CyCon 2024: Over the Horizon
16th International Conference on Cyber Conflict
C. Kwan, L. Lindström, D. Giovannelli, K. Podiņš, D. Štrucl (Eds.)
2024 © NATO CCDCOE Publications, Tallinn

Legal, Policy, and Compliance Issues in Using AI for Security: Using Taiwan's Cybersecurity Management Act and Penetration Testing as Examples

Wei-Che Wang
National Institute of Cyber Security
Taipei, Taiwan
wayne@nics.nat.gov.tw

Abstract: As artificial intelligence (AI) technology advances rapidly, integrating AI into cybersecurity practices poses new challenges for professionals. This paper focuses on the legal and policy implications of employing AI tools in penetration testing (PT). Key issues explored include liability in cases where AI tools cause damage and legal compliance challenges for organizations mandated to conduct PT. This paper argues that in the case of Taiwan, a comprehensive consideration of relevant laws, such as the Code of Civil Procedure, will be needed as AI products and tools become more widespread. The other issue concerns defining qualified PT, using Taiwan's Cybersecurity Management Act as an example. This paper concludes that, in addition to proper AI governance, governments should consider the legal frameworks necessary for the practical application of AI products or systems and develop appropriate AI safety testing methods to offer reference guidelines for public agencies to introduce risk-controllable AI tools, thus preparing for the transition into the AI era.

Keywords: *AI for cybersecurity, penetration testing, legal compliance, AI policy, product liability, Taiwan's Cybersecurity Management Act*

1. INTRODUCTION

As artificial intelligence (AI) technology advances, its relationship with cybersecurity has become increasingly noteworthy. Researchers have recently explored the potential of AI to automate cyber attacks. Some studies have focused on AI trained through machine learning methods, indicating that, while such AI might not currently revolutionize cyber attack techniques, it can effectively enhance the efficiency of each step in the Cyber Kill Chain.[1] Other researchers have used deep learning to investigate the performance and feasibility of AI tools in conducting automated penetration testing.[2] This suggests that AI can be not only a powerful tool to increase the threat of cyber attacks but also a potential means of bolstering an organization's cybersecurity defenses.

Market research shows that the value of AI cybersecurity products is still rising, a trend driven, not surprisingly, by the escalating severity of cyber attacks.[3] This increase in cyber threats, closely linked with the maturation of technologies like 5G cellular networks and IoT (the internet of things), compels government agencies, enterprises, and even individuals to allocate more resources for cybersecurity. For instance, large-scale data breaches can lead to significant financial losses and damage a company's reputation. Among various cybersecurity products, AI has emerged as a crucial technology in solutions, speeding up the identification of and response to cyber threats. Consequently, cybersecurity products augmented with AI technology are gaining popularity in the market.

However, throughout its development, AI has generated controversy. Issues raised include the lack of algorithmic transparency; vulnerability to cyber threats; potential discrimination in decision-making; contestability in AI decisions; the legal status of AI; intellectual property rights issues in AI; impact on labor, employment, and economic matters; privacy and data protection; accountability for damages caused; and lack of mechanisms for risk accountability.[4] These controversies have garnered significant attention and debate in the past. With the widespread adoption of neural network methodologies, also known as "black boxes," understanding how AI arrives at a specific answer or decision has become increasingly challenging. Today, as AI applications become more extensive and varied, the importance of these issues grows, necessitating more urgent attention and resolution.

1 Ben Buchanan et al., Automating Cyber Attacks (Center for Security and Emerging Technology 2020), https://doi.org/10.51593/2020CA002.

2 Zhenguo Hu et al., Automated Penetration Testing Using Deep Reinforcement Learning (2020), https://www.jaist.ac.jp/~razvan/publications/automated_penetration_testing_reinforcement_learning.pdf.

3 *Artificial Intelligence (AI) In Cybersecurity Market Size USD 102.78 BN by 2032*, NASDAQ OMX's News Release Distribution Channel (Jan. 23, 2023), https://www.proquest.com/wire-feeds/artificial-intelligence-ai-cybersecurity-market/docview/2768121329/se-2.

