SOCIAL SCIENCE REVIEW ARCHIVES
ISSN Online: 3006-4708
ISSN Print: 3006-4694
https://policyjournalofms.com

The Role of Artificial Intelligence in Tackling Cybersecurity Threats to Critical Infrastructure

Muhammad Zain[1], Muhammad Mahad[2], Waqas Nosher[3]

[1,2] BS-IR, Department of Political Science and International Relations (DPSIR), University of Management and Technology, Lahore. Email: mianzain270903@gmail.com, mmr31j@gmail.com
[3] BS-Computer Science, COMSATS University Islamabad, Lahore Campus.
Email: waqas.nosher2@gmail.com

DOI: https://doi.org/10.70670/sra.v3i4.1234

Abstract

Aimed at the ongoing wave of technological advancement, Critical infrastructure (CI) has become the cornerstone of modern security policy, societal functionality, national security, and economic stability of the state. However, the rapid integration of interconnected cyber systems has shifted from a reactive defense mechanism to a proactive one by exposing them to vulnerable threats such as ransomware, Advanced Persistent Threats (APTs), Distributed Denial of Services (DDoS) attacks, and cyber breaches. Therefore, traditional rule-based defense mechanisms have proven insufficient in countering these devastating threats due to over-reliance on autonomous systems that may fail under adversarial manipulation of systems. This research illustrates the transformative role of AI through a multi-stakeholder governance model, where government, the private sector, and cybersecurity agencies collaborate for AI's better effectiveness. Employing qualitative research methodology, by analyzing the secondary periodic resources to evaluate real-time threat detection, anomaly recognition, and automated incident responses for better resilient and intelligence defense mechanisms, is essential for preserving national security.

Keywords: Artificial Intelligence, Cybersecurity, Critical Infrastructure, Machine Learning, Predictive Analytics, Ransomware, Automated Response, National Security, Active, Proactive

Introduction:

Critical infrastructure forms the backbone of modern society by ensuring public safety, economic stability, national security, and progress. It refers to the system, assets, and network essential for the functioning of a society, economy, and government. These CI include energy systems providing electricity for homes, industries, emergency services, etc., water supply systems delivering clean water for industrial use, drinking, sanitation, etc., transportation networks for supplying goods, services & people, etc. Moreover, the financial system, including banking or the stock market, encourages global trade, investments, and daily transactions, telecommunication network supports businesses relying on data exchange, customer interactions. Therefore, CI plays a crucial role in ensuring national security by developing an energy-grid power defense system, transportation and communication network for military mobilization & disaster response. Thus, protecting these systems is a shared responsibility of the government, the private sector & communities with collaboration.

Moreover, these CIS are completely reliant on digital technologies, but they can also become severe victim of cyberattacks. The evolving nature of cyber threats to such digitized and interconnected infrastructure creates vulnerability, which can easily threaten its resilience & security. Those common attacks on CI through advanced tactics are advanced persistence attacks (APTs), Ransomware attacks, supply chain & insider attacks, etc. As in developing states, CI depend upon legacy technology like the integration of internet of things devices (IoT), which lack security protocols & often serve as an entry point for attackers. Sometimes, nation-states target CI of their rivalry under the constraint of geopolitical tension to create political instability or gain strategic advantages by causing widespread disruption, e.g., Stuxnet. Failure to mitigate these issues not only disturbs the state infrastructure but could also cause severe deaths, public safety risks & financial losses for both private & government institutions. The emergence of AI as a game-changer has become a transformative force in modern cybersecurity, offering innovative solutions for defending against cyber threats. It is considered the simulation of human intelligence in computer systems through learning, reasoning, problem-solving, decision-making, and language processing. Moreover, AI is vulnerable to threats such as data breaches and phishing schemes, and it can detect and respond to these threats through its advanced capabilities. In addition, it will be reactive through predictive analytics, automation, and technological advancement through faster incident response, predictive maintenance, resilience building, etc., ensuring a safer digital environment.

Problem Statement

In this contemporary era of rapid digital transformation, CI has become increasingly interconnected and dependent on cyber-physical systems. As this digital integration empowers efficiency, connectivity, and automation, it can also simultaneously provide CI to a growing spectrum of cybersecurity threats. Consequently, this will negatively affect national security, integrity, sovereignty, economic stability, and public safety. Therefore, traditional cybersecurity measures depend on manual monitoring, and rule-based systems are inadequate against advanced, AI-driven, and zero-day threats. However, its integration is still limited and has challenges such as a lack of empirical validation, limited transparency, ethical concerns, algorithm bias, and policy gaps limit the effective and responsible use of AI in several sectors. Thus, the core problem of this research is to examine how AI can be integrated into cyber cybersecurity framework to strengthen resilience against devastating digital threats.

