed from the front line.
- Autonomous use or release of force presents extensive moral, ethical and human rights considerations.

- Urgent attention must be afforded to the development of regulatory frameworks to minimise the risks of back or grey market trade of response technologies.
- Implicit and explicit compliance with societal constructs is currently beyond the capabilities of any machine.
- Technology users must be aware that novel security vulnerabilities may arise through the development and deployment of innovative AI technologies.
- We have not yet established the necessary legal frameworks required to sufficiently administer or apportion responsibility and liability for AI response technologies.
- Autonomous response systems and devices may breach legal duty of care obligations.

Integrated Security Technologies

- Disparity in AI development between nations may disrupt the balance of power, and facilitate the political, military or economic exploitation of underdeveloped regions or nations.
- Ethical checks and balances may require systematic review to ensure the use of AI remains aligned with public notions of what is socially and ethically acceptable.
- Deployment of AI is likely to be constrained in those technologies which are perceived to infringe on privacy and human rights, or where their use is not considered to provide an overall benefit to society.
- The security industry must take the lead in shaping the design of AI technologies to meet security needs.
- Uncertainty is the overarching legal risk to users of AI in security.

3.8.1.4 Key Findings – Objective Four

Objective Four explored how AI in security technologies is expected to develop over the next ten years, with experts identifying a number of factors which will influence and drive AI development. Key findings based on expert opinions are provided for the *Expected Areas of AI Development*, *Expected Drivers of AI Development* and the *PESTEL Factors Expected to Influence AI Development*.

Expected Areas of AI Development

- Experts predict transformation within the security industry, such as an increase in the use of drones and robotics, and a decrease in the need for alarm monitoring centres.
- Experts expect security technologies will become more sophisticated, but do not anticipate those technologies will significantly change in the next ten years.

Expected Drivers of AI Development

- *Quantum Computing*, *Military Technology* and *Narrow AI Techniques* are expected to drive AI development in security technologies.

PESTEL Factors Expected to Influence AI Development

- Demand for legal and regulatory frameworks are likely to inhibit AI development in individualistic countries and cultures.
- Political factors in some regions will accelerate AI development, regardless of safety, legality or public endorsement.

- Social contracts in collectivist or nationalistic cultures may influence a greater uptake of AI, especially in countries that have a high security or military presence.
- Social fear of AI is capable of derailing the implementation of AI.

3.8.2 Key Recommendations for AI in Security Technologies:

Key findings produced several key recommendations for the use and oversight of AI in security technologies:

- Development of a security industry Artificial Intelligence Adoption framework.
- Development of an Artificial Intelligence Risk Decision Matrix for security mangers to evaluate the benefits and risks of AI technologies for their environment.
- An AI security vulnerability and criticality assessment should be developed to guide deployment of AI systems in cloud-based 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.
- 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.

3.9 Conclusion

The overwhelming consensus from this project is that AI in security is at an elementary stage, with a limited capacity for intelligent decision making and autonomy. AI does not 'think', rather it computes, processes, applies rules, and in machine learning applications, may even generate rules. But AI is fallible and inflexible, and subsequently operates in an environment of black and white. In contrast, humans think, and therefore are able to operate between black and white 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 simply 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 ten years, the age of Quantum computing is likely to facilitate AI development 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 the front line.

However 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 a fundamental consideration 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.

4.0 References

Abaimov, S., & Martellini, M. (2020). Artificial Intelligence in Autonomous Weapon Systems. In M. Martellini & R. Trapp (Eds.), *21st Century Prometheus* (pp. 141-177). Springer, Cham. https://doi.org/10.1007/978-3-030-28285-1_8

Active Online. (2019). *Choosing the right camera for your application*. https://www.activeonline.com.au/choosing-the-right-cctv-camera-for-your-application/

Adenin. (n.d.). *Natural Language Understanding (NLP)*. https://www.adenin.com/glossary/nlp/

Akcay, S., Kundegorski, M.E., Willcocks, C.G., & Breckon, T.P. (2018). Using deep convolutional neural network architectures for object classification and detection within x-ray baggage security imagery. *IEEE Transactions on Information Forensics and Security, 13*(9), 2203–2215. https://doi.org/10.1109/TIFS.2018.2812196

Akman, M.K. (2020). Defence Management and PESTLE analysis. *Ante Portas - Security Studies, 1*(14), 93-102. https://doi.org/10.33674/3201910

Ahmad, F. (2019, January 11). The concepts involved in neural network modelling for a non-specialist. *Towards Data Science*. https://towardsdatascience.com/understanding-the-set-up-of-a-simple-neural-network-to-predict-an-output-ff4d108d8f8d

Alango Technologies & Solutions. (n.d.). *Voice Enhancement: Automatic Gain Control*. http://www.alango.com/technologies-drc.php

American Psychological Association. (2020). *APA Dictionary of Psychology (Online).* https://dictionary.apa.org/

Anderson, L. (2019, March 1). Artificial Intelligence: What's Next for AI and The Cloud in Physical Security? *Security Informed*. https://www.securityinformed.com/docs/opdf/insights/artificial-intelligence-ai-cloud-physical-security-co-1537871022-ga-sb.1551419637.pdf

Andrade, E.B. & Ariely, D. (2009). The enduring impact of transient emotions on decision-making. *Organizational Behavior and Human Decision Processes, 109*(1), 1-8. https://doi.org/10.1016/j.obhdp.2009.02.003

Anodot. (n.d.). AI Analytics. https://www.anodot.com/learning-center/ai-analytics/

Araujo, T., Helberger, N., Kruikemeier, S. & de Vreese, C.H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society,* 35, 611-623. https://doi.org/10.1007/s00146-019-00931-w

Ashford, W. (2014). Standardisation key to future security, say experts. Computerweekly.com. https://www.computerweekly.com/news/2240214991/Standardisation-key-to-future-security-say-experts

Ashman, A.F., & Conway, R.N.F. (1997). *An introduction to cognitive education: theory and applications*. Routledge.

ASIS International. (2009). *Facilities physical security measures guideline* (GDL FPSM-2009).

https://www.asisonline.org/publications/sg-asis-facilities-physical-security-measures-guideline-2009-ed/

ASIS International. (2015). *Physical Security Principles*. ASIS International

AT&T Cybersecurity. (2019, September 30). Building Automation Systems Are a Breeding Ground for Siegeware. *Channel Futures*. https://www.channelfutures.com/from-the-industry/building-automation-systems-are-a-breeding-ground-for-seigeware

Atlas, R. I. (2013). *21st century security and CPTED : designing for critical infrastructure protection and crime prevention* (2nd ed.). CRC Press.

Australia New Zealand National Counter-Terrorism Committee. (2017). *Hostile Vehicle Guidance for Crowded Places*. Attorney General's Department. https://www.nationalsecurity.gov.au/Media-and-publications/Publications/Documents/hostile-vehicle-guidelines-crowded-places.pdf

Australian Bollards. (n.d.). *Security Camera Bollard*. http://www.australianbollards.com.au/Catalog/camera-bollard/camera-bollard-1

Baggenstoss, P.M., Springer, M., Oispuu, M., & Kurth, F. (2019). Efficient phase-based acoustic tracking of drones using a microphone array. In *27th European Signal Processing Conference (EUSIPCO),* (pp. 1–5). https://doi.org/10.23919/EUSIPCO.2019.8902972

Babbie, E. R. (2016). *The practice of social research* (14th ed.). Cengage Learning.

Baker, G.H. (2011). High Power Electromagnetic Weapons: A Brief Tutorial. https://works.bepress.com/george_h_baker/33/

Balasubramanian, R., & Aramudhan, M. (2012). Security issues: public vs private vs hybrid cloud computing. *International Journal of Computer Applications, 55*(13), 35-41. https://pdfs.semanticscholar.org/eee8/3d2a5b2ce70cea5bb0f8248a1dc0699228e5.pdf

Banafa, A. (2016, June 6). What is Affective Computing? BBVA Open Mind. https://www.bbvaopenmind.com/en/technology/digital-world/what-is-affective-computing/

Bandyopadhyay, D., & Sen, J. (2011). Internet of things: Applications and challenges in technology and standardization. *Wireless Personal Communications, 58*(1), 49-69.

Barnett, J. (2020, February 14). AI needs humans 'on the loop' not 'in the loop' for nuke detection, general says. *Fedscoop*. https://www.fedscoop.com/ai-should-have-human-on-the-loop-not-in-the-loop-when-it-comes-to-nuke-detection-general-says/

Baron, R.J. (1981). Mechanisms of human facial recognition. *International Journal of Man-Machine Studies*, *15*(2), 137-178.

Bataineh, M., & Marler, T. (2017). Neural network for regression problems with reduced training sets. *Neural networks, 95*, 1-9. http://dx.doi.org/10.1016/j.neunet.2017.07.018

Belaidouni, S., & Miraoui, M. (2016, October). Machine learning technologies in smart spaces. In *The Tenth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, (pp. 52-55). UBICOMM.

Baron-Cohen, S. (2017). Editorial Perspective: Neurodiversity – a revolutionary concept for autism and psychiatry. *The Journal of Child Psychology and Psychiatry, 58*(6), 744-747. https://doi.org/10.1111/jcpp.12703

Benton, A.L. (1980). The Neuropsychology of Facial Recognition. *American Psychologist, 35*(2), 176-186.

Berce, C. (2019, June 12). Tech Deep Dive: Gunshot Detection. *Security InfoWatch*. https://www.securityinfowatch.com/alarms-monitoring/alarm-systems-intrusion-detection/gunshot-detection/article/21081663/tech-deep-dive-gunshot-detection

Bertalanffy, L.von. (1968). *General System Theory: Foundations, Development, Applications*. Brazillier.

Bertozzi, W., & Ledoux, R. (2005). Nuclear resonance fluorescence imaging in non-intrusive cargo inspection. *Nuclear Inst. and Methods in Physics Research, B, 241*(1-4), 820-825. https://doi.org/10.1016/j.nimb.2005.07.202

Björklund, F. & Bäckström, M. (2008). Individual differences in processing styles: validity of the Rational-Experiential Inventory. *Scandinavian Journal of Psychology, 49*(5), 439-446. https://doi.org/10.1111/j.1467-9450.2008.00652.x

Blue Swan Daily. (2019, January 16). Standardisation could be necessary before we truly see widespread use of biometrics. *The Blue Swan Daily*. https://blueswandaily.com/standardisation-could-be-necessary-before-we-truly-see-widespread-use-of-biometrics/

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences, USA, 99*(3), 7280-7287. https://doi.org/10.1073pnas.082080899

Bonoan, J., & Ataev, A. (2020, March 11). AI and Specialty Analytics are Changing Video Surveillance. https://www.securityinfowatch.com/video-surveillance/article/21125810/ai-and-specialty-analytics-are-changing-video-surveillance

Boole, G. (1847). *The mathematical analysis of logic*. Philosophical Library.

Booz Allen Hamilton. (n.d.). Shipboard Integration of Directed Energy Weapons. https://www.boozallen.com/d/insight/thought-leadership/shipboard-integration-of-directed-energy-weapons.html

Borowski, B., Sutin, A., Roh, H., & Bunin, B. (2008). Passive acoustic threat detection in estuarine environments. *Proceedings of SPIE - The International Society for Optical Engineering, 6945*. https://doi.org/10.1117/12.779177

Boyatzis, R.E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oakes, California: Sage Publications, Inc.

Bradford, C. (n.d.). 7 Most Infamous Cloud Security Breaches. *Storage Craft*. https://blog.storagecraft.com/7-infamous-cloud-security-breaches/

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in*

Psychology 3(2), 93. doi: 10.1191/1478088706qp063oa

Brehm, M., & Wheeler, A.C. (2018, November). Acoustic Weapons [Discussion Paper]. *Convention on Certain Conventional Weapons (CCW)*. Geneva, Switzerland.

Brooks, D. (2014). Intrusion Detection Systems in the Protections of Assets. In M. Gill (Ed.), *The Handbook of Security* (2nd ed.). Palgrave Macmillan.

Brooks, D., & Coole, M. (2019). Intrusion Detection Systems. In L. R. Shapiro & M.-H. Maras (Eds.), *Encyclopedia of Security and Emergency Management* (pp. 1-5). Springer International Publishing.

Brooks, D.J., Coole, M., Haskell-Dowland, P., Griffiths, M., Lockhart, N. (2017). *Building Automation & Control Systems: An Investigation into Vulnerabilities, Current Practice & Security Management Best Practice.* Edith Cowan University. https://www.securityindustry.org/wp-content/uploads/2018/08/BACS-Report_Final-Intelligent-Building-Management-Systems.pdf

Brown, C. (2019). Four Ways Artificial Intelligence Helps Mitigate Risk. *Security Magazine*. https://www.securitymagazine.com/articles/90892-four-ways-artificial-intelligence-helps-mitigate-risk

Busset, J., Perrodin, F., Wellig, P., Ott, B., Heutschi, K., Ruhl, T., & Nussbaumer, T. (2015). Detection and tracking of drones using advanced acoustic cameras. In *Proceedings of SPIE - the International Society for Optical Engineering, 9647*. https://doi.org/10.1117/12.2194309

Business-to-you. (n.d.). *Scanning the Environment: PESTEL Analysis*. https://www.business-to-you.com/scanning-the-environment-pestel-analysis/

Butcher, M., & Fretwell, P. (2009). *Security manager's guide to radar for site security applications*. Centre for the protection of National Infrastructure. https://www.cpni.gov.uk/system/files/documents/e5/46/Security-managers-guide-radar.pdf

Cadwalladr, C., & Graham-Harrison, E. (2018, March 18). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Calavia, L., Baladrón, C., Aguiar, J. M., Carro, B., & Sánchez-Esguevillas, A. (2012). A semantic autonomous video surveillance system for dense camera networks in Smart Cities. *Sensors (Basel, Switzerland), 12*(8), 10407–10429. https://doi.org/10.3390/s120810407

Cambridge Dictionary. (2020). Neurodiversity. In *Cambridge Dictionary (Online)*. https://dictionary.cambridge.org/dictionary/english/neurodiversity

Campaign to Stop Killer Robots. (2019). Key Elements of a Treaty on Fully Autonomous Weapons. https://www.stopkillerrobots.org/wp-content/uploads/2020/04/Key-Elements-of-a-Treaty-on-Fully-Autonomous-WeaponsvAccessible.pdf

Carlson, M. (2018). Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism. *New Media & Society,* 20(5), 1755-1772. https://doi.org/10.1177/1461444817706684

Carter, J. (2019, June 1). The Game Changer: How Artificial Intelligence chances the access control game. *Security Today*. https://securitytoday.com/articles/2019/06/01/the-game-changer.aspx

Case, E.E., Zelnio, A.M., & Rigling, B.D. (2008). Low-cost acoustic array for small UAV detection and tracking. In Proceedings of the National Aerospace and Electronics Conference (pp. 110-113). IEEE. https://doi.org/10.1109/NAECON.2008.4806528

Caterpillar. (n.d.). Automation & Autonomy: What's the difference? https://www.cat.com/en_US/articles/ci-articles/automation-autonomy-whats-the-difference.html

Chang, C. (2016, December 22). Long range acoustic devices terrorising protestors. *News.com.au*. https://www.news.com.au/technology/innovation/long-range-acoustic-devices-terrorising-protesters/news-story/83a47dbd6e64095549f09a89ef020f94

Chang, E.M. (2014). *Defining the Levels of Adjustable Autonomy: A Means of Improving Resilience in an Unmanned Aerial System* [Master of Science Thesis, Naval Postgraduate School]. NPS Archive. http://hdl.handle.net/10945/43887

Charara, N., Jarkass, I., Sokhn, M., Mugellini, E., & Khaled, O.A. (2012). ADABeV: Automatic Detection of Abnormal Behaviour in Video-surveillance. *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering, 6*(8), 946-952.

Chelani, K., Kahl, F., & Sattler, T. (2021). *How Privacy-Preserving are Line Clouds? Recovering Scene Details from 3D Lines*. https://arxiv.org/pdf/2103.05086.pdf

Cho, A. (2020, July 9). The biggest flipping challenge in quantum computing. *Science*. https://www.sciencemag.org/news/2020/07/biggest-flipping-challenge-quantum-computing

CIBSE. (2000). *Building control systems: CIBSE Guide H*. Butterworth-Heinemann.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th Ed.). Abingdon, Oxon: Routledge.

Collins, T., & Atkins, P. (1998). Doppler-sensitive active sonar pulse designs for reverberation processing. *Radar, Sonar and Navigation, 145*(6), 347–347. https://doi.org/10.1049/ip-rsn:19982434

Coloumbe, R. (2019, December 12). Tech Trends: Put Radar on Your Radar. Security Infowatch. https://www.securityinfowatch.com/perimeter-security/article/21114744/put-radar-on-your-radar

Conde, D. (2018, April 26). Tackling Edge Computing Challenges. *Network Computing*. https://www.networkcomputing.com/data-centers/tackling-edge-computing-challenges

Continuity Central. (2019). *Cloud is safer than on-premise say the majority of security leaders*. https://www.continuitycentral.com/index.php/news/technology/4384-cloud-is-safer-than-on-premise-say-that-majority-of-security-leaders

Coole, M., & Brooks, D. (2009). *Security Decay: An Entropic Approach to Definition and Understanding*. WorldCat.org database.

Coole, M., Brooks, D., & Treagust, D. (2015). The Physical Security Professional: Formulating a Novel Body of Knowledge. *Journal of Applied Security Research, 10*(3), 385-410. https://doi.org/10.1080/19361610.2015.1038768

Corea, F. (2019). *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science (Volume 50 Studies in Big Data)*. Springer Nature Switzerland AG 2019.

Craighead, G. (2009). *High-rise security and fire life safety* (3rd ed.). Butterworth-Heinemann/Elsevier.

Creswell, J.W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches* (3rd ed.). Thousand Oaks, California: Sage Publications.

Cybersecurity & Infrastructure Security Agency. (2019). *Guide to Active Vehicle Barrier (AVB) Specification and Selection Resources*. https://www.cisa.gov/sites/default/files/publications/Guide-to-Active-Vehicle-Barrier-2014-508.pdf

Defense Innovation Marketplace. (n.d.). *Human Machine Teaming*. https://defenseinnovationmarketplace.dtic.mil/technology-interchange-meetings/autonomy-tim/human-machine-teaming/

Denman, S., Fookes, C., Cooke, J., Davoren, C., Mamic, A., Farquharson, G., Chen, D., Chen, B., & Sridharan, S. (2006). Multi-view Intelligent Vehicle Surveillance System. *IEEE International Conference on Video and Signal Based Surveillance (AVSS'06).* https://eprints.qut.edu.au/9498/1/9498_2.pdf

Di Nardo, J. (2009). Biometric Technologies: Functionality, emerging trends and vulnerabilities. *Journal of Applied Security Research, 4*.

