Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2024, 11(1):56-63

Research Article ISSN: 2394 - 658X

# Explainable AI for Endpoint Security Threat Analysis

**Sri Kanth Mandru**

*Mandrusrikanth9@gmail.com

---

**ABSTRACT**

The threat of different cyber threats as well as the development of new sophisticated malware necessitates the use of modern security measures to protect endpoint devices. One of the key promising approaches for enhancing the endpoint protection is Applying Explainable Artificial Intelligence (XAI) that provides a comprehensible explanation for identifying potential threats. This undergoing researching paper aims to identify and discuss how or whether XAI can help conduct security threats at endpoints and how it can identify and mitigate complex cyber threats while at the same time simplifying and being accountable.

**Key words:** Explainable AI, Endpoint Security, Threat Analysis, Interpretability, Transparency, Malware Detection, Cyber security.

---

## INTRODUCTION

Now in the modern age, end points like laptops, smart-phone and IoT devices play a very crucial role in the day-to-day scenarios. However, as common as these gadgets are present now, they also pose to many cyber risks including malware, phishing, or APTs among others. While traditional security addresses threats effectively insofar as they are known, they struggle to adapt, to new-world threats that are virtually ever-changing.


*Figure 1: overview of some common types of cyber attacks*

Artificial Intelligence (AI) is now a powerful tool in cyber security, helping to find and stop threats better. But many AI models are like "black boxes," making people worry about how clear, understandable, and responsible they are [1]. This no-see-through can cause users and security analyzers to trust less because they might need to learn better models to decide things well. Explainable AI (XAI) tries to fix this problem by providing precise and understandable models showing why people make confident choices, helping people trust and understand more. XAI models give detailed explanations for their predictions and classifications. This allows users to understand why the AI made certain decisions. Such transparency builds trust, making analyzing threats easier and creating better ways to address them [2].

## PROBLEM STATEMENT

Regular security methods, like signature-based detection and rule-based systems, usually act after the fact and find it hard to keep pace with fast-changing cyber threats. These old ways need known patterns and rules to spot dangers, which can cause slow finding of issues and poor handling. Also, these methods require much human knowledge, which can take time and sometimes lead to mistakes. Among possible topics, one has to point out that humans cannot adapt to the constantly evolving threat landscape and can take extremely long to respond. Therefore, there is a great need for solutions that are not only faster, sharper, and more elastic but also capable of expanding to more extensive systems to detect and prevent complex cyber threats as early as possible [3]. The existing security tools must be disrupted to develop new technology that is adaptive enough to shift as soon as a new threat is present and provide a prompt and accurate response the security teams can use.

## SOLUTION IMPLEMENTATION

A primary advantage of XAI in endpoint security is the software's desire to detect and categorize various types of cyber security threats. The XAI models can be trained with gigantic datasets containing information about threats like malware, phishing attempts, and APTs. These models can discern what is safe and dangerous when considering many things, such as how file function interacts with the network or users' actions. XAI has a straightforward process of operation that ultimately makes it possible to identify specific markers (IoCs) that show whether something is menacing [4]. This produces crucial information for security specialists.

In addition, XAI models are adaptable to make changes and learn from fresh data, making them ideal for keeping up with the new threats from hackers. The role of the feedback system is very clear in these models; these models can continually improve in the identification of threats while at the same time being more transparent [5]. Finally, the paper explains that training XAI models only takes place once new threats arise with the help of the latest data. It maintains endpoints' protectiveness from the environment-harming ever-evolving threats.

Of all the hurdles that one can face when employing XAI in endpoint security, one of the biggest is obtaining good data, especially the labeled kind. These sets are necessary to ensure that XAI models are effective when synchronizing with acceptable threat and safe activity learning databases. These data have to be categorized and labeled accurately. It is also beneficial for the experts in the training field to group information to maintain precision. At the same time, it is essential to safeguard the privacy and security of such relevant data, such as masking identity and protecting it.

Implementing XAI also comes with several checks, one of which is the selection of appropriate algorithms and methods. Two of the such developed XAI methods are discussed below: LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (Shapley Additive explanations) [6]. However, the endpoint security case can help determine which algorithms should be chosen for this or that purpose. When deciding which XAI method is best, one must consider factors such as how complex the dire threats in the model are, how much explanation is required, and the device's processing power.