Research Objective

The core objective of this research is to provide empirical insight by evaluating the capabilities of AI in detecting, preventing, and mitigating cybersecurity attacks in the core CI, like Healthcare networks, energy grids, transportation systems, etc. Moreover, this will also illustrate the operational and organizational problems faced by cybersecurity professionals, policymakers, and infrastructure managers in adopting AI solutions. Therefore, it will also examine ethical, legal, and governance implications of deploying AI Software in any sort of critical infrastructure. It also covers issues related to accountability, transparency, data privacy, cost, data management, system integration, and algorithmic bias. As there is a crucial requirement for accessing the role of AI–driven automation and predictive analytics in emphasizing proactive threat management, incident response, and improving resilience within CI fields, with practical recommendations for secure, ethical, and effective implementation of AI technologies.

Significance of Study

This study holds significant importance in better understanding how Artificial intelligence works as a backbone for critical infrastructure by transforming cybersecurity strategies in this modern civilization. In this regard, there is a high frequency of threats such as ransomware, phishing attacks, zero-day exploits, and state-sponsored attacks that pose extremely harmful impacts for the economy, sustainability, and prosperity of states. This study is significant as it explores the transformative role in combating these devastating cyberattacks through machine learning algorithms, neural networks, and natural language analysis. AI has stronger capabilities to analyze large-scale network data, identify anomalies, and respond to potential data breaches far faster than a human analyst can. Furthermore, this study also holds practical significance for better policy development, strategic planning, and governance. In addition, this study also highlights some case studies relevant to cyber-attacks by illustrating. On a Broader level, this article emphasizes the need for international cooperation and standardized cybersecurity protocols between states. For this, an AI-enabled system must be aligned with global norms and accountability for ensuring better surveillance and information sharing among nations and industries to prevent cyberattacks.

Literature Review

This literature review demonstrates that AI holds significant promise for enhancing CI cybersecurity, as it is a powerful technology that reduces cybersecurity teams' burden through automated responses, accelerates threat detection and response, and improves the accuracy of their actions to strengthen security postures against devastating attacks. Therefore, cybersecurity involves implementing policies, procedures, and technical mechanisms to detect any modification or exploitation of information. In this era of enduring evolving technology and innovation, there is a crucial requirement to protect the CI of states from cyber threats for better sustainable development and economic stability (Ramanpreet Kaur, April 2023).

Overview of Cybersecurity Threats to Critical Infrastructure

As critical infrastructure is fundamental to the functioning of modern society, due to its increasing reliance on digital technology, states should prioritize its security, safety by protecting it from several cyber threats. There are multiple kinds of cyberattacks targeting critical infrastructure, like;

Advanced Persistent Attack (APTs)

APT's are dangerous because they cannot be easily detected. Attackers often employ devastating techniques to access and maintain access to security codes and computer systems for months or years. This leads towards the grabbing of intellectual property, financial data, customers' financial details, and other important credentials. APT is often state-sponsored, aimed at infiltration into adversaries' CI to steal sensitive data, conduct espionage to disrupt infrastructure, e.g., Stuxnet attack on Iranian nuclear system, Operation Aurora, etc. Stuxnet (2010) is the most dangerous APT attack, a massively integrated computer worm to targeted Iran's nuclear program. It distracted the Industrial control system in Uranium enrichment centrifuges. This weapon was funded by the USA and Israel, which invested billions of dollars to paralyze Iran's nuclear program. Aurora (2009), also known as Operation Aurora, attacked major technology companies like Google and several other organizations. Those Chinese hackers gained unauthorized access to sensitive data and intellectual property of the organization. APT28 (Fancy Bear) is a Russian APT group known for its multiple involvement in conducting espionage activities by attacking government agencies,

political documents, and media outlets worldwide. The Democratic National Committee (DNC) became a victim of this attack during the 2016 US election.

Ransomware Attacks

In this attack, malicious software decodes the critical data by making it inaccessible until a ransom/money is paid, particularly in cryptocurrency, e.g., an attack on the USA Colonial Pipeline in 2021 by Darkside Ransomware group & that company paid 4.4 $ million for regaining access. For CI, Ransomware can paralyze operations, forcing either downtime, an expensive payout, or both. This attack includes encryption of files or locking of system functions, also affecting the automation and speed of response. The Colonial Pipeline, an American oil pipeline system that presents in Taxes which used to carry gasoline and jet fuel mainly to the southeastern United States on May 7, 2021. Therefore, about 45% of all fuel consumed on the East Coast arrives via the pipeline system. It suffered with Ransomware cyberattack, shutting down a major USA fuel pipeline. The inspection by the FBI, the company paid the amount of 75 Bitcoin or $4.4 million USD within several hours as requested by the hacker group. Then, after paying on 9 May, Colonial stated that they started to restore and repair the pipeline operation fully within 1 week (Colonial Pipeline ransomware attack).

