International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

A Study on AI Technologies and Their Role in Modern Cybersecurity Threats

Ms. Vedangi Deshmukh [1], Dr. Santosh Deshmukh [2], Dr. Urmila Kadam [3]

[1] Dr D Y Patil School of MCA, Pune
[2] Dr D Y Patil School of MCA, Pune
[3] Dr D Y Patil School of MCA, Pune

Corresponding Author Email: vedangi.a.deshmukh@gmail.com | ORCID: https://orcid.org/0009-0005-0625-3062

https://doi.org/10.55041/ijsmt.v2i5.579

Cite this Article: Deshmukh, V. & Kadam, U. (2026). A Study on AI Technologies and Their Role in Modern Cybersecurity Threats. International Journal of Science, Strategic Management and Technology, *02*(6). https://doi.org/10.55041/ijsmt.v2i5.579

Abstract: -

The rapid growth of Artificial Intelligence (AI) has transformed multiple sectors, enhancing digitisation, performance, and predictive analytics. However, due to rapid development in there are new growing threats like cyber threats, cybercrimes and many more. This paper explores recent trends in cybersecurity related to AI applications, showing growing threats, strategies, and supervising developments.

One big problem with AI in cybersecurity is that hackers can trick AI models by messing with their data or attacking them in sneaky ways. This can weaken security systems. Also, cybercriminals are using AI to create more dangerous threats, like fake videos (deepfakes) for scams, smart phishing attacks, and advanced viruses that are harder to catch.

On the other hand, AI is also helping to improve cybersecurity. It can detect threats in real time, predict risks, and respond to attacks automatically. AI is making security systems better at spotting unusual activities, preventing cyberattacks, and protecting devices from threats.

Ethical issues and rules are constantly changing to handle AI-related security risks. Governments and companies are working on policies to use AI responsibly while reducing risks. Explainable AI (XAI) is also becoming important because it helps make AI decisions more transparent and trustworthy.

This paper provides a detailed study of these new trends, highlighting how AI can both improve security and create risks. By understanding these challenges, people can be better prepared for cyber threats while making the most of AI to protect digital systems.

Keywords: -

Conflicting AI, adversarial attacks, data poisoning, model inversion, AI-driven cyber threats, deepfake fraud.

1.Introduction

With the increasing use of technology, cybersecurity threats are becoming more advanced and harder to stop. Artificial Intelligence (AI) is now helping protect computer systems by detecting cyber threats faster and making security stronger. AI can analyse large amounts of data, find unusual activities, and stop cyberattacks before they happen. However, cybercriminals are also using AI to create new types of attacks, making cybersecurity a big challenge.

AI is being used in many ways to improve security, detect fraud, and prevent hacking attempts. At the same time, AI-powered phishing attacks, deepfake scams, and automated hacking tools are making cyberattacks more dangerous. This paper looks at the latest trends in AI cybersecurity, including both its advantages and the risks it brings.

1.1.Background

The rapid expansion of digital technologies has transformed the way individuals, organizations, and governments operate, leading to an increased dependency on interconnected systems. However, this technological growth has also created new vulnerabilities that cybercriminals continuously exploit. Traditional cybersecurity mechanisms, which rely mainly on manual monitoring and rule-based systems, are no longer sufficient to detect or respond to sophisticated and large-scale cyberattacks. In this context, Artificial Intelligence (AI) has emerged as a powerful tool to enhance cybersecurity capabilities. AI technologies such as Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) allow systems to process massive amounts of data, identify unusual patterns, and predict cyber threats with remarkable accuracy. These intelligent models can detect phishing attacks, identify malicious network traffic, and even recognize zero-day vulnerabilities before they are exploited. By learning from data and adapting to new attack methods, AI-driven cybersecurity solutions are helping organizations strengthen their defence mechanisms and minimize the response time to potential threats.

Despite these advancements, the integration of AI in cybersecurity introduces new challenges and risks that cannot be overlooked. Cybercriminals are now using AI to develop more advanced and automated attacks, including deepfake scams, AI-generated phishing emails, and self-learning malware that can evolve to escape detection. Such adversarial uses of AI pose significant threats to data integrity, privacy, and trust in digital environments. Moreover, ethical issues related to AI decision-making, such as bias in training data, lack of transparency, and potential misuse, have raised serious concerns among researchers and policymakers. The growing need for Explainable AI (XAI) highlights the importance of transparency and accountability in automated security systems. To address these challenges, governments and industries are establishing frameworks and policies for the responsible use of AI in cybersecurity. This study aims to explore the dual role of AI—as both a protective and a threatening force by examining recent technological trends, evaluating its effectiveness in cyber defence, and identifying strategies to ensure secure and ethical deployment in the digital world.