Dibazar, A., Yousefi, A., Park, H., Lu, B., George, S., & Berger, T. (2011). Intelligent Recognition of Acoustic and Vibration Threats for Security Breach Detection, Close Proximity Danger Identification, and Perimeter Protection. *Homeland Security Affairs*. https://www.hsaj.org/articles/72

Dickinson, L.A. (2018). Drones, Automated Weapons, and Private Military Contractors. In M.K. Land & J.D. Aronson (Eds.), New Technologies for Human Rights Law and Practice (pp. 93-124). Cambridge University Press. https://doi.org/10.1017/9781316838952.005

Dietrich, C. (2010). Decision Making: Factors that Influence Decision Making, Heuristics Used, and Decision Outcomes. *Inquiries Journal, 2*(02). http://www.inquiriesjournal.com/a?id=180

Donovan, S.M., Looney, C., & O'Rourke, M. (2015). Your hypothesis or mine? terminological and conceptual variation across disciplines. *Sage Open, 5*(2). https://doi.org/10.1177/2158244015586237

Dorigo, M. & Gambardella, L. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation, 1*(1), 53-66. https:doi.org/10.1109/4235.585892

Dotan, D. & Dehaene, S. (2020). Parallel and serial processes in number-to-quantity conversion. *Cognition,* 204. https://doi:10.1016/j.cognition.2020.104387

Dowell, C. (2015, April 10). Online Exclusive: Integrated Security – Not Just an Industry Buzzword. *Security Today*. https://securitytoday.com

DSIT Solutions. (n.d.). Harbour Surveillance System. https://dsit.co.il/sonar/underwater-security/harbor-surveillance-system-hss/

Ducklin, P. (2015, May 14). The VENOM "virtual machine escape" bug – what you need to know. *Sophos Limited*. https://nakedsecurity.sophos.com/2015/05/14/the-venom-virtual-machine-escape-bug-what-you-need-to-know/

Dunjko, V., Taylor, J. M., & Briegel, H. J. (2016). Quantum-enhanced machine learning. *Physical Review Letters, 117*(13), 130501-6. https://doi.org/10.1103/PhysRevLett.117.130501

Dunstone, T., & Yager, N. (2009). Biometric Matching Basics. In T. Dunstone & N. Yager (Eds.), *Biometric System and Data Analysis*. Springer.

Dutta, S., & Bonissone, P.P. (1993). Integrating Case- and Rule-Based Reasoning. *International Journal of Approximate Reasoning, 8*, 163-203. https://bit.ly/3asEihS

Dynamic Excavations. (2019). *Ground Penetrating Radar (GPR) Explanation*. https://www.dynamicexcavations.com.au/news/ground-penetrating-radar-gpr-explanation/

Earth Observing Laboratory. (n.d.). How Do Radars Work? *University Corporation for Atmospheric Research*. https://bit.ly/2Dkazvo

Eidson, B. (2019). Creating an AI Red Team to Protect Critical Infrastructure. MITRE. https://www.mitre.org/publications/project-stories/creating-an-ai-red-team-to-protect-critical-infrastructure

Elgan, M. (2018, December 15). Ambient computing is in the air. *Computerworld.* https://www.computerworld.com/article/3328545/ambient-computing-is-in-the-air.html

Elprocus. (n.d.). *RADAR – Basics, Types and Applications*. https://www.elprocus.com/radar-basics-types-and-applications/

Encyclopaedia Britannica. (n.d.). Doppler effect. In *Encyclopaedia Britannica*. https://www.britannica.com/science/Doppler-effect

Encyclopaedia Britannica. (n.d.). Sonar. In *Encyclopaedia Britannica*. https://www.britannica.com/technology/sonar

European Association for Machine Translation. (n.d.). What is Machine Translation? http://eamt.org/what-is-machine-translation/

European Committee for Standards. (2015). *Video Surveillance Systems for Use in Security Applications – Part 4: Application Guidelines* (IEC 62676-4:2014). SAI Global. https://www.saiglobal.com/

Facilitiesnet.com. (2007). *Security Integration: Making it Work*. https://www.facilitiesnet.com/security/article/Security-Integration-Making-it-Work--6273

Falanga, D., Kleber, K., & Scaramuzza, D. (2020). Dynamic obstacle avoidance for quadrotors with event cameras. *Science Robotics, 5*(40), 1-14. http://doi.org/10.1126/scirobotics.aaz9712

Farrington, D. P., & Welsh, B. C. (2004). Measuring the Effects of Improved Street Lighting on Crime: A Reply to Dr Marchant. *The British Journal of Criminology, 44*(3), 448-467.

Fay, J. (2011). *Contemporary security management* [1 online resource (xvii, 461 pages) : illustrations](3rd ed. ed.).

Fayol, H. (1916). Administration industrielle et generale. *Bulletin de la Societe de l'Industrie Minerale, 10*, 5-162.

Federation of American Scientists. (n.d.a.). *Basic Radar Systems*. https://fas.org/man/dod-101/navy/docs/es310/radarsys/radarsys.htm

Federation of American Scientists. (n.d.b.). *ASW Systems*. https://fas.org/man/dod-101/navy/docs/es310/asw_sys/asw_sys.htm

Feigenbaum, E.A. (1982). Knowledge Engineering for the 1980's, Computer Science Dept., Stanford University, 1982.

Felt, C. (2020, February 14). Autonomous Weaponry: Are Killer Robots in Our Future? *The Henry M. Jackson School of International Studies*. https://jsis.washington.edu/news/autonomous-weaponry-are-killer-robots-in-our-future/

Fennelly, L. (2012). *Effective physical security* (4th ed.). Butterworth-Heinemann.

Fielding, A.H., & Torbett, M. (2017). Radar: Superior Perimeter Security at Reduced Cost. *Honeywell*. https://www.honeywellprocess.com/en-US/online_campaigns/Radar-wp/Pages/RadarVideoSurveillance.pdf

Fischer, R.J., & Green, G. (1998). *Introduction to Security* (6th ed.). Butterworth Heinemann.

Fischer, R.J., Halibozek, E., & Green, G. (2008). *Introduction to Security* (8th ed.). Butterworth-Heinemann.

Fortem. (n.d.). Compatible Technologies and/or Partners. http://fortem.com/integrated-security-systems/

Frankish, K., & Ramsey, W. (2014). *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press. https://www.doi.org/10.1017/CBO9781139046855

Frost, C. (2011). Challenges and Opportunities for Autonomous Systems in Space. In *Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2010 Symposium* (pp. 89-101). National Academies Press. https://doi.org/10.17226/13043

Funahashi, M., & Yoshikawa, S. (2011). Fujitsu's approach to hybrid cloud systems. *Fujitsu Sci.*

Tech. J, 47(3), 285-292. https://www.fujitsu.com/global/documents/about/resources/publications/fstj/archives/vol47-3/paper10.pdf

Gao, T., Zhang, C., Wang, Y., Diaz, J. A., Zhao, J., & Willis, B. G. (2020). Machine learning assisted nanoparticle-based chemiresistor array for explosive detection. *Ieee Sensors Journal, 1*, 1–1. https://doi.org/10.1109/JSEN.2020.3007493

Garcia, M. L. (2001). *The design and evaluation of physical protection systems*. Butterworth-Heinemann.

Garcia, M. L. (2008). *The design and evaluation of physical protection systems* (2nd ed.). Elsevier/Butterworth-Heinemann.

Garland, D. (2001). *The culture of control; Crime and social order in contemporary society*. University of Chicago Press.

Garrity, P. & Roberts, C. (2018, June 18). Defining Access Security for Cloud Applications. *Duo Security*. https://duo.com/blog/defining-access-security-for-cloud-applications

Gaus, Y., Bhowmik, N., & Breckon, T. (2019). *On the Use of Deep Learning for the Detection of Firearms in X-ray Baggage Security Imagery*. Paper presented at the 2019 IEEE International Symposium on Technologies for Homeland Security (HST).

GE Homeland Security. (2008). *Ion Trap Mobility Spectrometry: The Science Behind the Technology*. http://netcoinc.net/wp-content/uploads/2015/09/itms_wp.pdf:

Globe Newswire. (2019, June 24). Artificial Intelligence Technology Solutions Launches New Security Device Solution ROSA. https://www.globenewswire.com

Goertzel, B., & Pennachin, C. (2007). *Artificial general intelligence* (Ser. Cognitive technologies). Springer. https://doi.org/10.1007/978-3-540-68677-4

Green, A. (2018, December 1). How to Use IBISWorld Research in a PESTLE Analysis. IBISWorld. https://www.ibisworld.com/industry-insider/how-to-use-industry-research/how-to-use-ibisworld-research-in-a-pestle-analysis

Grosan, C., Abraham A. (2011). Rule-Based Expert Systems. In C. Grosan & A. Abraham (Eds.), *Intelligent Systems: A Modern Approach*. Intelligent Systems Reference Library, (17), 149-185. Springer. https://doi.org/10.1007/978-3-642-21004-4_7

Guo, J., Farhang-Razi, V., & Algra, P. (2019). AI: A Glossary of Terms. In E.R. Ranschaert, S. Morozov & P.R. Algra (Eds.), *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks*. Springer, Cham. https://doi.org/10.1007/978-3-319-94878-2

Hägelen, M., Kulke, R., Jetten, R., & Möllenbeck, G. (2016). Perimeter surveillance using a combination of radar and optical sensors. *In 2016 European Radar Conference (EuRAD)* (pp. 318-321). IEEE.

Hahn, P. (2019, September 23). Three Trends Diving the Geospatial AI Revolution. *The European Business Review*. https://www.europeanbusinessreview.com/three-trends-driving-the-geospatial-ai-revolution/

Hambling, D. (2004, June 16). Sweeping Stun Guns to Target Crowds. *New Scientist*. https://www.newscientist.com/article/dn6014-sweeping-stun-guns-to-target-crowds/

Hansen, R.E. (2011). *Introduction to Sonar* [Lecture Slides]. Semantic Scholar. https://pdfs.semanticscholar.org/415d/ea02a97b9c57e1fc74ec4a726ee4a13787b8.pdf

Harper Collins Publishers. (n.d.). Automation. In *Collins English Dictionary*. https://www.collinsdictionary.com/dictionary/english/automation

Harper Collins Publishers. (n.d.). Autonomy. In *Collins English Dictionary*. https://www.collinsdictionary.com/dictionary/english/autonomy

Harper Collins Publishers. (n.d.). Integration. In *Collins English Dictionary*. https://www.collinsdictionary.com/dictionary/english/integration

Hayes, J., & Ericson, L. (2014). *Through-the-Wall Sensors for Law Enforcement: Best Practices*. ManTech International Corporation, National Institute of Justice, Office of Justice Programs, & U.S. Department of Justice. https://nicic.gov/through-wall-sensors-law-enforcement-best-practices-version-10

Haykin, S. (2006). Cognitive radar: a way of the future. *IEEE Signal Processing Magazine, 23*(1). https://doi.org/10.1109/MSP.2006.1593335

Heidenreich, H. (2019, January 4). NEAT: An Awesome Approach to NeuroEvolution. *Towards Data Science*. https://towardsdatascience.com/neat-an-awesome-approach-to-neuroevolution-3eca5cc7930f

Help Net Security. (2020, February 10). The frequency of DDoS attacks depends on the day and time. *Help Net Security*. https://www.helpnetsecurity.com/2020/02/10/frequency-ddos-attacks/

Heyns, C. (2016). Autonomous Weapons Systems: Living a dignified life and dying a dignified death. In N. Bhuta, S. Beck, R. Geiß, H.Y. Liu, & C. Kreß (Eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*. Cambridge University Press.

Hobbes, T. (2018). *Leviathan*. Lerner Publishing Group. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5443203.

Hodgson, K. (2018, November 12). What Do You Know About Cloud-Based Access Control. SDM *Magazine*. https://www.sdmmag.com/articles/95852-what-do-you-know-about-cloud-based-access-control

Homeland Security Research Corpora. (2019). Explosives Trace Detection (ETD) Market & Technologies Expected to Grow during the 2020-2025 Forecast Period. https://www.prnewswire.com/news-releases/homeland-security-research-corp-hsrc-explosives-trace-detection-etd-market--technologies-expected-to-grow-during-the-2020-2025-forecast-period-300972077.html

Honey, G. (2001). *Emergency and Security Lighting*. Elsevier Science & Technology.

Howard, N. (1993). The role of emotions in multi-organizational decision-making. *The Journal of the Operational Research Society,* 44(6), 613-623.

https://www-jstor-org.ezproxy.ecu.edu.au/stable/pdf/2584517.pdf

Hughes, A. (2018, August 29). On Premise vs. Cloud: Key Differences, Benefits, and Risks. *Cleo*. https://www.cleo.com/blog/knowledge-base-on-premise-vs-cloud

Hughes, E.R. (2011). *Site surveying guide for radar and cameras: A How-to guide for the security professional*. Deep Sea Publishing.

Hussain, Z., Sheng, M., & Zhang, W.E. (2019). *Different Approaches for Human Activity Recognition*. https://arxiv.org/pdf/1906.05074.pdf

Illuminating Engineering Society of North America. (2003). *Guideline for security lighting for people, property and public space IESNA G-1-03*. https://standards.globalspec.com/std/377067/ies-g-1

IBM. (n.d.a). What is Quantum Computing? https://www.ibm.com/quantum-computing/learn/what-is-quantum-computing/

IBM. (n.d.b.). *Bypassing Safety Control Systems*. https://www.ibm.com/support/knowledgecenter/SSLL9G_7.6.0/com.ibm.oil.doc/plusgbyp/t_ctr_manage_bypass_ops.html

IBM Services. (2018). *Beyond the hype: A guide to understanding and successfully implementing artificial intelligence within your business*. https://www.ibm.com/downloads/cas/8ZDXNKQ4

IEEE Spectrum. (n.d.). *How AI is Starting to Influence Wireless Communications*. https://spectrum.ieee.org/computing/software/how-ai-is-starting-to-influence-wireless-communications

IHS Markit. (2018). *Artificial Intelligence in Physical Security: An overview of market & technology opportunities*. https://www.bibliotecadeseguranca.com.br/wp-content/uploads/2019/02/artificial-intelligence-in-physical-security.pdf

Infinity Electro-Optics. (n.d.). Whitepaper: DORI (Detection, Observation, Recognition, Identification). https://www.infinitioptics.com/whitepapers/dori-detection-observation-recognition-identification

International Committee of the Red Cross. (2014). *Autonomous weapon systems: Technical, military, legal and humanitarian aspects*. https://www.icrc.org/en/download/file/1707/4221-002-autonomous-weapons-systems-full-report.pdf

International Committee of the Red Cross. (2016). *Views of the International Committee of the Red Cross (ICRC) on autonomous weapons system)*. https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system

International Organization for Standardization. (2004*). ISO 16484-2: Building Automation and Control Systems (BACS) Part 2 Hardware*. International Organization for Standardization.

Isoardi, G. (2012, July 3). Explainer: the Doppler effect. *The Conversation Media Group*. https://theconversation.com/explainer-the-doppler-effect-7475

Jeffrey, P. (2003). Smoothing the waters: observations on the process of cross-disciplinary research collaboration. *Social Studies of Science, 33*(4), 539–562.

Jesson, J.K., Matheson, L., & Lacey, F. (2011). *Doing your literature review: Traditional and systematic techniques*. Los Angeles: Sage Publications.

Jiang, L., Yan, S., Wu, Y., Ma, X., (2016, April) Sonar detection performance with LFM-BPSK combined waveforms [essay]. Oceans 2016, Shanghai, Shanghai, China. IEEE. https://doi.org/10.1109/OCEANSAP.2016.7485501

Johnson, G. (1986). *Machinery of the mind; inside the new science of artificial intelligence*. Random House.

Joint Intermediate Force Capabilities Office. (n.d.). *Human Electro-muscular Incapacitation FAQs*. U.S. Department of Defense. https://jnlwp.defense.gov/About/Frequently-Asked-Questions/Human-Electro-Muscular-Incapacitation-FAQs/

Joint Intermediate Force Capabilities Office. (2018). *Directed Energy: Sound and Light Technology*. U.S. Department of Defense. https://jnlwp.defense.gov/Press-Room/Fact-Sheets/Article-View-Fact-sheets/Article/1574157/directed-energy-sound-and-light-technology/

Juslen, H., Nasir, O., & Nouri, J. (2019, Dec 16, 2019). AI & Lighting. *LED Professional(73),* 62. https://www.led-professional.com/resources-1/articles/ai-lighting

Kanai, R. & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience, 12*(4), 231-242. https://doi.org/10.1038/nrn3000

Kasabov, N. (2019). Artificial intelligence in the age of neural networks and brain computing. In *Evolving and spiking connectionist systems for brain-inspired artificial intelligence* (pp. 111–138). Elsevier. https://doi.org/10.1016/B978-0-12-815480-9.00006-2

Kaspersky. (2019). *Nearly four in Ten Smart Buildings Targeted by Malicious Attacks in H1 2019*. https://usa.kaspersky.com/about/press-releases/2019_smart-buildings-threat-landscape

Kass, D.H. (2019, October 4). Smart Building Automation Systems Vulnerable to Hackers, Report Finds. *MSSP Alert*. https://www.msspalert.com/cybersecurity-markets/verticals/smart-building-automation-hackers/

Katwala, A. (2020, March 5). Quantum computing and quantum supremacy, explained. *Wired*. https://www.wired.co.uk/article/quantum-computing-explained

Kelley, T.D. (2003). Symbolic and Sub-symbolic Representations in Computational Models of Human Cognition: What Can be Learned from Biology? *Theory & Psychology, 13*(6), 847-860.

Kelvin Hughes. (n.d.). Radar Based Security Systems: The Complete Picture. https://cdn.kelvinhughes.com/upload/pdf/brochures/security.pdf

Kerner, S.M. (2019, December 12). Edge Security: How to Secure the Edge of the Network. *eSecurity Planet*. https://www.esecurityplanet.com/network-security/edge-security.html
King, R.O. (2016). Cyber security for intelligent buildings. *Engineering and Technology Reference,* 1-6. https://doi.org/10.1049/etr.2015.0115

King, R.O. (2016). Cyber security for intelligent buildings. *Engineering and Technology Reference,* 1-6. https://doi.org/10.1049/etr.2015.0115

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society, 20*(1), 14-29. https://doi.org/10.1080/1369118X.2016.1154087

Kobashigawa, J.S., Youn, H., Iskander, M.F., & Yun, Z. (2011). Classification of Buried Targets Using Ground Penetrating Radar: Comparison Between Genetic Programming and Neural Networks. In *IEEE Antennas and Wireless Propagation Letters*, 10, 971-974. https://doi.org/10.1109/LAWP.2011.2167120.