Insider attack

occurs when an individual with authorized access to the system & data, intentionally or unintentionally, exploits their access to the CI system. Therefore, these threats occur due to workers in a company who may use their access to facilitate hackers to provide harm to the organization, intentionally or unintentionally. So, the USA government is using modern, sophisticated intelligence capabilities against such a broader target to the critical infrastructure by initiating forces like the National Insider Threat Task Force (NITTF) and National Counterintelligence and Security Center (NCSC). The core aims of these forces are to identify anomalous behavior, protect the privacy and civil liberties of employees, and preserve the nation against any foreign cyber intelligence threat. (INS). The Department of Energy (DOE) received only three reports of cyberattack incidents at utilities, none of which affected the customers in 2016. All this happened due to the USA's advanced mitigation and preventive tactics under Cybersecurity Risk Information Sharing Programs (CRISP). The Electricity Information Sharing and Analysis Center (E-ISAC) mostly focuses on events and weather patterns, while pre-attack measures, by maintaining and exercising manual operations of the grid, planning and exercising recovery operations, can shorten the duration of any cyberattack. Such measures can reduce the economic and societal damage by tracing and locating the attack (Knake, April 2017). 2008 cyberattack on the U.S. power grid, the 2014 attack on a U.S nuclear facility, etc. The USA power grid is a target for major cyberattacks; such kind of attacks require proper inspection and proper planning, crucial resources, and a team with a broad range of expertise by terrorists and criminal organizations. In 2015, a Russian attacker hacked core parts of Ukraine's power grid. However, the USA power system has 3300 utilities to deliver power through 200000 miles of high voltage transmission and 5.5 million miles of distribution line that provide power to millions of homes and businesses. Threat by terrorists on such areas can damage the economic growth, as payment for ransomware malicious software for encryption of data, and will not provide a special code to unlock it until a ransom of at least $300 million is paid to terrorists (Knake, April 2017).

Distributed Denial of Services (DDoS)

It occurs when networks with excessive traffic cause them to crash & disrupt essential services, e.g., the 2016 attack on the U.S internet infrastructure (Dyn) by using Mirai botnets, the 2015 attack on Ukraine's power grid by Russia, etc. Unlike other cyberattacks, DDoS attacks do not exploit vulnerabilities in network resources of computers, but they use standard network connection protocols (HTTP) and Transmission Control Protocol (TCP) with more traffic than they bear or handle. For such attacks there require a botnet is required, a network of interconnected devices that is infected with malware that enables hackers to control the device remotely by spreading the malware throughout the whole system. Secondly, Hackers command the devices in the botnet to send the connection address or IP. In such a way, hackers rely on brute force, sending up multiple requests to eat up all the target's bandwidth. Some of the DDoS attack targets include online retailers, financial institutions, cloud service providers, governmental agencies, gaming companies, and military sites (Kosinski). 2016 attack on U.S internet infrastructure (Dyn) by using Mirai botnets, 2015 attack on Ukraine power grid by Russia, etc. Moreover, on December 23, 2015, Ukraine experienced unscheduled power outages affecting approximately 22500 customers. Although reports of malware found in Ukrainian companies in several critical infrastructure sectors by Russian Hackers. At that time, Ukrainian companies also noticed that Black Energy and Killdisk malware erase selected files on the target system, corrupt the master boot record by rendering systems inoperable (Cyber-Attack Against Ukrainian Critical Infrastructure, July 20, 2021).

Supply Chain Attacks

These attacks occur when cybercriminals target vulnerabilities within the supply chain of an organization in either software, hardware, or services. These attacks aim to exploit mutual relationships between companies, vendors, or service providers. Thus, such cyberattacks cause service disruption, which can penalize entire cities and regions, affecting the normal standard of living & economic activities (Lipps, June 2025). As economic consequences of such attacks adversely affect the businesses regarding financial losses, impacting health care, public safety risks, national security threats, etc. These supply chain attacks occur in any industry, such as the financial sector, the oil industry, to the government sector. Generally, supply chain attacks on information systems begin with an Advanced Persistent Threat (APT) (Supply chain attack, December 2020). In this regard, hackers do not directly attack areas like the government or Agencies; instead target the entity's software. As 3rd software is generally less protective, it becomes a victim of cyberattacks directly. According to an investigation produced by Verizon Enterprise, 92% of the cyber-attack incidents occurred among small firms. Most were vulnerable to supply chain attacks due to interconnected components. In October 2018, European law enforcement disclosed a highly sophisticated credit card fraud ring that stole customers' account details. In this regard, they gained access to account information and made repeated bank withdrawals and internet purchases, causing an estimated $100 million loss (Supply chain attack, December 2020).