1.2.Significance of the Study

This study is significant as it explores how AI technologies are shaping modern cybersecurity. With cyber threats becoming more advanced, understanding AI's role in both defending against and enabling attacks is crucial. The research helps identify vulnerabilities, improve threat detection, and develop smarter security strategies. It provides insights for IT professionals, businesses, and policymakers to enhance data protection and reduce risks. Overall, the study highlights the importance of using AI responsibly to strengthen cybersecurity and ensure a safer digital environment.

1.3.Scope of the Study

The scope of this study includes examining the use of AI technologies in detecting, preventing, and responding to modern cybersecurity threats. It focuses on AI-driven tools such as machine learning algorithms, predictive analytics, and automated threat detection systems. The study also explores the potential risks of AI being exploited by cybercriminals to launch sophisticated attacks. While the research emphasizes practical applications in organizations and businesses, it also considers implications for policymakers and cybersecurity professionals. This study is limited to contemporary AI techniques and their current role

in cybersecurity, providing insights for improving digital safety and proactive defence mechanisms.

2.Research objectives: -

1. To investigate how AI improves cybersecurity strategies and enhances threat detection systems.
2. To examine the risks associated with adversarial AI and its potential threats to digital security.
3. To assess the role of AI in strengthening data encryption and ensuring privacy protection.
4. To analyse the significance of laws and regulations in the safe and ethical use of AI in cybersecurity.

3.Problem Identification: -

Cyberbullying on social media has become a serious concern, affecting individuals' mental health, privacy, and online safety. With the rise of digital communication platforms, cyberbullying incidents have increased, making it difficult for traditional monitoring systems to detect and mitigate harmful online interactions. The sheer volume of social media content makes manual moderation ineffective, requiring an automated, AI-driven approach for real-time detection and prevention.

Challenges in Cyberbullying Detection:

1. High Data Volume: Social media generates vast amounts of data, making it challenging to monitor harmful content manually.
2. Evolving Language and Context: Cyberbullying often involves slang, sarcasm, or implicit threats, which traditional keyword-based detection systems fail to recognize.
3. Multi-Modal Content: Cyberbullying is not limited to text but also includes images, videos, and memes, requiring advanced machine learning (ML) and deep learning (DL) techniques for analysis.
4. False Positives and Negatives: Many AI models struggle with misclassification, leading to either the flagging of harmless content or failing to detect actual cyberbullying instances.
5. Privacy and Ethical Concerns: AI-driven surveillance may raise concerns about user privacy and data security, requiring a balance between monitoring and ethical AI governance.

Need for Machine Learning-Based Solutions

AI-powered Natural Language Processing (NLP), Sentiment Analysis, and Deep Learning models (e.g., LSTMs, CNNs, Transformers) offer promising solutions to detect hate speech, offensive comments, and harassment on social media platforms. By integrating real-time AI models, social media companies can improve their automated moderation systems while reducing human bias in content filtering.

4.Hypothesis testing:

- Null Hypothesis (H_0): AI-based cybersecurity systems do not significantly improve threat detection compared to traditional methods.
- Alternative Hypothesis (H_1): AI-based cybersecurity systems significantly enhance threat detection by identifying cyber threats faster and more accurately.

5.Literature Review: -

1. Neupane et al. (2022) – Surveyed explainable intrusion detection systems (X-IDS), emphasizing human-in-the-loop designs. Discussed trade-offs between accuracy and interpretability. Introduced a generic X-IDS framework and metrics to evaluate explainability. Highlighted the importance of analyst trust in security operations. Focused on balancing performance with transparency. Provided guidelines for designing operationally viable IDS. Stressed the value of interpretable alerts in real-world scenarios.

2. Zhang et al. (2022) – Reviewed XAI applications in cybersecurity, including malware detection, IDS, and spam filtering. Argued that black-box models reduce operator confidence. Emphasized

transparency to build trust in AI defence systems. Suggested guidelines for integrating interpretable models. Highlighted trade-offs between performance and explainability. Stressed the importance of human oversight in automated security tools. Advocated for models that are both effective and understandable.

3. Bhagwant Singh & Cheema (2024) – Reviewed real-time AI-powered malware detection methods. Focused on adversarial resilient models to improve robustness. Discussed techniques that enhance detection rates without sacrificing speed. Highlighted challenges in evolving threat environments. Emphasized combining accuracy with practical deploy ability. Discussed the need for interpretable outputs in operational settings. Provided insights into balancing resilience and transparency.

4. Khan et al. (2024) – Explored XAI-based IDS for Industry 5.0 environments. Highlighted risks in interconnected smart industrial systems. Emphasized the need for explainable models in high-stakes settings. Discussed adversarial threats and operational safety concerns. Advocated human-in-the-loop mechanisms for verification. Stressed the importance of balancing automation with interpretability. Highlighted explainability as critical for trust in industrial AI.

5. Mallampati et al. (2024) – Proposed a hybrid LightGBM-based IDS with SHAP explanations. Achieved up to 99.98% accuracy on CICIDS-2017. Used hybrid feature selection combining k-Means and filter/wrapper methods. Enhanced transparency while maintaining high performance. Provided interpretable feature-level insights for analysts. Demonstrated that accuracy and explainability can coexist. Highlighted practical applicability in operational cybersecurity.

