ARTIFICIAL INTELLIGENCE (AI) AND CYBERSECURITY

Amirudin Abdul WAHAB[*]

Declaration[*]

The Future of the Relationship Between Artificial Intelligence and Cybersecurity

Over the next 3–5 years, Artificial Intelligence (AI) will significantly transform cybersecurity, evolving from an emerging trend into a pivotal force and ally. AI will revolutionize cybersecurity roles, rather than eradicate them. The advancement of AI will transform cybersecurity employment functions, though it will not eliminate human roles. Gartner forecasts that automated systems will take control of more than 50% of SOC Level 1 analyst responsibilities by 2028, which includes alert prioritization and basic ticket resolution. AI systems will enhance human capabilities, rather than working as direct replacements for human experts. This means that AI will help people perform their jobs more effectively, rather than taking over. In cybersecurity, the workforce will spend more time focusing on evaluating AI-generated data through strategic investigations while handling model governance and validating system intention.

Collaborative efforts on the human-AI relationship are crucial in today's digital landscape. Rather than competing with machines or AI, the solution lies in forming a strategic and responsible partnership. Efforts and resources are needed for organizations to integrate AI systems in ways that enhance human capabilities while reducing human error where possible. Upskilling, reskilling and learning new skills is key. In addition to traditional threat management, security personnel must also learn ethical reasoning and AI literacy. To navigate the AI-driven cyber landscape, skills such as data interpretation, cross-team communication, and collaboration will be essential. Organizations must buck up and be ready to adapt, or risk falling behind. Leaders must proactively prepare for AI's impact by implementing robust AI training programs and clear usage policies. Cross-functional teams combining AI expertise with domain knowledge will ensure effective AI integration while preserving human oversight.

[*] Dato' Dr. CEO at CyberSecurity Malaysia

[*] In this study, AI tools such as ChatGPT were utilized for sentence editing. AI was used to translate the author's thoughts and ideas into a more academic framework for grammar and editing.

Artificial Intelligence Contributes To The Attackers' Advantage

According to the Global Risk Report 2024, there is a significant concern that emerging AI technologies will benefit cyber attackers more than defenders, potentially exacerbating the cyber threat landscape. The report states that 55.9% of respondents believe generative AI will give cyber attackers a competitive advantage, while only 8.9% feel it will give defenders a competitive advantage. Hence, there is a need to call for action for the cybersecurity community. Cyber defenders must seize this pivotal moment to continuously enhance their expertise and proficiency, diligently refining their knowledge and skills. Cyber defenders must not overlook the powerful opportunities that AI offers in strengthening cybersecurity and cyber resilience.

By leveraging AI for threat detection, automated response, and predictive analytics, cybersecurity professionals can shift from a reactive to a proactive defence strategy. To close the gap, defenders must harness AI's full potential, not only to keep pace with evolving threats, but to decisively tilt the advantage back toward security and trust. Not to mention, there is an example where AI tools that are most commonly used by cybercriminals are being used against them. The AI grandmother named Daisy, whose task is to waste scammers' time with meandering conversations (Thubron, 2024).

AI Could Not Eliminate Human Error

It was believed that AI could eliminate human error. However, AI capabilities are not yet fully developed due to issues such as AI hallucination, data poisoning, and the presence of low-quality data. Therefore, the most promising applications of AI are those that can be accomplished quickly. By examining vast amounts of unprocessed data, AI can identify trends and abnormalities that human analysts would overlook, improving threat detection and reaction times, for example, detecting deepfakes in videos or pictures, and the uncanny valley that AI can detect. In contrast, humans remain uncertain about whether to be suspicious or treat the material differently.

Additionally, AI automates repetitive tasks, such as handling notifications and monitoring network traffic, thereby freeing up cybersecurity experts for more strategic responsibilities. Furthermore, prospective cyber threats may be predicted using AI-driven predictive analytics, allowing for proactive defences and minimizing vulnerabilities before they can be exploited.

