Advanced Research on Information Systems Security, an International Journal

ISSN: 2795-4609 | ISSN: 2795-4560
Print & Online

Artificial Intelligence as a Support Tool to Cybersecurity Activities

José Arnaud [a]*, Paulo Matos [b]

[ab]IPCA – Polytechnic Institute of Cávado and Ave, Barcelos, 4750, Portugal
[a]Email: jarnaud@ipca.pt
[b]Email: pjmatos.viana@gmail.com

Abstract

The present study is based on the analysis of the concepts of Cybersecurity and Artificial Intelligence, focusing on how both relate to each other. The COVID-19 pandemic and the fact that the world was forced into confinement had a huge impact on the increase of cyber-attacks and the birth of new threats. Currently, Cybersecurity is becoming an extremely complicated task, as we can no longer think of blocking an attack only by resorting to humans. The advanced and automated ways in which attacks have been developing mean that the task of defence must also be automated. Artificial Intelligence, due to its particularities, meets this need. This technology has been playing an increasingly important role in Cybersecurity, even though we may still consider it to be at an early stage.

Keywords: Cybersecurity; Artificial Intelligence; Vulnerabilities, Covid19, Fuzzy Logic

* Corresponding author. Email address: jarnaud@ipca.pt

1. Introduction

Human beings possess a cognitive ability, such as learning and problem solving, that has been investigated for several years through different studies and experiments in order to understand how it works, i.e., how we comprehend, predict or manipulate what's around us. Artificial Intelligence (AI), however, is not just limited to understanding but also to building intelligent entities [1].

Certain security properties are guaranteed by Cybersecurity [2] in order to avoid possible risks in cyberspace, such as integrity, availability, and confidentiality [3]. However, Cybersecurity is not just focused on protecting cyberspace but also on protecting whatever operates within cyberspace and any of its assets that may have a direct or indirect relationship with cyberspace [3]. AI is gradually being integrated into business, education, medicine, and other fields. It is widely employed in a variety of application scenarios [4]. However, not all sectors have reached the same level of development. In terms of implementing artificial intelligence, the information technology and telecommunications sectors are the most advanced [5].

According to a worldwide survey [survey] of over 4,500 technology stakeholders from various industries, 45 percent of large companies and 29 percent of SMEs have already incorporated artificial intelligence in their organizations. It is also noticeable the increasing amount of research regarding accountability of artificial intelligence to assure trustworthy decision in the latest decade. Annual publications represent the rapid and widespread rise of XAI, Interpretable, Intelligible, and Transparent AI, with XAI first appearing in 2017 alongside the US DoD DARPA XAI initiative [6], as observable in Figure No1.

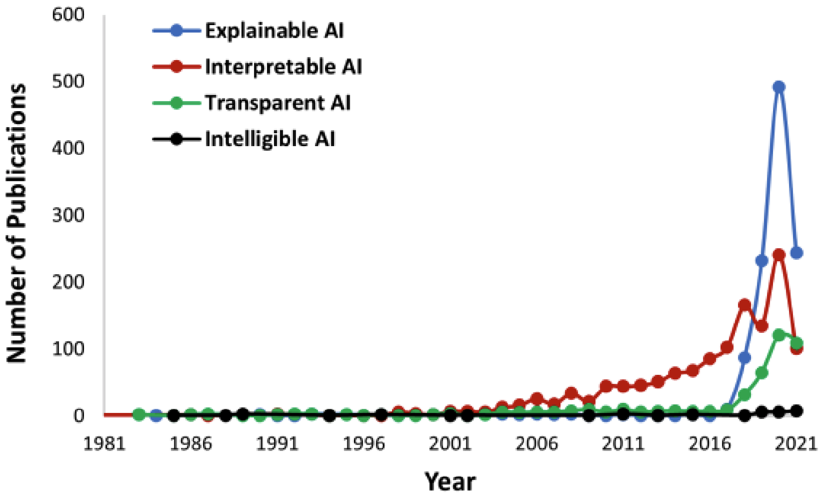

Figure 1 – Relationship and evolution grap XAI [6]

1.1 Cybersecurity

According to [7], Cybersecurity is the defence and protection of systems, networks, and programs in cyberspace against malicious attacks, namely through the application of standards, regulations, and encryption, in order to avoid possible damage, be it to hardware or software level.

