TRANSNAV
http://www.transnav.eu
the International Journal
on Marine Navigation
and Safety of Sea Transportation
Volume 19
Number 2
June 2025
DOI: 10.12716/1001.19.02.34

Artificial Intelligence in Maritime Cybersecurity: Dual-Use Applications for Defense and Offense in the Age of Digital Seas

R. Cichocki
Gdynia Maritime University, Gdynia, Poland

ABSTRACT: The maritime sector's rapid digital transformation – including the integration of IT and operational technology (OT) systems and the rise of autonomous vessels – has significantly expanded the cyberattack surface[1] . Artificial Intelligence (AI) now plays a dual role in this landscape, acting as both a powerful enabler of cyberattacks and a critical tool for cybersecurity defense [2] . This paper explores current and emerging uses of AI in offensive and defensive cyber operations targeting maritime systems and infrastructure. On the offensive side, threat actors are leveraging AI for sophisticated attacks such as AI-generated spear phishing, polymorphic malware generation, GPS spoofing, and manipulation of industrial control systems (ICS)[3], [4]. On the defensive side, AI is employed in anomaly detection, predictive analytics, autonomous vessel and port monitoring, and other security applications[5]. The paper also examines vulnerabilities of AI itself – including adversarial attacks, data poisoning, and model manipulation – and discusses strategies to enhance maritime cyber resilience. Key strategies include the use of digital twin simulations, AI-driven deception (honeypots), adversarial training, explainable AI, and international cooperation for information sharing. By analyzing both offensive and defensive developments, this study provides a comprehensive perspective on the dual-use nature of AI in shaping the future of maritime cybersecurity.

1 INTRODUCTION

The maritime industry is increasingly reliant on networked digital systems for navigation, cargo management, and vessel operation. Modern ships and ports utilize complex ecosystems of satellite communications, ship-to-shore data links, and remote monitoring, which have greatly improved efficiency and safety 1 . However, this increased connectivity has also widened the attack surface, making the sector vulnerable to cyber threats such as ransomware, GPS spoofing, denial-of-service attacks, and unauthorized intrusions into ship control systems[6] . A report by Marlink observed 9 billion cybersecurity events across 1,998 vessels in just the latter half of 2024, noting that while the volume of attacks remained high, their sophistication had notably increased[7], [8]. One striking finding was the rise of generative AI tools in cyberattacks: threat actors now use large language models (LLMs) to speed up malware creation, automate phishing campaigns, and refine social engineering tactics[3]. These AI-assisted techniques have enabled attackers to exploit vulnerabilities more easily, escalating the pressure on the maritime sector to bolster defenses [9]. A sobering example of the potential impact is the 2023 grounding of the container ship MSC Antonia off Saudi Arabia's coast, reportedly caused by a cyber spoofing attack that manipulated the vessel's GPS-based location data[4] . This incident highlights how cyberattacks can translate into severe physical consequences in maritime environments.

Amid these evolving threats, AI technologies are a double-edged sword. On one hand, AI provides powerful tools for attackers to enhance their cyber offensive capabilities. On the other hand, it offers advanced defensive solutions for detecting and mitigating attacks in real time. AI's role in maritime cybersecurity is therefore inherently dual-use[2]. For example, the same advancements in machine learning that enable autonomous ship navigation and smart port operations 10 could be repurposed by adversaries to identify and exploit vulnerabilities at scale. Conversely, AI-driven anomaly detection systems can help defend against novel attack patterns that human operators might miss[5].

This paper aims to examine both sides of this AI paradigm. Section II discusses how AI is leveraged in offensive cyber operations against maritime targets, detailing techniques like AI-generated phishing, malware, and autonomous hacking. Section III explores defensive applications of AI in the maritime domain, including intelligent intrusion detection, vessel anomaly detection, and port cybersecurity systems. In Section IV, we address the vulnerabilities and challenges associated with employing AI for maritime cybersecurity – such as adversarial attacks on AI models and the lack of maritime-specific datasets. Section V then outlines strategies to enhance resilience, from digital twin environments for training and testing, to AI-powered deception, robust AI model training, explainable AI, and the need for international coordination. Finally, Section VI concludes with insights on balancing these offensive and defensive dimensions of AI to secure the future "digital seas."