Together, these capabilities support cybersecurity efforts even if AI's error-free performance is currently limited.

AI is the greatest threat, but also the most excellent defence. It is a game-changer in cybersecurity defence. AI is capable of identifying anomalous login patterns, detecting suspicious network activity, reverse-engineering malware, and even forecasting potential vulnerabilities by analyzing historical data. Additionally, AI-driven automation is changing how businesses distribute their cybersecurity resources.

AI technologies offer the most potential in threat detection or response

Supervised and unsupervised Machine Learning (ML) and Generative AI (Gen AI) have emerged as transformative tools in cybersecurity. These technologies work way faster and better at detecting threats, all while taking some of the load off humans. Supervised ML uses labelled data to teach models how to recognize patterns or make predictions. This approach in cybersecurity helps catch known threats by learning from previous attacks. It is used in various ways to detect threats, such as malware classification, Intrusion Detection System (IDS), and real-time anomaly detection. Unsupervised ML does not rely on labelled data but instead identifies patterns and anomalies within datasets. This helps detect previously unfamiliar threats. This is the method used to identify threats, like anomaly detection, behavioural analytics, and entity resolution. Generative AI represents a significant leap forward by leveraging deep learning techniques to create predictive models and simulate scenarios. The capability of processing vast amounts of data and creating synthetic data makes it a powerful tool for threat detection and analysis.

i.Virtual assistance
ii. Threat contextualization
iii. Synthetic data generation

Combining supervised and unsupervised machine learning with generative AI improves cybersecurity. Each technology provides particular benefits. Supervised machine learning accurately identifies known dangers. Unsupervised machine learning uncovers unknown problems and new oddities. Generative AI adds background information and forecasts events. When people use them together, they also find that they respond to threats more quickly and with greater flexibility. As cyber threats become increasingly complex, utilizing AI technologies becomes necessary to stay ahead of attackers and build effective protective digital systems.

AI's Risk for Cybersecurity Ecosystems

We acknowledge the fact that AI both helps and harms cybersecurity. This idea holds as much importance as our adoption of AI's power in the field. The same capabilities that enable us to identify threats more quickly and precisely also allow bad actors to accelerate their attacks. AI brings a new era of cyber threats that adapt and act autonomously. One of the most concerning risks is the arrival of AI-powered malware, as well as automated attack systems. These can evolve on their own, bypass old defences, and initiate attacks with a speed and precision never seen before. Methods like poison injection as well as data manipulation harm training data, which spoils the base of AI models - this also lowers faith in automated systems.

Another common vulnerability is the use of AI to enable deepfakes and impersonation, which can fool both humans and security systems. These can be used for phishing, social engineering and even high-level fraud, blurring the lines between truth and manipulation in digital interactions.

As AI can be optimized for cybersecurity, it can also be utilized to counter cyberattacks, as machine learning algorithms are capable of identifying the most effective methods to gain access to systems or evade detection. This means more effective ransomware, brute force attacks and APTs that silently infiltrate and dwell in the network. Additionally, insider threat abuse, powered by AI, is becoming increasingly common. This is when behavioural analytics, which were meant for detection, are reversed and used to bypass internal controls. Moreover, AI can be used to launch more complex attacks on interconnected infrastructure in cyber-physical environments, resulting in real-world impacts. There is a growing concern that AI is being used more effectively to exploit vulnerabilities before cybersecurity experts can react, which may lead to a loss of trust, data breaches, and an increase in zero-day attacks.

The Role Of Regulation In Managing AI Use

The regulation is currently playing catch-up. Even the European Union (EU) AI Act is still facing concerns, with many EU leaders stating that there are still missing elements in place. Mostly concerns regarding the Act's capability to balance between innovation and security.

Those in the regulatory role have their hands full, as they need to consider the entire digital realm itself. Of course, this can be mitigated by focusing on parts rather than the overall view or creating a specialized and strategic thinking working group that can tackle the issue of playing catch-up, but that does not dismiss the fact that time is ticking and AI is not slowing down.