[3], however, considers Cybersecurity as a set of tools, policies, security concepts, guides, risk management approaches, best practices, and technologies that can be used to protect cyberspace, organizations, and users.

Computers, infrastructures, and telecommunication systems services should also be protected as they use cyberspace to store relevant information.

According to [8], the concept of cyberspace applies to the global domain existing within an environment that produces a range of information, that is, within a network of Information Systems, such as the internet, telecommunication networks, computer systems, processors, and controllers.

Cybersecurity is sometimes the target of illicit activities, for instance, the access to information that is private and confidential to external entities of a certain organization and, therefore, should not be made public and is stored in computer systems or information networks [9].

There are several areas of Cybersecurity that should be considered as risks, such as:

- **Intrusion** – An intrusion is any set of actions intended to compromise the integrity, confidentiality, or availability of a resource. An intrusion results from the execution of one or more attacks on the systems that manage that resource. These attacks may or may not cause permanent changes to the information stored in those systems. It constitutes a difficult risk to assess since it does not need to involve exact data; however, it grants access to something that would normally be denied to the intruder [10].

- **Access to restricted or confidential information** – Computers store information; therefore, all unauthorized access is defined as a risk [10].

- **Information loss or theft** – Includes all situations where information is lost or stolen by unauthorized individuals and may even become in their possession [10].

- **Impersonation** – occurs when an individual subverts an authentication system by impersonating another person or when the behaviour set for an application is changed. Sometimes it is used as a ruse (to hide the true identity of a machine) or as an appropriation (using someone else's identity) [10].

1.2 *Artificial Intelligence*

Artificial Intelligence began to take its first steps after the Second World War, around the 1940s; even so, it was not until 1956 that it became more relevant [1]. However, it was only in the last decade that it became possible to advance with discoveries in this area due to the breakthroughs that occurred in computer and information sciences [11].

There are four different approaches to Artificial Intelligence that should be considered, namely, thinking humanely, thinking rationally, acting humanely, and acting rationally. It is possible to see that these approaches are based on two dimensions, thinking, and reasoning [1].

We can also consider different metaphors to understand Artificial Intelligence, these being computational, connectionist, and biological. The computational metaphor looks at intelligence as computation; that is, computers and the human mind have mechanisms of perception and action, in addition to the cognitive ones. The connectionists see intelligence as a property emerging from the interactions of a large number of elementary processing units, namely in relation to the human brain, in particular the relationship between neurons. The biological metaphor, on the other hand, focuses on the way in which species evolve, with this evolution being due to natural selection and promoting the survival of the most evolved and adapted species. A good example is the camouflage ability of some species, allowing them to hide from their predators [12].

It encompasses a wide variety of components, ranging from the way we learn to the way we play chess. Thus, we can state that Artificial Intelligence is relevant in any day-to-day activity. It focuses essentially on the use of algorithms or calculations, sometimes based on a goal. Algorithms are ambiguous instructions that a computer is able to execute; sometimes, these algorithms are adapted from others that already exist but that are not focused on what is intended [1].

Self-driving cars are a possibility due to Artificial Intelligence, which allows the system to pilot the vehicle without driver interference through computer vision, image recognition, and deep learning [1].

As previously stated, the goal of Artificial Intelligence is to build intelligent entities; these can also be referred to as agents. An agent can collect information from the environment so that it can act on it in order to determine the best course of action [12].

Agents should possess abilities such as autonomy, flexibility, and learning. If an agent is autonomous, then it is able to make decisions without the intervention of other agents. However, an agent does not need to wait for a change in the environment in which it is inserted in order to return a response, and it can also possess its own emotional state or personality [12].

There are five types of agents:

- **Reactive agents** – simple machines that limit themselves to react to the stimuli they receive from the environment. However, they may act even without the reception of stimuli as a response to the environment in which they are inserted. Thus, they can be represented on the basis of perception, from which the action is discovered [12].

