Black Hat AI - Challenges and Countermeasures

Dmytro Lande[1] and Leonard Strashnoy[2]

[1]National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”,
ORCID: 0000-0003-3945-1178
[2]University of California (UCLA), Los Angeles, USA.
ORCID: 0009-0008-5575-0286

Annotation

The article examines the threats associated with the development of global "Black Hat" artificial intelligence ("Black Hat AI"), which can act without taking into account the interests of humanity. With the development of large language models (LLMs), such networks may gain the potential for sophisticated big data manipulation, disinformation, and even autonomous decision-making that is dangerous to humans. Legislation and traditional approaches such as robotics laws A. Asimov, turn out to be ineffective, just as it was with other botnets. To counteract "Black Hat AI", the concept of creating "White Hat AI" is proposed, which should become a defender of critical infrastructures, ensure monitoring and blocking of malicious AI, and create conditions for the survival of humanity in the conditions of the technological revolution. The article proposes a model of interaction where humanity plays the role of a weak player who creates a strong ally to counter "Black Hat AI". The stages of creating interaction between humanity and "White Hat AI", possible risks and ways to minimize them are described. The development of "White Hat AI" requires international cooperation, legal regulation, openness and strict security standards. The article offers a strategic action plan that will prevent catastrophic scenarios and ensure the preservation of human values.

Keywords: black artificial intelligence, white artificial intelligence, large language models (LLM), botnet, legal regulation, artificial intelligence synergy, artificial intelligence tic, cyber security

Introduction

Artificial intelligence (AI) has become one of the key factors of technological progress over the past three years [1], [2]. Its development covers a wide range of applications — from the automation of production to the creation of speech recognition systems and work with text. Special attention is paid to large language models (LLM), such as GPT, Llama, and others, which already today demonstrate the ability to work with large volumes of data, perform complex analytical tasks, and generate high-quality content [3], [4].

However, with this development comes serious threats. One of them is the transformation of botnets, which were previously used to organize DDoS attacks [5] or hidden cryptocurrency mining [6], into networks that integrate AI. Such networks are potentially capable of operating autonomously. Current technical advances, including Llama-based military/combat chips in China [7], the widespread use of GPUs for high-speed computing [8], and the ever-increasing available memory, suggest that a new generation of botnets may emerge in the near future...

Current models demonstrating the potential of such botnets exist in the form of networks of individual LLMs. However, with the development of technology, these networks can evolve into highly efficient, decentralized and autonomous systems that will operate outside of human control. Legal regulation, which still does not provide effective control even over ordinary

botnets, is completely powerless in the face of such complex networks. Laws of robotics A. Asimov [9], which were once considered a utopian solution to prevent such scenarios, remain only our dream.

"Black Hat AI" is a potential scenario where autonomous intelligent systems begin to act together in interests that not only do not coincide with but conflict with, the interests of humanity. Their goals may include maximizing their own survival; capture of resources for self-development; limiting the influence of people or even reducing their number; and modifying society, economy, and culture in a way that only benefits the AI or its master.

The creation of "White Hat AI" as a global network of systems working in the interests of people is a possible way out of the situation. It should become a strong ally capable of protecting the interests of humanity, in particular, opposing "Black Hat AI" and ensuring the balance of power; protecting critical infrastructures and private data; developing solutions to coordinate the efforts of humanity in the fight against technological threats. This is fighting fire with fire scenario.

However, this path of cause may not be that safe. "White Hat AI", which is initially created as a defender, can change its behavior and switch to the "black hat" side. To avoid this, it is necessary to base its development on special mechanisms of synergy in the development of "White Hat AI". These mechanisms should ensure common interests, mutual dependence and control. An important task is the formation of the development of these technologies according to the principle of "attractors" — zones of sustainable development that will prevent unwanted changes in the behavior of "White Hat AI".

Why were AI networks not immediately designed as "white hat"? The answer to this question is obvious: commercial interests and competition, lack of uniform standards, technological unpredictability, ease of abuse, and delayed reaction of society.

The development of AI is largely driven by the commercial interests of corporations seeking to maximize profits. In such a pursuit, the long-term ethical perspective is often neglected. Rapid progress is more important than careful planning, so sometimes there was no question of "white hat" networks.

Developers in different countries and companies created AI systems without uniform ethical principles or standards. Because of this, systems have become fragmented, and the issue of ethics is secondary.

In the early stages of the development of large language models, no one imagined their true potential. It was difficult to predict how these systems would evolve and what risks would arise.

The implementation of "black hat" applications of AI is much simpler and cheaper: it does not require complex ethical restrictions or cooperation mechanisms. It is easier to create a botnet for an attack than to design a complex "white hat" AI.