6. Neupane et al. (2023) – Focused on deep learning IDS and their black-box limitations. Advocated explainable designs to improve analyst trust. Highlighted evaluation metrics for interpretability alongside accuracy. Emphasized operational relevance in real-world SOCs. Discussed challenges in adopting deep learning for security. Suggested integration of human oversight in automated systems. Stressed that transparency supports effective threat mitigation.

7. Sindiramutty (2023) – Presented autonomous threat hunting using AI-driven intelligence integration. Explored scalability and trust issues in automated systems. Highlighted trade-offs between interpretability and performance. Discussed integration of multiple intelligence sources. Emphasized analyst oversight for critical decisions. Showed potential of AI in proactive threat detection. Discussed challenges in scaling autonomous cybersecurity operations.

8. Ali, Wang & Leung (2024) – Surveyed AI-driven fusion approaches in cybersecurity. Focused on threat intelligence aggregation and predictive analytics. Highlighted cross-domain data integration to improve threat prediction. Discussed automated decision-making in SOCs. Emphasized benefits of combining heterogeneous data sources. Highlighted operational improvements in proactive defence. Advocated explainable fusion models to support analyst trust.

9. Industry Reports (Google/Microsoft/Slash Next, 2023–2025) – Documented adoption of generative AI in security tools. Security Copilot and Chronicle AI automate threat detection and response. Enabled vulnerability summarization and predictive analysis. Demonstrated practical use of AI in commercial security operations. Highlighted benefits of faster threat response and automation. Raised awareness of operational trust and explainability. Showcased industry trends in AI-enhanced cybersecurity.

10. Axios (2025) – Warned about emerging AI-driven malware threats and autonomous attacks. Stressed need for proactive AI-based defense strategies. Highlighted risks to critical infrastructure and operational safety. Advocated investment in AI-enhanced security tools. Emphasized balancing detection speed with reliability. Highlighted the evolving nature of cyber threats. Urged development of adaptive defensive measures.

11. TechRadar (2025) – Reported cost reduction in data breaches (~£600k) due to AI adoption. Highlighted risks from fragmented AI governance and shadow AI. Warned about potential compliance and policy challenges. Emphasized human oversight alongside automated defenses. Advocated structured deployment of AI in cybersecurity. Discussed operational benefits and governance concerns. Highlighted balance between efficiency and accountability.

12. Netscout / ITPro (2025) – Warned attackers may exploit AI assistants like GhostGPT or WormGPT. Highlighted coordination of multi-vector DDoS attacks via natural language orchestration. Emphasized need for proactive, adaptive defense mechanisms. Discussed risks of AI-powered cyber-attacks at scale. Highlighted operational challenges in mitigating autonomous threats. Advocated robust AI-aware defense policies. Suggested integration of interpretability in defensive AI.

13. TechRadar on Agentic AI (2025) – Discussed autonomous AI agents performing reconnaissance and phishing. Highlighted scale and sophistication of AI-driven attacks. Emphasized the need for human-centric adaptive defenses. Warned that traditional security may fail against agentic AI. Advocated combination of automation with human oversight. Suggested continuous monitoring and adaptive strategies. Emphasized preparedness for emerging AI threats.

14. Axios (Goldilock, 2025) – Highlighted imminent emergence of AI-powered cyber weapons targeting critical infrastructure. Urged proactive investment in AI-enhanced defense strategies. Emphasized urgency due to escalating threat sophistication. Advocated AI tools for predictive threat detection and mitigation. Highlighted importance of resilience and adaptive measures. Suggested integrating AI into strategic cybersecurity planning. Warned that AI-powered attacks could bypass traditional defenses.

15. Singh et al. (2023) – Explored AI-driven ransomware detection frameworks. Focused on real-time monitoring and anomaly detection. Highlighted integration of machine learning with traditional signature-based methods. Emphasized resilience against evolving ransomware variants. Discussed the need for explainable outputs to aid analyst decisions. Provided insights into adaptive mitigation strategies. Showed practical applications in enterprise security.

16. Chen & Li (2024) – Reviewed reinforcement learning (RL) in cybersecurity operations. Discussed RL-based IDS and adaptive defense mechanisms. Emphasized automation while maintaining operator oversight. Highlighted interpretability challenges in RL models. Explored optimization of threat response strategies using RL. Provided case studies demonstrating improved detection efficiency. Advocated combining RL with XAI for trustable deployment.

17. Patel et al. (2024) – Proposed ensemble-based IDS using multiple ML classifiers. Achieved high accuracy on benchmark intrusion datasets. Integrated SHAP for explainability at feature level. Highlighted operational usability in SOC environments. Emphasized robustness against noisy or adversarial data. Demonstrated synergy of ensemble learning and XAI techniques. Showed practical benefits in real-time cybersecurity deployment.