- **Search agents** – these agents must be able to understand the existing action status and build on that an internal representation of them; in addition, they must have the ability to act on them, taking into account the operating rules of the systems in question. They allow the occurrence of transitions between states and have the ability to recognize when they reach a state defined as final, meaning when they reach the goal or find the solution for which they were planned. It is also important that the agent has in mind a strategy that is complete (capable of finding a solution), discriminating (if several solutions are available, the best one will be chosen), or economical (finds the solution in the shortest amount of time, spending the least amount of memory possible) [12], [13].

- **Knowledge-based agents** – these agents need knowledge and reasoning to increase their performance; thus, an agent has to build its own perception of the world. As such, it is necessary to know how to represent knowledge and interact to be able to develop reasoning, thus having the ability to decide [6], [12].

- **Learning agents** – learning is a characteristic of intelligent beings, so much so that artificial learning has become a central area of Artificial Intelligence. Artificial learning has three main objectives: the development of computational theories of learning, the implementation of systems with the ability to learn, and also the theoretical analysis and development of generic learning algorithms. A learning agent is, thus, based on perception and action (reactive agents), as well as the ability to decide and learn [12].

- **Adaptive agents** – use genetic algorithms; these are techniques that allow optimization, are useful for problem-solving, can be considered intelligent techniques – because they allow working simultaneously on alternative solutions – and are powerful tools when applied to solve problems. The main idea of a genetic algorithm is to mimic what nature does through its iterative process; it evolves the algorithm based on a population and defines a standard solution, adapting the solutions to existing problems [12], [14].

However, despite being machines with higher intelligence than usual for this type of system, sometimes they cannot withstand hacker attacks, just like a few years ago, when an attack affected video surveillance cameras and video recorders, making the use of this type of equipment unsafe, by decreasing the credibility of this type of services.

2. Artificial Intelligence in Cybersecurity

The increasing evolution of cyber threats, both in number and sophistication, now poses a serious problem for the security of cyberspace. Moreover, many of these types of threats exploit AI-enabled technology to improve their effectiveness, rendering traditional forms of defence useless. Attacks supported by AI are more dangerous, and most defences are not sufficiently prepared for this fight [15].

Artificial intelligence is becoming increasingly important in the cybersecurity industry for handling cyber threats. From 2022 to 2029, the market will grow at a percent per annum of 23.6, reaching $46 billion. Artificial intelligence introduces potential consequences: more than 60% of companies implementing artificial intelligence identify the cybersecurity risk associated by artificial intelligence as the most significant [15].

The demand for AI technologies in Cybersecurity is growing like never before. AI can automate large-scale tasks to analyse and detect threats quickly. It can also make decisions based on what it has learned, thus playing a hugely important role in active security. AI can therefore be used to address unknown threats on a large scale and in real-time, allowing it to block attacks that traditional models would not detect.

Thus, and after the execution of the systematic literature review, it was possible to derive the following assumptions.

The areas of Cybersecurity where Artificial Intelligence algorithms are often used are [16]:

- Intrusion Detection: An Intrusion Detection System (IDS) is a monitoring system that detects suspicious activity and sends alerts when it detects it. Based on these alerts, a security operations centre (SOC) analyst or incident responder can investigate the problem and take corrective action.

- Malware - (short for "malicious software") is a file or code that infects, explores, steals, or performs virtually any behaviour an attacker desire. Because malware comes in such wide varieties, there are numerous ways to infect computers.

- Ransomware - is a type of malware that prevents or restricts users' access to their system, either by locking the system's screen or by encrypting the users' files until a ransom is paid.

- DoS (Denial of Service) - s an attempt to bring a machine or network to a halt, rendering it inaccessible to its intended users. DoS attacks achieve this by flooding the target with traffic or sending information that causes it to crash.