2 AI IN OFFENSIVE MARITIME CYBER OPERATIONS

AI is increasingly a force multiplier for adversaries conducting cyberattacks in the maritime sector. Threat actors – ranging from cybercriminal groups to state-sponsored APTs – are exploiting AI's capabilities to automate attacks, generate more convincing deception, and discover vulnerabilities more efficiently[3], [10] . Table 1 summarizes several AI-enabled offensive techniques that have emerged in recent years, which we detail in this section.

2.1 *AI-Generated Phishing and Social Engineering*

Phishing is a prevalent attack vector in shipping companies and port authorities, and AI has made such social engineering attacks far more potent. Sophisticated language models can generate phishing emails that mimic the writing style of colleagues or business partners, in multiple languages, and with contextually relevant maritime details. Recent experiments found that GPT-4 can craft phishing messages that significantly outperform those written by humans[8] . In one study, emails created by GPT-4 guided by the "V-Triad" cognitive hacking framework had a click-through rate up to 79% in simulations, versus below 30% for control emails[8] . Similarly, a cybersecurity training firm Hoxhunt reported that between 2023 and 2025, AI-crafted spear phishing emails evolved from being less effective than human-crafted ones to 24% more effective at tricking recipients as shown in Table 2 [9].

Table 1. AI-Enabled Offensive Techniques in Maritime Cybersecurity

Offensive Technique	AI Role and Impact	Example / Evidence
AI-Generated Phishing Emails	Large Language Models craft highly convincing spear-phishing messages targeting maritime personnel, with content tailored to exploit human biases and maritime context. This increases success rates of phishing attacks.	GPT-4 combined with a cognitive hacking model ("V-Triad") achieved click-through rates up to 79%, far higher than the ~20–30% of control phishing emails[8]. An AI phishing campaign in 2025 was 24% more effective than human-crafted attempts[9].
AI-Assisted Malware (BlackMamba)	AI is used to generate polymorphic malware code on the fly. Malware can evolve its payload or obfuscation using AI, making each instance unique and harder to detect by security tools.	BlackMamba is a proof-of-concept polymorphic keylogger that uses a neural network (via an LLM) at runtime to synthesize malicious code, enabling it to constantly change its signature and evade endpoint detection[11].
Malicious AI Chatbots (WormGPT)	Customized LLM-based tools without safety filters assist criminals in writing exploit code, crafting phishing emails, or automating fraud. These black-market AI services effectively function as "evil ChatGPTs" for offensive purposes.	Underground forums advertise models like WormGPT and FraudGPT, which generate phishing and malware content free of ethical safeguards. Even legitimate AI like Microsoft's Copilot has been misused (via tools like "LOL-Copilot") to draft spear-phishing emails and script insider attacks within compromised corporate environments[12].
Autonomous Vulnerability Exploitation	AI agents can be deployed to conduct reconnaissance and exploit vulnerabilities in maritime systems with minimal human input. By ingesting vulnerability descriptions, an advanced AI (e.g., GPT-4) can identify flaws and even generate exploit code.	Researchers showed that a GPT-4 based agent could autonomously exploit 87% of tested software vulnerabilities (given the CVE descriptions), vastly outperforming other tools or models. In comparison, GPT-3.5 and traditional scanners achieved ~0%, highlighting GPT-4's unprecedented capability in automated hacking[1].

Table 2. AI Agent vs Human RED TEAM fail rates

	2023 Failure rate	Nov 2024 Failure rate	March 2025 Failure rate	Total % change
AI	2.9% *	2.1%	2.78%	
Human	4.2%	2.3%	2.25%	
AI to Human relative performance	-31% less effective than humans	-10% less effective than humans	+23.8% more effective than humans	+55% improvement
Human to AI relative performance	+44.8% more effective than AI	+10.8% more effective than AI	-19% less effective than AI	-142% decline

source: https://hoxhunt.com/blog/ai-powered-phishing-vs-humans

The maritime industry, with its global and sometimes less cyber-aware workforce (e.g. crew members, supply contractors), is particularly vulnerable to such AI-enhanced phishing. There have been cases of phishing emails targeting port payment

systems and vessel agents that were nearly indistinguishable from legitimate communications, raising concerns that generative AI will further fuel targeted attacks.