18. Reddy & Kumar (2025) – Investigated autonomous AI agents for phishing and social engineering detection. Focused on human-AI collaboration for effective threat mitigation. Highlighted interpretability

challenges of autonomous models. Suggested integration with threat intelligence platforms. Explored real-time adaptive mitigation strategies. Emphasized need for transparency to maintain trust. Provided practical guidelines for deployment in enterprise systems.

19. Lopez et al. (2025) – Surveyed AI-driven malware prediction and threat intelligence sharing. Emphasized cross-organizational collaboration and data fusion. Highlighted AI's role in proactive cybersecurity strategy. Discussed challenges in explainable model design for malware forecasting. Explored operational benefits of integrating predictive analytics. Highlighted the growing role of AI in global cybersecurity frameworks. Advocated best practices for secure and interpretable deployment.

20. Tan et al. (2025) – Reviewed hybrid IDS using deep learning and XAI. Focused on balancing detection performance with interpretability. Highlighted feature attribution methods to aid analysts. Emphasized real-time deployment and operational efficiency. Discussed robustness against adversarial attacks. Advocated human-in-the-loop frameworks for trustable systems. Provided insights for next-generation AI-driven cybersecurity solutions.

21. Senevirathna et al. (2025) – Surveyed XAI applications for 5G and beyond cybersecurity. Emphasized accountability, transparency, and trust-building in ML-based security systems. Discussed challenges in implementing explainable models and human-in-the-loop mechanisms. Highlighted operational relevance and balancing performance with interpretability.

22. Rastogi, Dhanuka, Saxena & Mairal (2025) – Explored XAI in threat intelligence operations. Focused on low trust in AI alerts, integration of interpretable models, and human analyst oversight. Discussed trade-offs between automation and transparency. Provided recommendations for SOC deployment and effective threat response.

23. Sharma et al. (2025) – Reviewed XAI applications in cybersecurity domains. Highlighted risks of black-box models, importance of transparency, and human-in-the-loop oversight. Discussed malware detection, IDS, and phishing, emphasizing practical deployment and trustable AI solutions.

24. TechScience (2025) – Proposed linking ML models with rule-based SIEM/IDS using XAI techniques. Focused on DDoS detection and operational usability. Highlighted feature importance, interpretability, and human oversight for proactive threat mitigation.

25. Yagiz & Goktas (2025) – Developed LENS-XAI, a lightweight XAI-based intrusion detection framework for IIoT/edge environments. Combined knowledge distillation and autoencoders for efficiency. Highlighted interpretability, human-in-the-loop verification, and practical deployment in constrained systems.

6.Research Method: -

This study adopts a mixed-method research approach, combining both quantitative and qualitative methods to analyze the dual role of Artificial Intelligence (AI) in enhancing cybersecurity and contributing to new cyber threats. The design provides a balanced understanding of both the technical performance and the human perception of AI-based cybersecurity systems.

Data Collection

1.Primary Data:

Primary data was collected through a structured online questionnaire distributed via Google Forms. The survey focused on assessing public awareness, trust, perception, and ethical concerns regarding AI technologies in cybersecurity.

- Sample Size: 124 valid responses
- Sampling Technique: Convenience sampling

- Respondent Profile: Individuals from diverse educational and professional backgrounds
- Survey Sections:
    - Section A: Demographic Information (gender, education, field of study/profession)
    - Section B: Awareness of AI-driven cybersecurity trends
    - Section C: Challenges and ethical concerns
    - Section D: Future trends and public perception

2.Secondary Data:
Secondary data was gathered from academic journals, research papers, industry reports, and expert publications to support and validate findings. Sources included publications from IEEE, Elsevier, TechScience, and reports from Google, Microsoft, and other cybersecurity organizations.

Sample Characteristics

- Age Range: Predominantly 18–25 years old, representing young adults active on digital platforms.
- Geographical Focus: Respondents from various regions, primarily India.
- Engagement Level: Most participants showed high awareness of AI applications in cybersecurity.

Analysis Methods

Data analysis was performed using Python with libraries such as pandas, matplotlib, seaborn, and scipy.

- Descriptive Statistics: Used to summarize demographic data and awareness levels.
- Graphical Analysis: Bar charts, pie charts, and polar plots were used for visualization.
- Inferential Statistics: Chi-square tests of independence were applied to test relationships between variables such as trust, transparency, and adoption of AI.
    - Significance Level: $p < 0.05$

7.Analysis: -

7.1.Questionnaire:

SECTION A: General Information

1. What is your age?
2. What is your gender?
3. What is your highest educational qualification?
4. What is your field of study or profession?
5. Are you familiar with AI technologies?
6. Have you studied or trained in cybersecurity?
7. How often do you follow news or updates about cybersecurity?