The Role & Application of AI in Securing Critical Infrastructure

Undoubtedly, one of the most prominent technologies for strengthening cybersecurity in CI is Artificial intelligence. An AI system can exponentially understand, analyze large data sets, recognize patterns, and make informed decisions with minimal human effort. Critical infrastructure is increasingly becoming a victim of cybercriminals due to its importance in national

security, economic prosperity & public safety (Bansemer, october 2024). The key roles of AI in securing critical infrastructure include the following key aspects:

1. Threat Detection & Prediction

AI's primary applications are to provide signaling through threat detection & prediction before they cause significant harm. As in traditional security systems depend on predefined signs to detect known threats, in contrast, AI systems use machine learning (ML) algorithms to detect, analyze & predict the vast database from multiple sources, which further identifies patterns indicative of potential threats. However, AI-powered systems can detect abnormal behavior like unusual network traffic patterns, unusual login attempts, or irregular system operations, which are considered indicators of threat. AI detects those attacks that may be undetected by traditional signature-based systems. AI can detect & analyze traffic, irregularities in data flow in the power grid by identifying any sort of DDoS or ransomware attack. In 2020, Google Chronicle Security introduced AI-based threat detection tools that include Machine Learning to detect cyberattacks in industries, including various CI (Ahmad, April 2025).

2. Automated Response and Mitigation

Another key advantage of AI is the ability to initiate an automated response after a threat is detected. In such regards, traditional processes often involve interventions that may be slow, non-consistent. Whereas, AI can automatically isolate infected devices, block malicious traffic, and protect from data breaches to vulnerable system without the need for human intervention. Such automation can reduce the response time in developing a counter response more quickly than the traditional way (Jones, April 2025). In the energy sectors, AI can be used to protect power grids in the event of DDoS attacks. In such circumstances, an AI-driven system can automatically reroute the power distribution by minimizing the risk of widespread outage. In 2017, ransomware attacks named Wannacry affected several companies and organizations, such as the UK's National Health Service (NHS). If an AI-powered system had been there, then there would be fewer chances of such damage to various infrastructures.

3. Anomaly Detection and Behavior Analytics

AI can also use these methods to identify unusual patterns, behaviors, and activities within critical infrastructure systems by using historical data to analyze the normal changes in behavior. Such capability is very beneficial, particularly in the healthcare system, as AI can monitor access to patient or hospital records. If any employee tries to breach, data, it will automatically alert the security team. Similarly, in, financial sector, AI can detect an unauthorized transaction by monitoring user behavior and signaling the fraud prevention team promptly (Jones, April 2025). DarkTrace platform/company uses AI for such methods of cyber threat detection, which is very helpful in creating unique, innovative solutions for detecting & stopping any sort of insider attack by preventing data theft and operational degradation through flagging abnormal user behavior within CI. (Challenges of threat detection in cybersecurity)

4. Vulnerability Management & Patch Automation

CI often builds up with legacy software & hardware, which are very difficult to protect, so under vulnerability management, AI performs the process of identifying, evaluating & mitigating security weaknesses within the system and network. While patch automation involves the continuous updating of software for which often includes fixing security flaws, bringing

upgradation & improvement by minimizing human efforts, reducing errors & delays in securing the system (Marry, March 2025). Most of the healthcare centers use IoT-enabled medical equipment like infusion pumps and MRIs that are more vulnerable to cyber-attack. Therefore, AI automatically detects vulnerability & applies security patches without manual intervention in these devices, reducing the risk of devastating impacts on Health CI.

6. AI in securing Industrial Control Systems

Industrial control systems (ICS) are used to control & monitor in some CI centers, such as energy, water management & telecommunication. These sectors are enduring targets of cyberattacks, because most of those ICS are designed without considering modern cybersecurity challenges, especially in developing & underdeveloped states. Therefore, AI can boost the security of ICS by providing continuous monitoring, anomaly detection & mitigation response. Moreover, AI can also implement preventive measures such as shutting down affected systems or alerting cybersecurity teams promptly (GuL, 21 August 2025). During the **Stuxnet attack**, the Iranian nuclear system caused much devastating impacts on ICS. If primary AI were deployed, then Iran might have detected an anomaly in their system behavior and mitigated its consequences.

7. AI-Powered Defense Strategies

AI has significantly transformed defense-related cybersecurity in various sectors like military, border security, public safety, Maritime, air, and Space defense. Therefore, these sectors play a pivotal role in establishing well-developed security, sovereignty & integrity of any state in social, economic, and political terms. From a military perspective, AI focuses on decision-making, surveillance, autonomous operations, etc. For example, Ai AI-driven drone for reconnaissance & targeted strike, as MQ-9 Reaper drones use AI for real-time threat detection & autonomous navigation. Moreover, AI tools are also used for operational readiness as the USA Army uses IBM's Watson AI to explore potential issues in their vehicles & machinery. Moreover, AI-powered technologies are also used for strengthening the border security in various ways, e.g. AI AI-based facial recognition systems are deployed at checkpoints for better identification of individuals on the whitelist. As, USA customs & broader protection use AI to identify passengers. (Shiva, December 2022). Furthermore, AI also proves helpful in protecting civilians during disasters or emergency response for better public safety. For example, an AI system like Predpol analyzes the historical crime data to predict the locations & times of potential incidents. Tokyo developed AI surveillance to monitor the crowd during large events to detect unusual activities. However, it also has situational awareness & response capabilities in maritime, air, and space. For example, AI systems in satellites & drone imagery help to identify illegal activities like smuggling, AI in unmanned Aerial systems like Boeing's MQ-25 aircraft supporting reconnaissance, in space debris tracking- AI systems like LeoLab detect & predict the trajectory of space orbits for preventing collision with other satellites (Chavula, Feb 2025).