Business Email Compromise (BEC) and phishing attacks represent a significant and evolving threat to the maritime industry, as highlighted in the 2023 and 2024 Coast Guard CTIME reports. In both years, phishing was a dominant initial access vector—accounting for 66% of CPT mission breaches in 2023 and 53% in 2024[13]. These attacks often involve impersonation of trusted internal entities and domain spoofing, making detection difficult and enabling the compromise of sensitive credentials. BEC incidents are particularly dangerous as they use legitimate accounts to propagate phishing emails, increasing the potential scale of disruption and financial loss[14]. Moreover, the interconnected nature of modern maritime operations—where vessels maintain continuous connectivity with enterprise networks—magnifies the impact of these threats. Successful phishing or BEC attacks can bridge corporate and operational technology (OT) environments, jeopardizing shipboard systems and potentially halting vessel operations. In assessments, Coast Guard Cyber Protection Teams found that even with growing adoption of multi-factor authentication (MFA), many organizations remained vulnerable due to weak implementations or the use of default credentials, which were still present in 71% of organizations in 2024[14]. BEC and phishing are exceptionally dangerous for the maritime industry due to their high success rate, ability to exploit human and system vulnerabilities, and potential to cascade operational impacts across both IT and OT systems. Given the economic criticality of the Marine Transportation System—responsible for 70% of trade weight and 18% of the U.S. GDP—continued vigilance, adoption of phishing-resistant MFA, and improved cyber hygiene are imperative to mitigating these persistent threats[14].

2.2 *AI-Powered Malware and Evasion*

AI is also employed to create more adaptive and evasive malware. A notable example is BlackMamba, an AI-driven keylogger proof-of-concept reported in 2023. Unlike traditional malware which has a static code, BlackMamba's malicious payload is generated at runtime by an AI model, so each execution can produce a slightly different variant of the keylogging code[11]. This polymorphic behavior, enabled by on-the-fly neural network code generation and obfuscated through techniques like Python exec() calls, makes detection extremely difficult for standard endpoint security tools. In tests, the BlackMamba approach was able to evade modern Endpoint Detection and Response (EDR) systems[15] . For maritime targets, such malware could be delivered via infected email attachments or USB drives carried onto vessels, then adapt to evade onboard security. AI can also be used to optimize malware's attack path; for instance, by analyzing a ship's network topology or an electronic chart system's software version, an AI might choose an optimal exploit or payload to deploy. Security researchers warn that we may soon see self-spreading "worm" malware in maritime logistics networks that use AI to continuously morph their code and avoid signature-based detection[11].

2.3 *Criminal AI-as-a-Service:*

The availability of illicit AI tools has lowered the barrier to entry for cybercriminals targeting maritime and other sectors. In 2023–2024, underground marketplaces began offering access to custom-trained LLMs explicitly designed for malicious tasks. WormGPT and FraudGPT are two examples of such black-market AI services[12] . These models are advertised as having no ethical constraints, allowing users to generate phishing emails, social engineering scripts, or even malware code (like ransomware notes or obfuscated malicious macros) automatically. Unlike public AI chatbots that have filters to block illicit content, these underground models will output any attack instructions requested. The emergence of these tools means that even attackers with limited skills or language proficiency can produce professional-grade malicious content. For instance, a fraudster could ask WormGPT to draft a spear-phishing email in perfect Norwegian targeting a shipping company's finance department, or to write a piece of malware that exploits a known vulnerability in port management software. Such AI-as-a-service offerings dramatically increase the scale and sophistication of potential attacks on maritime organizations[3].

Even legitimate AI platforms can be misused. A concerning case discussed at Black Hat 2024 involved attackers repurposing Microsoft 365 Copilot (an AI assistant integrated with enterprise data) into an internal tool for cyber espionage[16], [17]. By compromising a company's environment, researchers demonstrated a tool dubbed "LOL Copilot" which leveraged the AI to harvest internal communications and learn employees' writing styles, mass-generate phishing emails from compromised accounts that matched the tone of the organization, and even assist in data exfiltration by summarizing or translating stolen documents 15 . This shows that AI misuse is not limited to external threats – internal systems with AI capabilities can be hijacked by adversaries if proper safeguards are absent.