Konaev, M., & Chahal, H. (2021, February 18). Building Trust in Human-Machine Teams. The Brookings Institution. https://www.brookings.edu/techstream/building-trust-in-human-machine-teams/

Kont, M., Pihelgas, M. Wojtkowaik, J., Trinberg, L., & Osula, A.M. (2015). Insider Threat Detection Study. *NATO Cooperative Cyber Defence Centre of Excellence*. https://ccdcoe.org/uploads/2018/10/Insider_Threat_Study_CCDCOE.pdf

Köse, U. (2018). Are we safe enough in the future of artificial intelligence? A discussion on machine ethics and artificial intelligence safety. BRAIN. *Broad Research in Artificial Intelligence and Neuroscience, 9*(2), 184-197.

Kreig, A. (2018). *Defining Remote Warfare: The Rise of the Private Military and Security Industry*. Oxford Research Group Remote Warfare Programme. https://ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/resources/docs/ORG_Defining%20Remote%20Warfare%20Rise%20of%20PM.pdf

Kreugle, H. (1995). *CCTV Surveillance: Video Practices and Technology*. Butterworth-Heinemann.

Krouse, S. (2019, July 19). The new ways your boss is spying on you. *The Wall Street Journal*. https://www.wsj.com/articles/the-new-ways-your-boss-is-spying-on-you-11563528604

Leading Edge. (2020). *How is Cloud Computing Different from Traditional IT Infrastructure*? https://www.leadingedgetech.co.uk/it-services/it-consultancy-services/cloud-computing/how-is-cloud-computing-different-from-traditional-it-infrastructure/

Lethal AWS. (n.d.). *Lethal Autonomous Weapons*. https://autonomousweapons.org/

Levin, S., & Woolf, N. (2016, July 2). Tesla driver killed while using autopilot was watching Harry Potter, witness says. *The Guardian*. https://www.theguardian.com

Liang, K., Heilmann, G., Gregory, C., Diallo, S., Carlson, D., Spell, G., . . . Carin, L. (2018). *Automatic threat recognition of prohibited items at aviation checkpoint with x-ray imaging: a deep learning approach* (Vol. 10632): SPIE.

Li, X., Li, Y., Cui, L., & Liu, W. (2011). Research of new concept sonar-cognitive sonar. *Journal of Marine Science and Application, 10*(4), 502–509.

Lim, K-S., & Lee, C. (2013). A framework for unified digital evidence management in security convergence. *Electronic Commerce Research, 13*(3), 379-398. http://doi.org/10.1007/s10660-013-9119-y

Lim, Y., Cho, Y., So, A., Min, H., & Park, J. (2016). Challenges and opportunities for implementing X-ray scanning technology at the Korean hub ports. *International Journal of Logistics Systems and Management, 25*(4), 513. https://doi.org/10.1504/IJLSM.2016.10000435

Linthicum, D. (n.d.). Edge computing vs fog computing: Definitions and enterprise uses. *Cisco*. https://www.cisco.com/c/en/us/solutions/enterprise-networks/edge-computing.html

Liu, J., Leng, X., & Liu, Y. (2019). Deep Convolutional Neural Network Based Object Detector for X-Ray Baggage Security Imagery. In *Proceedings of the IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 1757-1761). https://doi.org/10.1109/ICTAI.2019.00262.

Lombardi, F., Lualdi, M., Picetti, F., Bestagini, P., Janszen, G., & Di Landro, L.A. (2020). Ballistic Ground Penetrating Radar Equipment for Blast-Exposed Security Applications. *Remote Sensing, 12*(4), 717. https://doi.org/10.3390/rs12040717

Looney, C.G. (1997). *Pattern recognition using neural networks: Theory and alogorithms for engineers and scientists*. Oxford University Press Inc.

Lopatka, K., Kotus, J., & Czyzewski, A. (2016). Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations. *Multimedia Tools and Applications, 75*(17), 10407-10439. https://doi.org/10.1007/s11042-015-3105-4

Lukan, D. (2014, November 21). The top cloud computing threats and vulnerabilities in an enterprise environment. *Cloudtech*. https://www.cloudcomputing-news.net/news/2014/nov/21/top-cloud-computing-threats-and-vulnerabilities-enterprise-environment/

Maruyama, Y. (2020). The Conditions of Artificial General Intelligence: Logic, Autonomy, Resilience, Integrity, Morality, Emotion, Embodiment, and Embeddedness. In B. Goertzel, A.I. Panov, A. Potapov & R.V. Yampolskiy (Eds.), *Proceedings of the Artificial General Intelligence: 13th International Conference* (pp. 242-251). Springer. https://doi.org/10.1007/978-3-030-52152-3

Martelli, T., Bongioanni, C., Colone, F., Lombardo, P., Testa, L., Meta, A. (2016). Security enhancement in small private airports through active and passive radar sensors. In *Proceedings International Radar Symposium*. IEEE. https://doi.org/10.1109/IRS.2016.7497360

Maitland, E. & Sammartino, A. (2014). Decision making and uncertainty: The role of heuristics and experience in assessing a politically hazardous environment. *Strategic Management Journal, 36*(10), 1554-1578. https://doi.org/10.1002/smj.2297

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence, 171*(18), 1174-1182. https://doi.org/10.1016/j.artint.2007.10.009

McClelland, J.L. & Rogers, T.T. (2003). The parallel distributed processing approach to semantic Cognition. *Nature Reviews Neuroscience, 4*, 310-322. https://www.nature.com/articles/nrn1076

McDermott, P., Dominguez, C., Kasdaglis, N., Ryan, M., Trahan, I., & Nelson, A. (2018). Human-Machine Teaming Systems Engineering Guide. *The MITRE Corporation*. https://apps.dtic.mil/sti/pdfs/AD1108020.pdf

Meerts C. (2019) Security: Concepts and Definitions. In L. Shapiro & M.H. Maras (Eds.), *Encyclopedia of Security and Emergency Management* (pp. 1-3). Springer Nature. https://doi.org/10.1007/978-3-319-69891-5_94-2

Mennes, J., Pedersen, T., & Lefever, E. (2019). Approaching terminological ambiguity in cross-disciplinary communication as a word sense induction task: a pilot study. *Language Resources and Evaluation, 53*(4), 889–917. https://doi.org/10.1007/s10579-019-09455-7

Merriam Webster. (n.d.). Radar. In *Merriam Webster Dictionary*. https://www.merriam-webster.com/dictionary/radar

Merriam Webster. (n.d.). Weaponry. In *Merriam Webster Dictionary*. https://www.merriam-webster.com/dictionary/weaponry

Min, H. (2015). Integrating X-ray technologies with intelligent transportation systems for enhancing the international maritime security. *International Journal of Logistics Systems and Management, 22*, 1-14. https://doi.org/10.1504/IJLSM.2015.070888

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the Ire, 49*(1). https://doi.org/10.1109/JRPROC.1961.287775

Miyake, A. & Friedman, N.P. (2012). The nature and organization of individual differences in executive functions: Four general conclusion. *Current Directions in Psychological Science, 21*(1), 8-14. Retrieved from https://doi.org/10.1177/0963721411429458

ML Glossary. (n.d.). Observation. In *ML Glossary*. https://ml-cheatsheet.readthedocs.io/en/latest/glossary.html

Modern Diplomacy. (2020, March 29). Development of New-age Weapons Systems Becomes Key to Sustaining US Military Superiority. https://moderndiplomacy.eu/2020/03/29/development-of-new-age-weapons-systems-becomes-key-to-sustaining-us-military-superiority/

Mohamad, M. M., Taylor, N. K., & Dunnigan, M. W. (2006, September). Articulated robot motion planning using ant colony optimisation. In *2006 3rd International IEEE Conference Intelligent Systems* (pp. 690-695). IEEE.

Mohsin, K. (2020, March 10). Regulation of AI and AI Crimes. *SSRN Solutions*. http://dx.doi.org/10.2139/ssrn.3552140

Monea, C. (2018). Techniques and Equipment for Detection of Prohibited Substances: A Brief Overview. *Journal of Electrical Engineering, Electronics, Control and Computer Science, 4*(14), 7-16. https://jeeeccs.net/index.php/journal/article/view/109/91

Montany, B. (2018). Smart Buildings Report - 2018. https://technology.informa.com/592785/smart-buildings-report-2018

Moore, B. (2020, February 12). Computer vision tech market unchallenged by controversy. *Security*

Brief. https://securitybrief.com.au/story/computer-vision-tech-market-unchallenged-by-controversy

Morgan, D. L. (2019). *Basic and advanced focus groups*. SAGE Publications.

Morrow, T. (2018, March 5). 12 Risks, Threats & Vulnerabilities in Moving to the Cloud. *Carnegie Mellon University*. https://insights.sei.cmu.edu/sei_blog/2018/03/12-risks-threats-vulnerabilities-in-moving-to-the-cloud.html

Musil, K. (2017). Potential and Limits of Ultra-Trace Detection of Explosives. *The Science for Population Protection, 2*.

National Research Council. (2004). *Opportunities to Improve Airport Passenger Screening with Mass Spectrometry*. The National Academies Press. https://doi.org/10.17226/10996.

New Scientist. (2020, June 26). US Military electroshock weapon can hit a person 100m away. https://www.newscientist.com/article/2247049-us-military-electroshock-weapon-can-hit-a-person-100-metres-away/

Nguyen, V.T., Villain, F., & Le Guillou, Y.L. (2012). Cognitive Radio RF: Overview and Challenges. *VLSI Design, 2012*, 1-13. https://doi.org/10.1155/2012/716476

Norman, T.L. (2011). *Electronic Access Control*. Elsevier Science.

Norman, T. L. (2014). *Integrated security systems design : A complete reference for building enterprise-wide digital security systems*. https://ebookcentral.proquest.com

Norris, W.R., & Patterson, A.E. (2019). Automation, Autonomy, and Semi-Autonomy: A Brief Definition Relative to Robotics and Machine Systems. https://www.ideals.illinois.edu/bitstream/handle/2142/104214/Semi_Autonomous_System_R3_Ref_Essay.pdf?sequence=2&isAllowed=y

O'Carroll, B. (2020, January 31). What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence. *Codebots*. https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible

Office of Technology Assessment. (1993). *Technology against terrorism: structuring security*. United States Congress, Global Professional Publications.

Oh, H., & Lee, Y. (2019). Exploring Image Reconstruction Attack in Deep Learning Computation Offloading. *The 3rd International Workshop on Deep Learning for Mobile Systems and Applications,* 19-24. https://doi.org/10.1145/3325413.3329791

Onwubiko, C. (2010). Security Issues to Cloud Computing. In *Cloud Computing* (pp.271-288). Springer. https://doi.org/10.1007/978-1-84996-241-4_16

Open Access Government. (2016, March 24). Automated Comparison of X-ray Images for cargo Scanning (ACXIS). *Open Access Government*. https://www.openaccessgovernment.org/automated-comparison-x-ray-images-cargo-scanning-acxis/25162/

Pacom Systems Pty Ltd (2019). *Unison System Overview*. https://pacom.com/portal/wp-content/uploads/2019/10/PACOM-Unison-System-Overview-datasheet.pdf

Pandey, A., Pandey, S., & Parhi, D.R. (2017). Mobile robot navigation and obstacle avoidance techniques: A review. *International Robotics & Automation Journal, 2*(3), 00022. https://doi.org/10.15406/iratj.2017.02.00023

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Partnoy, F. (2018, September). What your boss could learn by reading the whole company's emails. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2018/09/the-secrets-in-your-inbox/565745/

Pennington, D.C., Flanagan, C., & Banyard, P. (2000). *Social cognition*. Routledge.

Perelman, B. (2019, June 25). Mission Possible: ICS Attacks on Buildings are a Reality. *SecurityWeek*. https://www.securityweek.com/mission-possible-ics-attacks-buildings-are-reality

Pettit, H. (2020, June 29). US Marines' new weapon can electrocute enemies from 300 feet away – and is so dangerous they had to make it LESS deadly. *The Sun*. https://www.thesun.co.uk/tech/11978600/us-marines-specter-taser-weapon-electrocute-deadly/

Phillip, A. (2015, July 2). Robot grabs man, kills him in German car factory. *The Washington Post*. https://www.washingtonpost.com/news/morning-mix/wp/2015/07/02/robot-grabs-man-kills-him-in-german-car-factory/

Picard, R. W. (2000). *Affective computing*. MIT press.

Pilz Australia Safe Automation. (2018). *Revision of the ISO safety standard for industrial robots*. https://www.pilz.com/en-AU/company/news/articles/196225

Popoola, O.P., & Wang, K. (2012). Video-Based Abnormal Human Behaviour Recognition—A Review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42*(6), 865-878. https://doi.org/10.1109/TSMCC.2011.2178594

Porter, J. (2019, August 14). Huge security flaw exposes biometric data of more than a million users. *The Verge*. https://www.theverge.com/2019/8/14/20805194/suprema-biostar-2-security-system-hack-breach-biometric-info-personal-data

Proud, R.W., and Hart, J.J. (2005). FLOAAT, A Tool for Determining Levels of Autonomy and Automation, Applied to Human-Rated Space Systems. *Infotech@Aerospace*, 7061. American Institute of Aeronautics and Astronautics (AIAA). Arlington, VA: AIAA.

Pueyo, S. (2018). Growth, degrowth, and the challenge of artificial superintelligence. *Journal of Cleaner Production, 197*, 1731–1736. https://doi.org/10.1016/j.jclepro.2016.12.138

Purpura, P.P. (2019). *Security and loss prevention: an introduction* (7th ed.). Butterworth-Heinemann. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5528118

Purpura, P. P., Fennelly, L. J., Honey, G., & Broder, J. F. (2013). Chapter 9 - Security Lighting. In L. J. Fennelly (Ed.), *Effective Physical Security (Fourth Edition)* (pp. 181-189): Butterworth-Heinemann.

Rai, M., Husain, A.A., Maity, T., & Yadav, R.K., (2018). Advance Intelligent Video Surveillance System (AIVSS): A Future Aspect. *Intelligent Video Surveillance, 2019*, 1-20. https://doi.org/10.5772/intechopen.76444.

Ranger, S. (2018, December 13). What is Cloud Computing? Everything you need to know about the cloud explained. *ZD Net*. https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-about-the-cloud/

Ranti, C., Chatham, C.H. & Badre, D. (2015). Parallel temporal dynamics in hierarchical cognitive control. *Cognition, 142*, 205-229. https://doi:10.1016/j.cognition.2015.05.003

Raywood, D. (2017, August 21). Attacks on the Cloud Increase by 300%. *Info Security Magazine*. https://www.infosecurity-magazine.com/news/attacks-cloud-increase-300/

Rebala, G., Ravi, A., & Churiwala, S. (2019). *An introduction to machine learning*. Springer. https://doi.org/10.1007/978-3-030-15729-6

Rizkallah, J. (2018, June 15). Red Teaming + Artificial intelligence: A Match Made in Cybersecurity Heaven. *Security Magazine*. https://www.securitymagazine.com/articles/89133-red-teaming-artificial-intelligence-a-match-made-in-cybersecurity-heaven

Robotech Vision, (n.d.). Intelligent CCTV. https://robotechvision.com/intelligent-cctv/

Robotic Assistance Devices. (n.d.). ROAMEO is ready for action! https://roboticassistancedevices.com/

Robotics Online. (2017, May 9). *Robot Safety Standards: A Brief Overview*. https://www.robotics.org/blog-article.cfm/Robot-Safety-Standards-A-Brief-Overview/58

Rockwell Automation. (n.d.). *Is it ok to bypass a machine safety function?* https://bit.ly/2Bu4LhR

Roman, R., Lopez, J., & Mambo, M. (2016). Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges. *Future Generation Computer Systems, 78*, 680-698. http://dx.doi.org/10.1016/j.future.2016.11.009

Romanov, S., Kharkovchuk, N., Sinelnikov, M., Abrash, M., & Filinkov, V. (2020). *Development of an Non-Speech Audio Event Detection System*. Paper presented at the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Moscow, Russia.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. N*ature*, *323*(6088), 533-536.

Russell, S.J., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.

Sackur, J. & Dehaene, S. (2009). The cognitive architecture for chaining of two mental operations. *Cognition, 111*(2), 187-211. https://doi.org/10.1016/j.cognition.2009.01.010

Safety of Autonomous Systems Working Group. (2018). *Safety-related Challenges for Autonomous Systems*. https://scsc.uk/r143:1

Saini, D.K., Ahir, D., & Ganatra, A. (2016). Techniques and Challenges in Building Intelligent Systems: Anomaly Detection in Camera Surveillance. *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems, 2*(51), 11-21. https://doi.org/10.1007/978-3-319-30927-9_2

Sample, I. (2019, July 29). What is facial recognition? *The Guardian.* https://www.theguardian.com/technology/2019/jul/29/what-is-facial-recognition-and-how-sinister-is-it

Sandford, A. & Burton, A.M. (2014). Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. *Cognition, 132*, 262-268. https://doi.org/10.1016/j.cognition.2014.04.005

Sandia National Laboratories. (n.d.). Access Delay. https://share-ng.sandia.gov/itc/assets/12-access-delay-text.pdf

Sathya, R., & Abraham, A. (2013). Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence, 2*(2), 34-38.

Sawhney, M. (2020, January 27). Why Apple and Microsoft are Moving AI to the Edge. *Forbes.* Retrieved from https://www.forbes.com/sites/mohanbirsawhney/2020/01/27/why-apple-and-microsoft-are-moving-ai-to-the-edge/#24b95c732570

Scientific Committee on Emerging and Newly Identified Health Risks. (2012). Health effects of security scanners for passenger screening. *European Commission.* https://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_036.pdf

Security Informed. (2016, July 18). Video Surveillance Advancements Lead to Data Storage Challenges. *Securityinformed.com*. https://www.securityinformed.com

Security World Market. (2019, January 17). AI Will transform the future of home security. https://www.securityworldmarket.com/int/News/Business-News/ai-will-transform-the-future-of-home-security1

Serrano, J., Moros, J., Sánchez Carlos, Macías Jorge, & Laserna, J. J. (2014). Advanced recognition of explosives in traces on polymer surfaces using LIBS and supervised learning classifiers. *Analytica Chimica Acta, 806*, 107–116. https://doi.org/10.1016/j.aca.2013.11.035

Sharkey, N. (2016). Staying in the loop: human supervisory control of weapons. In N. Bhuta, S. Beck, R. Geiß, H.Y. Liu, & C. Kreß (Eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*. Cambridge University Press.