Nevertheless, it is up to the national leader to handle how regulation will manage AI, including cybersecurity. This example can be seen in Singapore, Japan, China, and almost the whole world with responsible leadership. Governments or industry prioritize when regulating the integration of AI into critical digital infrastructure. When focusing on cybersecurity, the integration of AI must prioritize secure-by-design, resiliency, zero-trust, adaptivity, proactivity, and holism.

Thus, the crucial thing that both governments and industry bodies need to be concerned with is striking a balance between security, ethics, and innovation. Providing clear guidelines for reporting any cybercrime incidents, ethical standards and secure personal data while still promoting innovation and healthy competition among industries. Guard rails, human-in-the-loop, backup, and many more are essential to avoid data poisoning, data bias or data hallucinations,

Legacy systems are an issue that will undoubtedly arise when discussing critical digital infrastructure, given the constant evolution of technology. As such, any policy, guideline, or regulation related to AI must be adaptable and constantly one step ahead to ensure that no loopholes can be found or abused later on. Digital literacy, awareness, and training are essential to reduce skill gaps among employees. This can be achieved through initiatives such as Safer Internet Day, Cybersecurity Awareness Month, or regular biweekly brief meetings to exchange knowledge.

The issue of AI sovereignty is slowly gaining traction as more nations have begun to focus on developing their own AI models. As such, this matter needs to be handled as soon as possible to avoid unwanted conflict with other nations while still reaping the benefits of knowledge sharing and maintaining, or at least improving, the relationship.

Adopt AI In Cybersecurity Operations

The introduction of AI into cybersecurity will indeed be a game-changer. However, ensuring its adoption is safe, fair, and effective goes beyond providing tools that are in demand; it also requires the right mindset and skill sets. Six key practices establish the foundation for responsible AI adoption in cybersecurity. This is a human-centred design where AI systems are based on human principles. Such as inclusivity, ethics, and responsibility. The Issac

Asimov three laws of robotics can also be applied here if those have read or know about AI or robot culture (Becher, 2024):

i. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
ii. A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law.
iii. A robot must protect its existence as long as such protection does not conflict with the First or Second Law.

These three robotic laws can be applied to the current AI ethical dilemma, particularly in light of the numerous specialized AIs that exist today, providing the help and assistance needed to reach both technical goals and public trust.

Second, accuracy alone is insufficient to identify the types of AI models, as well as their fairness, transparency, robustness, and real-world applicability. A holistic view of the performance could help avoid any loopholes, alongside ensuring that the AI model itself operates reliably.

Third, the quality of an AI model's output is directly linked to the quality of its raw data, which it processes. The organization responsible for the AI model should continuously examine the data fed into the model. If the data that was fed is skewed, the result itself will be biased and essentially skewed. The organization is responsible for understanding the data proactively and practically to ensure that the data accurately represents the environment the organization aims to defend.

Fourth, understanding the model limits and its datasets. The current AI still has its limitations, as it is not yet capable of solving every complex problem or predicting future trends. There is still a need for backup plans, as well as a strong foundation of human-in-the-loop analysis to determine which models are suitable and which ones are not.

Five, constantly, always, and never stop testing. Resilience is ensured through ongoing testing in various scenarios. AI models that not only function in theory but also survive the complexity of today's cyber threat landscape are needed. This can be achieved by simulating real-world attacks, particularly those involving technology that incorporates AI. AI Regulations or policies can run through the regulatory sandbox to determine which areas still need improvement and identify weaknesses.

Moreover, constant monitoring and patching of the systems after they go live. AI is not just a one-time buy. Models get old, dangers change, and info moves. Good use requires steady care, adjustment, re-teaching, checking, and changing AI setups to ensure they align with safety objectives and moral principles.