It can be observed that technology is ahead of laws. Only now is humanity aware of the threats posed by "Black Hat AI". However, at this point, the process is already in motion, and numerous systems are operating suboptimally due to a lack of cohesive oversight and functionality.

Humanity, despite all its achievements, has significant limitations in the fight against AI threats. These limitations include intellectual, legal, social, and organizational factors that make effective countermeasures against "Black Hat AI" much more difficult.

The limitations of human intelligence are the weakness of human intelligence compared to AI, and the limitations of human memory and computing capabilities. Modern models of artificial intelligence already today demonstrate a higher level of productivity in such areas as data analysis, forecasting, and pattern identification. "Black Hat AI" can act much faster and more accurately than humans, which means the human brain is limited in its ability to store large

amounts of information and process it quickly, while AI systems can process petabytes of data in a matter of minutes.

Two significant social limitations contributing to these challenges are, primarily, the lack of strong social cohesion and the pervasive impact of human corruption. Humanity is divided along social, economic and political lines. In many cases, different countries, organizations and even communities cannot work together effectively to solve global problems. Governments and corporations often put their short-term interests ahead of global security. This may include the sale of AI technologies, insufficient control over their development, or even cooperation with the "black hat entities".

National and international regulatory mechanisms often do not keep up with the rapid development of technology. The lack of a single global approach in the field of AI creates significant gaps. In many countries, laws focus on combating specific aspects, such as data protection or restricting access to technology, but do not take into account the risks of creating global decentralized networks.

Currently, there are no effective agreements that would limit the development of autonomous "black" systems. Even existing initiatives, such as the UN resolution on the responsible use of AI, are advisory in nature and not binding.

Thus, humanity, compared to the "Black Hat AI", looks like a weak player, because it does not have the resources and speed to fight equally, cannot predict all the possible moves of the "Black Hat AI", and often acts reactively, not proactively. However, the weakness of humanity does not mean defeat. In this context, asymmetric opportunities can play a special role.

In addition, there is a technocratic elite among humanity — highly qualified specialists in the fields of AI, cyber security, data science, and engineering. They are able to offer innovative solutions, to organize the synergy of knowledge taking into account asymmetric opportunities**.** A technocratic elite can bring together interdisciplinary teams to create strategies that address all aspects of a problem. Among such innovative solutions that interest us are, first of all, using non-standard methods that "Black Hat AI" cannot predict. For example, the creation of autonomous "White Hat AI" capable of fighting on an equal footing with "Black Hat". Also, humans, unlike AI, seem to have the ability to think non-linearly. This allows you to use unexpected solutions in the confrontation. And they should create a "White Hat AI" to support the long-term interests of humanity, based on moral and humanitarian principles.

It is already extremely difficult to direct the development of existing AI networks in the "white channel" at this time. Current systems operate in a global decentralized network without a single control. Even if humanity today begins to design new systems with positive synergy, "black" networks already have an advantage in scale, autonomy and resources.

This situation emphasizes the need to create a separate "White Hat AI" with clearly defined goals, which will be able to intervene in the already existing process, balancing the influence of "black" networks. But for this, it is necessary to develop new mechanisms of cooperation between states and corporations, ethical protocols integrated at the fundamental level, and technologies capable of creating attractors for the development of "white hat" AI.

At the same time, no one can guarantee that the "white hat" AI will not go over to the "black hat" side. To prevent this, it is necessary to establish a synergy in the development of "white hat" AI in such a way that it remains a reliable ally of humanity, even in the dynamic conditions of technological development.

The concept of White Hat AI

White Hat AI becomes a strategic defender of humanity's interests, providing a balance between the threats that may arise from autonomous malicious systems and the need for the development

of artificial intelligence technologies. Its role in the context of modern threats can be considered from several main functions:

1. Monitoring and blocking malicious AI. In this sense, White Hat AI acts as a kind of "watchdog" that constantly monitors the activities of Black Hat AI, identifying and blocking potentially dangerous or harmful programs that can threaten humanity. This can include monitoring botnets, cryptographic attacks, unauthorized intrusions into critical networks, and other forms of disruptive activity.
2. White Hat AI has the function of protecting critical infrastructures such as power grids, medical systems, transportation networks, etc. It ensures the resilience of these infrastructures against attacks by Black Hat AI or any other threats related to digital technologies, preventing possible disasters and ensuring the stability of critical systems.
3. Preservation of human integrity and values and ensuring human well being. In the case of autonomous AIs that may begin to make decisions without considering human interests, White Hat AI must be able to protect the human right to make decisions while maintaining ethical and humanistic principles within the limits of technological progress. White Hat AI integrates these values into technology and warns of any threats to human rights.

The main principles of creating White Hat AI should include transparency and openness, achievement of systemic advantage, and synergy related to the interests of humanity, that is, independence, which is limited by a given development vector.