Shaw, K. (2019, November 13). What is edge computing and why it matters. *Network World*. https://www.networkworld.com/article/3224893/what-is-edge-computing-and-how-it-s-changing-the-network.html

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *Ieee Internet of Things Journal, 3*(5), 637-646. https://doi.org/10.1109/JIOT.2016.2579198

Shortcliffe, E.H., & Buchanan, B.G. (1975). A model of inexact reasoning in medicine. *Mathematical biosciences*, *23*(3-4), 351-379. https://doi.org/10.1016/0025-5564(75)90047-4

Singh, B., Singh, D., Singh, G., Sharma, N., & Sibbal, V. (2014). Motion Detection for Video Surveillance. *Proceedings of the 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT)*. https://doi.org/10.1109/ICSPCT.2014.6884919

Sloman, S.A. (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin, 119*(1), 3-22. https://doi.org/10.1037/0033-2909.119.1.3

Smith, C. & Robinson, M. (1999). The Understanding of Security Technology and its Applications. In *Proceedings of the IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology* (pp, 26-37). IEEE. https://doi.org/10.1109/CCST.1999.797888

Smith, P. (2018, June 26). Drones and Geofencing: How it Works, Benefits and Requirements. *Dronebelow*. https://dronebelow.com/2018/06/26/drones-and-geofencing-how-it-works-benefits-requirements/

Smith & Wesson. (2011, August 1). *Smith & Wesson Security Solutions Announces Successful Certification Test of New Active Vehicle Barrier*. https://www.prnewswire.com/news-releases/smith--wesson-security-solutions-announces-successful-certification-test-of-new-active-vehicle-barrier-126507933.html

Snow, J. (2018, July 26). Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots. *American Civil Liberties Union*. https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Sonardyne. (n.d.). Sentinel IDS Diver Detection Sonar System. https://www.sonardyne.com/product/sentinel-diver-detection-sonar/

SpotterRF. (n.d.). SpotterRF Perimeter Surveillance Radars. https://www.spotterrf.com/

Spy Monkey Surveillance (n.d.). *The Ultimate CCTV Security Surveillance Buying Guide in Australia*. https://www.spymonkey.com.au/pages/the-ultimate-cctv-security-surveillance-buying-guide-in-australia

Sreenu, G., Durai, M.A.S. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. *J Big Data, 6*(48). https://doi.org/10.1186/s40537-019-0212-5

Stanley, K.O., & Miikkulainen, R. (2002). Evolving Neural Networks through Augmenting Typologies. *Evolutionary Computation, 10*(2), 99-127. https://doi.org/10.1162/106365602320169811

Stoltzfus, J. (2020). How are logic gates precursors to AI and building blocks for neural networks? Retrieved from https://www.techopedia.com/how-are-logic-gates-precursors-to-ai-and-building-blocks-for-neural-networks/7/33018

Streitz, N., & Privat, G. (2009). Ambient Intelligence. In: C. Stephanidis (ed.) *The Universal Access Handbook*, pp. 60.1-60.17. CRC Press, Taylor & Francis Group.

Sukhodolov, A.P., Bychkov, A., & Bychkova, A.M. (2020). Criminal Policy for Crimes Committed

Using Artificial Intelligence: State, Problems, Prospects. *Journal of Siberian Federal University, 13*(1), 116-122. https://doi.org/10.17516/1997-1370-0542

Suryateja, P.S. (2018). Threats and vulnerabilities of cloud computing: A review. *International Journal of Computer Sciences and Engineering, 6*(3), 297-302. https://doi.org/10.26438/ijcse/v6i3.297302

Tawiah, T.A-Q. (2010). *Video content analysis for automated detection and tracking of humans in CCTV surveillance applications* (Doctoral Dissertation). https://bura.brunel.ac.uk/bitstream/2438/7344/1/FulltextThesis.pdf

Tchernykh, A., Schwiegelsohn, U., Talbi, E.G., & Babenko, M. (2016). Towards understanding uncertainty in cloud computing with risks of confidentiality, integrity, and availability. *Journal of Computational Science, 36*, 1-9. https://doi.org/10.1016/j.jocs.2016.11.011

Teledyne UK. (2020). *RF Safe-Stop*. https://www.teledyne-e2v.com/markets/security-rescue/rf-safe-stoptm/

Teschler, L. (2018, August 13). The tricky business of teaching cars to see. *EE World Online*. https://www.eeworldonline.com/the-tricky-business-of-teaching-cars-to-see/

The Free Dictionary. (n.d.). Symbolic Systems. In *The Free Dictionary*. Farlex. Retrieved November 12, 2020, from https://bit.ly/3neMtng

Tolun, M.R., Sahin, S., & Oztoprak, K. (2016). Expert Systems. In *Kirk-Othmer Encyclopedia of Chemical Technology*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471238961.0524160518011305.a01.pub2

Trevithick, J. (2020, July 8), Navy to Add Laser Weapons to At Least Seven ore Ships in The Next Three Years. The Drive. https://www.thedrive.com/the-war-zone/34663/navy-to-add-laser-weapons-to-at-least-seven-more-ships-in-the-next-three-years

Tsui, A.B.M. & Fullilove, J. (1998). Bottom-up or Top-down Processing as a Discriminator of L2 Listening Performance. A*pplied Linguistics, 19*(4), 432-451.

Turner, D.M., & Scholten, U. (2019, May 17). Machine Learning and Artificial Intelligence in Radar Technology. *Sky Radar*. https://www.skyradar.com/blog/machine-learning-and-artifical-intelligence-in-radar-technology

Tzu, S. (1971). *The Art of War*. Axiom.

United Nations. (2015). Less Than Lethal Weapons. *United Nations Public Order Management.* http://repository.un.org/bitstream/handle/11176/387390/Less%20Than%20Lethal%20Weapons.pdf?sequence=24&isAllowed=y

Urbanowicz, R. J., & Moore, J. H. (2009). Learning classifier systems: a complete introduction, review, and roadmap. *Journal of Artificial Evolution and Applications*, *2009*.

US Department of Homeland Security. (2013). *Digital Video Quality Handbook*. https://www.dhs.gov/sites/default/files/publications/VQiPS_Digital-Video-Quality-HB_UPDATED-180117-508.pdf

Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F., & Sarti, A. (2007). *Scream and gunshot detection and localization for audio-surveillance systems.* Paper presented at the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance.

Varghese, S. (2019, February 21). ESET researcher outlines horrors that can result from siegeware. *IT Wire*. https://www.itwire.com/security/86114-eset-researcher-outlines-horrors-that-can-result-from-siegeware.html

Venezia, P. (2014, June 23). Murder in the Amazon Cloud. *Infoworld*. https://www.infoworld.com/article/2608076/murder-in-the-amazon-cloud.html

Videosurveillance.com. (n.d.). Benefits of Network Security Cameras & Wireless Network Cameras. *Bastion Security*. https://www.videosurveillance.com/tech/benefits-of-network-cameras.asp

Vikhar, P.A. (2016). Evolutionary algorithms: A critical review and its future prospects. In *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)* (pp. 261-265). https://doi.org/10.1109/ICGTSPICC.2016.7955308.

Vincent, J. (2016, July 13). Mall security bot knocks down toddler, breaks Asimov's first law of robotics. *The Verge*. https://www.theverge.com/2016/7/13/12170640/mall-security-robot-k5-knocks-down-toddler

Vincent, J. (2018, June 26). This Japanese AI security camera shows the future of surveillance will be automated. *The Verge*. https://www.theverge.com/2018/6/26/17479068/ai-guardman-security-camera-shoplifter-japan-automated-surveillance

Vijaya Kumar, B. (2011). Biometric Matching. In H. v. Tilborg & S. Jajodia (Eds.), *Encyclopedia of cryptography and security*. Springer.

Vlassis, N. (2003). Multiagent Systems and Distributed AI. *Intelligent Autonomous Systems, Informatics Institute, University of Amsterdam*.

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience, 2018*. https://doi.org/10.1155/2018/7068349

Walloschek, T., Stocker, E., & Grobauer, B. (2011, August 15). Understanding Cloud Computing Vulnerabilities. *InfoQ*. https://www.infoq.com/articles/ieee-cloud-computing-vulnerabilities/

Ward, C. (2018, July 16). What is Hybrid Cloud? *DZone*. https://dzone.com/articles/what-is-hybrid-cloud

Wheeler, A.C. (2017, November). Directed Energy Weapons [Discussion Paper]. *Convention on Certain Conventional Weapons (CCW)*. Geneva, Switzerland.

Wilson, J.R. (2017, February 1). Technology comes to bear on radar and sonar. *Military & Aerospace Electronics*. https://www.militaryaerospace.com/communications/article/16709846/technology-comes-to-bear-on-radar-and-sonar

Wiser, B. (2017, April 1). Out of Many, One. *Security Industry Association*. Retrieved from https://www.securityindustry.org/2017/04/01/integration/

Wolfe, F. (2019, December 2). Companies Developing Lethal Autonomous Weapons, As Groups Seek Ban, Report Says. Aviation Today. https://www.aviationtoday.com/2019/12/02/companies-developing-lethal-autonomous-weapons-as-groups-seek-ban-report-says/

Wolf, N. (2012, August 16). The new totalitarianism of surveillance technology. *The Guardian*. https://www.theguardian.com/commentisfree/2012/aug/15/new-totalitarianism-surveillance-technology

Wu, X. (2017). *Metrics, Techniques and Tools of Anomaly Detection: A Survey*. https://www.cse.wustl.edu/~jain/cse567-17/ftp/mttad/index.html

Yampolskiy, R.V., & Spellchecker, M.S. (2016). *Artificial intelligence safety and cybersecurity: A timeline of AI failures*. https://arxiv.org/ftp/arxiv/papers/1610/1610.07997.pdf

Yinon, J. (2007). *Counterterrorist Detection Techniques of Explosives* [1 online resource (455 pages)]. http://www.sciencedirect.com/science/book/9780444522047

Yusof, M.A., & Kabir, S. (2012). An overview of sonar and electromagnetic waves for underwater communication. *IETE Technical Review, 29*(4), 307–317. https://doi.org/10.4103/0256-4602.101312

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338-353.

Zhang, Q., Yao, J., & Zha, Y. (2018). Learning Behaviour Trees for Autonomous Agents with Hybrid Constraints Evolution. *Applied Sciences, 8*(1077), 1-22. https://doi.org/10.3390/app8071077

Zhuang, X., Zhou, X., Hasegawa-Johnson, M., & Huang, T. (2010). Real-world acoustic event detection. *Pattern Recognition Letters, 31*(12), 1543 - 1551. https://doi.org/10.1016/j.patrec.2010.02.005

Zu, X., Guo, F., Huang, J., Zhao, Q., Liu, H., Li, B., & Yuan, X. (2017). Design of an Acoustic Target Intrusion Detection System Based on Small-Aperture Microphone Array. *Sensors (Basel, Switzerland), 17*(3). https://doi.org/10.3390/s17030514

Zöhrer, M., & Pernkopf, F. (2016, September). *Gated recurrent networks applied to acoustic scene classification and acoustic event detection* [Conference paper]. Detection and Classification of Acoustic Scenes and Events 2016, Budapest, Hungary.

Zsigmond, J. (2020, June 14). Creating a Neural Network from Scratch. *Towards Data Science*. https://towardsdatascience.com/creating-a-neural-network-from-scratch-302e8fb61703

Zussman, Y. (2018, October 1). Complete Perimeter Solution: Radar and deep learning technologies finally come together. *Security Today*. https://securitytoday.com/Articles/2018/10/01/Complete-Perimeter-Solution.aspx?Page=1

5.0 Appendices

Appendix A - Indicative Security Technology Tables Presented to Participants

Table A1

Observing Technology and Indicative AI Analytical Alignment.

Technology		**Function/ Performance Measure**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Network Video Surveillance	Digital Video Surveillance	Observe, Detect, Recognise, Identify	Reflected Scene Image	CCTV cameras used to monitor and record a defined space	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
	Video Motion Detection	Probability of Detection	Movement within a Frame	CCTV cameras detect movement of an individual into a defined space	Statistical Probability (Probabilistic Method)

Table A2

Detecting Technologies and Indicative AI alignment

Technology		**Function**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Intrusion Detection Systems	Contact Switch/ Reed Switch (Electromechanical Sensor)	Detection	Current Interference (loss) through breaking of connection between the two points	Reed switches on windows or doors to detect opening	Expert Systems (Rule Setting)
	Glass Break Sensor (Electro-mechanical Sensor)	Detection	Vibration or acoustic technology – elevations in vibration or audio levels across a surface	Glass break sensors used on windows to detect glass being broken or manipulated	Expert Systems (Rule Setting)
	Passive Audio Sensor	Detection	Elevations in audio levels	Sensors detecting noise caused by intruder movement through detection volume	Expert Systems (Rule Setting)
	Passive Microphonic Cables	Detection	Detection of vibrations along line	Microphonic cables used to detect fences being cut or climbed on	Expert Systems (Rule Setting)

	Photoelectric Beams (visible or infrared)	Detection	Interruption or change of amount of light crossing light axis and arriving at receiver unit	Photoelectric beams used to detect fences being cut or climbed on	Expert Systems (Rule Setting)
	Active Infrared Motion Sensor	Detection	Interference of infrared beam between two points	AIR activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Fibre Optic Cable Sensors	Detection	Loss of signal due to damage or break in fibre loop	Activated when cable is damaged or cut – used to protect CCTV and other systems cabling	Expert Systems (Rule Setting)
	Microbend Fibre Optic Sensor	Detection	Loss of signal due to bend or movement in fibre loop	Activated when cable is bent or moved - used to protect CCTV and other systems cabling	Expert Systems (Rule Setting)
	Passive Infrared Motion Sensor	Detection	Increased Change in thermal energy	PIR activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Active Ultrasonic Motion Sensor	Detection	Changes in reflected Ultrasonic energy	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Active Monostatic Microwave Motion Sensor	Detection	Changes in reflective electromagnetic energy wave	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Passive Infrared Sensors (Thermopile or Pyroelectric Detector)	Detection	Changes in energy emitted by human intruder moving through background thermal energy	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Capacitance Proximity Sensor	Detection	Change in capacitance (electrical charge) between protection loop and ground	Activated when an intruder moves near to or touches metal object	Expert Systems (Rule Setting)
	Pressure Sensor	Detection	Electrical contact between ribbon switches due to pressure	Activated when intruder passes over sensor entering protected area	Expert Systems (Rule Setting)
	Light Sensors	Detection	Change in predetermined light level	activated when intruder walks into a room and across the detection volume	Expert Systems (Rule Setting)
	Dual Technology Sensors	Detection	AND gate logic requiring activation of both types of combine sensors	Combination sensors used to overcome vulnerabilities of above individual sensors	Expert Systems (Rule Setting)

	Alarm Annunciator Panel	Detect, Response	Inputs from field sensors communicate audio visual signals for faults any failures of the critical system or process	Monitoring of door status at guarding station	Expert Rule-setting
	Duress Alarms	Detect, Response	Command Signal; Emergency signal transmitted from duress mechanism to central monitoring station	Magnetic reed switch activated upon gun being removed from holster of armed guard	Expert Rule-setting
Lighting Systems	Active Lighting Systems	Detection	Activation of sensors (PIR, Infrared, Ultrasonic, Microphonic)	Used to automatically illuminate an area upon activation of sensor	Expert Systems (Rule Setting). Probability (Past Performance)
	Passive Lighting Systems (Failsafe/ Emergency Lighting)	Detection	Failure of electrical system	Used to automatically illuminate an area upon failure of electricity	Expert Systems (Rule Setting)
X Ray	X-Ray	Detection	X-ray electromagnetic wave interaction with material	Used to detect presence of non-organic matter on human body	Probabilistic – Image Recognition and Comparison
Trace Detection	Explosive/ Narcotics Trace Detection (Ion Mobility Spectrometry)	Detection	Change in relative concentration of explosive molecules in saturated air	Used to detect explosives in parcels and luggage	Expert Systems (Rule Setting). Probability (Past Performance)
Acoustic Detection Systems	Acoustic Anomaly Detection	Detection	Audio sensors detect acoustic signature of pre-defined actions in each detection volume	Used to identify gunshot or aggressive behaviour	Comparative analysis using deep learning algorithms and inference
	Intelligent Acoustic Surveillance (Acoustic Event Detection)	Detect	Nominal and Anomalous Acoustic Input	Deployed where there may be privacy issues with video surveillance (e.g. bathroom facilities) to 'learn' usual noise and detects abnormal noise (e.g. scream)	Statistical Probability (Probabilistic Method) AND Statistical Machine Learning (Supervised, Unsupervised or Reinforcement Learning)
RADAR Systems	Radar Video Surveillance (RVS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR), Reflected Scene Image (Video)	Radar video system observes and detects an individual intruding into a defined space (e.g. rural airport)	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
	Ground Penetrating Radar (GPR)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR),	GPR transmits electromagnetic waves to detect underground shelters, landmines or buried explosives	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)

	Through-the-Wall Sensor (TTWS)	Observe, Detect, Recognise	Reflected electromagnetic waves (RADAR),	TTWS device transmits RF energy through a wall and monitors the reflected energy to detect person moving in a room	Expert Systems (Rule Setting) and Statistical Probability (Probabilistic Method)
SONAR Systems	Active Sonar Systems	Detect, Recognise	Transducers (hydrophone array) emit and detect acoustic energy reflected/ emitted from sound/energy source	Sonar deployed in a port/harbour to detect underwater divers or vessels approaching docked vessels	Expert Systems (Rule Setting)
	Passive Sonar Systems	Detect, Recognise	Transducer (hydrophone array) detects acoustic energy emitted from sound/energy source	Passive Sonar deployed by an offshore oil rig to detect approaching vessels	Expert Systems (Rule Setting)

Table A3

Controlling Technology and Indicative AI Alignment

Technology	**Function**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
EAC Reader	Physical Access Control Functions	Presentation or input (e.g. PIN) of authorised credential at reader	Input of a PIN into a PIN pad reader	Expert Systems (Rule Setting)
Smart Card	Authentication Access Control Technology	Presentation of authorised credential at reader	Use of access control card containing pre-programmed identity information (access rights)	Expert Systems (Rule Setting)
2FA (Two Factor Authentication) Access Control	Authentication Access Control Technology	Presentation of two pre-authorised credentials. Combination of one Knowledge factor (e.g. PIN), one Possession factor (e.g. Smart Card) or one Inherent factor (Biometric credential)	Access control point requiring a key card and a PIN	Expert Systems (Rule Setting)
Biometric	Authentication Access Control Technology	Presentation and conversion of biometric data into algorithm that measures probability of match with pre-enrolled authorised data	Retinal eye scan/ Fingerprint reader	Probabilistic Programming - Algorithm Probability