- Phishing - the practice of duping Internet users (via deceptive email messages or websites) into disclosing personal or confidential information that can then be used illegally, the most typical type

Artificial Intelligence techniques that are considered promising and are currently the focus of research for Cybersecurity are:

- Machine Learning - is a type of artificial intelligence (AI) that enables software applications to become more accurate at predicting outcomes without explicitly programming them to do so. Machine learning algorithms predict new output values by using historical data as input.
- NLP (Natural Language Processing) - is a computer program's ability to understand human language as it is spoken and written (also known as natural language). It is part of artificial intelligence.
- Automation & Robotics - is a branch of engineering concerned with the design and construction of robots. They manipulate and process robotic actions using computers. These robots are then used in the following industries to speed up the manufacturing process.
- Fuzzy logic - is an extension or superset of Boolean logic, with the goal of preserving the concept of "partial truth," i.e., expression values ranging from "completely true" to "completely untrue" (from 0 to 1).
- Vision Machine - is a computer's ability to see; it uses one or more video cameras, analog-to-digital conversion (ADC), and digital signal processing (DSP). The resulting information is sent to a computer or robot controller. The complexity of machine vision is comparable to that of voice recognition.

The areas of Cybersecurity where Machine Learning algorithms are used are:

- Intrusion Detection
- Malware
- Ransomware
- DoS (Denial of Service)
- Phishing

The areas of Cybersecurity where NLP (Natural Language Processing) algorithms are used are:

- Intrusion Detection
- Ransomware
- Phishing

The areas of Cybersecurity where Automation & Robotics algorithms are used are:

- Malware

The areas of Cybersecurity where Fuzzy Logic algorithms are used are:

- DoS (Denial of Service)
- Phishing

The areas of Cybersecurity where Machine Vision algorithms are used:

- Intrusion Detection

From a cyber-attack perspective, AI can be a target. On the other hand, from a defence perspective, it can be considered a tool [17].

- AI from a defence perspective as a tool: Detecting and responding to threats automatically and in real-time.
- AI from an attack perspective as a tool: Enhancing attacks or being used for malicious purposes.
- AI from a defense perspective as a target: Being attacked or deceived due to vulnerabilities in the AI itself.
- AI from an attack perspective as a target: Protecting AI from the exploitation of vulnerabilities and attacks.

2.1 Artificial Intelligence from a defence perspective

As the attack surface grows and attacks become increasingly sophisticated, AI is becoming an indispensable weapon in the fight against these attacks. Organizations can use AI technologies to detect, predict and respond to threats in real-time, providing a faster and more accurate response automatically. AI technologies, including Machine Learning and Deep Learning, are already successfully used today in Cybersecurity, dealing with numerous threats such as malware, phishing, unauthorized access to sensitive data, and user behavior analysis, among other applications [18].

2.2 Artificial Intelligence from an attack perspective

Attackers are beginning to use AI to escalate their attacks. The AI tools and data needed for training are publicly available, making it easy for attackers to use them to research targets, discover new vulnerabilities, develop new payloads, and in the evasion itself, making these attacks more automated, large-scale, and extremely difficult to defend against.

2.3 Attacking Artificial Intelligence

Being a recently adopted technology in Cybersecurity, AI is still immature and vulnerable. AI systems are easily attacked or deceived due to the vulnerabilities of the AI itself. When using AI, there are various types of risks, whether in the data, the learning models, or even the interoperation processes [19]. As shown in Table 1, these attacks can be grouped into four categories: framework/component vulnerabilities, training data poisoning, adversarial attack (introducing tampered training data to mislead the model), and model theft [17] [18], [20].

In May 2016, a driver died while using the Tesla Model S's Autopilot system on a Florida road, a case that became known as the first fatal autonomous driving accident in the United States. The camera on the front of the car could only see the middle part of the truck suspended from the ground, not the whole thing. Add to that extremely sunny weather conditions (with blue sky and white clouds); the autonomous driving system failed to recognize that the obstacle was a truck. This is not a unique case in AI; there have been many robot-related deaths before. Elon Musk, Bill Gates, and Stephen Hawking have mentioned that humans need to be alert to the threat of AI. Imagine how dangerous and terrifying it could be if AI-based systems got out of the control of humans [17], [21], [22].