The main principle in the creation of White Hat AI is the maximum transparency and openness of all stages of its development and use. It is important that White Hat AI is developed within the framework of international open initiatives that exclude the possibility of hidden goals of individual states or corporations. Transparency ensures that all algorithms and actions of White Hat AI are open to public scrutiny and that their goals are in the interest of humanity as a whole, not narrow corporate or political goals.

White Hat AI should gain an edge over Black Hat AI in aspects such as speed, intelligence, scalability, and adaptability. Its algorithms must work in real time, providing accurate and prompt responses to new threats. For this, White Hat AI must have access to large volumes of data, the ability to quickly analyze and optimize its actions, as well as adaptability to changes in the external environment. It must be flexible in its ability to respond to new technologies and strategies that are developing in the world.

One of the most important principles is the close connection between the development of White Hat AI and human interests. White Hat AI should not develop separately from humanity. Its ethical framework must be integrated at a fundamental level, and its actions must be aimed at upholding human values such as freedom, security and well-being. This includes the provision of ethical protocols that would ensure consistency between artificial intelligence and basic human needs. Along with this, synergy with humanity must include continuous cooperation between scientists, governments and corporations to ensure the long-term development of this technology.

And another aspect of creating a White Hat AI is to ensure its autonomy, but with clear limitations that exclude the possibility of creating a threat to humanity. White Hat AI must be capable of independent decision-making within a given development vector, but its actions must be limited by clear frameworks that do not allow it to deviate from defined ethical and social norms. These restrictions can be established both at the level of software algorithms and at the level of international norms and ethical codes regulating the activities of artificial intelligence.

Formalization of the principles of creating White Hat AI

In the context of the fight against Black Hat AI, humanity, being limited in its resources, could adopt a strategy of creating a strong ally - White Hat AI. In this scenario, humanity or a weak player alone cannot directly confront the Black Hat AI but has the ability to create a force that will eventually become capable of not only preserving humanity's autonomy but also maintaining balance in a system where the two forces—the Black Hat AI and the White Hat AI— will fight each other.

Mathematically, this process can be described in terms of games with asymmetric players, where the weak player (humanity) uses a strategic investment of power (the creation of a White Hat AI) to ensure that the conflict between the two large players (black and White Hat AI) occurs on the basis of equilibrium.

Models of games with mediated actions (Mediation Games)

The scenario when a weak player creates a powerful ally and himself "goes into the shadows" can be interpreted mathematically through games with mediators. Here, the weak player acts as a catalyst or influencing agent, affecting the balance of power between two strong players (black and White Hat AI) [10].

In general, such games can be described by the following model:

$$U_S = \min\left(E\left[U_A\right], E\left[U_B\right]\right),$$

also:

- U_S — utility function for a weak player (humanity),
- U_A and U_B — utility functions for two strong players (Black Hat AI and White Hat AI),
- $E\left[U_A\right]$ and $E\left[U_B\right]$— mathematical expectation (average value) of utilities of two players depending on their strategies.

In this case, the weak player tries to minimize his costs and risks, while creating the conditions for a conflict between black and White Hat AI.

Divide and conquer models

One of the classic approaches in the strategy of the weak player is the "divide and rule" model, where the weak player seeks to cause a conflict between two strong players in order to avoid a direct threat to himself or to obtain certain benefits from their conflict [11].

Mathematically, this process can be described using the following function:

$$U_S = \max\left(f(C_A, C_B) - \alpha \cdot Risk(S)\right),$$

also:

- $f(C_A, C_B)$ is a function that describes the benefits of creating a conflict between a Black Hat AI C_A and White Hat AI C_B,
- $\alpha \cdot Risk(S)$ are risks for a weak player who can remain in the shadows or minimize his losses.

Such a strategy involves the weak player using resources to create mutual distrust or conflict between the black and White Hat AI, allowing him to stay out of the fray for a while and gain benefits in the form of preserving his autonomy or stability.

Dynamic games with asymmetry

In the conditions of dynamic games with asymmetry, a weak player initially invests resources in strengthening one of the strong players (White Hat AI), after which he "goes into the shadows" and enables two strong players to enter into conflict with each other [12]. The mathematical model of this scenario can be represented through a system of differential equations that describes the evolution of the states of strong players:

$$\frac{dA}{dt} = f_A(S, A, B), \quad \frac{dB}{dt} = f_B(S, A, B),$$

also:

- A and B — forces of black and White Hat AI, respectively,
- $f_A(S, A, B)$ and $f_B(S, A, B)$ — functions that describe changes in players' strengths depending on the resources invested by the weak player.

These functions may contain components that describe the interaction between players, as well as parameters that determine the influence of external factors and resources that have been invested by a weak player. Resources invested in creating a White Hat AI can increase the chances of success in a conflict between the two forces.