Case-Studies & Real-World Examples

AI has become a cornerstone of cybersecurity for protecting critical infrastructures against any sort of cybercriminal activity. However, there are some case studies & real-life examples about several cyberattacks & the exponential role of AI, given below;

1. Energy Sector: Ukraine power grid attack (2015)

In 2015, a devastating cyberattacks happened on Ukraine's power grid that affected almost 230,000 people. This attack was deployed with Black Energy malware that disrupted the industrial

control system (ICS). After that attack, Ukraine applies an AI system like Siemens Mindsphere for better analysis of unusual patterns, traffic, so that grids could be able to respond abruptly with an anomaly detection system (Pollard, 4-23-2024). Therefore, in response to such threats, AI has been deployed in modern power grids to enhance cybersecurity through the use of a platform like DarkTrace. While in real real-world example, the European Network for Cyber Security (ENCS) has implemented such AI tools & tactics to prevent the whole of Europe from such attacks.

2. Health Care Sector: Wannacry Ransomware Attack on NHS (2017)

There was a disruption in the UK's National Health Service due to the WannaCry Ransomware attack affected an outdated system. This attack led to the crippling of hospital operations, forcing the cancellation of surgeries & appointments at that time. Then AI-based solution like Cylance Project is adopted by healthcare sectors to predict & block malware by identifying vulnerabilities & patches in the system. In real real-world example, Mayo Clinic, a non-profit organization committed to conducting clinical research, practice, etc., deployed Aidoc (AI orchestration engine) for detecting unauthorized access attempts and safeguarding electronic health records (KCB, 24 April 2018).

3. Transportation Sector: FAA DDoS Attack Attempt 2020

A DDoS attack was conducted on the U.S. Federal Aviation Administration (FAA), targeting its air control traffic system, which aimed to weaken the whole network, causing delays in the delay of flights. Later on, AI platforms like Palantir and Splunk were deployed in transportation for better detection of air traffic & allowing uninterrupted operation of the communication channel. Although in real-time examples, Airbus (a leader in designing, manufacturing, and delivering aerospace products to customers all over the globe) has used AI AI-based system in its aviation infrastructure for detecting malicious intrusion & preventing the air traffic management system from destruction (Case study: Distributed Denial of Service attacks (DDoS), June 2020).

4. Financial Sector: Bangladesh Bank Heist (2016)

In this attack, Hackers used malware in the Bangladesh Central Bank to steal $ 81 million through the manipulation of the SWIFT Financial Messaging network. Later on, after this attack, financial institutions started deploying AI systems such as Feedzai that analyze real-time transaction data in unusual locations or with abnormal patterns. Therefore, in real real-time example, JP Morgan (a leader in investment, banking, and financial transaction processes) adopted AI-based tools like IndexGPT, etc. to detect transactions, fraud patterns, improving detection accuracy. (Bangladesh Bank robbery, 2016)

AI Future in Cybersecurity for CI

Undoubtedly, the AI role in cyberspace will continue to spread with updated innovations & precision for threat detection to CI. AI is evolving to provide threat intelligence by analyzing traffic patterns & implementing countermeasures before a threat happens. Therefore, AI-powered future behavior analysis will make it vulnerable to potential insider threats. It will also work alongside advanced technologies like blockchain for enhancing data integrity. 5 G will respond to more potential threats. Moreover, AI will empower security operations centers, which will be helpful for human operators for strategic decision-making rather than routine monitoring. Thu in the future, AI will provide much developed predictive & accurate solutions for ensuring cybersecurity in CI. (Teslim, oct0ber 2024)

Challenges & Ethical Considerations

Implementing AI in protecting Critical infrastructure also brings several challenges & considerations. As hackers also have access to AI in this modern era, hackers can exploit vulnerabilities by introducing false data into an AI model that causes it to misidentify that data. There is also an issue of complexity & cost in implementing AI integrations within a legacy system that is not ready for a modern cybersecurity solution. Such actions require technical expertise & significant financial investment that multiple organizations lack, making it difficult to deploy, develop an AI solution effectively. While in ethical concern, there may risk of privacy issues as AI systems capture a vast amount of sensitive data, e.g., errors in a healthcare center's databases may create devastating impacts regarding medical treatments. Therefore, various governments & organizations also use AI for their extensive surveillance by threatening human rights, e.g., drone attacks. In addition, AI can be used as a dual technology, e.g., hackers for malicious activities might manipulate an AI tool that works for threat detection. Furthermore, it might create limitations, like AI might start lacking contextual understanding & intrusion of human experts by over-reliance on it. To cope with such challenges requires data governance by ensuring data privacy regulation, collaboration between private, govt & academia can further boost knowledge sharing & proper training for effective & regular AI deployment with continuous updating of AI models.