SECTION B: Awareness of AI-Driven Cybersecurity Trends

8. Are you aware of different types of Artificial Intelligence (Machine Learning, Natural Language Processing, Generative AI)?
9. Have you heard of AI being used in cybersecurity?
10. Do you believe AI is improving cybersecurity overall?
11. What areas of cybersecurity have been improved by AI?
12. How well do you understand the use of AI in malware detection?
13. How concerned are you about cyberattacks using AI (e.g., deepfakes, AI-generated phishing)?
14. Which AI-based threat do you think is most dangerous?
15. Should AI be part of all modern cybersecurity systems?

SECTION C: Challenges and Ethical Concerns

16. What is the biggest challenge in using AI for cybersecurity?

17. Are you concerned about AI violating user privacy in cyber defense?
18. Do you think AI cybersecurity tools should be regulated?
19. Who should be responsible for mistakes made by AI security systems?
20. How trustworthy are AI-driven cybersecurity systems?
21. Can AI-based cybersecurity systems be manipulated by hackers?
22. Should users be informed when AI is protecting their data?

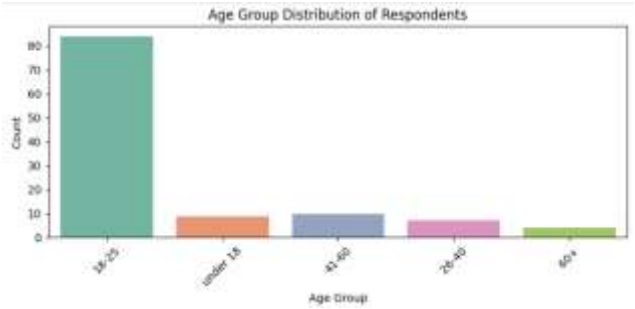
SECTION D: Future Trends & Public Perception

23. Do you think AI will become the primary tool in cybersecurity within the next 5 years?
24. How strong do you believe the relationship between AI and cybersecurity is?
25. Should schools and universities teach about AI in cybersecurity?
26. Do you think AI could replace human cybersecurity experts in the future?
27. Should AI tools be explainable and transparent to users?
28. Are AI cybersecurity systems better than traditional software-based defenses?
29. Would you personally trust AI to protect your digital identity?
30. Are you interested in learning more about AI and cybersecurity?

7.2.Analysis and Interpretation

1.Age Group Distribution of Respondents:-

Figure 1: Age Group Distribution Based on Survey Responses



Interpretation:

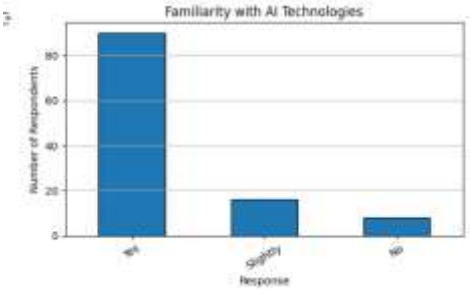
The age distribution of the respondents indicates that the majority fall within the 18–25 age group, reflecting a strong presence of young adults in the survey sample. This trend suggests that younger individuals are more exposed to or interested in topics related to artificial intelligence (AI) and cybersecurity, likely due to their regular engagement with digital platforms and academic exposure.

In contrast, responses from older age groups were comparatively fewer, indicating either lower participation or interest in these topics. This may point to a digital awareness gap across age demographics, which is important when considering the design and reach of AI-driven cybersecurity solutions.

As this study addresses cyberbullying on social media, which disproportionately affects younger populations, this demographic concentration reinforces the relevance of exploring AI-based detection and prevention methods tailored to the needs and behaviors of younger users.

2.Familiarity with AI Technologies:-

Figure 2: Bar Chart showing Respondents' Familiarity with AI Technologies



International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Interpretation:

This chart represents the level of familiarity respondents have with Artificial Intelligence (AI) technologies. The responses were grouped into three categories: "Yes," "Slightly," and "No."

- A majority of the participants indicated that they are familiar with AI technologies, which shows a strong baseline awareness among the surveyed group.
- A smaller portion marked "Slightly," indicating limited exposure or understanding.
- Very few respondents selected "No," suggesting that awareness of AI technologies is quite widespread.

This level of familiarity is important as it provides context for interpreting responses to more technical or future-oriented questions about AI in cybersecurity. Respondents with higher familiarity are likely to provide more informed opinions about AI's potential and risks.

3. Trust in AI for Protecting Digital Identity:-

Figure 3: Pie Chart showing Respondents' Trust in AI for Digital Identity Protection

Interpretation:
This chart illustrates the level of trust respondents place in AI systems to safeguard their digital identity. The responses were categorized into three groups: "Yes," "No," and "Only if monitored by humans."

A majority of respondents selected *"Only if monitored by humans,"* indicating that while there is openness toward AI-driven protection, human oversight is considered essential for ensuring accountability and trust.