4 Rowena Rodrigues, Legal and Human Rights Issues of AI: Gaps, Challenges, and Vulnerabilities (2020), https://doi.org/10.1016/j.jrt.2020.100005.

As AI applications become more widespread, an increase in related legal disputes is anticipated. According to a study published by the Stanford Institute for Human-Centered Artificial Intelligence, in 2022, the United States saw 110 AI-related litigation cases, 6.5 times more than in 2016.[5] Of these, 29% were civil law cases, 19% were related to intellectual property rights, and 13.6% were contract law. Currently, civil cases greatly outnumber criminal or national security-related cases.[6] As AI increasingly impacts people's lives, the number of ensuing legal disputes will rise accordingly.

For instance, the penetration testing operations discussed in this article inherently carry certain security risks for the tested systems. If AI-driven testing leads to property damage or, more gravely, endangers human life, determining the legal relationships involved and allocating responsibility becomes a critical issue. This paper explores these aspects by examining AI policies and legal frameworks in major countries today.

Different governance methods can be chosen depending on the required purpose or degree of enforcement, typically including policies, regulations, and reference guidelines. Legislation for emerging technologies must consider various aspects, such as the law's purpose, the subjects and scope under its regulation, how the law will be implemented, and its societal impact. Therefore, countries often allow new technologies to function and develop in society for a period of time to ascertain their ramifications before legislating. In the meantime, countries usually outline their approach to these technological issues through policies (such as national investment in development or encouraging public–private collaboration) and practical reference guidelines that provide interpretations and practical examples. This approach mitigates the impact of emerging technologies while allowing them to create more possibilities for overall technological advancement, transitioning society and legal systems smoothly from a regulatory vacuum to an established framework.

The legal and policy issues surrounding AI are currently in this transitional phase. As AI capabilities and applications continue to mushroom, many countries and international organizations have extensively discussed how to impose more legal obligations on AI (and its developers, trainers, or users). While many similar conclusions have been drawn, such as the need for trustworthy AI and algorithmic transparency, the question of whether to incorporate these conclusions into legal regulations and the potential impacts of such legislation are still being considered. Currently, National governments' approaches to AI legal policies can be broadly categorized into two types: "active legislative regulation" and "guiding free development." The former, exemplified by the European Union's AI Act, directly regulates subjects, AI systems, requirements, and legal effects (penalties). By contrast, the latter approach, observed

5 Artificial Intelligence Index Report 2023 (Stanford Institute for Human-Centered Artificial Intelligence 2023) 291.

6 *Id*. at 294.

in countries like the United Kingdom and the United States, guides the development of new technologies in a government-friendly direction without imposing restrictions on industrial and technological growth through significant policy documents and technical reference guidelines.

Irrespective of the approach taken in policy regulation, autonomous tools developed through AI training have substantially impacted people's lives. Hence, this paper focuses on the context of "conducting penetration testing," discussing the legal issues that may arise when autonomous AI tools make it more convenient to carry out such testing. Furthermore, it examines compliance challenges that institutions, organizations, or enterprises might face when incorporating AI products or services, using the requirement for regular penetration testing under Taiwan's Cybersecurity Management Act as an example.

2. PENETRATION TESTING WITH AI

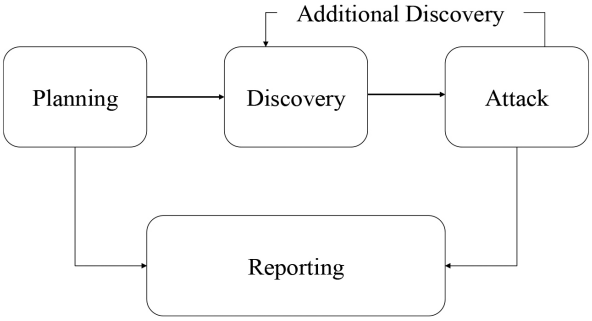
A. Introduction

Penetration testing is a method for assessing the security of information systems and detecting vulnerabilities. Tests are typically conducted by experienced cybersecurity experts. Testers simulate attacks using the same techniques and tools as attackers, often involving finding combinations of vulnerabilities on one or more systems. These combinations can grant more access privileges than would be possible through a single vulnerability, helping organizations (such as government agencies or businesses) improve their security defenses. By simulating attackers' behaviors, penetration testing can uncover vulnerabilities and weaknesses that many organizations are unaware of, thereby providing recommendations for improvement. Organizations can then use the tests' results and reports to better understand their system's security status, strengthen previously undiscovered weaknesses, and further protect their critical data and business operations.