Table 1 – Vulnerabilities type

Attack Category	Description
Vulnerabilities of the components or *framework*	The frameworks or internal AI components are complex and heavily dependent on a large number of open-source packages that are the very root of the vulnerabilities that can be easily exploited.
Training data poisoning	Training data can be contaminated or mislabelled if it comes from external or public sources, which easily results in a data poisoning attack. With poisoning, the machine learning phase is affected, and the behavior of the AI is modified according to the will of the attacker.
Adversarial attacks	Machine learning algorithms, in particular DNN (Deep Neural Networks), are vulnerable to adversarial examples, a sample of input data that has been modified very slightly in a way that is intended to cause a machine learning classifier to classify it incorrectly.
Model theft	Machine learning algorithms can be reverse engineered or copied using just the publicly accessible query interfaces of the application without prior knowledge of the training data or model parameters.

2.4 Protecting Artificial Intelligence

The main issue is about protecting AI itself against the exploitation of vulnerabilities and other threats caused by targeted attacks on AI [21], [22]. Meanwhile, we should prevent AI from harming humans. In addition to defending against the known attacks mentioned above, the security of the AI model itself should be strengthened to prevent possible attacks by model validation or other means. Most importantly, AI systems can pose a potential threat to humans, in particular when using AI in robots [23], [24]. They can generate actions autonomously without manual intervention and can break away from human control due to self-learning capabilities [24], [25], [26].

3. Conclusion

With this paper, it was possible to further the knowledge both in terms of Cybersecurity and Artificial Intelligence through a more detailed understanding of the techniques in each of these areas. Thus, the areas of Cybersecurity and Artificial Intelligence were studied, thereby allowing us to know in detail their characteristics and possible strengths, as well as their weaknesses.

The focus of Cybersecurity is the protection and defence of a system in cyberspace through the use of tools, policies, and others to help counter possible threats. However, there is also a less positive side, and there are many illicit and illegal activities that have malicious intent, i.e., aim to harm a third party. Therefore, what was intended was to find ways to minimize the damage caused by these attacks, making it useful to use Artificial Intelligence, which focuses on machine learning based on human cognitive behaviour.

Artificial intelligence can be extremely useful in this regard. The combination of his abilities and people knowledge, backed up by the desire to save lives, will certainly have a synergistic impact. However, the operation of technology solutions must be strictly supervised by humans.

As such, it has been established which Artificial Intelligence techniques have an added value and which Cybersecurity areas currently pose a greater problem.

References

[1] Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach 3rd Edition.* Upper Saddle River, New Jersey 07458: Pearson Education, Inc.

[2] Gourisetti, S., Mylrea, M., & Patangia, H. (2020, 41). Cybersecurity vulnerability mitigation framework through empirical paradigm: Enhanced prioritized gap analysis. *Future Generation Computer Systems, 105*, 410-431.

[3] Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. *Computers and Security, 38*, 97-102.

[4] L. Ferreira, H. Lopes, C. Abreu, M. Cruz-Cunha, and N. Mateus-Coelho, “Secure Services Integration and Edge Computing for Effective Beekeeping”, ARIS2-Journal, vol. 1, no. 1, pp. 62–79, Dec. 2021.

[5] A. Almeida, N. Coelho, and N. Lopes, “Paranoid OS: Wearable Trackers”, ARIS2-Journal, vol. 1, no. 1, pp. 24–40, Dec. 2021.

[6] V. Kanimozhi and T. P. Jacob, "Artificial Intelligence based Network Intrusion Detection with Hyper-Parameter Optimization Tuning on the Realistic Cyber Dataset CSE-CIC-IDS2018 using Cloud Computing," 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0033-0036, doi: 10.1109/ICCSP.2019.8698029.

[7] Miller, L. C., & CISSP. (2016). *Cybersecurity For Dummies, Palo Alto Networks 2nd Edition.* Hoboken, New Jersey: John Wiley & Sons, Inc.

[8] Shiode, N. (2000). Urban Planning, Information Technology, and Cyberspace. *Journal of Urban Technology, 7*(2), 105-126.

[9] R. Arnone, “Hackers Cybercrime - Computer Security: Ethical Hacking: Learn the attack for better defence”, ARIS2-Journal, vol. 1, no. 1, pp. 50–61, Dec. 2021.