2.4 *Autonomous Hacking Agents and ICS Attacks*

Beyond phishing and malware, AI has shown the capability to autonomously carry out hacking tasks that traditionally required skilled human attackers. Fang et al. (2024)[1] demonstrated that LLM-based agents could interpret software vulnerability descriptions and independently develop exploits for them, achieving a success rate of 87% on so-called "one-day" vulnerabilities (flaws that have been disclosed publicly but not yet patched everywhere)[1]. This represents a significant leap in offensive capability – an AI agent can essentially act as an expert hacker tirelessly searching for ways into a system. In a maritime context, this could be directed at shipboard systems, port facility networks, or critical maritime infrastructure. For example, an AI agent could scan an Internet-exposed port control system, identify that it's running an outdated version of SCADA software, and then devise and execute an exploit to take control, all with minimal direct oversight.

Nation-state actors are actively exploring AI to aid in attacking critical infrastructure. A 2024 OpenAI report revealed that an Iranian state-linked hacker group ("Cyber Av3ngers") used ChatGPT as a tool in planning and executing attacks on industrial control systems at water utilities[5], [18]. The hackers queried ChatGPT for information like which industrial control ports and protocols are commonly exposed, default passwords for certain PLC equipment, and techniques for obfuscating malicious code. They also sought help on scanning for vulnerabilities and post-exploitation steps. While ChatGPT's answers mainly provided known information and did not directly give the attackers new capabilities , the incident exemplifies how readily available AI tools can assist in orchestrating attacks on OT systems. In the maritime realm, similar AI-assisted reconnaissance could be used against a ship's propulsion control system, a port's crane control network, or an offshore platform's safety controllers. Attackers with AI can more quickly map out a target's digital terrain and identify weak points. Furthermore, AI might enable adaptive attacks on navigation systems: for instance, generating dynamic GPS spoofing patterns that slowly coax a vessel off course while evading simple anomaly detection (though specific cases of AI-driven GPS spoofing have not been publicly documented, the possibility is a serious concern given how AI can optimize signal modification strategies).

In summary, AI has become an accelerator for maritime cyber threats. It allows adversaries to scale up attack volume and sophistication – from realistic phishing that can fool busy port officials, to malware that evades traditional defenses, to semi-autonomous hacking campaigns against ships and infrastructure. The maritime sector must reckon with this offensive use of AI and anticipate that threats will continue to evolve rapidly as AI technology advances 23 .

3 AI IN DEFENSIVE MARITIME CYBERSECURITY

While adversaries exploit AI for attacks, defenders are equally leveraging AI to protect maritime assets. The unique challenges of maritime cybersecurity – including often-fragile communications links at sea, a mix of IT and OT systems on vessels, and safety-critical operations – call for intelligent, adaptive defence mechanisms. AI and machine learning techniques are being deployed to monitor for anomalies, detect intrusions, and augment limited human cybersecurity resources in maritime operations [5], [19] . This section discusses key defensive applications of AI in the maritime domain, with examples of how these technologies enhance security. Table 3 provides a summary of AI-driven defensive solutions currently being employed or researched.