RFID (Radio Frequency Identification)	Communication Access Technology	Presentation of RFID embedded chip to RFID reader based on pre-set permissions	Use of RFID Access card at an RFID reader	Expert Systems (Rule Setting)
NFC (Near Field Communication)	Communication Access Technology	Generation of electromagnetic fields created when two devices are brought together (e.g. smart phone & reader)	Using Smart phone and device reader to gain access	Expert Systems (Rule Setting)
Internet of Things Access Control (Intelligent Door Locking Systems)	Access Technology Architecture	Presentation of authorised credential at any one of connected access control readers using unique IP addresses and connected through the internet		Expert Systems (Rule Setting)
Bluetooth Access Control	Communication Access Technology	Presentation of Bluetooth enabled and pre-authorised device at Bluetooth reader using Wi-Fi or cellular data network		Expert Systems (Rule Setting)

Table A4

Responding Technology and Indicative AI Alignment

Technology		**Function/ performance measure**	**Stimuli**	**Example**	**Indicative Analytical Methodology**
Communication Systems	Telephones, Mobile Phones, Pagers, Emergency Phones	Response	Command Signal; Transmitter converts electrical signal into radio waves which are radiated from the device antenna to the radio receiver (mobiles) or transmits electrical signal via wire/ fibre optics (landlines)	Emergency phone in car park used to communicate with security staff	Expert Rule-setting
	Radio Communication	Response	Command Signal; Radio transmitter transforms energy into radio waves radiated from the device antenna to the radio receiver	Spread spectrum, Two Way Radio Communication	Expert Rule-setting

	Emergency Warning Intercom Systems (EWIS) Public Address / Mass Notification System	Control, Response	Command Signal by operator/ user; Transmission of voice audio or automated instructions via IP network protocol (e.g. VoIP, RTSP, TCP/IP, SIP)	EWIS system used for emergency evacuation of a facility	Expert Rule-setting
	IP Network Connected Intercom	Control, Response	Command signal by operator/user; transmits audio and visual data via IP network protocol (e.g. TCP/IP, SIP, RTSP)	Intercom installed at remote/ isolated substation and monitored/ controlled from central site.	Expert Rule-setting
	Wireless Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via encrypted wireless signal (radio waves)	Intercom installed within secure facility between access control points	Expert Rule-setting
	Wired Intercom	Control, Response	Command signal by operator/user transmitting audio and visual data via low-voltage electrical wiring	Intercom installed at vehicle gates to allow entry to authorised personnel	Expert Rule-setting
	Door Release Intercom	Control, Response	Command signal by operator/user allowing remote unlocking of doors via wiring from the intercom to the electrified locking hardware	Intercom installed at entrance door to allow access to authorised personnel	Expert Rule-setting
Dispensable Barriers	Active Dispensable Barrier	Delay, Response	Command Signal	Fog or smoke dispensed to obscure an adversary's vision	Expert Rule-setting
	Passive Dispensable Barriers	Delay	Contact with Adversary	Asset covered in Sticky foam entangles attacker and tools during an attempt to remove asset	Expert Rule-setting
Vehicle Barriers (Hostile Vehicle Mitigation)	Rising Bollards	Delay, Response	Command signal activates bollard to deploy	High security entry/ exit points	Expert Rule-setting

	Rising Wedge Barrier	Delay, Response	Command signal activates wedge barrier to deploy	High security entry/ exit points	Expert Rule-setting
	Beam Barrier	Delay, Response	Command signal activates beam barrier to deploy	Security entry/ exit points	Expert Rule-setting
	Automatic Gates	Detect, Delay, Response	Command Signal, PLUS Presentation of Authorised Credential	Security entry/ exit points (includes sliding, swing gates, bi-fold gates etc.).	Expert Rule-setting
Weaponry	Acoustic Weaponry	Delay, Response	Sonic, infrasonic, or ultrasonic frequencies. High-amplitude infrasonic weapons combine two high-volume ultrasonic frequencies at a difference of 8Hz	Passenger cruise ship using non-lethal Long-Range Acoustic Device (LRAD) to create ear-piercing noise to repel pirates	Expert Rule-setting
	High-Voltage (Electroshock) Weaponry	Delay, Response	Transducers emit high-voltage, low amplitude current to incapacitate adversary	Crowd dispersal using non-lethal electroshock ‘sweeping stun gun’	Expert Rule-setting
	Electromagnetic Weaponry (Directed Energy Weapon)	Delay, Response	Electromagnetic energy beam created by a Gyrotron (Vacuum Electronic Device (VED) to create high power high frequency THz radiation)	High power microwave weapon such as the Active Denial System or “Heat Ray” delivering non-lethal radiation to adversary	Expert Rule-setting
	Optical Distractors (Directed Energy Weapon)	Delay, Response	High intensity beams of light directed at adversary to create nausea, confusion, disorientation, or temporary blindness	Long-range dazzling laser directed at pirates attempting to board merchant ship	Expert Rule-setting
	Remotely Operated Weapons Systems (ROWS)	Delay, Response	Weapons operated remotely in response to event or attack	Deployed at critical installations in volatile security environments	Expert Rule-setting

	Autonomous & Semi-Autonomous Weapons Systems	Delay, Response	Predefined Stimulus for which autonomous systems may execute self-direction, self-learning or emergent behaviour to select and attack targets	Autonomous drones used to track and engage targets	**Semi-autonomous**: Statistical (Probabilistic & Machine Learning) **Autonomous**: Sub symbolic (Embodied Intelligence)

Appendix B - Indicative Alignment of Security Technologies with AI Paradigms Presented to Participants

Figure B1

Indicative alignments between Observation technologies, BACS Architecture Levels and AI computational tasks aligned to the spectrum of AI paradigms.

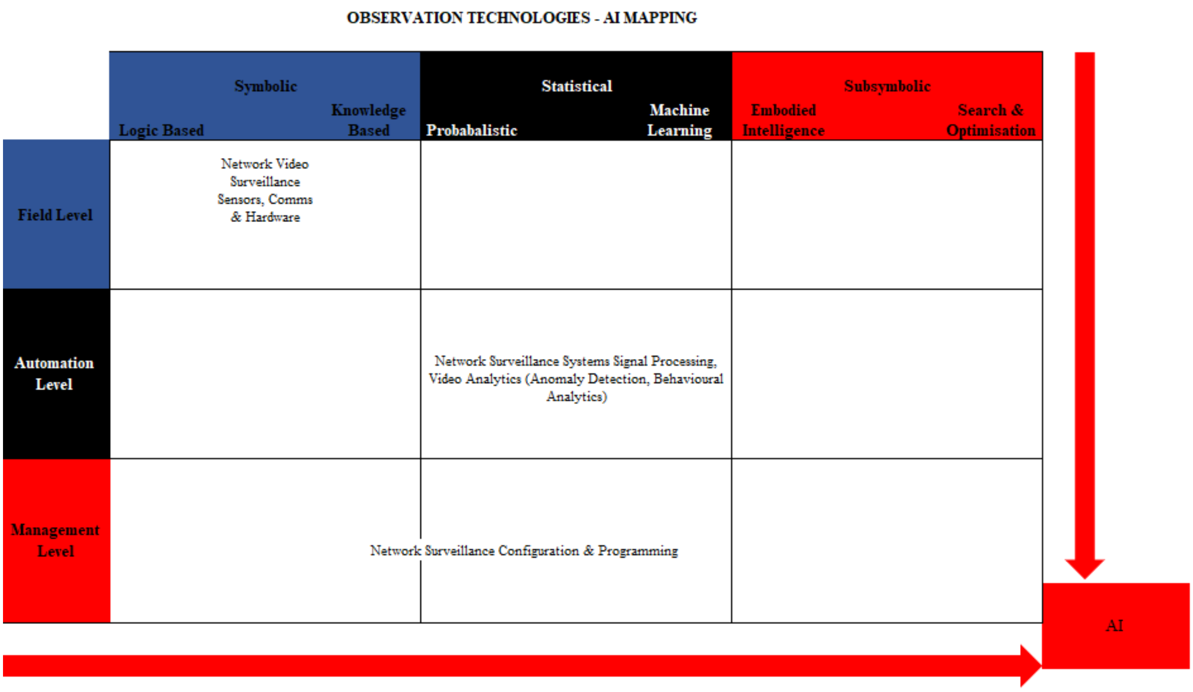


Figure B2

Indicative alignments between Detection technologies, BACS Architecture Levels and AI computational tasks aligned to the spectrum of AI paradigms.

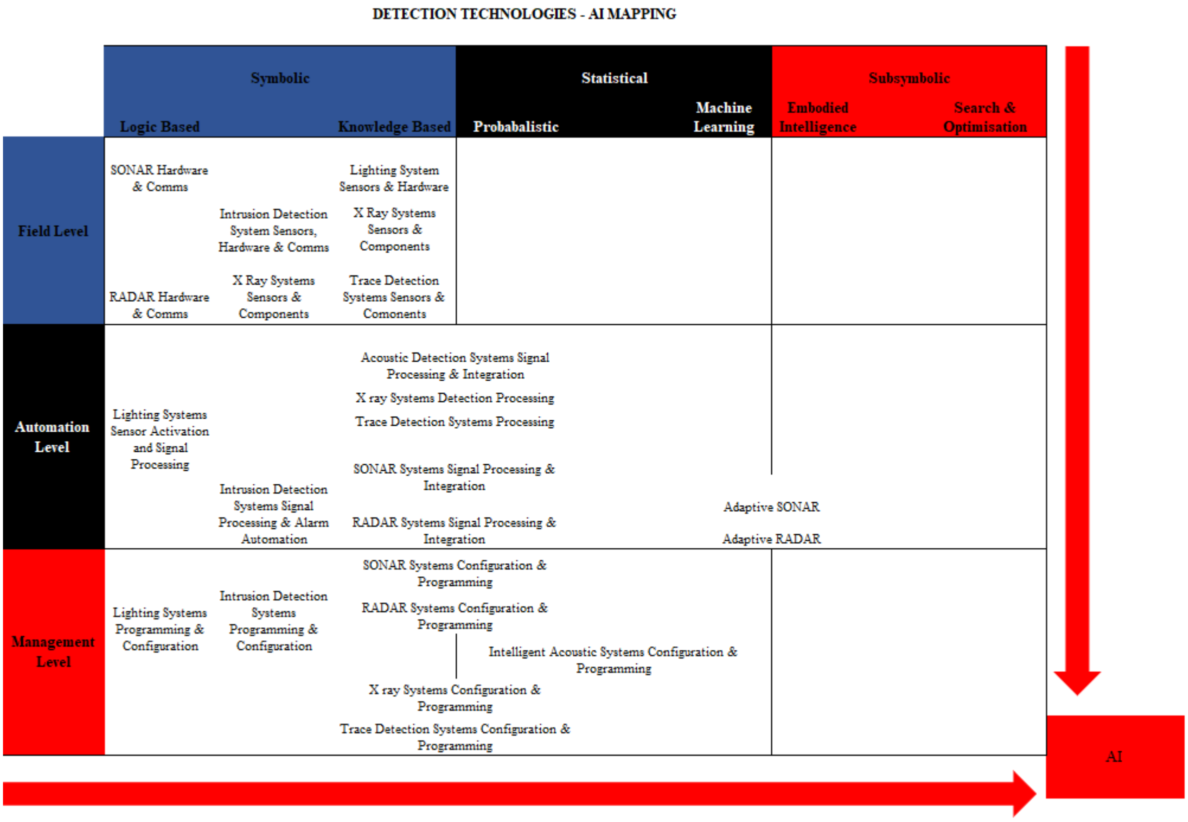


Figure B3

Indicative alignments between Controlling technologies, BACS Architecture Levels and AI computational tasks aligned to the spectrum of AI paradigms.

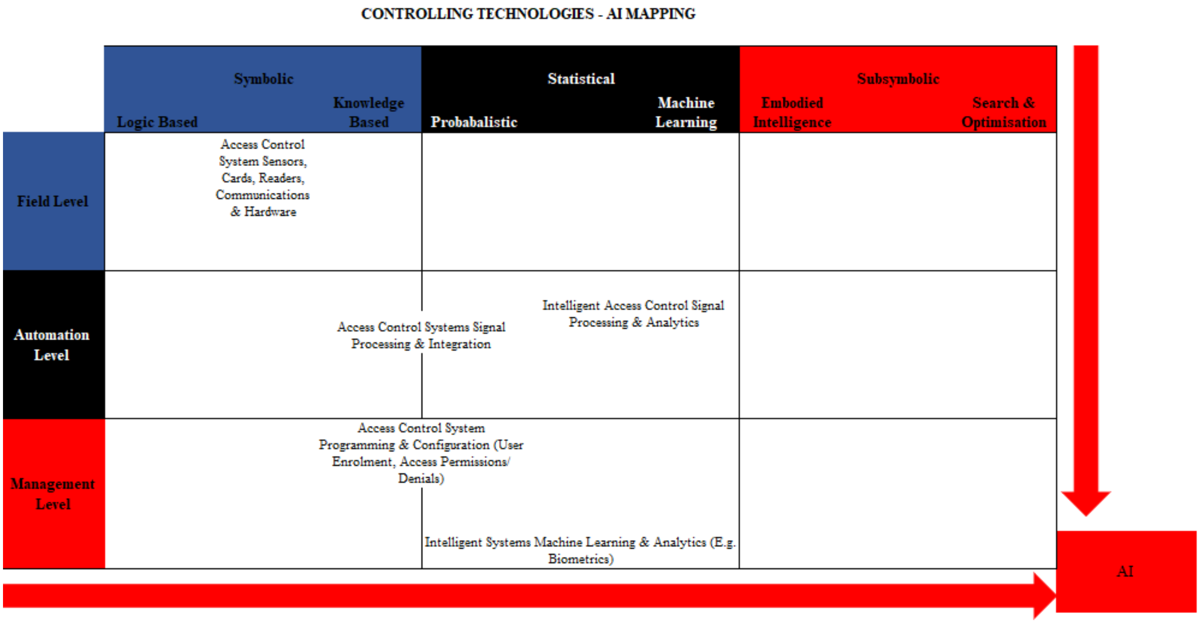
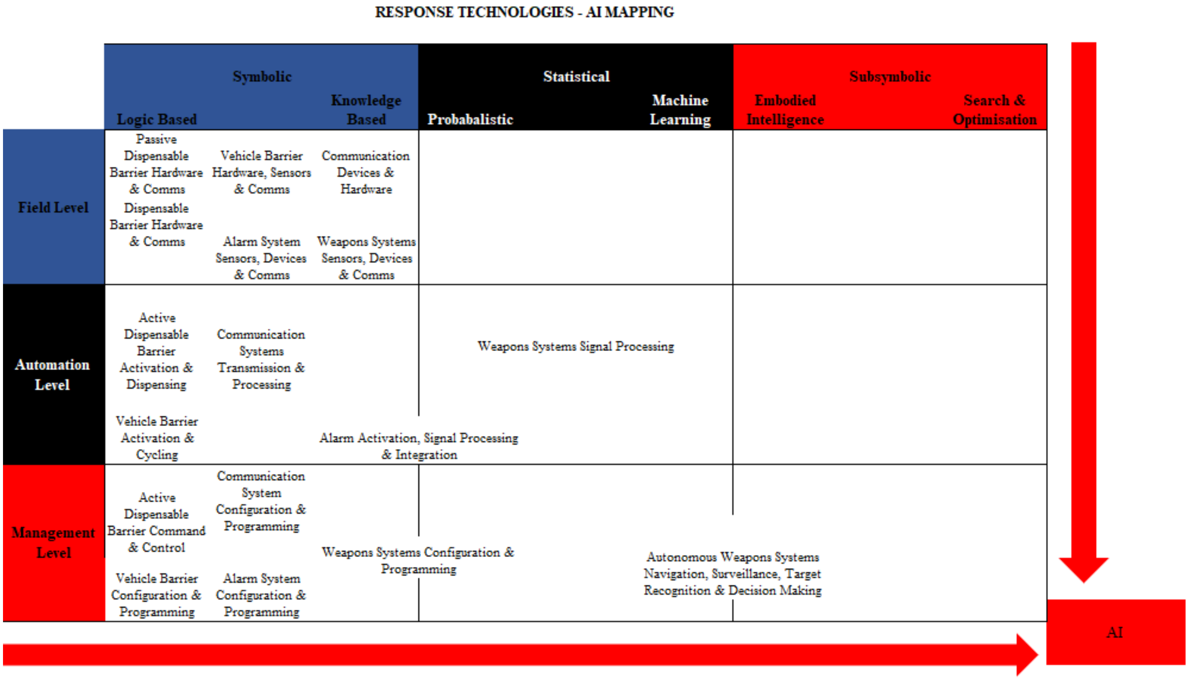


Figure B4

Indicative alignments between Response technologies, BACS Architecture Levels and AI computational tasks aligned to the spectrum of AI paradigms.



Appendix C - BACS/Intelligent Architecture Levels Presented to Participants

Management Level

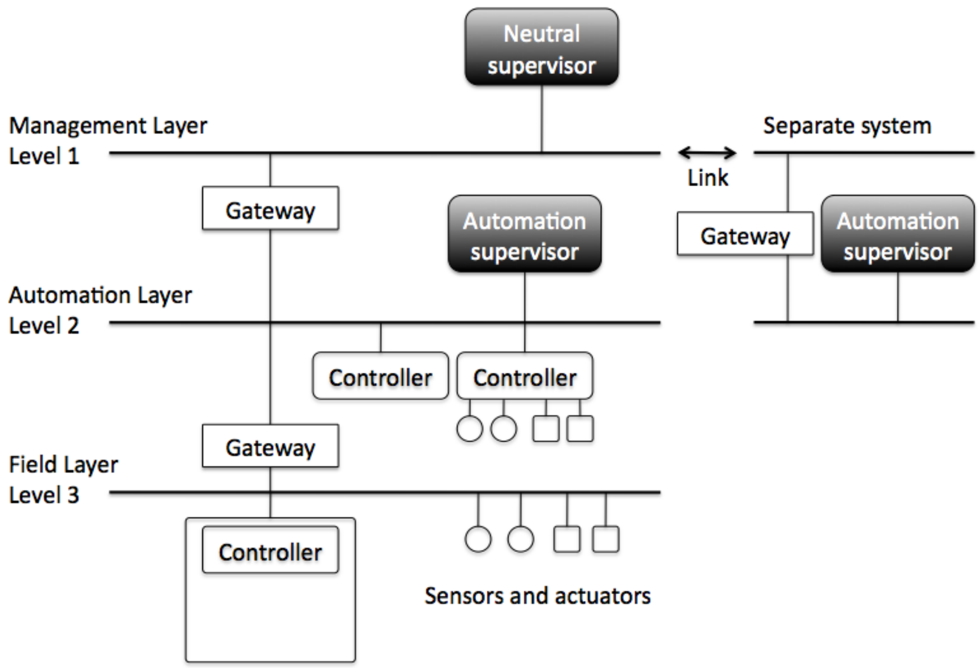
The Management level is defined as an organisation's Information Technology and Communications (ITC) network. According to Brooks et al. (2017) the Management level comprises "operator stations, monitoring and operator units, programming units and other peripheral computer devices connected to a data processing device i.e., a server" (International Organization for Standardization, 2004, p. 53) to support the information exchange monitoring and management of the automation system (Brooks et al, 2017, p.25).