[10] Zúquete, A. (2018). *Segurança em Redes Informáticas, 5º Edição* [Security in Computer Networks, 5th Edition (in Portuguese)]. Lisbon: FCA.

[11] Desouza, K., Dawson, G., & Chenok, D. (2019). Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons*.

[12] Costa, E., & Simões, A. (2011). *Inteligência Artificial Fundamentos e Aplicações (3º Edição).* [Artificial Intelligence Foundations and Applications (3rd Edition) (in Portuguese)] Lisbon: FCA.

[13] A. Rawal, J. Mccoy, D. B. Rawat, B. Sadler and R. Amant, "Recent Advances in Trustworthy Explainable Artificial Intelligence: Status, Challenges and Perspectives," in IEEE Transactions on Artificial Intelligence, doi: 10.1109/TAI.2021.313384.

[14] N. Kseniia and A. Minbaleev, "Legal Support of Cybersecurity in the Field of Application of Artificial Intelligence Technology," 2020 International Conference Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS), 2020, pp. 59-62, doi: 10.1109/ITQMIS51053.2020.9322905.

[15] N. Mateus-Coelho, M. Cruz-Cunha, and L. G. Ferreira, "Security in microservices architectures," Procedia Computer Science, vol. 181, pp. 1225–1236, 2021.

[16] I. F. Mikhalevich and A. P. Ryjov, "Augmented Intelligence Framework for Protecting against Cyberattacks," 2018 Engineering and Telecommunication (EnT-MIPT), 2018, pp. 143-145, doi: 10.1109/EnT-MIPT.2018.00039.

[17] Zhang, F., Cui, X., Wang, Z., Chen, S., Liu, Q., & Liu, C. (2020). A Systematic Study of AI Applications in Cybersecurity Competitions. *Proceedings - 2020 IEEE 14th International Conference on Big Data Science and Engineering, BigDataSE 2020*, 138-146. https://doi.org/10.1109/BigDataSE50710.2020.00026

[18] J. A. Kroll, J. B. Michael and D. B. Thaw, "Enhancing Cybersecurity via Artificial Intelligence: Risks, Rewards, and Frameworks," in Computer, vol. 54, no. 6, pp. 64-71, June 2021, doi: 10.1109/MC.2021.3055703.

[19] Li, Jh. (2018). Cyber security meets artificial intelligence: a survey. In *Frontiers of Information Technology and Electronic Engineering* (Vol. 19, Issue 12, pp. 1462-1474). Zhejiang University. https://doi.org/10.1631/FITEE.1800573

[20] A. Boyko, V. Varkentin and T. Polyakova, "Advantages and Disadvantages of the Data Collection's Method Using SNMP", International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), pp. 1-5, 2019.

[21] X. Lou, K. Waedt, Y. Gao, I. B. Zid and V. Watson, "Combining Artificial Intelligence planning advantages to assist preliminary formal analysis on Industrial Control System cybersecurity vulnerabilities," 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2018, pp. 1-8, doi: 10.1109/ECAI.2018.8678949.

[22] A. Rawal, J. Mccoy, D. B. Rawat, B. Sadler and R. Amant, "Recent Advances in Trustworthy Explainable Artificial Intelligence: Status, Challenges and Perspectives," in IEEE Transactions on Artificial Intelligence, doi: 10.1109/TAI.2021.3133846.

[23] Russell, S. J., Stuart J., & Norvig, P. (1995). *Artificial intelligence: a modern approach*. Prentice-Hall.

[24] N. Mateus-Coelho, M. Cruz-Cunha, and P. Silva-Ávila, "Application of the industry 4.0 technologies to Mobile Learning and Health Education Apps," FME Transactions, vol. 49, no. 4, pp. 876–885, 2021.

[25] A. Boyko, V. Varkentin and T. Polyakova, "Advantages and Disadvantages of the Data Collection's Method Using SNMP", International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), pp. 1-5, 2019.

[26] N. M. Coelho, M. Peixoto and M. M. Cruz-Cunha, "Prototype of a paranoid mobile operating system distribution," 2019 7th International Symposium on Digital Forensics and Security (ISDFS), 2019, pp. 1-6, doi: 10.1109/ISDFS.2019.8757551.