Penetration testing can be divided into several key stages. This article explains these stages based on the guidance document of the National Institute of Standards and Technology (NIST) in the United States.[7] This document divides penetration testing into four stages: the Planning stage, the Discovery stage, the Attack stage, and the Reporting stage. The relationship between each stage is illustrated in Figure 1.

7 NIST, Technical Guide to Information Security Testing and Assessment (SP 800-115) (2008), https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-115.pdf.

FIGURE 1: STAGES OF PENETRATION TESTING



In penetration testing, the Planning stage sets the operational groundwork, including establishing rules, securing managerial consent, and defining clear objectives. This is crucial to safeguard testers, whose actions resemble legal hacking, and for organizations to differentiate between testing and actual cyber threats. The Discovery stage involves data collection and scanning, such as gathering the target's basic details (host names, IP addresses, system information, etc.), followed by vulnerability analysis, which uses this information to identify potential security gaps. Central to the testing, the Attack stage uses collected data to attempt system breaches, confirming vulnerabilities and their impact on system security. The approach adapts based on attack outcomes and additional information gathered, highlighting the interplay between the Discovery and Attack stages. The Reporting stage develops during other stages, culminating in a final report that outlines detected vulnerabilities and suggests reinforcement measures, building on the test plan's objectives and norms.

While penetration testing contributes significantly to enhancing system security, it can also potentially cause harm. For instance, simulated attacks might inflict actual damage on the system. Therefore, before conducting tests, testers should thoroughly understand the system architecture and develop contingency plans to reduce the likelihood of actual harm. Additionally, the testing process could expose sensitive information due to improper data or tool management or be exploited by hackers. Consequently, strict security measures must be implemented when the tests are carried out in order to ensure the tests' safety and confidentiality. Beyond technical controls, it is also essential to ensure that testers possess adequate professional skills and ethical integrity before testing. Testers should assess and manage potential risks to minimize their impact. In addition to controlling risks of actual harm, understanding the legal allocation of responsibility in the event of actual harm is crucial. This legal liability for any real harm should be managed through contractual arrangements or insurance.

B. The Theory and Practice of Autonomous Penetration Testing

Many tools now automate penetration testing through simulated environments by scanning and analyzing network structures and deployment environments and attempting attacks on known vulnerabilities. Manual penetration testing relies heavily on the tester's knowledge, experience, and skills, requiring significant time, effort, and resources. Autonomous AI tools, however, can streamline this process by autonomously exploring potential paths and weaknesses, analyzing intrusion strategies, and adapting to new information during the test. This automation can significantly lower the barrier to penetration testing and enhance efficiency.

Penetration testing includes various aspects such as software, hardware, environments, and personnel. While certain aspects, such as exploiting human errors or specific habits, may still rely on human experts, much of the testing, such as identifying vulnerabilities in software versions or system configurations, can be done autonomously. Key steps of penetration testing—target scanning, strategy formulation, and attack execution—can now be handled by autonomous tools shown in Table I, suggesting the feasibility of an integrated AI tool capable of conducting a complete penetration test with a single command. Thus, the concept of fully autonomous penetration testing is increasingly becoming a reality.

TABLE I: KEY STEPS AND CORRESPONDING TOOLS IN AUTOMATED PENETRATION TESTING

Steps	Description	Tools/methods used	Purpose/outcome
Reconnaissance	Scanning and detecting network topology	Nmap	Identify network structure and potential targets
Simulation	Simulating network architecture	CyberBattleSim	Understand the network environment and potential vulnerabilities
Strategic Planning	Planning penetration strategies	AutoPentest-DRL	Develop a strategic approach to penetration testing
Execution	Actual penetration testing operations	Metasploit	Carry out the attack to identify vulnerabilities
Reporting	Generating the test report	ChatGPT	Provide a detailed analysis and findings of the penetration test

With autonomous tools that can scan and detect the target network topology and the rapid advancement of AI technology, autonomous target penetration can be achieved. Even with current tool capabilities, autonomous penetration testing can integrate various AI tools for different tasks: Nmap for reconnaissance, CyberBattleSim for

network architecture simulation, AutoPentest-DRL for strategic penetration planning, Metasploit for actual operations, and large language models like ChatGPT for report generation. This concept of using a single AI tool for penetration testing is gradually becoming a reality.