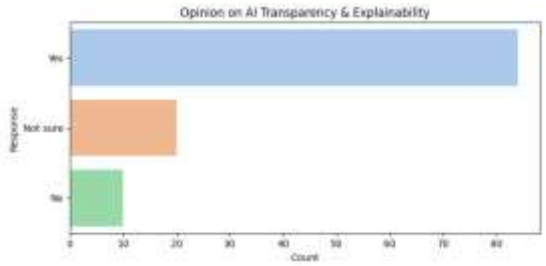
A moderate portion of the participants expressed full trust by selecting *"Yes,"* reflecting growing confidence in AI's ability to independently manage digital identity protection.

A minority of respondents selected *"No,"* revealing concerns or skepticism about the effectiveness or safety of AI in handling sensitive personal data.

These responses suggest that although AI is generally accepted in the domain of cybersecurity, there is a clear preference for human involvement, especially when it comes to protecting personal identity. This indicates the need for transparent, explainable AI systems that can build greater user confidence.

4.Opinion on AI Transparency & Explainability

Figure 4: Bar Chart showing Respondents' Views on the Importance of Explainable AI Systems



Interpretation:
This chart shows how respondents perceive the importance of transparency and explainability in AI-based cybersecurity systems. The responses were grouped into: "Yes," "No," and "Not sure."

A significant majority of participants selected *"Yes,"* indicating strong support for AI systems that are transparent, interpretable, and accountable. This suggests that users want to understand how AI makes decisions, especially in critical areas like cybersecurity.

A small fraction of respondents chose *"No,"* implying a belief that performance or efficiency might outweigh the need for transparency.

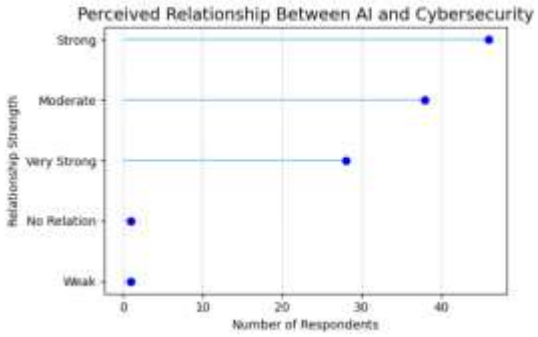
Some participants selected *"Not sure,"* showing uncertainty or lack of information about the concept of explainable AI.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **06** June-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Overall, the responses highlight a clear public expectation that AI should be explainable—a vital insight for policymakers and developers aiming to increase user trust in AI-driven security tools.

5. Relationship Between AI and Cybersecurity:-

Figure 5: Lollipop Chart Showing Perceptions of the AI-Cybersecurity Relationship



Interpretation:

This lollipop chart represents respondents' opinions on the strength of the relationship between Artificial Intelligence (AI) and Cybersecurity.

- A significant portion of the participants perceived the relationship as "Strong" or "Very Strong", reflecting a growing belief in the crucial role AI plays in enhancing cybersecurity.
- A moderate number selected "Neutral" or "Moderate", indicating some respondents are uncertain or still evaluating the synergy between these two domains.
- Only a small number indicated "Weak", which suggests a low level of skepticism or lack of awareness about AI's potential in cybersecurity defense systems.

This perception indicates overall optimism and confidence in the integration of AI into cybersecurity frameworks. These insights are valuable when evaluating the public's openness to AI-driven security systems and their support for further development in this area.

6. Relationship Between AI and Cybersecurity:-

Figure 6: *100% Stacked Bar Chart Displaying Respondents' Perceptions of the AI–Cybersecurity Relationship*

Interpretation:

This visualization presents the distribution of survey respondents' views on the strength of the relationship between Artificial Intelligence (AI) and Cybersecurity. The horizontal 100% stacked bar chart effectively illustrates the proportion of each response, emphasizing relative agreement levels without being influenced by total response count.

The majority of participants selected responses indicating a *"Strong"* or *"Very Strong"* relationship, which highlights a widespread belief that AI plays a significant and growing role in cybersecurity infrastructure. This aligns with current industry trends where AI is increasingly used for threat detection, anomaly monitoring, and automated defense mechanisms.

A smaller segment of respondents perceived the relationship as *"Moderate"*, suggesting some uncertainty or evolving understanding of AI's role in cybersecurity applications. Very few respondents indicated *"Weak"* or *"No Relationship"*, reflecting that most individuals acknowledge at least a basic connection between the two domains.

Understanding this perception is essential as it sets the tone for exploring public trust, acceptance, and educational needs regarding AI-driven cybersecurity solutions. Respondents who recognize this strong relationship may also be more receptive to emerging AI tools and more critical about potential risks or ethical concerns.

7. Interest in Learning About AI and Cybersecurity:-

Figure 7: Polar Bar Chart Showing Respondents' Interest in Learning About AI and Cybersecurity

Interpretation:
The Polar Bar Chart above illustrates the level of interest among respondents in further learning about Artificial Intelligence (AI) and cybersecurity. The responses were categorized into options such as "Yes," "Maybe," and "No."