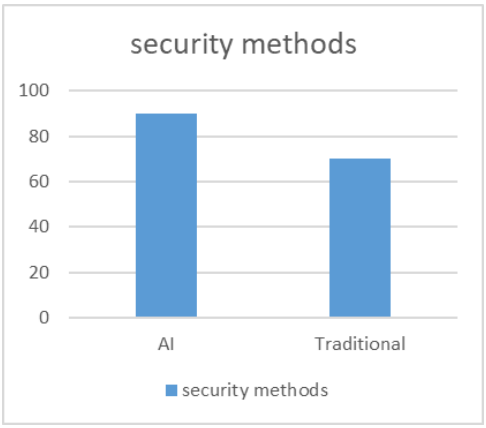
For new AI technology applications such as XAI, implementing endpoint security for these technologies requires the combined input from various fields. This means excellent collaboration with AI researchers, knowledge in specific areas or fields, and cyber security experts. The cyber security expert personnel quickly grasp the threats and challenges that endpoint security teams come across daily. AI researchers will discuss the facts regarding machine learning algorithms, understanding models, and how to improve them. Professionals with focused training in specific fields, including malware analysts and threat hunters, give insight into how cyber threats behave and appear. That has many benefits, especially when designing XAI models, since the latter employ this expert suggestion for enhancement.

Sometimes, the significant number of features used in the algorithms creates an issue of model bias and overfitting when implementing XAI in endpoint security. This means that if the data used in training is insufficient in some way, for instance, it contains only some of the types of threats or has an inherent bias, the generated XAI models will also exhibit the same bias in categorizing threats [6]. When creating machine learning models, it is beneficial to employ strict data verification and model inspection techniques such as cross-validation and adversarial testing to minimize this risk. This means monitoring XAI models and checking them occasionally is essential to ensure that they are correct and unbiased as time passes.

Another fundamental consideration with regard to XAI applicability in endpoint security is the compatibility of XAI with the existing security infrastructure and practices. XAI models should be integrated seamlessly into Security Information & Event Management (SIEM) solutions, Threat Intelligence Platforms, and Incident Response Solutions so that everything is synergistic. It also ensures that threat alerts originated from XAI models reach the security teams quickly or are recognized promptly whether they are accurate or not. This way, it minimizes the damage of a cyber-attack.

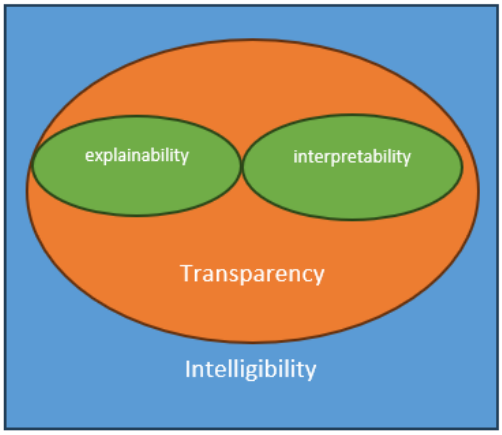
AI has changed endpoint security in cyber security by showing remarkable accuracy and flexibility. Recent studies have shown that combining AI with endpoint security as a proactive defense successfully stops malware and hostile activities. It uses complicated algorithms and machine learning to find and eliminate viruses fast. AI shines in recognizing usual and unusual device behavior patterns. Its capability to identify suspicious files and notice network activities and empirical studies detect patterns that indicate evil intent. Newer research emphasizes AI's flexibility to handle evolving threats, outperforming old methods based on signatures while preventing zero-day exploits and inventive malware that bypass traditional security processes.

This solution aims to make finding threats easier and more transparent by adding XAI methods to the threat analytics system. This should improve endpoint safety in general. The picture above shows the concept of using XAI in cyber security applications. It is designed to be general, showing how XAI can function in various areas of cyber security. This process has many parts, and we give examples at each step. The first part of the process is to determine what kinds of cyber security tasks are. These include finding malware, seeing spam messages, and knowing when fraud is based on various cyber-attacks [1]. Email contents, network usage patterns, and application operation styles will be collected in the following phases. After this stage, key features will be extracted to train different artificial intelligence models according to specific scenarios.