Research Gap

Although the existing literature provides a detailed discussion of AI capabilities regarding threat detection, predictive threat analytics, etc, through advanced tools such as machine learning, automation. Moreover, existing studies are largely descriptive, focusing on technological overviews rather than data-driven and experimental research within the critical infrastructure network, with several case studies like the Stuxnet attack of 2010, the Bangladesh Bank heist of 2016, etc.

Theoretical Framework

Critical Infrastructure cybersecurity is inherently socio-technical physical systems are defended by technical control but operated by humans within agencies, organizational, regulatory, and economic contexts. Therefore, AI is a technical innovation that changes detection, response, and decision pathways, while its real-world implications are solely dependent on human acceptance, investment choices, directions, and system resilience. Thus, A single theory approach purely economic, technical, and behavioral is insufficient for this outlay. However, A **hybrid Framework** that combines NIST cybersecurity, NIST at Risk management, socio-technical systems, resilience systems, etc., plays a crucial role in organizing security functions that identify, detect, protect, respond, and recover through operational security goals. In this way, these govern AI-specific risks include reliability, transparency, privacy, and robustness across AI. These aim to emphasize human, organizational, and technical interactions by promoting system resilience for absorbing, adopting, and recovering from cyber incidents in Critical infrastructure.

Research Question

Here are some of the core sub-questions about this topic:

1. How effectively can Artificial Intelligence detect, predict, and mitigate cybersecurity threats targeting Critical Infrastructure compared to traditional defense mechanisms?

2. How does AI contribute to proactive versus reactive cybersecurity postures with national critical infrastructure protection policies?

Research Methodology

This study employs a qualitative Research method to understand how Artificial Intelligence contributes to mitigating cybersecurity threats targeting critical infrastructure across various sectors, energy, healthcare, finance, and transportation. This qualitative approach is chosen to gain a better understanding of human experiences, organizational practices, and policy perspectives regarding the ethical and operational integration of AI in the cybersecurity framework. This research follows an interpretivist paradigm that encourages the subjective implementation of real-world phenomena through several secondary sources. This nature of data assists in identifying emerging patterns, challenges, and implications of AI adaptation for securing CI.

Data Collection

Secondary data used in this proposal were obtained from highly reputable periodic resources such as newspapers, journals, peer-reviewed articles, research papers, and government publications. In this regard, this study contains authentic and verifiable information through key resources, including Science Direct, IEEE Xplore, Springer link, MDPI, etc. All of these resources offer massive insights into AI-driven cybersecurity mechanisms, risk management models, and the ethical governance of digital systems. Moreover, Reports from international organizations such as the European Union Agency for Cybersecurity (ENISA,2022), the US Department of Homeland Security (DHS, 2023) also illustrate the understanding of global cybersecurity. This academic literature, technical reports, and government databases provide a triangulated view of several events related to AI in CI. Thus, this also includes analyzing the existing empirical and policy-based data for better compilation of material. Hence, this research paper not only addresses the ethical limitations of direct field research but also supports the study aims of exploring AI‘s transformative potential in securing critical infrastructure against external digital threats.

Discussion and Analysis

The immense paradigm shift in the Defense system from traditional reactive to a digitally proactive mechanism for tackling cyberattacks on CI. Consequently, AI-powered mechanisms outperform the conventional rule-based security systems for addressing any complex and destructive cyberattack.

1. Comparative Analysis of AI and Traditional Systems in Cyberattack Mitigation:

AI software possesses excessive abilities to process massive data, like Network Intrusion Detection System (NIDS) are essential tools for monitoring and analyzing the network traffic, irregular patterns, network behavior, and malicious activities & breaches. In this regard, this software can also identify potential threats like unauthorized access, data exfiltration, etc., through real-time monitoring by protecting organizations before any significant damage occurs. A traditional system comprises signature-based detection, which works by comparing network traffic against a database of predefined attack patterns, but such systems are unable to identify attacks that do not have any known signature. Moreover, such systems require constant updates, which can be time-consuming and require a significant amount of resources. In addition, traditional Anomaly-based detection has a high false positive rate and difficulties in defining normal behavior. Therefore, according to recent statistics, the number of data breaches has dramatically risen in the recent decade, with 60%

companies experiencing breaches in 2022, with $4.35 million, a 12% increase from previous years. Recent analysts predict that 70% of customers would stop doing business with companies in the future, if they do not control through AI AI-based system. For example, a Gartner report illustrates that AI is expected to increase by 25% in the next 2 years, with 60% of organizations approaching AI-based security solutions (AI vs Traditional Security: A Comparative Analysis of Protecting Customer Data in the Digital Age, 2025). Contrary to this, AI-based techniques have a machine learning system that has supervised methods for better classification of known and unknown threats by splitting the data into subsets based on the most informative attributes. Moreover, they also use support vector machines by using high-dimensional space to separate different classes, which enhances their effectiveness in detecting binary-class intrusions, either malicious or non-malicious. Consequently, such tools have high accuracy by using labelled datasets with well-defined features can effectively detect known attack patterns. In addition, unsupervised learning methods don't require labelled data, and only possess k-means clustering algorithms that group data into clusters based on similarity. Thus, Reinforcement learning in AI is used for blocking irregular traffic, raising alerts based on the observation of network behavior, which can directly lead towards self-improving capabilities by allowing the system to continuously adapt to new and existing threats (Olabiyi, March 2025).