The advantage of autonomous penetration testing tools is that they allow smaller organizations or businesses with limited resources to conduct thorough cybersecurity defense checks. This can significantly benefit national cybersecurity. However, the adoption of new technological tools should be approached with caution to avoid unforeseen risks.

C. Discussion of Legal Issues in Autonomous Penetration Testing

As previously explained, a consensus has yet to emerge on the regulatory framework for AI systems (or products, services, etc.). Given the global scale of the free market and the significance of major AI companies like Microsoft, Google, and Meta, it is crucial to closely monitor the regulatory developments in major economics like the European Union and the United States. These developments are likely to shape the direction of legal compliance for the entire AI industry. In response, Taiwan should closely observe international regulatory and policy trends, explore various emerging legal issues, and propose relevant legal and policy recommendations. This proactive approach aims to ensure a smoother integration of Taiwan's legal system with AI regulations and standards as AI becomes more widespread and its regulatory framework begins to take shape.

1) Exploring Liability for Damages Caused by AI Products

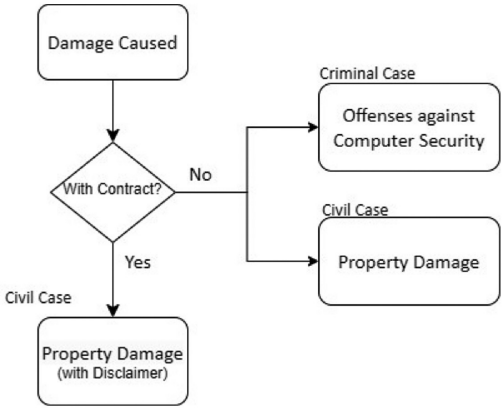
Over the past few years, AI has been extensively applied in numerous fields, such as autonomous vehicles and medical diagnostics. There has been considerable debate over whether AI can be considered a subject of tortious acts, that is, whether AI itself can be held responsible for damages caused. The prevailing opinion is that, under current legal principles, the answer remains unclear.

Accordingly, in the current legal system, until a new type of legal personhood for AI is defined, making AI systems responsible for the damages they cause, the primary subjects of liability are still those recognized as legal personalities under current law, namely, natural persons or legal entities. In the context discussed in this paper, corporations mainly involve the developers, providers, or suppliers of the AI tool, while natural persons are typically those using, operating, or carrying out the autonomous AI tools, issuing commands to carry out one or more stages of penetration testing.

When an actor uses the aforementioned autonomous tools and they result in damage to an organization, the relevant civil and criminal liabilities must be discussed. The

content involved in the planning stage of penetration testing plays a crucial role in making such a determination. That is, if written documents from the planning stage prove that the organization consented to the actor's hacking (testing) actions, it is easier to establish a contractual relationship to conduct penetration testing between the actor and the organization. This relationship, in addition to being used to determine whether the computer crimes under criminal law are "without cause," may also potentially exclude damages within a certain range from the compensation scope based on the agreement's substantive terms. The legal assessment is shown in Figure 2. In Figure 2, when damage is caused, the type of responsibility can be divided according to whether there was a contract. In a scenario that included a contract, the main responsibility would concern any damages that were caused outside the scope of the contract. In the absence of a contractual agreement, the victim can pursue both civil and criminal charges.