Balance Between Innovation and Risk in the Context of AI And Cybersecurity

In the field of cybersecurity, AI has the potential to be both a source of previously unheard-of risk and a spur for innovation. Finding the ideal balance between innovation and risk is a strategic necessity that requires foresight, effective governance, and international collaboration. It is not only a technological problem.

We must strike a balance between convenience and security, as well as innovation and risk, when utilizing AI in cybersecurity, particularly in the context of AI versus human decision-making. This might demonstrate that the company has a robust, flexible, and comprehensive governance framework and strategy that addresses people, process, and technology. Some organizations also view cybersecurity from an administrative, physical, and technical perspective. In addition to complying with existing laws, such as the PDPA and new AI-specific regulations, AI systems must also uphold fundamental ethical principles, including fairness, accountability, and privacy. Effective governance ensures that AI systems operate under clear rules, are closely supervised, and align with societal values rather than operating independently.

There are no options in the development of responsible AI. An organization must possess the right skills, knowledge, and steps to develop a responsible AI. In addition to security by design, an organization must employ a human-centred design approach, identifying multiple metrics to assess AI/ML training and monitoring. It is also recommended to directly examine your raw data and understand the limitations of your dataset and model. Furthermore, always test and retest the AI/ML data. Please continue to monitor and update the system even after it has been deployed.

There are several AI-related issues and challenges that we must face. AI may cause hallucinations and bias. At times, the datasets contain unfair risk grading or faulty threat identification that can cause these issues of bias. AI systems should be transparent, easy to understand, and fair. We need AI that can explain how it makes decisions, so people can check, trust, or question them if needed.

However, we need to remember that even though cyber defenders use AI to enhance and strengthen cyber defences, cyber-criminals or perpetrators can also use it as a weapon to conduct illicit criminal activities, such as AI-powered attacks, spreading phishing campaigns, launching much more sophisticated malware attacks, exploiting system vulnerabilities, or generating realistic deepfakes. To prevent the misuse of AI while still encouraging innovation, we must set clear ethical limits. Trust in AI should be built into its design; it cannot be assumed. That is why many experts support a 'zero-trust' approach, where AI systems are constantly checked and tested, not just when they are launched but throughout their use.

It is also crucial to highlight the importance of collaboration in cybersecurity. No organization can work alone. Everyone must be involved and responsible for cybersecurity. There must be cybersecurity collaboration among government agencies, industry, civil society, and academia. These collaborations will include knowledge sharing, threat intelligence sharing, the exchange of best practices, joint workshops, and joint cyber exercises, all aimed at promoting transparency. Public-private partnerships and global forums are also essential in aligning diverse perspectives and ensuring that AI adoption is both secure and ethical.

It is not just engineers, IT personnel, or innovators who need to understand AI; policymakers, regulators, and the public as a whole must also understand how it works and its implications for society. AI literacy goes beyond basic digital skills. It requires continuous learning because technology is evolving at an unprecedented rate. We must go beyond merely discussing ethical principles and start putting them into practice. That means using industry-specific guidelines that provide real, practical steps from identifying threats and fixing weaknesses to defending against attacks targeting machine learning. These tools transform good intentions into tangible protection, ensuring that AI systems are not only advanced but also secure and reliable.

Innovation and risk should not be viewed as mutually exclusive or separate from each other; they must be managed together, hand in hand. We can unlock the full power of AI while safeguarding digital trust, our institutions, and the people we serve. This requires strong governance, ethical design, robust security, collaboration, and continuous learning.

Reference

Brooke Becher, The Three Laws of Robotics: What Are They?, November 18, 2024, Built in, retrieved from, https://builtin.com/articles/3-laws-of-robotics.

Rob Thubron, Phone network employs AI "grandmother" to waste scammers' time with meandering conversations, November 14, 2024, Techspot, retrieved from, https://www.techspot.com/news/105571-phone-network-employs-ai-grandmother-waste-scammers-time.html.