2. Proactive and Reactive Cyber Defense: The Strategic Impact of AI

Historically, most of the cybersecurity measures focused on mitigation after a data breach happens; that approach resulted in delayed responses with exponential operational and financial loss. AI works through a proactive hunting technique in modern cybersecurity by deliberately searching for malicious activities within any critical infrastructure. Unlike the traditional reactive method, which is compatible and responds to threats post-incident, proactive approaches do work through hybrid AI models. In this regard, it utilizes several tools such as machine learning for anomaly detection, natural language processing for malware & spear-phishing attacks analysis, along with a neural network for predictive analysis by identifying evolving attack patterns (Gebremeskel, 3 Feb 2025). Moreover, AI has the most advanced cybersecurity network that is widely recognized and implemented globally. Among them most common is the NIST cybersecurity framework, developed by the National Institute of Standards and Technology focuses on 5 core functions (identify, protect, detect, respond, and recover). Moreover, ISO/IEC 2700 has international standards for information security systems that provide a systematic approach to risk assessment, treatment, and continuous improvement. In addition, the General Data Protection Regulation (GDPR) has been enforced by the European Union, establishing rights for individuals regarding personal data. Thus, automated compliance platforms such as LogicGate and IBM provide real-time dashboards for minimizing the risk of non-compliance (Marapu, 2022).

Research Limitations

Despite the depth and relevance of this research, certain limitations persist regarding generalizability and the scope of the findings. Therefore, these constraints primarily arise from the complex, dynamic, and sensitive nature of AI in Cybersecurity. Due to the classified and sensitive nature of the Ci cybersecurity framework, obtaining first-hand operational data or direct access to real-time AI security systems was not feasible. As limitations include ethical and governance considerations, as CI and AI both deal with sensitive data and algorithmic decision making, limited transparency and restricted disclosure of ethical violations create knowledge gaps. Most organizations restrict the sharing of detailed cybersecurity practices to protect from exploitation

by malicious software. Consequently, the research relied on secondary resources such as periodic, published case studies and institutional analyses.

Conclusion

Taking everything into consideration, AI is at the forefront of the fight against cybersecurity challenges to Critical Infrastructure. Its core ability to analyze, predict, detect anomalies & instant response makes it a crucial component of modern security strategy. As protection of any state brings integrity, economic prosperity, public & border safety for civilians better standard of living. However, the adaptation of AI into such areas demands careful consideration of issues like data dependency, integration with legacy systems, etc. Ultimately, institutions must adopt a balanced approach with ethical governance to fully leverage the potential of AI. Thus, in this contemporary digital environment, adopting AI in CI is a crucial requirement for every state to ensure its protection from adverse cybercriminal attacks. Although AI adaptation in a well-organized, however structured manner will enhance the state's security, sovereignty, with an exponential position in the cyber-security landscape.

Bibliography

Ahmad, A. B. (April 2025). Cybersecurity In Industrial Control Systems: A Systematic Literature Review On AI-Based Threat Detection For SCADA And IOT Networks. 16. Retrieved from https://www.researchgate.net/publication/392035437_Cybersecurity_In_Industrial_Control_Systems_A_Systematic_Literature_Review_On_AI-

(2025). *AI vs Traditional Security: A Comparative Analysis of Protecting Customer Data in the Digital Age.* Super AGI. Retrieved from https://superagi.com/ai-vs-traditional-security-a-comparative-analysis-of-protecting-customer-data-in-the-digital-age/

Artificial Intelligence Risk Management. (january 2023). 48. Retrieved from https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf?utm_source=chatgpt.com

Bangladesh Bank robbery. (2016). *From Wikipedia, the free encyclopedia*. Retrieved from https://en.wikipedia.org/wiki/Bangladesh_Bank_robbery

Bansemer, K. C. (october 2024). *Securing Critical Infrastructure in the Age of AI.* CSET. Retrieved from file:///C:/Users/Lenovo/Downloads/CSET-Securing-Critical-Infrastructure-

(June 2020). *Case study: Distributed Denial of Service attacks (DDoS).* Retrieved from https://www.theengineroom.org/wp-content/uploads/2020/08/OrgSec-Case-study-DDoS-attacks-June-2020.pdf