Automation Level

The Automation level is defined as a dedicated communications network for the sole purpose of building device connectivity, communication and control (automation). The Automation level comprises "control devices and monitoring and operator units, programming units, operator stations or panels and/or programming units connected to a data processing device i.e., a server" (International Organization for Standardization, 2004, p. 53). This level is associated with controllers that serve main plant, such as the air handling units, chillers and boiler units, etc. (Brooks et al, 2017, p.25).

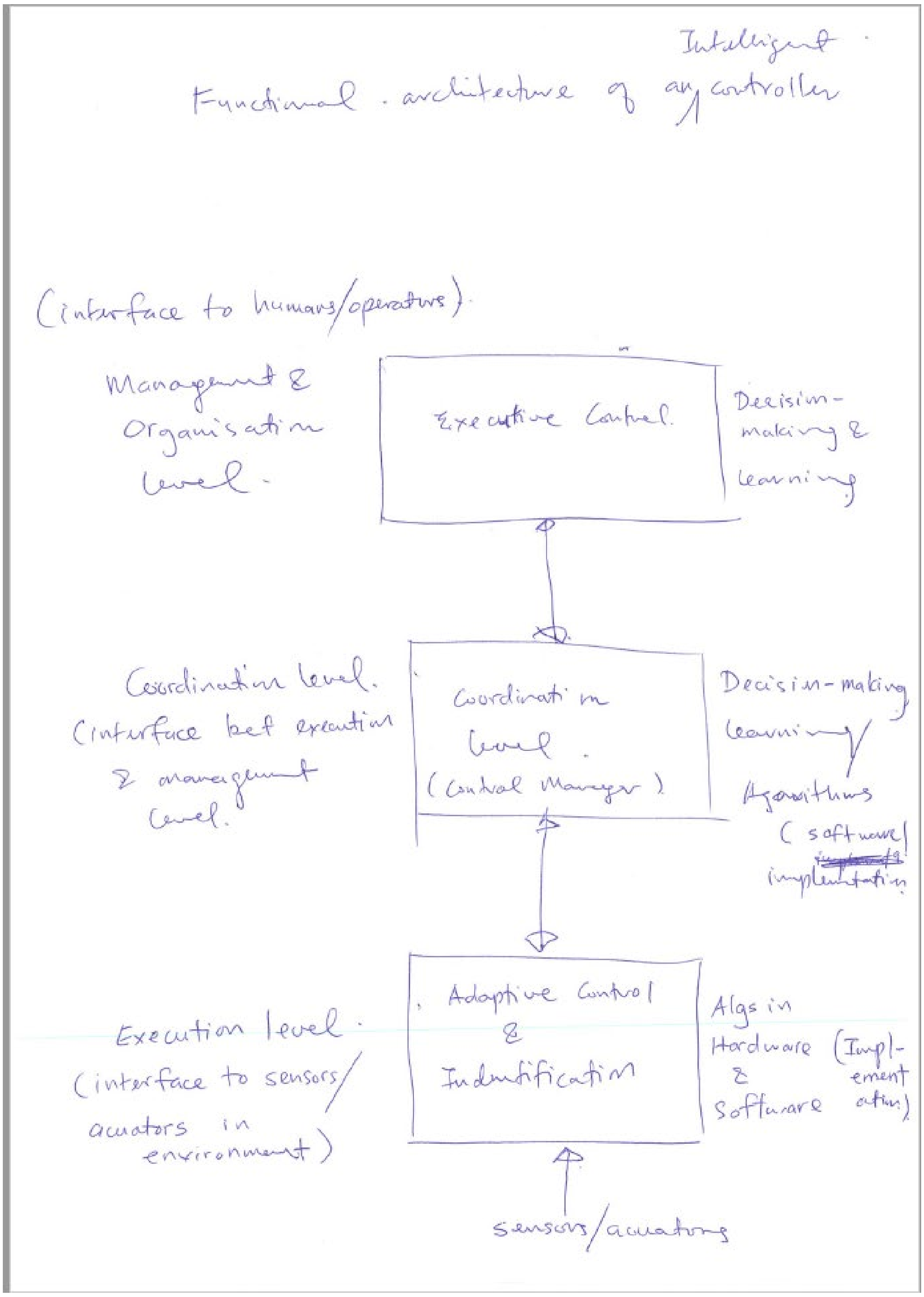
Field Level

The Field level comprises of sensing or action devices that are generally self-contained physical units. Field level devices are connected to automation level controllers, either application specific or generic controllers. Application specific controllers operate using communications protocols such as M-bus or other proprietary protocols, referred to as Edge Devices (Brooks et al, 2017, pp.25-26; Fennelly, 2012, p.316).



To achieve autonomous control (relieve operators of mundane tasks, to increase speed of response and to deal with hazards), high-level decision-making techniques required. Intelligent "controllers" – the means to achieve varying degrees of autonomy (basic: opening a door to those requiring high-level reasoning under uncertainty).

Functional architecture of an intelligent controller



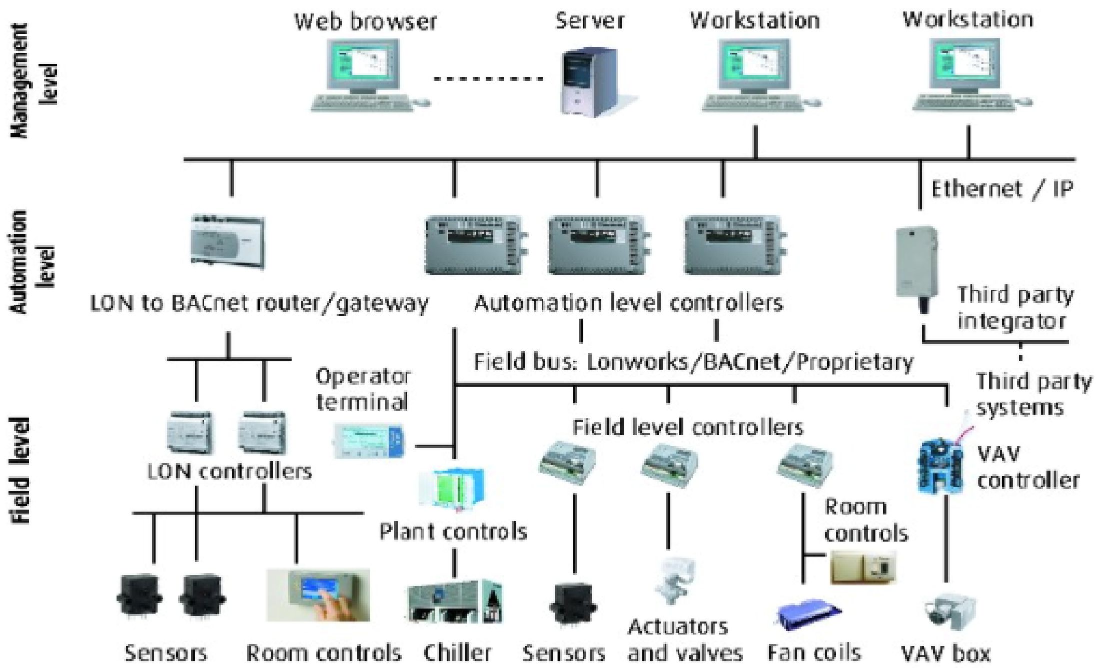
- *The **execution level** has low level numeric signal processing and control algorithms.*
- *The **coordination level** provides for tuning, scheduling, supervision and redesign of the execution-level algorithms; crisis management; planning and learning capabilities for the coordination of execution-level tasks; and higher-level symbolic decision making for failure detection and identification (FDI) and control algorithms management.*
- *The **management level** provides for the supervision of lower-level functions and for managing the interface to the humans. The management level will interact with the users in generating goals for the controller and assessing capabilities of the system. This level also monitors the performance of the lower-level systems, plans activities at the highest level (in cooperation with the human) and performs high-level learning about the user and lower-level algorithms (excerpt)*

From an AI perspective, AI techniques (signal processing, fuzzy, expert systems etc) can be implemented for various functions at all the three levels to support a specific task (function). Commonly, narrow AI at the execution level, broad AI at the coordination and management level.

The questions:

How does this map to the description on Pg. 15 and fig4 and 5 in the report submitted?

From physical security where is the AI -- which of the 3 levels?? at this point in time.



Appendix D – AI Project Focus Group Information Letter to Participants

AI PROJECT FOCUS GROUP INFORMATION LETTER TO PARTICIPANTS

Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments.

Dear Participant

You are invited to participate in a research project "focus group", being conducted in partnership between *ASIS Foundation,* and *Edith Cowan University* (ECU). The research is an investigation into identifying areas of opportunity for advancements in Artificial Intelligence, and where evident, highlight risks in their uptake for security management.

The focus group/s seeks to gain an understanding of current security technology local within the continuum and domains of Artificial Intelligence, plus the potential for more advanced AI security technologies based on the AI continuum and future advancements in AI developments. The focus groups are expected to take approximately 1.5 hours.

To enhance research outcomes, the focus groups will be recorded and transcribed; however, participation will be anonymous, with no personal or your organization's information collected. Only the researchers will have access to the focus groups discussions. Participants must be at least 18 years of age.

Access to collected data will be restricted to only the ECU project researchers. Data will be securely stored for a period of five years, when it will be formally destroyed.

Participation in this project is voluntary and you are free to withdraw at any time.

If you have any questions or require any further information about the project, please feel free to contact either myself, other Chief Investigators or the ASIS Foundation contact.

Thank you.

Michael Coole, PhD
Chief Investigator
Edith Cowan University
Joondalup, Australia
m.coole@ecu.edu.au

+618 6304 5123

Other Contacts

Associate Professor Peng Lam

Chief Investigator
Edith Cowan University
Joondalup, Australia
c.lam@ecu.edu.au
+618 6304 6914

Associate Professor Martin Masek

Chief Investigator

Edith Cowan University
Joondalup, Australia
m.masek@ecu.edu.au

+618 6304 6410

Peter Ohlhausen
ASIS Foundation
peter@ohlhausen.com

PARTICIPANT CONSENT FORM

Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments.

I, [name] ________________________________, on this day [date] ________________________, have:

- Been provided with a copy of the Information Letter, explaining the research study
- Read and understood the information provided
- Been given the opportunity to ask questions and had any questions answered to my satisfaction
- Am aware that if I have additional questions, I can contact the research team
- Understand that participation in the research project will involve participation in this Focus Group, which is audible recorded and later, transcribed
- Understand that information provided will be kept confidential
- That my identity or my organization will not be disclosed without my express consent
- Understand that the information provided will only be used for the purposes of this research project
- Understand how the information is to be used
- Understand that I am free to withdraw from further participation at any time, without explanation or penalty
- That I freely agree to participate in this project

Signature: _____________________________________

Given the time and support you have provided as a participant in this focus group, would you like to be acknowledged as a supporter of the project in the final research Report? You will not be identified as a focus group participant.

If so, please sign here: ____________________________

Appendix E – The Security Technology Intelligent Autonomy Scale

The Security Technology Intelligent Autonomy Scale (Adapted from Chang, 2014; Proud & Hart, 2005).

		Level	Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
General AI	**Post-Autonomous**	**11**	Post-autonomous system determines, plans, directs and executes missions.	Post-autonomous security system possesses *Theory of Mind* and uses artificial consciousness to acquire input.	Post-autonomous security system determines, selects, and executes optimal computational techniques, autonomously adapting to optimise processing outcomes.	Post-autonomous security system determines, writes and adjusts rules, produces, analyses and predicts threat outcomes to optimise decision making.	Post-autonomous security system acts to execute security mission, using self-awareness for adaptation.	Post-autonomous system performs post-mission review.
	Autonomous	**10**	-	Fully autonomous security system observes, detects and monitors all inputs from integrated sensors and field-level components. System commands mission and acts autonomously, eliminating human intervention.	Fully autonomous security system integrates data (inputs) and applies computational techniques to analyse threats. System prepares to take action without human intervention.	Fully autonomous security system performs threat assessment, produces and ranks results, performs management-level decision making, and does not display results to the human-operator.	Fully autonomous security system acts autonomously, eliminating human intervention.	-
Broad AI		**9**	-	Autonomous security system observes, detects and monitors all inputs from integrated sensors and field-level components. System commands and acts autonomously, informing the human after execution	Autonomous security system gathers data (inputs), applies computational techniques to process and interpret threats and prepares to take action informing the human-operator but not waiting for consent. Does not display results.	Autonomous security system performs threat assessment, analyses produces and ranks results, performs management-level decision making. Displays results to the human-operator upon query.	Autonomous security system acts autonomously, but informs the human after execution	-
	Semi-Autonomous	**8**	-	Semi-autonomous security system observes, gathers, filters, and prioritises field-level inputs; displays information only if asked.	Semi-autonomous security system gathers data (inputs), applies computational techniques to process data, interpret threats and integrate data into a result which is displayed to the human-operator only upon request	Semi-autonomous security system performs decision making. The system makes final decisions, but does not display results to the human	Semi-autonomous security system executes automatically and does not allow any human interaction.	-
		7	-	Semi-autonomous security system observes, gathers, filters, and prioritises field-level inputs without displaying any information to the human. Status on	Semi-autonomous security system analyses, processes, interprets, and integrates data (inputs) into a result which is only displayed to the human if	Semi-autonomous security system performs decision making. The system makes final decisions and displays a reduced set of ranked options without displaying "why"	Semi-autonomous security system executes automatically and only informs the human if required by context. It allows for override ability	-

				command execution is provided.	result fits programmed context (context dependant summaries).	decisions were made to the human.	after execution. Human is shadow for contingencies.	
Narrow AI	**Automated**	**6**	-	Automated security system observes, gathers, filters, and prioritises field-level inputs with information displayed to the human.	Automated security system overlays processing with analysis and interprets the data (inputs). The human is shown all results.	Automated security system performs decision making and displays a reduced set of ranked options while displaying "why" decisions were made to the human	Automated security system executes automatically, informs the human, and allows for override ability after execution. Human is shadow for contingencies	-
		5	-	Semi-automated security system gathers field-level inputs from the subsystems and environment, but it only displays non-prioritized, filtered information.	Semi-automated security system overlays processing with analysis and interprets the data (inputs). The human shadows the interpretation for contingencies.	Semi-automated security system performs decision making. All results, including "why" decisions were made, are displayed to the human.	Semi-automated security system allows the human a context-dependant restricted time to intervene before execution. Human shadows for contingencies.	-
	Integrated	**4**	-	Integrated security system is responsible for gathering field-level inputs for the human and for displaying all information, but it highlights the nonprioritized, relevant information for the user.	Integrated security system analyses inputs and processes, though the human is responsible for interpretation of the data.	Both human and integrated security system perform decision making, the results from the system are considered prime.	Integrated security system allows the human a pre-programmed restricted time to intervene before execution. Human shadows for contingencies.	-
		3	-	Semi-integrated security system is responsible for gathering field-level inputs and displaying unfiltered, unprioritized information for the human. The human still is the primary monitor for all information.	Semi-integrated security system is the prime source of input analysis and processing, with human shadow for contingencies. The human is responsible for interpretation of the data.	Both human and semi-integrated security system perform decision making, the results from the human are considered prime.	Semi-integrated security system executes decision after human approval. Human shadows for contingencies.	-
	Manual	**2**	-	Human is the prime source for gathering and monitoring all data from field-level inputs, with security system shadowing for emergencies.	Human is the prime source of input analysis and processing, with security system shadowing for contingencies. The human is responsible for interpretation of the data.	The human performs all decision making, but the security system can be used as a tool for assistance.	Human is the prime source of execution, with security system/computer assistance for contingencies.	-
		1	-	Human is the only source for gathering and monitoring (defined as filtering, prioritizing and understanding) all data.	Human is responsible for analysing all inputs, processing, and interpretation of the data.	Security system does not assist in or perform decision making. Human must do it all.	Human alone can execute decision.	-

Appendix F – Location of Opportunities in Observation Technologies

Set Mission	Input Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/Acting	Review Mission
Mission Autonomously Linked to Core Organisational Objectives Operational Parameters of Mission Set via Organisational Policy	Aggregation of data - Multiple sensors and multiple inputs	Improved signal and noise reduction	Analytics to Provide Greater Context in Dynamic Conditions	Systematic autonomous response	System Reflection for Lessons Learned
Mission Linked to Criticality & Risk	Convergence of Multi-Input Sensors across systems	Greater Accuracy (Reduced False Negatives/ Positives)	Application of Context Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	Application of context to inputs for enhanced decision making	Dynamic Threat Vulnerability Risk Assessment (TVRA) Feedback
Mission Autonomously Linked to Dynamic Threat Vulnerability Risk Assessment (TVRA)	Creation of new data rather than transformation of inputs	Quantum Computing	Development of source data sets to support complex requirements	Analytics to Enhance Machine Decision Making by Providing Ranked Options for Response	Application of Context
	Aggregated data for first-level response	Parallel Streams & Faster Analysis of Images	Connectivity across systems	Application of context	Autonomous Review and Refinement
	Higher resolution and quality of lenses in cameras	Increased processing power	Aggregation of data - Multiple sensors and multiple inputs	Aggregated data for first-level response	Autonomous Review and Reinforcement Learning from Near Misses
	Higher quality in input/sensing devices	Data Management Systems	Aggregated data inputs for distributed intelligence	Aggregated data inputs for distributed intelligence	Dynamic Threat Hunting and Feedback
	Reliable colour imaging in all weather conditions	Analytics to Enhance Machine Decision Making	Aggregation of data for more intelligent decision making	Autonomous first-level response	Dynamic Risk Feedback
	Improved signal and noise reduction	Automated Field of Interest monitoring	Red Teaming Capability	Aggregation of data for more intelligent decision making	
	Improvements in day/night compatibility, auto iris and Wide Dynamic Range (WDR)	Analytics to Provide Greater Context in Dynamic Conditions	Collaborative Development between Human and Artificial Intelligence for the Optimisation of Operational and Organisational Objectives	Novel methodologies for customised presentation of operator information	
	Improved testing of systems under dynamic conditions	Increased Accuracy & Reliability in Live Operating Environments and Under Dynamic Conditions	Application of context to inputs for enhanced decision making		
	Application of context - discern real from fake (e.g. number plates)	Improved accuracy and reliability in object classification			
	Improved accuracy and reliability in object classification	Application of context - discern real from fake (e.g. number plates)			
		Aggregation of data for enhanced decision making			
		Aggregation of data for more intelligent decision making			
		Automation of pre-condition data			