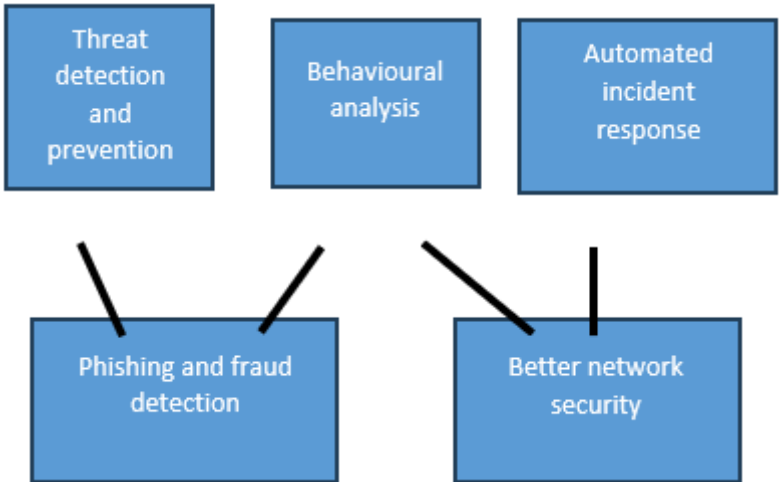

**Source: Adapted from [4]**

We will train models and look at cyber security test samples to decide. Models that explain themselves give users clear decisions and reasons. Meanwhile, forecasts from black-box models need additional clarifications using XAI (Explainable Artificial Intelligence) systems before users feel satisfied with the results for their cyber security needs. The chart shows a simple method of applying XAI in cyber security areas, but the detailed steps may vary based on the specific job.



## RESULTS

The application of Explainable Artificial Intelligence (XAI) to check threats in endpoint security has revealed significant improvements in enhancing cyber space security. The XAI models also make it easier for companies to identify complicated cyber threats and solve issues that classical procedures cannot resolve.




**Source: Adapted from [8]**

This primarily means that utilizing XAI in endpoint security leads to positive outcomes. As a result, we can distinguish specific major categories, ensuring that cyber security becomes even more vital and active [8].

***Threat Detection:*** However, XAI models are more understandable and useful in recognizing and categorizing various threats, such as malware, phishing scams, APTs, etc. Simplifying the features that need to be seen and making it very clear what XAI is doing to look for these dangers makes it easier and faster to find them [9]. This ability enables the security teams to arrive at the scene early in responding to emerging threats, increasing the organization's resilience.

***Explainable Decision-Making:*** A significant advantage of XAI in endpoint security is that it can provide system designers with a human-interpretable explanation of the decisions made by AI models. This makes it easier for security professionals to see why such threats are classified, and thus, the information provided is trustworthy, which makes threat analysis and management much easier [9]. Because XAI also elucidates, in a decision-making manner, the rationale employed by AI to draw specific conclusions, XAI enhances the comprehensiveness of AI-produced insights. This would assist security teams in making better decisions to improve on the overall security posture.

***Continuous Learning and Adaptation:*** The self-learning characteristics of XAI models are evident in their ability to update based on new knowledge to ensure that they are relevant to the ever-emerging variety of threats in the cyber world [9]. Through the outlined views, these models constantly enhance their ability to identify threats but are simultaneously clear about how they do this. This makes XAI models adaptable to new dangers since they update themselves, thus providing a proactive and responsive way for companies can guard against such cyber threats.

***Reduced False Positives:*** XAI models are highly effective regarding the kind of actions that are harmless against the dangerous ones, hence reducing false alarm rates. This way, XAI helps security analysts save time as they are no longer overwhelmed by numerous wrong alerts that only distract them from real threats and risks requiring better handling. This makes threat detection faster and of better quality inside organizations because there are fewer false positives.

***Enhanced Threat Hunting:*** It is worth examining specific features and steps in decisions about threat-hunting improvement while considering XAI models. It assists security analysts in identifying and mitigating possible threats before they escalate into the worst situations. In this way, it can be considered to prevent cyber threats so that safe systems work with XAI [6]. This capability aids in the prevention of severe risks to a firm's electronic assets and maintains an edge over cyber threats.