FIGURE 2: LEGAL ASSESSMENT OF THE DAMAGE TO INSTITUTIONAL INFORMATION SYSTEMS CAUSED BY AI TOOLS



In this scenario, the acting subject should be a natural person (for example, the engineer carrying out the testing project), and the autonomous AI program serves as a tool for conducting the test. The engineer is expected to supervise and intervene as necessary during the operation of the testing tool. For instance, actions such as immediate cessation, restoration, or repair should be taken if the autonomous tool successfully breaches a system, as this could lead to sensitive data leakage. Therefore, according to Taiwan's civil law provisions, tort liability is established if the autonomous AI tool causes property damage under the supervision and use of that natural person. Moreover, if it can be proven that the actor was negligent, they can be held responsible.

However, in most civil litigation cases, the allocation of the burden of proof substantially influences the outcome of the lawsuit. For example, according to Article 277 of Taiwan's Code of Civil Procedure, when a party asserts facts that are favorable to them, they have the responsibility to provide evidence for those facts. In other words, in the case scenarios discussed in this article, the party that suffered damage must prove that the AI tool caused the system's damage. Furthermore, to establish tort liability, they may also need to prove that the cybersecurity personnel responsible for supervising the use of the AI tool were negligent or worse.

Beyond civil liability, since penetration testing inherently involves acts of computer crime against information communication systems, relevant provisions can refer to Article 358 and subsequent articles of Taiwan's Criminal Code. Examples of such acts include "entering another's account and password, cracking computer protection measures, or exploiting computer system vulnerabilities to intrude into another's computer or its related equipment," "accessing sensitive information by obtaining, deleting, or altering the electromagnetic records of another's computer or its related equipment, causing damage to the public or others," or "using computer programs or other electromagnetic methods to interfere with another's computer or its related equipment, causing damage to the public or others." All these can constitute elements of computer crimes. At this point, whether the related intrusion actions have a legitimate reason for committing this "criminal act" is very important. This is also why the first penetration testing phase discussed in this article emphasizes the importance of project authorization documents.

2) EU's Product Liability Directives with AI

To develop trustworthy AI, the European Commission proposed a draft AI Liability Directive (AILD) in 2022,[8] which, along with the aforementioned AI Act, shapes the EU's legislative framework for AI. The primary purpose of the AI Liability Directive is to ensure that if users suffer harm due to AI products, the burden of proof for claims against AI is reduced. Additionally, clarifying how responsibility is allocated helps companies providing AI products or services to assess risks and reduce legal uncertainties. In line with the AI risk classification structure established by the AI Act, the new AILD applies in two scenarios. First, in claims for civil liability for negligence in non-contractual relationships, it requires disclosure of evidence concerning high-risk AI. Second, it adjusts the burden of proof in EU (member states') courts for compensation claims for damages caused by AI systems under non-contractual civil law.

In the section related to high-risk AI liability, this directive grants courts the power, under specific circumstances, to require relevant personnel of the high-risk AI (such as service providers) to disclose evidence related to the AI. To ensure fairness between

8 *Liability Rules for Artificial Intelligence*, European Commission, https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en.

both parties regarding evidence and information disclosure, the directive also sets many restrictions on the circumstances mentioned above. For example, the plaintiff must have already requested evidence from the AI-related personnel and been refused, and at the same time, the plaintiff must present sufficient facts and evidence to support the claim. On the other hand, the court should limit the scope of evidence disclosure to ensure that what is disclosed is relevant to the claim.

Although the regulations are more detailed in their applicability and conditions for high-risk AI, both the content of the regulations and the discussions during the legislative process indicate that the EU anticipates an increasing number of cases where AI products will conflict with human rights. Regardless of whether the current litigation system and the allocation of the burden of proof can properly handle such disputes, this AILD proposal may provide valuable references for Taiwan's legal and policy considerations.

At the same time, the EU is also considering amending the existing Product Liability Directive.[9] Since 2018, the European Commission has been amending the existing Product Liability Directive. The currently proposed revision has primarily updated three parts of the directive. First, it addresses the legally unclear concepts in the application of the law to emerging technologies. Second, it addresses the burden of proof that works against victims in cases involving products of emerging technologies, such as self-driving cars and AI products. Third, the previous Product Liability Directive had a threshold of €500 for claims, meaning that damages not reaching this amount could not be claimed; in the current proposal, this threshold has been removed.