Challenges of threat detection in cybersecurity. (n.d.). *Darktrace Threat Detection*. Retrieved from https://www.darktrace.com/cyber-ai-glossary/darktrace-threat-detection

Chavula, F. K. (Feb 2025). AI-Powered Satellite Imagery Processing for Global Air Traffic Surveillance. *Researchgate*. Retrieved from https://www.researchgate.net/publication/389753062_AI-Powered_Satellite_Imagery_Processing_for_Global_Air_Traffic_Surveillance

Colonial Pipeline ransomware attack. (n.d.). *Wekipidia the free encyclopedia*. Retrieved from https://en.wikipedia.org/wiki/Colonial_Pipeline_ransomware_attack

Cyber-Attack Against Ukrainian Critical Infrastructure. (July 20, 2021). *American Cyber Defense Agency*. Retrieved from https://www.cisa.gov/news-events/ics-alerts/ir-alert-h-16-056-01

Gebremeskel, B. (3 Feb 2025). Proactive vs. Reactive AI in Cybercrime: Fighting Cyber Threats with Modern AI Strategies. Retrieved from https://teckpath.com/proactive-vs-reactive-ai-in-cybercrime-fighting-cyber-threats-with-modern-ai-strategies/

GuL, M. M. (21 August 2025). Artificial intelligence for secure and sustainable industrial control systems - A Survey of challenges and solutions. *Springer Nature Link*, 86. Retrieved from https://link.springer.com/article/10.1007/s10462-025-11320-9

Insider Threatmitigation for U.S. Critical Infrastructure Entities:Guidelines from an Intelligence Perspective. (n.d.). *National Counterintelligence and Security Center*, 20. Retrieved from https://www.dni.gov/files/NCSC/documents/nittf/20240926_Insider-Threat-Mitigation-for-

Jones, R. (April 2025). *AI Automated Incident Response and Threat Mitigation Using AI*. doi:10.4018/979-8-3373-3296-3.ch007

KCB, S. A. (24 April 2018). *Investigation: WannaCry cyber attack and the NHS.* Department of Health. Retrieved from https://www.nao.org.uk/wp-content/uploads/2017/10/Investigation-

Knake, R. K. (April 2017). A Cyberattack on the U.S. Power Grid. *Counsil on Foriegn relations*. Retrieved from https://www.cfr.org/report/cyberattack-us-power-grid

Kosinski, J. h. (n.d.). What is a distributed denial-of-service (DDoS) attack? *IBM*. Retrieved from https://www.ibm.com/think/topics/ddos

Lipps, S. S. (June 2025). Critical Infrastructure Security and the Role of AI: An Overview. *Researchgate*, 9. Retrieved from https://www.researchgate.net/publication/393048665_Critical_Infrastructure_Security_and_the_Role_of_AI_An_Overview

Marapu, N. R. (2022). Future-Proofing National Cybersecurity: The Role of AI in Proactive Threat Hunting and Framework Optimization. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 3*(4), 27-37. doi:https://doi.org/10.63282/3050-9262.IJAIDSML-V3I4P104

Marry, B. J. (March 2025). AI for Automated Vulnerability Assessment and Patch Management in Cloud Systems. *Researchgate*. Retrieved from https://www.researchgate.net/publication/389717403_AI_for_Automated_Vulnerability_Assessment_and_Patch_Management_in_Cloud_Systems

Olabiyi, E. C. (March 2025). Comparative Study of Traditional vs. AI-Based Techniques in Network Intrusion Detection Systems. Retrieved from https://www.researchgate.net/publication/389717078_Comparative_Study_of_Traditional_vs_AI-Based_Techniques_in_Network_Intrusion_Detection_Systems

Pollard, M. (4-23-2024). A Case Study of Russian Cyber-Attacks on the Ukrainian Power. *16*, 25. Retrieved from https://digitalcommons.pepperdine.edu/cgi/viewcontent.cgi?article=1216&context=ppr

Ramanpreet Kaur, T. K. (April 2023). Artificial Intelligence for Cybersecurity: Literature Review and Future Research Directions. *Research Gate*, 30. doi:10.1016/j.inffus.2023.101804

Shiva, R. (December 2022). The Role of AI in Securing Critical Infrastructure: A Data-Driven Approach to Cyber Defense. *Researchgate*. Retrieved from https://www.researchgate.net/publication/388525378_The_Role_of_AI_in_Securing_Critical_Infrastructure_A_Data-Driven_Approach_to_Cyber_Defense

Supply chain attack. (December 2020). *Wikipedia The free Encyclopedia*. Retrieved from https://en.wikipedia.org/wiki/Supply_chain_attack?utm_source=chatgpt.com

Teslim, B. (oct0ber 2024). *The Future of AI in Cybersecurity: Trends and Predictions.* Retrieved https://www.researchgate.net/publication/384986932_The_Future_of_AI_in_Cybersecurity_Trends_and_Predictions