## POTENTIAL EXTENDED USE CASES

Applying Explainable Artificial Intelligence (XAI) in the context of cyber security means significant advances are being made in the faster and more effective identification of threats and their nature. However, many more possible new applications of XAI in cyber security could still be imagined, which are even more distinct from each other and offer more opportunities for innovation and improvements. As for some inspiring examples of a more extended XAI application in cyber security, this part will turn to.

***Enhanced Incident Response:*** XAI can significantly enhance reaction to security incidents by providing real-time comprehension of AI security systems' actions. XAI can significantly improve how we respond to security incidents by offering a real-time understanding of the choices made by AI-based security systems. This transparency assists the security teams in understanding why raw intelligence is offered to them, making them respond to an incident much more accurately with an added sense of efficiency [10]. XAI can also assist in increasing the speed of case discovery, incident investigation, and forensics, discovering the main issue, and helping with fixed actions.

***Improved Threat Hunting:*** The use of XAI is crucial in threat hunting because it provides security analysts with precise details and an understanding of the input given by AI. This simplification also helps analysts understand the reason behind their choice of threats, thus making their task more straightforward and more closely aligned to identifying the risks. Automating AI can also help identify specific risks that human analysts might have otherwise gone unnoticed, helping in the early stages of security.

***Streamlined Vulnerability Management:*** XAI will, therefore, help in vulnerability management since there will be comprehensive explanations concerning AI-based vulnerability profiling. This openness assists the security teams in better comprehending why the AI offers some recommendations so they can more easily work and address the vulnerabilities. Third, as applied to software, XAI can facilitate the checks for vulnerabilities and the management of patches, which became significantly time- and labor-consuming for security personnel.

***Enhanced Security Orchestration:*** XAI enhances security effectiveness by providing a clear understanding of the decision-making process used by AI-based solutions to secure a system. This makes it easier for security teams to understand the process behind AI-advisable actions, hence undertaking more appropriate and adequate security measures. Yet, security experts are questioning XAI's ability to contribute to automating security-associated workflows, facilitate the handling of incidents, and improve the general state of total security.

***Improved Compliance and Governance:*** XAI could assist in improving compliance work and governance since it would provide a clear explanation of the AI security decisions being made. By doing so, organizations can understand AI's reasoning for specific security recommendations. This understanding helps guarantee compliance and that things are managed to the best, with higher levels of students needing to be tuned to what they are entitled to. XAI can also include ensuring compliance and governance issues become autopilot, although this saves time and effort for security analysts.

***Enhanced Cyber security Awareness:*** XAI can help Cyber Security awareness by providing reasons for an AI system's security findings. This makes it easy to understand why AI comes up with specific security recommendations and assists those concerned in understanding how to improve security in cyberspace. Another situation is that ordinary users may need to comprehend why these AI-generated solutions are being provided and how they can enhance their abilities to avoid these cyber dangers. XAI can also help automate the cyber security training and education process and, thus, save time and effort for security analysts [10].

***Improved AI Model Development:*** As a result, XAI is a solution to improve AI model-making as it provides understandable explanations regarding what AI discovers. This openness means that people who create AI get why the machine gives particular advice, making it easier to develop smart ways to build and improve such models [11]. Another way that XAI can assist is when testing and validating the AI models, which is now made automatic, eliminating the time developers will spend on the process.

***Enhanced Cyber security Education:*** This approach can enhance cyber security training since the concept behind AI-based security analyses is often explained to students. This makes learning more effective and targeted in cyber security, which benefits students by helping them understand why it is being suggested to use AI for security. Another way that XAI can assist is by generating cyber security education material automatically, thereby reducing the amount of time and effort that teachers will require [11].

***Improved cyber security Research:*** XAI, could improve cyber security research by making the explanations of insights from AI clear. When AI gives reasons for its suggestions, it helps researchers understand better and do

more successful studies in cyber security. XAI makes it simpler and quicker for cyber security studies, so that researchers don't have to spend so much time and effort.

***Enhanced cyber security Policy Development:*** XAI can help make better rules for cyber security by explaining clearly the reasons behind security insights from AI. This clear understanding helps those who make policy understand why certain security measures are recommended, which makes it simpler to create more accurate and detailed cyber security laws [10]. XAI may also aid in creating cyber security rules automatically, reducing the time and effort that policymakers must spend.