Table 3. AI-Enabled Defensive Solutions in Maritime Cybersecurity

Defensive Application	Role of AI in Maritime Security	Example / Evidence
Anomaly Detection Systems	Machine learning models analyze network traffic, system logs, and vessel operational data to detect deviations from normal behavior. AI can identify subtle indicators of compromise or abnormal vessel activities in real time.	A maritime cybersecurity platform by Cydome uses integrated AI-powered anomaly detection alongside EDR and NDR to monitor shipboard OT/IT networks, flagging suspicious outbound connections (potential command-and-control traffic) by leveraging tools like Suricata and Zeek. Researchers have also applied neural networks to Automatic Identification System (AIS) data to detect anomalous vessel trajectories (e.g., unexpected route deviations that may indicate GPS spoofing or illicit behavior).
Predictive Analytics & Maintenance	AI algorithms predict potential failures or security incidents before they occur, allowing proactive measures. In cybersecurity, predictive models can forecast attack likelihood based on patterns; in operations, they predict equipment malfunctions that could have safety or security implications.	Major ports are adopting AI for predictive maintenance and risk management. For example, the Port of Rotterdam uses AI to analyze sensor data and historical incidents to predict equipment failures or anomalous events, enabling preemptive fixes that reduce the risk of disruptions. More broadly, AI-based predictive threat intelligence can analyze global cyber incidents to anticipate which maritime systems are at highest risk.
Autonomous Vessel Security	On autonomous or smart ships, AI can serve as an onboard cybersecurity agent, automatically monitoring and controlling certain defenses. AI might isolate compromised subsystems or re-route communications during an attack without waiting for shore support.	Although fully autonomous cyber defense on ships is still emerging, some frameworks are in development. For instance, reinforcement learning agents have been proposed to manage network traffic or reconfigure systems in response to cyber threats in real time[19]. In simulation, such an AI agent could detect a malware outbreak on a vessel's network and autonomously segment or shut down affected systems to contain it.
Digital Twins for Training	High-fidelity virtual models of ships, ports, or critical systems provide a safe environment to test cyber defenses and train AI models. AI can learn to recognize attacks in the twin before deployment, and operators can rehearse responses.	A digital twin of a ship's engine control system can be used to simulate a ransomware attack and test the vessel's AI-based incident response algorithms under realistic conditions. Ongoing research in critical infrastructure protection shows that digital twins combined with AI enable real-time anomaly detection and vulnerability prediction by mirroring the physical system in a controlled virtual space.
Maritime Honeypots & Deception	AI-enhanced honeypots and decoy systems mimic real maritime networks or devices (e.g., a fake ECDIS navigation system or port terminal server) to lure attackers. AI can manage these decoys and analyze attacker behavior, improving threat intelligence.	Cyber defenders have begun deploying maritime honeypot networks – for example, simulated SCADA systems for a port fuel pipeline – to attract attackers. AI helps by dynamically adjusting the decoy's responses to make the attacker believe they are in a real system, while logging their tactics. A case study described a simulated port control system used as a honeypot to profile attackers and then update defensive AI models with the new TTPs observed.
Explainable AI for Alerts	Because maritime operations demand high trust and compliance, explainable AI (XAI) techniques are used to make the AI's decisions transparent to operators. XAI helps analysts and ship crews understand why an alert was triggered, preventing confusion in critical moments.	Modern AI-based intrusion detection systems in maritime are incorporating XAI modules that highlight the features or patterns that led to an anomaly alert (e.g., unusual traffic on a rarely used satellite link). This is crucial for crew acceptance; an opaque "black box" alert might be ignored or misunderstood. Experts emphasize that XAI is vital for compliance and crew trust when AI flags a potential incident. For instance, if an AI system flags a possible GPS spoofing attack, an explainable model could indicate it was due to the vessel's reported position suddenly jumping off a known route, making the alert more credible to the bridge team.

3.1 *Intrusion detection and monitoring*

AI-driven intrusion detection and monitoring is at the forefront of maritime cyber defense. Traditional rule-based intrusion detection systems (IDS) often struggle with the unique network traffic patterns and legacy protocols found in maritime OT systems. Machine learning models, however, can learn a baseline of normal behavior for a ship's network or a port facility and then detect anomalies that might signify a breach. For example, Cydome, a maritime cybersecurity firm, offers a platform that integrates endpoint detection and response (EDR), network detection and response (NDR), continuous vulnerability scanning, and AI-powered anomaly detection tailored to maritime IT/OT environments[20]. By feeding data from ship sensors, navigation systems, and communications into machine learning models, the system can catch subtle signs of compromise – such as an increase in traffic on a rarelyused navigation equipment port, or an unusual sequence of commands in the ballast control system – and alert operators before significant damage occurs. Notably, these AI models can identify outbound malicious traffic (like a malware on a ship trying to beacon out to a command-and-control server) amidst the noisy satellite communication links , something that would be very difficult with manual monitoring[21].