The Product Liability Directive imposes responsibility on the economic operator of a product if a natural person suffers harm due to a defect in that product. In terms of enhancing the clarity of legal concepts, the new directive draft explicitly includes in the definition of "product" items commonly seen in the digital age, including digital files and software. Thus, AI-related products may also fall within this scope.

Regarding the requirements for the burden of proof, in addition to putting forward several presumptions where the causal relationship of damages is established, the new directive draft adds that if a plaintiff faces difficulty in proving the causal relationship between the product and the damage due to "technical or scientific complexity," the court may, under certain conditions (for example, if the plaintiff has provided sufficient evidence that the product is likely defective), acknowledge the causal relationship between the product and the damage for compensation. This legislative proposal echoes the approach taken in the AI Liability Directive, addressing the challenge of establishing tort liability after presenting strong evidence due to the high complexity

9 *New Liability Rules on Products and AI to Protect Consumer*, European Commission (Sep. 28, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5807.

of technology, which is difficult to handle with traditional legal concepts. Table II compares the two proposals.

TABLE II: COMPARISON OF THE EU'S AI LIABILITY LEGISLATION PROPOSALS

Aspect	AI Liability Directive	Product Liability Directive Amendment
Scope	High-risk AI and non-contractual civil liability claims	Includes products of emerging technologies, such as AI and self-driving cars
Legal Clarity	Addresses liability issues specific to AI, including complex AI systems	Enhances clarity of legal concepts for products in the digital age
Burden of Proof	Adjusts burden of proof in favor of victims in AI-related cases	Adds presumptions for the causal relationship in cases of technical/scientific complexity
Damage Threshold	None	Removes the €500 damage threshold for claims
Inclusion of Digital Products	Broader scope to AI-related products	Directly includes digital files and software

3) Legal Regulations Related to AI Liability in Taiwan

On the other hand, when it comes to Taiwan's liability-related regulations, discussions regarding property damage caused by the use of autonomous AI products can be approached from both substantive and procedural law perspectives. Taiwan's legal system is deeply influenced by the civil law tradition. The provisions of substantive law mainly involve civil law and consumer protection law. Additionally, when the damage pertains to the use or leakage of personal data, the Personal Data Protection Act might apply. On the procedural law side, because the burden of proof comes into play in litigation, it is necessary to consider whether civil litigation laws should be adapted to reduce the plaintiff's burden of proof in lawsuits concerning disputes over AI product liabilities.

In Taiwan's civil law, the issue of compensation for property damage needs to be addressed. When AI tools are used for penetration testing and this results in damage to the tested host system, given the frequent information asymmetry between product manufacturers and consumers in terms of economics, knowledge, and product performance, Taiwan has established the Consumer Protection Act[10] to safeguard consumer interests. This act specifically addresses business operators involved in designing, producing, and manufacturing goods or providing services. When these goods enter the market or services are provided, business operators must ensure the safety of these goods or services, according to the reasonably expected standards of

10 Consumer Protection Act, art. 7, https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=J0170001.

current technology and professional expertise at that time. The burden of proof for claims regarding these facts also rests with the business operators.

Under the current provisions of Taiwan's civil litigation law, a party who asserts facts favorable to their case bears the burden of proof (Article 277). However, recognizing that not all types of disputes can be adequately addressed by this general rule, the law provides a flexible adjustment mechanism with the provision: "This limitation does not apply if there are specific legal stipulations to the contrary, or if adhering to this rule would manifestly be unfair."

A classic example of a type of dispute in which the burden of proof is allocated differently is medical disputes. Other types include environmental pollution and traffic incidents; there, Taiwan's legislation also explicitly includes the responsibility of product manufacturers. In medical litigation, due to the high level of expertise, uncertainty, unpredictability, and information asymmetry inherent in medical practices, and given that the general public lacks relevant professional knowledge, it is often challenging to prove negligence in medical acts by hospitals or physicians, especially since medical records and equipment are predominantly controlled by them. Therefore, a shift in the burden of proof is applicable in these cases.