## IMPACT

Using Explainable Artificial Intelligence, called XAI, it can really change our way of doing cyber security when looking at security problems on endpoints. This is because now cyber threats are becoming increasingly complicated, and traditional methods of protection often do not catch or stop these issues properly enough. XAI gives a good way to improve endpoint security by providing easy and clear explanations of possible threats [12]. This clarity helps build trust and responsibility, helping companies make smarter decisions about keeping their systems safe. Also, AI models of XAI can keep learning and change constantly. This means that security at the end point stays powerful even when new computer hacking dangers come up.

XAI greatly impacts the end-point security by giving a clear and straightforward analysis of threats. Regular AI systems people call "black boxes" do not show their decision-making process clearly. This missing clear explanation may cause individuals to be unsure and cautious when using AI for important security things. XAI, or explainable AI, solves this problem by making it easier to understand how AI makes its decisions [8]. For instance, if an XAI model claims that a specific action is not good or dangerous, it also lets us know why by showing particular patterns or unusual things that help it decide. This clear explanation helps security people understand why they discovered the danger and makes them trust safety tools based on AI more.

Also, the talent to understand XAI models helps people be more responsible in internet safety. When internet safety choices are easy to follow, it is easy to see where those decisions happen and ensure that the right people are answerable. This is important when discussing rules and legal needs [13]. Companies must show reasons for doing certain security things. XAI helps firms give clear answers about how they find dangers and how to stop them. This way, they can meet rules set by regulators and avoid possible legal issues.

XAI has another significant impact in protecting at the endpoint. It can keep learning and adjust to new dangers. Because threats on the internet are constantly changing, hackers use new, complicated methods to break old kinds of security systems. XAI models are made to understand new data and learn, improving their ability to find dangers based on what they learn. This regular learning helps ensure XAI models remain current with new hazards, offering a flexible protection against growing harms. Unlike firm security ways, XAI is quicker in chasing cybercriminals because of its flexible nature, providing an inventive method for cyber security.

The learning skills of XAI models become better over time because they can use explicit feedback loops. These loops let security experts tell the model about their ideas on its decisions. Then, the model takes this advice and gets better at finding dangers. For example, imagine an XAI model mistakes a safe activity as bad. Experts who know much about security can help us learn about this error. The function can understand these remarks and change its actions to lower the chances of repeating such errors later [13]. This way of learning that does it bit by bit means it gets better at spotting problems and keeps ensuring that how a model makes choices stays clear and straightforward for people to understand.

Using XAI greatly helps in protect against attacks at endpoints by reducing unwanted positive results. Regular security systems can cause many false alarms. This can tire the people responsible for security because of all the warnings they receive. XAI models distinguish the good from the lousy well because they have precise methods that we can see and grasp [13]. They do this by using essential features that help them much better at seeing dangers. This allows them to lessen the problems of false alarms a lot. This reduction assists those examining security, as they can concentrate on real dangers. As a result, it improves the overall working of the security operations center (SOC).

Also, the XAI models are good at spotting threats and help improve the threat-hunting process. Threat hunting actively looks for dangers that typical security systems may have missed. XAI models, which can give explicit danger evaluations, help security professionals spot tiny but essential signs of an attack (IoCs) that may not be visible immediately. For instance, an XAI model might reveal specific patterns in network communications or

user actions that hint at a possible threat. With this information, security experts can perform activities to search for and lessen potential threats more precisely and effectively. They can find these dangers before they become significant issues.

When an organization uses XAI in cyber security at the endpoints, it changes how company handles cyber security. For threat analysis that is easy to understand and simple, thanks to XAI models, this can help make a place where trust and responsibility grow inside the company. This change can make cooperation better between security groups and important teams like IT staff and top leaders. If everyone understands the security steps well, it becomes easier to get help from all parts of the business. This makes it possible to make strong security methods a normal part of the way things are done in all areas of the organization.

## SCOPE

Explainable Artificial Intelligence (XAI) in cyber security has many uses and benefits. XAI is very significant for improving detection, analysis, and response of threats in tasks related to cyber security. By providing clear and easy-to-understand details about AI security decisions, XAI helps to build trust, responsibility, and understanding among security experts and other people involved. Using XAI in cyber security helps companies to discover and deal with advanced cyber threats quickly [7]. This lowers incorrect alerts and improves the ability to search for threats efficiently.