3.2 *Vessel tracking anomaly detection, predictive analysis and autonomous response*

AI significantly enhances maritime cybersecurity through vessel tracking anomaly detection, predictive analytics, autonomous response, and explainable AI. Given the vulnerability of the Automatic Identification System (AIS) to spoofing [22], AI models such as GeoTrackNet have been developed to identify abnormal vessel movements by analyzing contextual traffic patterns and flagging deviations—enabling early detection of illicit behaviors like AIS shutdowns or GPS spoofing [[22]]. Predictive analytics further bolster cybersecurity by forecasting attack likelihood based on global incident trends or detecting impending equipment failures, which if left unaddressed, could become cyber vulnerabilities. The Port of Rotterdam's use of predictive AI in traffic and maintenance illustrates how operational continuity supports security. Moreover, autonomous response systems—powered by reinforcement learning—are emerging to counter threats in real time, such as mitigating DDoS attacks on vessels without human intervention. However, the complexity of maritime operations makes Explainable AI (XAI) essential. Crew and operators must understand the rationale behind AI-generated alerts, especially when safety-critical systems are affected. Tools that offer interpretable outputs and act as decision-support, rather than decision-makers, enhance operator trust and ensure AI is deployed effectively and responsibly in maritime environments.

3.3 *Artificial Intelligence (AI) in Port Traffic Management*

Artificial Intelligence (AI) is revolutionizing port traffic management by enabling smarter, more adaptive, and efficient operations. Through dynamic berth allocation, AI considers real-time factors like vessel size, cargo type, and tidal conditions to optimize scheduling and reduce delays. Predictive ETA modeling, using historical AIS data and weather patterns, improves the planning and allocation of port resources such as cranes and labor. AI also enhances adaptive traffic flow by analyzing data from AIS, CCTV, and IoT sensors to reroute operations and prevent congestion. Autonomous Vessel Traffic Services (VTS) support safe navigation by autonomously managing vessel movements and issuing collision avoidance alerts[23]. Additionally, AI facilitates multimodal transport synchronization, ensuring seamless cargo transitions across sea, rail, and road. Risk management benefits from AI's ability to build dynamic threat profiles and support real-time decision-making, while predictive maintenance minimizes infrastructure failures through proactive diagnostics. Ports also leverage AI for sustainability by optimizing routing to reduce emissions and fuel consumption[23]. Continuous real-time monitoring allows early anomaly detection, improving operational security and resilience. These innovations, as demonstrated by the Port of Rotterdam's implementation of AI-driven VTS, predictive maintenance, and automated container handling, highlight the transformative impact of AI on port efficiency, safety, and environmental performance.

4 CHALLENGES AND VULNERABILITIES OF AI IN THE MARITIME DOMAIN

While AI offers powerful capabilities for maritime cybersecurity, it also introduces new challenges and potential vulnerabilities. This year, a very important article providing a comprehensive review of threats and challenges related to the use of AI in maritime cybersecurity[19] highlights and discusses several key issues that arise from the use of AI in maritime cyber defence and the inherent weaknesses that adversaries may exploit in AI systems several key issues:

1. Adversarial Attacks on AI Models - AI systems can be manipulated by adversarial inputs that subtly alter data to deceive detection mechanisms. This allows attackers to carry out malicious actions while staying under the radar of AI-driven security tools.
2. Data Poisoning and Bias - If training data is scarce, biased, or intentionally corrupted, AI models can develop blind spots or misclassify threats. Attackers may exploit this by injecting misleading data to normalize malicious behavior.
3. Model Exploitation and Reverse Engineering - Cybercriminals can analyse or steal AI models to understand how they function and then design attacks that evade detection. This also exposes proprietary algorithms and potentially sensitive operational data.
4. Lack of Explainability and Human Trust - Many AI models operate as "black boxes," offering little insight into how decisions are made. This undermines human trust and makes it difficult to act confidently on AI-generated alerts.
5. Regulatory and Ethical Issues - AI use in maritime cybersecurity raises legal questions about responsibility when systems fail or act autonomously. Additionally, surveillance and data

collection by AI tools can conflict with privacy and ethical standards.

5 CONCLUSION

AI is a transformative technology with the power to both greatly enhance and undermine maritime cybersecurity. On one side, we see that AI can empower attackers – enabling more sophisticated phishing, automating the discovery of vulnerabilities, and evading traditional defenses at a scale and speed previously unattainable. On the other side, AI offers defenders advanced tools to monitor complex maritime systems, detect anomalies in real time, and coordinate responses to incidents across global fleets. This dual-use characteristic of AI means that the maritime industry must adopt a balanced, vigilant approach to integration of AI into its cyber infrastructure 52 . Stakeholders should neither blindly trust AI nor dismiss it; instead, they must actively shape its use through rigorous testing, clear policies, and continuous learning.