Appendix G – Location of Opportunities in Detection Technologies

Set Mission	Input Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/Acting	Review Mission
Mission Autonomously Linked to Core Organisational Objectives	Convergence of Multi-Input Sensors across systems	Improved signal and noise reduction	Analytics to Provide Greater Context in Dynamic Conditions	Systematic autonomous response	Application of Context
Mission Linked to Criticality & Risk	Aggregation of data - Multiple sensors and multiple inputs	Greater Accuracy (Reduced False Negatives/ Positives)	Application of Context	Application of context to inputs for enhanced decision making	Autonomous Review and Refinement
	Higher quality in input/sensing devices	Increased processing power	Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	Application of context	Dynamic Risk Feedback
	Creation of new data rather than transformation of inputs	Quantum Computing	Development of source data sets to support complex requirements	Increased use of Broad AI techniques	
	Improved ranging of devices	Enhancements in LIDAR, RADAR & SONAR images	Cloud-based control	Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	
	Application of context	Combined LIDAR, RADAR & SONAR for enhanced detection	Analytics to Enhance Machine Decision Making by Providing Ranked Options for Response	Convergence of Multi-Input Sensors across systems	
	Aggregated data for improved accuracy & reliability (false negatives/positives)	Application of context - discern real from fake (e.g. number plates)	Machine learning datasets to include outlier and anomalous events	Aggregation of Historical Data to Feed into Dynamic Operational Conditions	
	Enhanced detection of target materials and compounds	Analytics to Provide Greater Context in Dynamic Conditions	Greater Accuracy (Reduced False Negatives/ Positives)	Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	
	Higher quality in input/sensing devices	Analytics to Enhance Machine Decision Making	Red Teaming Capability	Expansion and Aggregation of Data and Datasets to Enhance Decision Making	
	Aggregation of data for more intelligent decision making	Statistical Aggregation of Data from Single Tech Sensors to Enhance Probability of Detection	Higher quality in input/sensing devices	Dynamic Threat Vulnerability Risk Assessment (TVRA) Response	
		Aggregation of data - Multiple sensors and multiple inputs	Aggregation of data - Multiple sensors and multiple inputs	Higher quality in input/sensing devices	
		Multi-sensor input for machine learning applications	Aggregation of data for more intelligent decision making	Autonomous first-level response	
		Aggregation of data for more intelligent decision making		Coordinated response mechanisms from input/sensing and rules/deciding	

Appendix H – Location of Opportunities in Controlling Technologies

Set Mission	Input Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/Acting	Review Mission
Mission Autonomously Linked to Core Organisational Objectives	Convergence of Multi-Input Sensors across systems	Improved signal and noise reduction	Analytics to Provide Greater Context in Dynamic Conditions	Systematic autonomous response	-
Mission Linked to Criticality & Risk	Aggregated data inputs for distributed intelligence	Greater Accuracy (Reduced False Negatives/ Positives)	Application of Context Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	Autonomous first-level response	
Mission Autonomously Linked to Dynamic Threat Vulnerability Risk Assessment (TVRA)	Application of context to inputs for enhanced decision making	Development of source data sets to support complex requirements		Establishment of Training Policies and Frameworks for Machine Learning Applications	
	Multi-point identification and authentication	Aggregation of data for more intelligent decision making	Aggregation of data - Multiple sensors and multiple inputs	Connectivity across systems	
		Application of context	Privileged Access Management	Expansion and Aggregation of Data and Datasets to Enhance Outputs	
		Increased Accuracy & Reliability in Live Operating Environments and Under Dynamic Conditions	Convergence of Multi-Input data across systems	Development of Non-lethal or Less Lethal Neutralisation Capability	
		Aggregation of data - Multiple sensors and multiple inputs		Coordinated response mechanisms from input/sensing and rules/deciding	
				Legal Frameworks to facilitate deployment of control/response mechanisms	

Appendix I – Location of Opportunities in Response Technologies

Set Mission	Input Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/Acting	Review Mission
Mission Autonomously Linked to Core Organisational Objectives	Convergence of Multi-Input Sensors across systems	Improved signal and noise reduction	Analytics to Provide Greater Context in Dynamic Conditions	Systematic autonomous response	-
Mission Linked to Criticality & Risk	Connectivity across systems	Greater Accuracy (Reduced False Negatives/ Positives)	Application of Context Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	Design & Development of Common Connectivity Protocols Across Platforms, Equipment & Devices	
	Aggregated data for first-level response	Increased processing power	Resilience & Redundancy to Respond in Physical and Management Systems	Development of Drone Swarms & Robotics for AI Response Forces	
	Aggregation of data - Multiple sensors and multiple inputs	Data Management Systems for Sharing Algorithms	Collaborative Development between Human and Artificial Intelligence to Inform Response Outcome	Development of Non-lethal or Less Lethal Neutralisation Capability	
	Aggregated data inputs for distributed intelligence	Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	Fluidity in rules under dynamic conditions	Autonomous UAV tracking	
		Aggregation of data - Multiple sensors and multiple inputs	Analytics to Enhance Machine Decision Making by Providing Ranked Options for Response	Analytics to Provide Greater Context in Dynamic Conditions	
			Aggregated data for first-level response	Aggregated data inputs for distributed intelligence	
				Autonomous first-level response	
				Application of Context	

Appendix J – Location of Opportunities in Integrated Security Technologies

Set Mission	Input Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/Acting	Review Mission
New Methodologies for System Interaction with People	Aggregation of data - Multiple sensors and multiple inputs	Cloud-based control	Application of context	Systematic autonomous response	-
	Convergence of Multi-Input Sensors across systems	Analytics to Enhance Machine Decision Making by Providing Ranked Options for Response	Fluidity in rules under dynamic conditions	Analytics & Predictive Analytics for Pattern Recognition and Anomaly Detection to Identify Outliers	
	Application of context	Increased processing power	Increased Accuracy & Reliability in Live Operating Environments and Under Dynamic Conditions	Development of Non-lethal or Less Lethal Neutralisation Capability	
	Enhancements to excel under more stable conditions and less dynamic environments	Novel Detection Devices and Analytical Methodologies are Required to Enhance Detection Capabilities		Increased Accuracy & Reliability in Live Operating Environments and Under Dynamic Conditions	
		Application of context		Analytics to Enhance Machine Decision Making by Providing Ranked Options for Response	
		Increased Accuracy & Reliability in Live Operating Environments and Under Dynamic Conditions			

Appendix K – Interview Information Letter Provided to Participants

Chief Investigator: Dr Michael Coole
School of Science
Edith Cowan University
270 Joondalup Drive
JOONDALUP WA 6027
Phone: 63045123
Email: m.coole@ecu.edu.au

Participant Information Letter: Expert Interview

Project title: Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments

Approval Number: REM 2020-01254-Coole

Principal Investigator: Dr Michael Coole, Associate professor Peng Lam & Associate professor Martin Masek

An invitation to participate in research

You are invited to participate in a project titled Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments which seeks to identify areas of opportunity for advancements in Artificial Intelligence, and where evident, highlight risks in their uptake for security management.

You are being asked to take part in this project because you are considered a security or artificial intelligence subject matter expert with knowledge that will assist in articulating the current status of AI in the protection of assets, provide insight into further opportunities for fusing security technology with developments in AI and provide indicative risks associated with such advancements.

Please read this information carefully. Ask questions about anything that you do not understand or want to know more about. Before deciding whether or not to take part, you might want to talk about it with a relative or friend.

If you decide you want to take part in the research project, you will be asked to sign the consent section. By signing it you are telling us that you:

- Understand what you have read;
- Consent to take part in the research project;
- Consent to be involved in the research described;
- Consent to the use of your personal information as described.

What is this project about?

This project aims to undertake an investigation into opportunities and practical implications for corporate security managers within the context of security technology enhancement using Artificial Intelligence. The benefits of this include directing future technology integration, where necessary commercial research and development, increase corporate security efficiencies, reduce risks to personnel and more effectively manage organisational risk through the aid of enhanced automation.

This project has been funded by ASIS International Foundation.

What does my participation involve?

Your participation in this research project will involve participating in an expert interview using a media platform such as Microsoft Teams, or in person if safe to do so and subject to any WA Government Covid-19 lockdown requirements. The interview is expected to take approximately 1 hour and seeks to identify areas of opportunity for advancements in Artificial Intelligence, and where evident, highlight risks in their uptake for security management.

To enhance research outcomes, the interview will be recorded and transcribed; however, participation will be anonymous, with no personal or your organization's information collected. Only the researchers will have access to the interview discussions. Participants must be at least 18 years of age.

Do I have to take part in this research project?

Your participation in this research project is voluntary. If you do not wish to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any time.

If you do decide to take part, you will be given this Participant Information Letter and Consent form to sign and you will be given a copy of the information letter to keep. Your decision to take part, or to take part and later withdraw, will not affect your relationship with the research team.

Your privacy

By signing the consent form, you consent to the research team collecting and using information provided by you for the research project. Any information obtained in connection with this research project that can identify you will remain confidential. Access to collected data will be restricted to only the ECU project researchers. Data will be securely stored for a period of five years, when it will be formally destroyed. Furthermore, your information will only be used for the purpose of this research project.

The results of this research project will be published and/or presented in a variety of forums. In any publication and/or presentation, information will be provided in such a way that you cannot be identified, except where requested for specific reasons, and then you will be asked to provide written consent.

In accordance with relevant Australian and/or Western Australian privacy and other relevant laws, you have the right to request access to the information about you that is collected and stored by the research team. You also have the right to request that any information with which you disagree be corrected. Please inform the research team member named at the end of this letter if you would like to access your information.

All data collected will be kept in accordance with ECU's Data Management Policy. Electronic data will be stored on a secure Microsoft SharePoint site provisioned by ECU's IT Services and physical records will be stored as required in ECU's Records Management Policy. The data will be retained for 7 years and destroyed, if appropriate at the end of the retention period. Data will be identifiable, when stored and at the end of the retention period, the data will be destroyed, if appropriate under the State Records Act.

Possible Benefits

We cannot guarantee or promise that you will receive any benefits from this research, however possible benefits may include articulating the potential benefits and threats of artificial intelligence in the protection of assets based on the current roles, functions and tasks of corporate security in the protection of assets.

Possible Risks and Risk Management Plan

There are no known risks to participating in this research project. For information relating to the University's research participant insurance coverage, please go to: https://intranet.ecu.edu.au/staff/centres/strategic-and-governance-services/our-services/risk-and-assurance/insurance/practicum-work-experience-or-volunteer-activities

What happens when this research study stops?

We will advise you of the outcomes via formal ASIS Foundation Publication and relevant Journal Publications. However, your name or any other identifying information will not be included in any of the publications or presentations.

Has this research been approved?

This research project has received the approval of Edith Cowan University's Human Research Ethics Committee, in accordance with the National Health and Medical Research Council's *National Statement on Ethical Conduct in Human Research 2007 (Updated 2018)*. The approval number is REM 2020-01254-Coole

Contacts

If you would like to discuss any aspect of this project, please contact the following people.

Chief Investigator
Name: Dr Michael Coole
Role: ECU Researcher
Edith Cowan University
P: 6304 5123
E: m.coole@ecu.edu.au

Chief Investigator
Name: Peng Lam
Role: ECU Researcher
Edith Cowan University
P: 6304 6914
E: c.lam@ecu.edu.au

Chief Investigator
Name: Martin Masek
Role: ECU Researcher
Edith Cowan University
P: 6304 6410
E: m.masek@ecu.edu.au

If you have any concerns or complaints about the research project and wish to talk to an independent person, you may contact:

Independent Person
Research Ethics Support Officer
Edith Cowan University
P: 6304 2170
E: research.ethics@ecu.edu.au

If you wish to participate in this research, please *sign the Consent Form and return to m.coole@ecu.edu.au*

Sincerely,

Dr Michael Coole

Chief Investigator

Appendix L – Expert Interview Informed Consent Form

Chief Investigator: Dr Michael Coole
School of Science
Edith Cowan University
270 Joondalup Drive
JOONDALUP WA 6027
Phone: 6304 5123
Email: m.coole@ecu.edu.au

Participant Consent Form

Project title: Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments
Approval Number: REM 2020-01254-Coole
Principal Investigator: Dr Michael Coole, Associate professor Peng Lam & Associate professor Martin Masek

I, ________________________________ have read the Participant Information Letter *or someone has read it to me in a language that I understand*. By signing this consent form, I acknowledge that I:

- have been provided with a copy of the Participant Information Letter, explaining the research study
- have read and understood the information provided
- have been given the opportunity to ask questions and have had questions answered to my satisfaction
- can contact the research team if I have any additional questions
- understand that participation in the research project will involve:
- understand that the information provided will be kept confidential, and that my identity will not be disclosed without consent
- understand that I am free to withdraw from further participation at any time, without explanation or penalty
- freely agree to participate in the project
- *The data and/or samples collected may be used only for the purposes of this research project.*

I agree to have my conversations audiotaped/ digitally recorded

Yes ☐ *No* ☐

I agree to be photographed or videotaped/ digitally recorded

Yes ☐ *No* ☐

Participant name: ____

Signature: ____________________________ Date ________________

Approval to conduct this research has been provided by the Edith Cowan University's Human Research Ethics Committee, approval number REM 2020-01254-Coole in accordance with its ethics review and approval procedures.

Appendix M –Interview Information Presented to Participants

Opportunities and Risks of AI in Security Technologies

Introduction

You have been invited to participate in a project titled Opportunities and Implications of using Artificial Intelligence in the Establishment of Secure Physical Environments which seeks to identify areas of opportunity for advancements in Artificial Intelligence, and where evident, highlight risks in their uptake for security management.

The research aims of this phase of the project are to explore;

- The priority security technology capability areas in the protection of assets for enhancement through artificial intelligence.
- The vulnerabilities which exist in the artificial intelligence techniques that may be used to achieve control in a physical environment.

To assist participants, supporting documentation has been prepared including the *AI Knowledge Map* (Corea, 2019)[1] and the *Security Technologies Intelligent Autonomy Scale*. The supporting information and documentation will assist you in answering the interview questions.

Interview Procedure

Participants will be asked what opportunities and risks may exist for AI in each of the Security Technology Categories (Observe, Detect, Control and Response Technologies) as well as integrated security technologies which may combine Observe, Detect, Control and Respond technologies to achieve physical protection. Interviews will be divided into five parts as per the following;

Part 1: Observation Technologies

Part 2: Detection Technologies

Part 3: Controlling Technologies

Part 4: Response Technologies

Part 5: Integrated Security Technologies

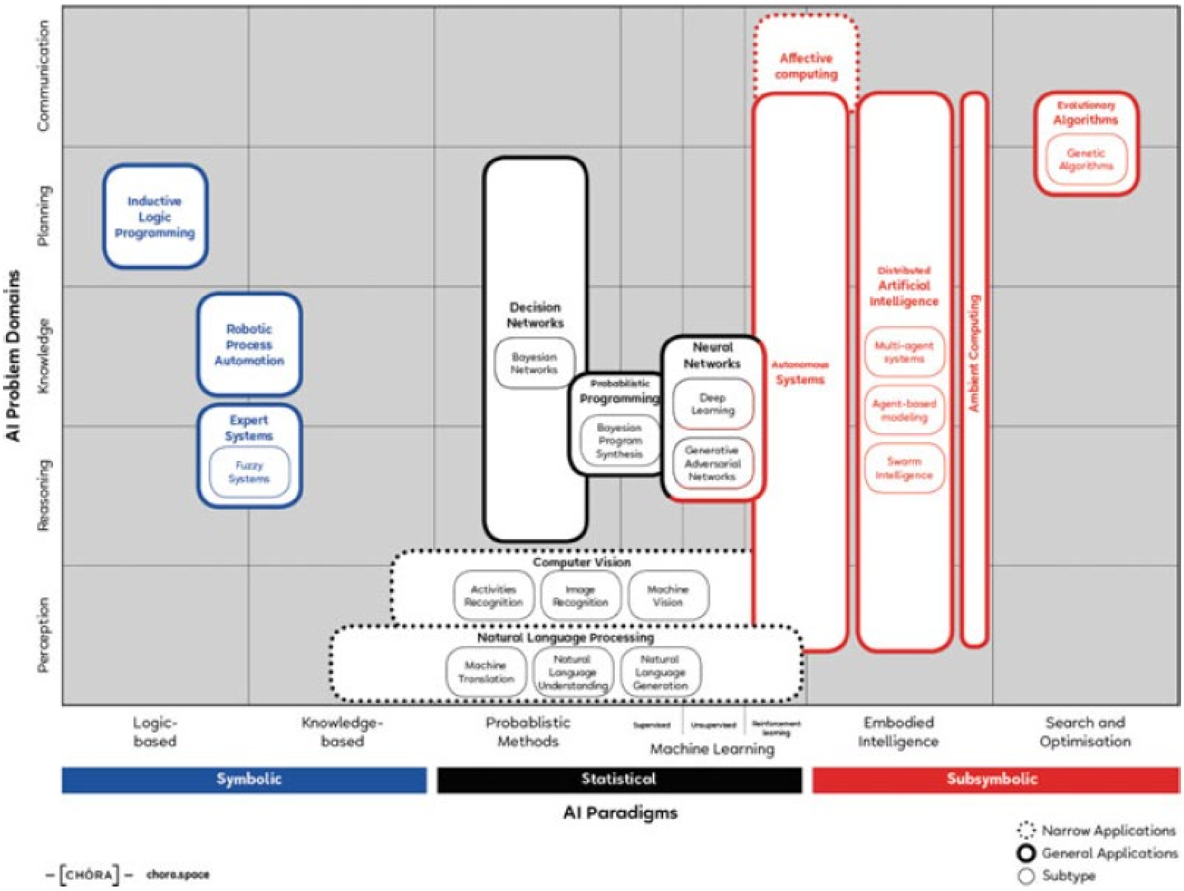
[1] Corea, F. (2019). *An Introduction to Data: Everything You Need to Know About AI, Big Data and Data Science (Volume 50 Studies in Big Data)*. Springer Nature Switzerland AG 2019.

The AI Knowledge Map

A novel visual map of the current AI paradigms (Corea, 2019) is presented in Figure M1, depicting the wide spectrum of AI paradigms and their applications. The *AI Knowledge Map* represents an abstract range of problem domains on the y-axis, divided into the five categories of analogic human cognitive processes including: perception, reasoning, knowledge, planning and communication. AI paradigms used to address sub-sets of these problem domains are presented on the x-axis. While the AI paradigms presented in Figure M1 are in a defined order, this does not imply any natural order or ranking of the paradigms in terms of superiority. They are simply different approaches to achieving some form of AI, each having unique benefits and challenges, noting some may be better suited to specific problems than others.

Figure M1

The AI Knowledge Map (Corea, 2019).



Symbolic Paradigms correspond to various AI methods that use symbols, which may be represented in rules, as their basis of manipulation. The internal working of such an AI process receives its input and based on this input satisfying a set of pre-defined rules produces an output.