Damages caused by AI products share similar characteristics. In autonomous penetration testing, because AI models function like a "black box" (meaning that the reasons for their decisions cannot easily be discerned), operators, despite their own IT or cybersecurity expertise, may rely on the AI's judgment. Critical information about the AI model's training, judgment, or decision-making logic is usually held by the operators or suppliers who trained and adjusted the AI product.

D. Legal Implications of AI in Penetration Testing

The legislative model of the EU demonstrates that the governance policy for AI and the subsequent design of the responsibility structure must be considered simultaneously. That is, the categorization of AI systems must precede the question of whether to impose specific responsibilities on certain categories of AI.

As the above discussion shows, in terms of legal liability arising from damages caused by using AI tools for penetration testing, Taiwan already has applicable provisions for both civil and criminal liabilities. However, it may still be necessary to adjust the relevant regulations based on the characteristics of AI tool products, such as the issue of the burden of proof in civil damage compensation lawsuits. Additionally, in terms of administrative responsibility, we can refer to the legislative model of the EU AI Act concerning roles such as manufacturers or suppliers of AI tools, products, or services.

After establishing the basic legislation for Taiwan's Artificial Intelligence Act, we can then continue with special legislation to regulate these important obligations.

3. DISCUSSION OF LEGAL COMPLIANCE ISSUES OF AUTONOMOUS PENETRATION TESTING

A. The Cybersecurity Management Act and the Role of Penetration Testing

Given the practices in the cybersecurity industry and the provisions of Taiwan's Cybersecurity Management Act, it is not difficult to see that penetration testing is a relatively high-standard requirement in current cybersecurity defense testing. For example, the current Cybersecurity Management Act (and its related subsidiary laws, collectively called the CMA) requires organizations with a cybersecurity responsibility Level-C or above[11] to regularly conduct penetration testing on their core information communication systems. Additionally, for core information communication systems classified as "high" in protection level, the CMA also requires penetration testing during the development and acquisition stages of the Secure Software Development Life Cycle (SSDLC). Furthermore, according to the Enforcement Rules of the CMA,[12] agencies are required to include provisions for penetration testing when outsourcing their customized information and communication systems to ensure compliance and enhance security measures. This demonstrates the importance of penetration testing in compliance with Taiwan's cybersecurity-related legal requirements.

However, as this paper mentioned, a common issue agencies face in practice is that penetration testing requires considerable expenses and resources. If an agency decides to use autonomous tools for penetration testing and report generation, additional issues need to be addressed, such as whether these reports meet legal compliance requirements.

B. The Legal Effect of Penetration Testing

To answer the aforementioned question, it is necessary to explore what practical effects policymakers hope to achieve. Generally, if the process only involves simple scanning of a system for known vulnerabilities, it is usually referred to as a vulnerability scan. On the other hand, there may be exceptional cases in which professionals are hired for penetration testing but—due to negligence or other reasons—the results of their testing do not differ significantly from those obtained through mere tool scanning.

11 The cybersecurity responsibility levels of government agencies and specific non-government agencies are classified from high to low into Level-A, Level-B, Level-C, Level-D, and Level-E. Agencies rated Level-C and above are defined by regulations as those that maintain and operate, or outsource the establishment and development of, their cybersecurity systems. Regulations on Classification of Cyber Security Responsibility Levels, https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=A0030304

12 Enforcement Rules of Cyber Security Management Act, art. 4, https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=A0030303.

Therefore, this paper holds that, in defining what constitutes legally compliant penetration testing, one should observe whether the effects of the testing meet the needs of the agency's cybersecurity defense rather than merely determine whether tools were used for the test or whether there was human involvement.

As for what content meets an agency's cybersecurity protection needs, the "Solicitation Document for Government Agency Penetration Testing Service Outsourcing Proposal (Template),"[13] issued by the National Institute of Cyber Security in Taiwan, provides guidance. Besides listing all the necessary security testing items, the document makes two main points. The first is that the testers must have qualifications, such as cybersecurity certifications like CEH (Certified Ethical Hacker), CPENT (Certified Penetration Testing Professional), and so on. The second is that the submitted test report must be comprehensive. It should not only detail the methods used to discover vulnerabilities and the attack techniques employed and assess the risk level of the vulnerabilities but also—and this is most important for the agency—provide actionable improvement recommendations so that the agency can follow these to strengthen protection after the test.