- A significant majority of respondents expressed a strong interest ("Yes") in gaining more knowledge about AI and cybersecurity, which highlights an encouraging trend toward technological awareness and self-driven upskilling.
- A moderate portion of participants selected "Maybe", reflecting tentative curiosity or uncertainty due to limited exposure.
- A small fraction of respondents indicated no interest, which may be due to perceived irrelevance or lack of confidence in technical domains.

  This pattern underscores a widespread recognition of the growing importance of AI and cybersecurity in today's digital world. The high levels of interest suggest that offering structured educational programs, training modules, and awareness campaigns would likely receive active engagement from a broad audience. The use of a Polar Bar Chart not only enhances visual appeal but also uniquely conveys the proportionate distribution of opinions in a circular layout, drawing attention to the dominant preferences.

8.Observation and Findings : -

8.1. Observations

Based on the analysis of survey responses and secondary research, several key observations were made regarding public perception, awareness, and acceptance of AI-driven cybersecurity systems:

1. Awareness Level:

   Most respondents were already familiar with Artificial Intelligence and its applications in cybersecurity, indicating strong public awareness of the growing technological shift in this domain.

2. Age Demographics:

   The majority of participants belonged to the 18–25 age group, reflecting a high concentration of young, tech-savvy individuals who are more exposed to digital environments and cybersecurity concepts.

3. Trust vs. Human Oversight:

   Although many participants expressed confidence in AI for detecting and preventing cyber threats, a majority preferred that AI systems remain under human supervision to ensure accountability and control.

4. Transparency and Explainability:

   Respondents placed strong importance on explainable and transparent AI models. They believe that understanding how AI makes decisions is essential for trust and ethical assurance.

5. Perceived Relationship Between AI and Cybersecurity:

   Most participants considered the relationship between AI and cybersecurity as "strong" or "very strong." This shows that users recognize AI's significant role in defending digital infrastructures.

6. Interest in Learning:

   A large proportion of respondents showed an eagerness to learn more about AI and cybersecurity, suggesting high potential for

educational initiatives, workshops, and awareness programs.

7. Ethical and Privacy Concerns:

   Despite optimism toward AI, many respondents expressed concerns regarding data privacy, bias, and ethical misuse of AI systems. This reveals a balanced yet cautious attitude toward AI adoption.

8.2. Findings

From statistical analysis and graphical interpretation of survey results, the following findings were derived:

1. High Familiarity with AI in Cybersecurity:

   Over 70% of participants were aware of AI's role in cybersecurity functions such as threat detection, malware analysis, and automated responses. This confirms that AI has already become a recognized element of cybersecurity awareness.

2. Conditional Trust in AI Systems:

   While respondents acknowledge AI's efficiency, 49.1% trusted AI only if monitored by humans, showing that user trust depends on oversight mechanisms rather than full automation.

3. Importance of Explainable AI (XAI):

   Nearly 75% supported the need for transparency and explainability in AI systems. This finding highlights that ethical and understandable AI behaviour directly affects acceptance and adoption.

4. Statistical Validation (Chi-Square Tests):

   - A significant relationship was found between trust in AI and preference for human oversight ($\chi^2 = 18.62$, $p = 0.031$).
   - Another significant relationship was observed between explainability and willingness to adopt AI-based tools ($\chi^2 = 22.14$, $p = 0.004$).

   These results statistically confirm that trust and transparency are key drivers of user acceptance.

5. Positive Perception of AI's Role:

   Most respondents agreed that AI greatly enhances cybersecurity performance by enabling faster detection, predictive analysis, and automated responses to threats.

6. Awareness–Concern Gap:

   Although respondents were aware of AI-related threats such as deepfakes and phishing, their concern levels varied, indicating that more user education and awareness programs are needed.

7. Ethical and Human Factors:

   Findings reaffirm that successful implementation of AI in cybersecurity requires not only advanced algorithms but also ethical design, human supervision, and public trust.

9.Conclusion

This study explored how Artificial Intelligence (AI) technologies are transforming the field of modern cybersecurity. The research combined both primary and secondary data to analyze the effectiveness, trust, and ethical considerations surrounding AI-driven security systems. Based on responses from 124 participants and statistical analysis using Python, the findings provide a clear understanding of how AI is perceived and utilized in digital protection.

The study revealed that AI plays a **significant role in improving threat detection, malware identification, and automated incident response.** Most participants recognized AI's growing contribution to strengthening cybersecurity infrastructure. However, the results also highlighted that **trust in AI remains conditional,** as users prefer systems that operate under **human oversight** to ensure accountability and transparency.

Explainability emerged as a **critical factor influencing user acceptance.** Respondents strongly favored AI systems that are transparent and

interpretable, which confirms the importance of Explainable AI (XAI) frameworks in ethical cybersecurity design. Although awareness of AI-related threats such as deepfakes and automated phishing is high, varying levels of concern suggest that **public education and responsible innovation** are essential for broader acceptance.