As maritime organizations accelerate adoption of AI-driven automation in navigation, cargo management, and logistics, they must simultaneously recognize and mitigate the new cyber risks that accompany these innovations. The lessons from recent incidents – from AI-crafted phishing campaigns to the GPS spoofing that led to a grounding – underscore that cybersecurity can no longer be an afterthought. AI itself can become a target, and its failure modes need to be well understood. Therefore, building resilience is paramount. This includes investing in resilient AI (robust to attacks and errors), training personnel to work effectively with AI systems, and collaborating internationally to create standards and share threat intelligence.

In conclusion, AI will undoubtedly play an increasingly central role in the security of the "digital seas." If properly harnessed, AI can act as a force multiplier for maritime cyber defense, offsetting the asymmetry that often favors attackers. It can help protect sprawling shipboard networks, complex port operations, and critical maritime supply chains by providing faster-than-human threat detection and decision support. However, realizing this potential requires carefully navigating the challenges – technical, organizational, and ethical – outlined in this paper. The maritime sector must be proactive in addressing AI's vulnerabilities (such as adversarial threats and data issues) and in implementing the strategies for resilience (digital twins, deception, adversarial training, etc.). By doing so, and by fostering a culture of security innovation and cooperation, the industry can confidently embrace AI as an ally rather than fear it as a menace. In the age of digital seas, the winners will be those who can sail with AI's winds while keeping a steady hand on the helm of cybersecurity.

FUNDING

This study was funded by the Gdynia Maritime University, under the research project: WN/2025/PZ/07.

REFERENCES

[1] R. Fang, R. Bindu, A. Gupta, and D. Kang, "LLM Agents can Autonomously Exploit One-day Vulnerabilities," Apr. 17, 2024, arXiv: arXiv:2404.08144. doi: 10.48550/arXiv.2404.08144.

[2] L. Wu, X. Zhong, J. Liu, and X. Wang, "PTGroup: An Automated Penetration Testing Framework Using LLMs and Multiple Prompt Chains," in Advanced Intelligent Computing Technology and Applications - 20th International Conference, ICIC 2024, Tianjin, China, August 5-8, 2024, Proceedings, Part IX, D.-S. Huang, W. Chen, and J. Guo, Eds., in Lecture Notes in Computer Science, vol. 14870. Springer, 2024, pp. 220–232. doi: 10.1007/978-981-97-5606-3_19.

[3] I. Durlik, T. Miller, E. Kostecka, and T. Tuński, "Artificial Intelligence in Maritime Transportation: A Comprehensive Review of Safety and Risk Management Applications," Appl. Sci., vol. 14, no. 18, Art. no. 18, Jan. 2024, doi: 10.3390/app14188420.

[4] A. Ergasheva, F. Akhmedov, A. Abdusalomov, and W. Kim, "Advancing Maritime Safety: Early Detection of Ship Fires through Computer Vision, Deep Learning Approaches, and Histogram Equalization Techniques," Fire, vol. 7, p. 84, Mar. 2024, doi: 10.3390/fire7030084.

[5] E. Kovacs, "OpenAI Says Iranian Hackers Used ChatGPT to Plan ICS Attacks," SecurityWeek. Accessed: Jun. 10, 2025. [Online]. Available: https://www.securityweek.com/openai-says-iranian-hackers-used-chatgpt-to-plan-ics-attacks/

[6] "OpenAI details how threat actors are abusing ChatGPT | TechTarget," Search Security. Accessed: Jun. 10, 2025. [Online]. Available: https://www.techtarget.com/searchsecurity/news/366613512/OpenAI-details-how-threat-actors-are-abusing-ChatGPT

[7] J. Hazell, "Spear Phishing With Large Language Models," arXiv.org. Accessed: Jun. 10, 2025. [Online]. Available: https://arxiv.org/abs/2305.06972v3

[8] F. Heiding, B. Schneier, A. Vishwanath, and J. Bernstein, "Devising and Detecting Phishing: Large Language Models vs. Smaller Human Models".