Therefore, this paper holds that in the legal compliance issue of autonomous penetration testing, the judgment should be based on the effectiveness of the testing rather than merely on whether it is completed by autonomous tools. As for whether the use of autonomous tools (such as utilizing language models like ChatGPT to write vulnerability analysis reports) could lead to the leakage of the agency's sensitive information (for example, by inputting important core system configurations or internal network architecture information), that is another regulatory issue that needs consideration.

C. Balancing Automation and Risk

The impact of AI development is less about replacing humans than about assisting them—that is, making tasks that originally required many resources and had higher barriers easier to conduct or access. The penetration testing discussed in this paper is an example. If in the future, agencies start effectively using autonomous tools for penetration testing, that would be a positive development. However, this phenomenon requires attention from both the agencies themselves and the higher-level units conducting audits.

For the agencies themselves, although they might use AI tools for testing, they still need to be aware of the related risks, including damage to the information communication systems during testing or the leakage of sensitive information to AI tools, as mentioned earlier. To that end, this paper suggests that agencies should not

13 National Institute of Cyber Security, Solicitation Document for Government Agency Penetration Testing Service Outsourcing Proposal (Template), https://download.nics.nat.gov.tw/UploadFile/attachfilespmo/%E6%BB%B2%E9%80%8F%E6%B8%AC%E8%A9%A6%E6%9C%8D%E5%8B%99RFP%E7%AF%84%E6%9C%ACv5.0_1100915.pdf.

only thoroughly assess AI tools before selection and choose autonomous tools with lower risks but also hire operators with professional knowledge or qualifications.

Furthermore, auditors reviewing penetration testing reports provided by agencies will need to clearly understand the related background information of the report, including methods of execution, the tools used, test items, methods, scope, and improvement suggestions. That is, whether the testing complies with regulations is independent of whether it was conducted manually or automatically. Judgment should still be based on the substantive content of the testing.

4. CONCLUSION

The legislative model of the EU demonstrates that the governance policy for AI and the subsequent design of the responsibility structure must be considered simultaneously. That is, the categorization of AI systems must precede the question of whether to impose specific responsibilities on certain categories of AI. On the other hand, Taiwan's experience with cybersecurity legislation shows that the increasingly powerful performance of AI tools will also affect compliance with existing regulations or the auditing of standards.

In light of the associated cybersecurity risks that may accompany the use of AI automation tools, this paper posits that while governments contemplate AI governance, they must also be attentive to ancillary approaches. For instance, with respect to the product liability of AI, as discussed in this paper, clear legal norms are needed that delineate product responsibilities. Moreover, for legal and compliance issues related to AI automation tools in various types of legal operations, standards or auxiliary guidelines should be established based on practical scenarios to address the impact of AI. Below, this paper also offers three recommendations.

First of all, AI tools should be accepted as assistance. AI's role is predominantly to simplify and make accessible tasks that were previously resource-intensive and complex. An illustration of this can be seen in the penetration testing discussed in this paper. Should agencies begin to effectively deploy autonomous tools for penetration testing in the future, it would represent a significant advancement. However, such a shift demands vigilant oversight from both the agencies involved and the higher-level authorities responsible for auditing their activities.

Secondly, a list of usable AI tools should be established. While these tools are beneficial for testing, agencies must remain cognizant of potential risks, such as possible damage to information communication systems or unintended exposure of

sensitive data during the testing phase. Consequently, this paper recommends that agencies rigorously evaluate AI tools to select those with minimal risks and also ensure that they employ skilled operators who have the necessary expertise and credentials.

Finally, the validation of penetration test reports must prioritize the depth and quality of the content over superficial elements. It is imperative that auditors who review these reports from various agencies gain a comprehensive understanding of the detailed context provided within them. This includes not only the methodologies and tools employed but also the specific areas tested, the scope of the tests, and any recommendations for improvements. Importantly, the compliance of these tests with regulatory standards should be judged independently of the methods used, whether manual or automated. Decisions should be rooted in a thorough assessment of the actual findings and outcomes of the tests, emphasizing the importance of substance over form in these evaluations.