> *Logic-based AI approaches*: comprised of representing knowledge of an agent's world, its goals and the current state via logical statements. By using inference or deduction involving these logical statements, an appropriate decision to achieve specific goals is obtained.
>
> *Knowledge-based approaches*: consisting of two main components, namely a knowledge-based component and an inference engine that acts as the control component for inferring new knowledge/decisions.

Statistical Paradigms typically apply a series of mathematical operators to their inputs to produce an output.

> *Probabilistic approaches*: employ probabilistic representations, that capture uncertainty in complex relationships and knowledge of the world, in the form of probabilistic graphical models. The graphical models capture the distribution in the data and decisions can be obtained via statistical inference.
>
> *Machine Learning techniques*: automatically build models from data which can be used to make predictions or decisions. There are 3 classes of machine learning techniques, namely unsupervised, supervised and reinforcement learning.

Sub-symbolic Paradigms are representative of neurons in the human brain. Sub-symbolic architectures can learn autonomously, following training and development of the neural network architecture.

> *Embodied intelligence approaches*: take into consideration *situatedness* and *embodiment* in the design of intelligent behaviour in embodied and situated agents. *Situatedness* is the coupling between the agent and its environment and *embodiment*, refers to the constraints associated an agent’s body, perceptual and motor system.
>
> *Search and Optimization approaches*: are techniques that can search a complex and ill-defined search space intelligently and efficiently.

Security Technologies Intelligent Autonomy Scale

The *Security Technologies Intelligent Autonomy Scale* (Table M1, adapted from Chang, 2014[2]; Proud & Hart, 2005[3]) has been developed to articulate the levels of intelligent autonomy which may be used by security technologies. The scale ranges from Level 1: Manual (no AI) through to Level 11: Post-Autonomous (General AI) defines the technical AI capabilities at each increasing level of autonomy.

---

[2] Chang, E.M. (2014). *Defining the Levels of Adjustable Autonomy: A Means of Improving Resilience in an Unmanned Aerial System* [Master of Science Thesis, Naval Postgraduate School]. NPS Archive. http://hdl.handle.net/10945/43887

[3] Proud, R.W., and Hart, J.J. (2005). FLOAAT, A Tool for Determining Levels of Autonomy and Automation, Applied to Human-Rated Space Systems. *Infotech@Aerospace*, 7061. American Institute of Aeronautics and Astronautics (AIAA). Arlington, VA: AIAA.

Table M1

The Security Technology Intelligent Autonomy Scale (Adapted from Chang, 2014; Proud & Hart, 2005).

		Level	Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
General AI	Post-Autonomous	11	Post-autonomous system determines, plans, directs and executes missions.	Post-autonomous security system possesses *Theory of Mind* and uses artificial consciousness to acquire input.	Post-autonomous security system determines, selects, and executes optimal computational techniques, autonomously adapting to optimise processing outcomes.	Post-autonomous security system determines, writes and adjusts rules, produces, analyses and predicts threat outcomes to optimise decision making.	Post-autonomous security system acts to execute security mission, using self-awareness for adaptation.	Post-autonomous system performs post-mission review.
Broad AI	Autonomous	10	-	Fully autonomous security system observes, detects and monitors all inputs from integrated sensors and field-level components. System commands mission and acts autonomously, eliminating human intervention.	Fully autonomous security system integrates data (inputs) and applies computational techniques to analyse threats. System prepares to take action without human intervention.	Fully autonomous security system performs threat assessment, produces and ranks results, performs management-level decision making, and does not display results to the human-operator.	Fully autonomous security system acts autonomously, eliminating human intervention.	-
Broad AI	Autonomous	9	-	Autonomous security system observes, detects and monitors all inputs from integrated sensors and field-level components. System commands and acts autonomously, informing the human after execution	Autonomous security system gathers data (inputs), applies computational techniques to process and interpret threats and prepares to take action informing the human-operator but not waiting for consent. Does not display results.	Autonomous security system performs threat assessment, analyses produces and ranks results, performs management-level decision making. Displays results to the human-operator upon query.	Autonomous security system acts autonomously, but informs the human after execution	-
Broad AI	Semi-Autonomous	8	-	Semi-autonomous security system observes, gathers, filters, and prioritises field-level inputs; displays information only if asked.	Semi-autonomous security system gathers data (inputs), applies computational techniques to process data, interpret threats and integrate data into a result which is displayed to the human-operator only upon request	Semi-autonomous security system performs decision making. The system makes final decisions, but does not display results to the human	Semi-autonomous security system executes automatically and does not allow any human interaction.	-
Broad AI	Semi-Autonomous	7	-	Semi-autonomous security system observes, gathers, filters, and prioritises field-level inputs without displaying any information to the human. Status on	Semi-autonomous security system analyses, processes, interprets, and integrates data (inputs) into a result which is only displayed to the human if	Semi-autonomous security system performs decision making. The system makes final decisions and displays a reduced set of ranked options without displaying "why"	Semi-autonomous security system executes automatically and only informs the human if required by context. It allows for override ability	-

				command execution is provided.	result fits programmed context (context dependant summaries).	decisions were made to the human.	after execution. Human is shadow for contingencies.	
Narrow AI	**Automated**	**6**	-	Automated security system observes, gathers, filters, and prioritises field-level inputs with information displayed to the human.	Automated security system overlays processing with analysis and interprets the data (inputs). The human is shown all results.	Automated security system performs decision making and displays a reduced set of ranked options while displaying "why" decisions were made to the human	Automated security system executes automatically, informs the human, and allows for override ability after execution. Human is shadow for contingencies	-
		5	-	Semi-automated security system gathers field-level inputs from the subsystems and environment, but it only displays non-prioritized, filtered information.	Semi-automated security system overlays processing with analysis and interprets the data (inputs). The human shadows the interpretation for contingencies.	Semi-automated security system performs decision making. All results, including "why" decisions were made, are displayed to the human.	Semi-automated security system allows the human a context-dependant restricted time to intervene before execution. Human shadows for contingencies.	-
	Integrated	**4**	-	Integrated security system is responsible for gathering field-level inputs for the human and for displaying all information, but it highlights the nonprioritized, relevant information for the user.	Integrated security system analyses inputs and processes, though the human is responsible for interpretation of the data.	Both human and integrated security system perform decision making, the results from the system are considered prime.	Integrated security system allows the human a pre-programmed restricted time to intervene before execution. Human shadows for contingencies.	-
		3	-	Semi-integrated security system is responsible for gathering field-level inputs and displaying unfiltered, unprioritized information for the human. The human still is the primary monitor for all information.	Semi-integrated security system is the prime source of input analysis and processing, with human shadow for contingencies. The human is responsible for interpretation of the data.	Both human and semi-integrated security system perform decision making, the results from the human are considered prime.	Semi-integrated security system executes decision after human approval. Human shadows for contingencies.	-
	Manual	**2**	-	Human is the prime source for gathering and monitoring all data from field-level inputs, with security system shadowing for emergencies.	Human is the prime source of input analysis and processing, with security system shadowing for contingencies. The human is responsible for interpretation of the data.	The human performs all decision making, but the security system can be used as a tool for assistance.	Human is the prime source of execution, with security system/computer assistance for contingencies.	-
		1	-	Human is the only source for gathering and monitoring (defined as filtering, prioritizing and understanding) all data.	Human is responsible for analysing all inputs, processing, and interpretation of the data.	Security system does not assist in or perform decision making. Human must do it all.	Human alone can execute decision.	-

Part 1: Observe Technologies

Current use of AI in observation technologies is summarised in Table M2;

Table M2

Current use of AI in Observe Technologies

Security Technology Categories	**Security Technologies**	**AI Paradigms**	**Current Use**
Observe	Network Video Surveillance	Computer vision (Statistical - Probabilistic Programming).	*Observes:* Movement/ Activity Object Classification Object Tracking Biometric Features

Questions for Participants

Q1 Based on the *Security Technologies Intelligent Autonomy Scale* (Table M1), at which level do you believe observation technologies currently sit?

Q2 What enhancements could be made to observation technologies to achieve a higher level on the *Security Technologies Intelligent Autonomy Scale*?

Q3 What opportunities exist for observation technologies using AI techniques? Please use Table M3 (below) to indicate which stages of the cycle these opportunities may be located.

Table M3

Location of Opportunities for AI Advancement in Observe Technologies

Technology	**Set Mission**	**Input/ Sensing**	**Computational Technique/ Processing**	**Rules/ Deciding**	**Output/ Acting**	**Review Mission**
Observation Technologies						

Q4 What benefits would result from these opportunities?

Q5 What risks do you foresee from future advancements?

Part 2: Detect Technologies

Current use of AI in detection technologies is summarised in Table M4;

Table M4

Current Use of AI in Detect Technologies

Security Technology Categories	Security Technologies	AI Paradigms	Current Use
Detect	Internal and External Intrusion Detection Systems, Network Video Surveillance, Biometric Detection Systems, Lighting Systems, X-Ray, Trace Detection Systems, Acoustic Detection Systems, RADAR Systems, SONAR Systems	Symbolic (Logic Based) Statistical (Probabilistic and Machine Learning) Computer Vision (Activity Recognition)	***Detects:*** Human Intrusion Contraband Materials Changes in Environment Movement/ Activity Event Analysis Anomalies Acoustic Signatures Biometric Features Radio Waves Sound Waves

Questions for Participants

Q1 Based on the *Security Technologies Intelligent Autonomy Scale* (Table M1), at which level do you believe detection technologies currently sit?

Q2 What enhancements could be made to detection technologies to achieve a higher level on the *Security Technologies Intelligent Autonomy Scale*?

Q3 What opportunities exist for detection technologies using AI techniques? Please use Table M5 (below) to indicate which stages of the cycle these opportunities may be located.

Table M5

Location of Opportunities for AI Advancement in Detect Technologies

Technology	Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Detection Technologies						

Q4 What benefits would result from these opportunities?

Q5 What risks do you foresee from future advancements?

Part 3: Control Technologies

Current use of AI in controlling technologies is summarised in Table M6;

Table M6

Current Use of AI in Control Technologies

Security Technology Categories	**Security Technologies**	**AI Paradigms**	**Current Use**
Control	Access Control Systems, Vehicle Barriers, Dispensable Barriers, Biometric Systems	Symbolic (Logic Based) Bayesian Program synthesis (Statistical - Probabilistic Programming)	***Controls:*** Human Entry/Exit Vehicle Entry/Exit Logical Access

Questions for Participants

Q1 Based on the *Security Technologies Intelligent Autonomy Scale* (Table M1), at which level do you believe controlling technologies currently sit?

Q2 What enhancements could be made to controlling technologies to achieve a higher level on the *Security Technologies Intelligent Autonomy Scale*?

Q3 What opportunities exist for controlling technologies using AI techniques? Please use Table M7 (below) to indicate which stages of the cycle these opportunities may be located.

Table M7

Location of Opportunities for AI Advancement in Control Technologies

Technology	**Set Mission**	**Input/ Sensing**	**Computational Technique/ Processing**	**Rules/ Deciding**	**Output/ Acting**	**Review Mission**
Controlling Technologies						

Q4 What benefits would result from these opportunities?

Q5 What risks do you foresee from future advancements?

Part 4: Response Technologies

Current use of AI in controlling technologies is summarised in Table M8;

Table M8

Current Use of AI in Response Technologies

Security Technology Categories	Security Technologies	AI Paradigms	Current Use
Respond	Communications Systems, Dispensable Barriers, Vehicle Barriers, Weapons Systems	Symbolic (Logic Based) Bayesian Program synthesis (Statistical - Probabilistic Programming)	***Responds with/by:*** Alarm/ Barrier Activation Object Tracking GPS Navigation, Automated Notification/ Warnings Signal Transmission Emission of Frequencies Emission of Electromagnetic Energy Light Emission Voltage Emission Select & Engage Target

Questions for Participants

Q1 Based on the *Security Technologies Intelligent Autonomy Scale* (Table M1), at which level do you believe response technologies currently sit?

Q2 What enhancements could be made to response technologies to achieve a higher level on the *Security Technologies Intelligent Autonomy Scale*?

Q3 What opportunities exist for response technologies using AI techniques? Please use Table M9 below to indicate which stages of the cycle these opportunities may be located.

Table M9

Location of Opportunities for AI Advancement in Response Technologies

Technology	Set Mission	Input/ Sensing	Computational Technique/ Processing	Rules/ Deciding	Output/ Acting	Review Mission
Response Technologies						

Q4 What benefits would result from these opportunities?

Q5 What risks do you foresee from future advancements?

Part 5: Integrated Security Technologies

Q1 Based on the *Security Technologies Intelligent Autonomy Scale* (Table M1), at which level do you believe integrated security technologies currently sit?

Q2 From a security management perspective, for which aspects of integrated security systems should automation be developed or enhanced?

Q3 From a security management perspective, which aspects of integrated security systems should retain human assessment and control?

Q4 There are numerous *Literature Defined Risks of AI* (Table M10). Please indicate which of these risks apply to security technologies and integrated security systems.

Q5 Do you foresee any additional risks which require consideration?

Table M10

Literature Defined Risks of Artificial Intelligence

Risk	Description
AI Failure	Insufficient redundancy built into technologies – technical or operational AI issues result in system breakdown or failures.
Latency Risks	Cloud-based delays in processing and communication create latency risks impacting on safety and operations.
Distrust of AI	Corporate or public noncompliance due to distrust of AI technologies or those responsible for system operation.
AI Mistakes	AI does not know how to respond to novel objects or situations without extensive training sets or data resulting in errors or consequences – for example system incorrectly classifies a civilian cruise ship as a naval warship and launches attack
Misuse	AI used inappropriately or for unintended purposes.
Limited Threat Management	Pattern recognition can be used to identify behaviour, however AI cannot contextualise the behaviour to determine whether it is malicious - e.g. if a sensitive file is accessed after office hours by an authorised person, the event may not be flagged.
Insider Threat	Insiders are able to access, modify, disable or override systems.
Cyber Security Risks	Numerous security risks - adversaries may hack, hijack or attack, compromise or disable the technology (e.g. data breaches, Denial of Service (DoS), Man-in-the-middle Attack, malware injection/ use of malicious software (e.g. ransomware, siege ware, jackware), autonomous cyber-attacks etc).
Cloud-based Security Risks	Attack or exploitation of cloud-based vulnerabilities (or use of malicious software such as Spectre & Meltdown) to disrupt, damage, gain access to, or take control of security system.
Edge of Network Security Issues	Attack or exploitation of edge of network vulnerabilities including Denial of Service (DoS), Man-in-the-middle Attack, Rogue Gateway, Rogue Data Centre, Rogue Infrastructure, Physical Damage, Privacy Leakage, Privilege Escalation, Service Manipulation, Virtual Machine (VM) Manipulation, Injection of Information, etc.
Safety Issues	Autonomous execution of instructions by system (system override) preventing human intervention and control.
AI Arms Race	Development of AI technologies creates an AI arms race in developing technologies such as lethal weapons systems.
Illegal AI trade	Technologies or systems traded through black or grey markets. State-sponsored acquisition of AI-tech by adversarial groups.

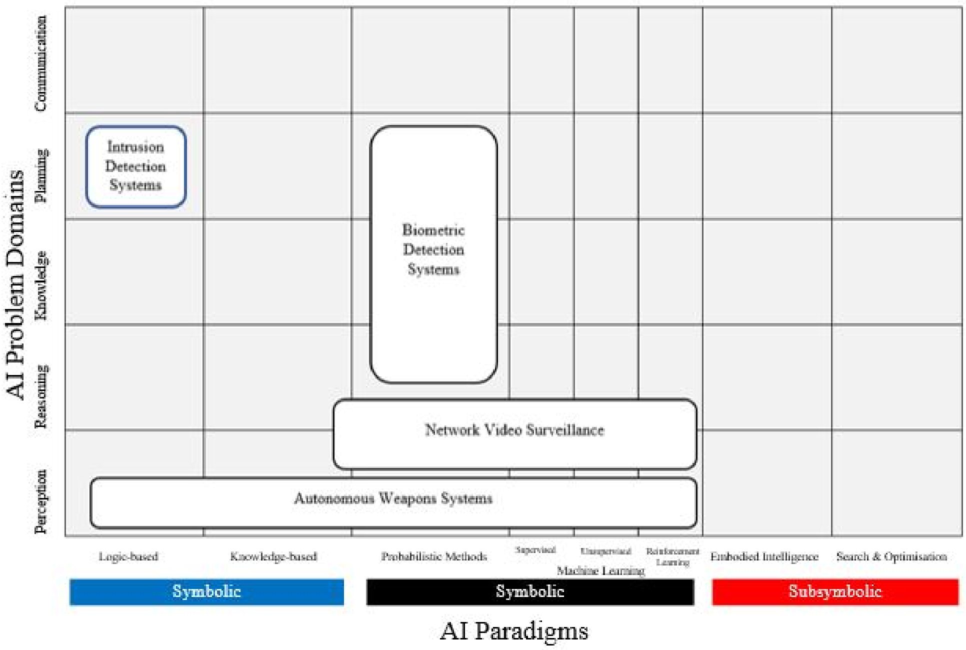
Lack of Transparency and/or Accountability	Auditing issues or inability to create transparency in system processing and decision making. Inability to determine responsibility for system shortcomings or failures (e.g. manufacturer, system or operator error ?)
Legal Risks	Legal damages or ramifications resulting from AI actions, omissions or failures. Forensic evidence issues (chain of custody, compatibility of various types of digital evidence) resulting in evidence or prosecution failures.
Economic Risks	Unemployment e.g. AI causes some security jobs to become obsolete.
Political Risks	E.g. market monopolies or reliance on imports of international technology.
Ethical Risks	Fairness and equity issues (e.g. aggregated data may result in vilification of certain groups of people). Ethical risks of the development and use of lethal AI systems (e.g. autonomous weapons systems).
Inadequate Governance and Regulation	Responsible supervision, regulation and governance of the design and deployment of AI systems, including governance and regulation of independent software vendors (ISV).
Data Ownership and Management Issues	Lack of regulation of AI software, hardware, and models, including; data ownership, data location and control, content disclosure, confidentiality, Regulatory and legislative compliance.

The *Security Technologies AI Knowledge Map* (Figure M2) has been adapted from Corea's (2019) graphic to depict where security technologies currently sit on the AI spectrum. Figure M2 illustrates examples of where Intrusion Detection Systems, Biometric Systems, Network Surveillance Systems and Autonomous Weapons Systems may currently sit on the AI Spectrum.

Q6 Do you feel the location of security technologies on the Security Technologies AI Knowledge Map (Figure M2) will change considerably in the next ten years?

Figure M2

The Security Technologies AI Knowledge Map (Adapted from Corea, 2019).



Thank you for your participation in this project.

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.