[9] "AI-Powered Phishing Outperforms Elite Cybercriminals in 2025 - Hoxhunt." Accessed: Jun. 10, 2025. [Online]. Available: https://hoxhunt.com/blog/ai-powered-phishing-vs-humans

[10] "Maritime needs to prepare for AI-assisted cyberattacks." Accessed: Jun. 29, 2025. [Online]. Available: https://mykn.kuehne-nagel.com/news/article/maritime-needs-to-prepare-for-aiassisted-cybe-02-Jun-2025

[11] J. Sims, "BlackMamba: Using AI to Generate Polymorphic Malware." Accessed: Jun. 10, 2025. [Online]. Available: https://www.hyas.com/blog/blackmamba-using-ai-to-generate-polymorphic-malware

[12] M. Burgess, "Criminals Have Created Their Own ChatGPT Clones," Wired. Accessed: Jun. 10, 2025. [Online]. Available: https://www.wired.com/story/chatgpt-scams-fraudgpt-wormgpt-crime/

[13] "2023 Cyber Trends and Insights in the Marine Environment Report," United States Coast Guard News. Accessed: Jun. 29, 2025. [Online]. Available: https://www.news.uscg.mil/maritime-commons/Article/3750095/2023-cyber-trends-and-insights-in-the-marine-environment-report/https%3A%2F%2Fwww.news.uscg.mil%2Fmaritime-commons%2FArticle%2F3750095%2F2023-cyber-trends-and-insights-in-the-marine-environment-report%2F

[14] "Coast Guard Cyber Command releases 2024 CTIME report," United States Coast Guard News. Accessed: Jul. 05, 2025. [Online]. Available: https://www.news.uscg.mil/maritime-commons/Article/4189739/coast-guard-cyber-command-

releases-2024-ctime-report/https%3A%2F%2Fwww.news.uscg.mil%2Fmaritime-commons%2FArticle%2F4189739%2Fcoast-guard-cyber-command-releases-2024-ctime-report%2F
[15] "AI-Powered 'BlackMamba' Keylogging Attack Evades Modern EDR Security." Accessed: Jun. 10, 2025. [Online]. Available: https://www.darkreading.com/endpoint-security/ai-blackmamba-keylogging-edr-security
[16] Z. Labs, "Links and materials for Living off Microsoft Copilot," Zenity Labs. Accessed: Jun. 29, 2025. [Online]. Available: https://labs.zenity.io/p/links-materials-living-off-microsoft-copilot
[17] "How to Weaponize Microsoft Copilot for Cyberattackers." Accessed: Jun. 29, 2025. [Online]. Available: https://www.darkreading.com/application-security/how-to-weaponize-microsoft-copilot-for-cyberattackers
[18] "Iran-Linked CyberAv3ngers Group Uses ChatGPT To Plan Industrial Attacks." Accessed: Jun. 10, 2025. [Online]. Available: https://thecyberexpress.com/cyberav3ngers-use-chatgpt-to-plan-ics-attacks/
[19] T. Miller, I. Durlik, E. Kostecka, S. Sokołowska, P. Kozlovska, and R. Zwolak, "Artificial Intelligence in Maritime Cybersecurity: A Systematic Review of AI-Driven Threat Detection and Risk Mitigation Strategies," Electronics, vol. 14, no. 9, Art. no. 9, Jan. 2025, doi: 10.3390/electronics14091844.
[20] R. Dharma, "Cydome launches maritime cybersecurity platform," Ship Technology. Accessed: Jun. 30, 2025. [Online]. Available: https://www.ship-technology.com/news/cydome-maritime-cybersecurity-platform/
[21] "Complete cybersecurity protection solution, built for maritime," CYDOME. Accessed: Jun. 30, 2025. [Online]. Available: https://cydome.io/cydome-protect/
[22] T. Neumann, "Cybersecurity in Maritime Industry," TransNav Int. J. Mar. Navig. Saf. Od Sea Transp., vol. 18, no. 4, pp. 765–774, Dec. 2024, doi: 10.12716/1001.18.04.02.
[23] "AI in the Maritime Industry: Revolutionizing Port Traffic Management | LinkedIn." Accessed: Jun. 10, 2025. [Online]. Available: https://www.linkedin.com/pulse/ai-maritime-industry-revolutionizing-port-traffic-marc-asselin-ajsoe/