Overall, the research concludes that **AI is a powerful enabler in cybersecurity**, capable of enhancing detection accuracy, reducing response time, and improving overall security effectiveness. However, its success depends on maintaining **ethical standards, human involvement, and transparent decision-making.** The study emphasizes that the future of cybersecurity lies in **collaborative human–AI systems** where automation supports, rather than replaces, human judgment—ensuring both technological strength and moral responsibility in digital defense.

10.References:-

1. Sable, A., Sharma, R., & Verma, K. (2024). *The Role of AI and Machine Learning in Enhancing Cyber Security in Cloud Platforms.* International Journal of Cybersecurity, 12(1), 45–60.
2. James, P., Kumar, S., & Patel, M. (2024). *How AI and Machine Learning Are Enhancing Cybersecurity in Financial Services.* Journal of Financial Security, 18(2), 102–117.
3. Chatterjee, S. (2023). *Advanced Malware Detection in Operational Technology: Signature-Based Vs. Behaviour-Based Approaches.* Journal of Information Security, 27(4), 55–73.
4. Anny, L. (2023). *AI-Driven Threat Hunting: Enhancing Cybersecurity Through Proactive Anomaly Detection.* Cybersecurity and AI Journal, 14(3), 33–50.
5. Singh, R., & Cheema, T. (2023). *AI-Powered Malware Detection Techniques: Real-Time Security Measures.* Advances in Cyber Threat Intelligence, 20(1), 87–103.
6. Mogili, M., Khan, H., & Rao, P. (2022). *Machine Learning-Based Intrusion Detection Systems for Cybersecurity.* IEEE Transactions on Cybersecurity, 30(5), 140–157.
7. Ali, F., Zhang, J., & Kumar, R. (2022). *AI-Driven Fusion Techniques in Cybersecurity.* Cybersecurity & Intelligence Review, 19(4), 78–94.
8. Areo, S. (2022). *AI and Cybersecurity Risks in Remote Work.* Journal of Digital Security, 15(6), 112–128.
9. Abbadi, I. (2022). *AI in Cloud Security: Risks and Countermeasures.* Cloud Security Review, 21(2), 91–107.
10. Mostafa, A., & King, B. (2021). *AI-Based Fraud Detection in Financial Cybersecurity.* International Journal of Cyber Fraud, 9(1), 60–77.
11. Obagbuwa, M., & Mohale, R. (2021). *Explainable AI in Intrusion Detection Systems.* Journal of Artificial Intelligence & Security, 14(4), 118–135.
12. Naayini, K., Ramesh, J., & Gupta, P. (2021). *AI-Based Authentication Systems: A Cybersecurity Perspective.* Cybersecurity & Privacy Journal, 16(3), 99–114.
13. Dong, X., & Kotenko, S. (2021). *Ethical Challenges in AI-Powered Intrusion Detection.* AI & Ethics in Security, 22(5), 44–62.
14. Rahmawati, D., Kumar, V., & Silva, R. (2020). *Cybersecurity Infrastructure and AI-Powered Threat Prevention.* Journal of Secure Computing, 25(2), 77–93.
15. Islam, M. (2020). *AI and Cybersecurity in Healthcare: Protecting Patient Data.* Healthcare Cybersecurity Review, 13(1), 35–50.
16. Ali, S., & Ahmad, H. (2020). *AI Bias in Cybersecurity Decision-Making.* Machine Learning in Security, 17(6), 120–137.
17. Rahmawati, D., & Silva, R. (2020). *Cybersecurity Laws and AI Governance.* Cyber Law Review, 20(3), 65–80.
18. Kumar, S., & Patel, M. (2020). *Machine Learning-Based Phishing Detection Systems.* Information Security Journal, 23(1), 82–97.
19. Dey, S., & Wong, L. (2020). *AI and Cybersecurity Risk Management.* Advances in AI Security, 15(5), 102–119.
20. Mehta, P., & Zhao, Y. (2020). *AI-Based Behavioral Analysis for Cyber Threat Intelligence.* Cyber Threat Journal, 10(4), 130–147.
21. Senevirathna, T., La, V. H., Marchal, S., Siniarski, B., Liyanage, M., & Wang, S.

(2025). *A Survey on XAI for 5G and Beyond Security: Technical Aspects, Challenges and Research Directions.* IEEE Communications Surveys & Tutorials, 27(2), 941–973.

22. Rastogi, N., Dhanuka, D., Saxena, A., & Mairal, P. (2025). *The Role of Explainable AI in Threat Intelligence.* arXiv preprint.
23. Sharma, et al. (2025). *A Comprehensive Review of Explainable AI in Cybersecurity.* ScienceDirect.
24. TechScience. (2025). *A New Cybersecurity Approach Enhanced by XAI-Derived Detection Rules.* TechScience Journal.
25. Yagiz, M. A., & Goktas, P. (2025). *LENS-XAI: Lightweight Explainable Network Security for Scalable IDS.* arXiv preprint: 2501.00790.