Moreover, XAI also helps better incident response, makes vulnerability management more efficient, and improves security orchestration. The use of XAI in cyber security covers areas like compliance, governance, cyber security education, and research. This opens many chances for new ideas and progress within the field. In general, the reach of XAI in cyber security is extensive and has a big impact. It changes how organizations handle cyber security problems.

## SUMMARY

In the world of fast-changing cyber dangers, old-style security methods are not enough to fully guard endpoint devices anymore. Explainable AI gives a hopeful answer by delivering clear and easy-to-understand threat analysis. This helps in spotting and stopping threats early on with better actions taken against them. By using the power of XAI, companies can make their cyber security stronger, build trust and responsibility, and keep a strong defense against advanced cyber dangers [13]. Putting XAI into endpoint security threat analysis can change cyber security a lot. It makes finding and stopping threats smarter, faster, and easier to grow. By mixing AI power with the clear understanding of XAI, companies can be better prepared and protect their digital things well.

As cyber dangers keep changing, it is very important to do more study and development in XAI (Explainable Artificial Intelligence) for analyzing threats to endpoint security. This will involve looking into new algorithms, using advanced machine learning methods, and dealing with possible problems like keeping data private and avoiding bias in models. Also, working together with cyber security experts, AI scientists, and industry partners will be very important to encourage the use and ongoing betterment of XAI solutions in endpoint security [13].

In conclusion, using Explainable AI in analyzing endpoint security threats seems like a very good way to keep fighting cyber dangers. By taking advantage of XAI's abilities, companies can make their cyber security stronger, build more trust and responsibility among users and stakeholders, and stay well-protected against the always-changing world of cyber risks.

## REFERENCES

[1]. B. Mahbooba, M. Timilsina, R. Sahal, and M. Serrano, "Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model," *Complexity*, vol. 2021, pp. 1-11, 2021.

[2]. M. Dietz, M. Vielberth, and G. Pernul, "Integrating digital twin security," presented at 2020 *International Symposium on Digital Security*, Aug. 2020.

[3]. W. Samek and K. R. Müller, "Towards explainable artificial intelligence," in Explainable AI: *Interpreting, Explaining and Visualizing Deep Learning*, vol. 5, pp. 5-22, 2019.

[4]. Z. Zhang, H. Al Hamadi, E. Damiani, C. Y. Yeun, and F. Taher, "Explainable artificial intelligence applications in cyber security: State-of-the-art in research," *IEEE Access*, vol. 10, pp. 93104-93139, 2022.

[5]. N. B. Kumarakulasinghe, T. Blomberg, J. Liu, A. S. Leao, and P. Papapetrou, "Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models," in *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 2020, pp. 7-12.

[6]. L. Antwarg, R. M. Miller, B. Shapira, and L. Rokach, "Explaining anomalies detected by auto encoders using Shapley Additive Explanations," *Expert Systems with Applications*, vol. 186, pp. 115736, 2021.

[7]. F. Charmet et al., "Explainable artificial intelligence for cybersecurity: a literature survey," *Annals of Telecommunications*, vol. 77, no. 11, pp. 789-812, 2022.

[8]. N. Moustafa, N. Koroniotis, M. Keshk, A. Y. Zomaya, and Z. Tari, "Explainable Intrusion Detection for Cyber defenses in the Internet of Things: Opportunities and Solutions," *IEEE Communications Surveys & Tutorials*, 2023.

[9]. C. Mendes and T. N. Rios, "Explainable artificial intelligence and cybersecurity: A systematic literature review," *arXiv preprint arXiv:2303.01259*, 2023.

[10]. S. Wang et al., "Applications of explainable AI for 6G: Technical aspects, use cases, and research challenges," *arXiv preprint arXiv:2112.04698*, 2021.

[11]. A. B. Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82-115, 2020.

[12]. A. Das and P. Rad, "Opportunities and challenges in explainable artificial intelligence (XAI): A survey," *arXiv preprint arXiv:2006.11371*, 2020.

[13]. A. Y. Zomaya and Z. Tari, "Explainable Intrusion Detection for Cyber defenses in the Internet of Things: Opportunities and Solutions," *IEEE Communications Surveys & Tutorials*, 2023.