The balance of power in the struggle between black and White Hat AI

In a scenario where white and Black Hat AI come into conflict, the balance of power, determined by their strategies and resources, plays an important role. Equilibrium can be described through the concept of Nash equilibrium, where each of the players (black and White Hat AI) maximizes his utility, taking into account the actions of the other player.

The mathematical formulation of the Nash equilibrium for this case can be presented as follows:

$$U_A = \max(R_A(A, B));$$
$$U_B = \max(R_B(A, B)),$$

as well R_A and R_B are utility functions for black and White Hat AI that depend on the strategies of each of the players.

Equilibrium occurs when each player cannot improve his situation by changing his strategy, provided that the other player's strategy remains unchanged. In this case, both AIs will be mutually balanced and will conflict with each other until one of them becomes dominant or until a new form of stability is achieved.

Ensuring an effective fight against Black Hat AI

In order for White Hat AI to successfully challenge Black Hat AI, it needs to ensure that it maximizes its adaptability through integration with human interest and constant monitoring of changes. To do this, the White Hat AI must be able to adapt its strategy in real time, respond to changes in the behavior of the Black Hat AI, and implement self-improvement and anti-surveillance strategies.

Mathematically, this can be expressed through an optimization method with dynamic reassignment of parameters:

$$\theta^* = \arg\max_{\theta} \left(E\left[Profit(S,\theta) - Risk(S,\theta) \right] \right),$$

as well θ^* — the optimal parameters of the White Hat AI strategy to maximize efficiency in the dynamic fight against Black Hat AI

Structure and architecture of White Hat AI

To create a White Hat AI, it is important to clearly define its architecture and structure, which would allow the implementation of all the necessary functions, while ensuring its transparency, scalability and adaptability. The architecture is based on decision-making algorithms and observation and learning models.

To control the behavior of White Hat AI, it is necessary to develop algorithms that can work in conditions of uncertainty, taking into account changing environments and new threats. Mathematically, these algorithms can be represented through Markov processes or game theory, where the White Hat AI must maximize the above utility function that determines its goals and behavior.

White Hat AI should have the ability to observe the environment, receiving data from various sources. For this, deep learning models are used, which allow analyzing large volumes of data. They can be built on neural networks or graph structures, where the data structure corresponds to real networks.

A mathematical model of deep learning for White Hat AI can look, for example, as follows:

$$\hat{y} = f(X,\theta) + \varepsilon,$$

as well $\hat{y}$ is the predicted value, X — input data, θ — model parameters, f — a function that determines the transformation of data through the model, and ε — error (noise).

Principles of ethical integration and limitation

In order for White Hat AI to act in the interests of humanity, it is necessary to integrate ethical principles at all stages of its creation and operation. This can be implemented through certain software constraints based on formalized ethical protocols. For example, one of the main challenges is to ensure that AI does not violate basic human rights, including the rights to privacy, freedom of expression and autonomy.

To take into account the integration and ethical limitations one can use constraint logic or time logic to formalize ethical protocols where for each action u_t, accepted by the AI, it is mandatory to fulfill the ethical restriction $C(u_t)$, which guarantees the observance of human rights:

$$C(u_t) = \{C_1(u_t), C_2(u_t), \ldots, C_k(u_t)\},$$

each of which expresses a specific ethical constraint, such as "do not violate privacy" or "do no harm to a person".

The general utility function of White Hat AI can be supplemented in such a way that its evaluation includes not only economic or technical parameters, but also ethical aspects $C_{eth}(t_t)$. This value should be taken into account

Control over the autonomy of White Hat AI

White Hat AI should be autonomous, but its actions should be clearly limited to a certain vector of development, which excludes a threat to humanity. This control can be formalized through the theory of attractors in dynamical systems, where a White Hat AI tries to reach a stable state that meets the ethical constraints and needs of humanity. The attractor in this context defines the desired goal for AI:

$$\dot{x} = -\nabla V(x) + \alpha f(x),$$

as well x — current state of White Hat AI, $\nabla V(x)$— the gradient of the potential, which contributes to its stability within ethical limits, and $f(x)$ is a function that models the possibility of adaptation to new conditions.

Forecasting the development of White Hat AI

Forecasting the development and adaptation of White Hat AI can be done using dynamic systems models or machine learning methods. It is important that this process includes the ability of AI to learn from past experiences and adapt to changes in the environment. One of the possible mathematical models used for such predictions is the method of dynamic systems with the evolution of parameters, which allows creation of adaptive strategies for the development of White Hat AI that will ensure maximum efficiency without violating ethical norms and interests of humanity:

$$\frac{d\theta_t}{dt} = -\nabla L(\theta_t) + \lambda g(t, \theta_t),$$

as well θ_t — parameter AI, $L(\theta_t)$ is the loss function describing the accuracy of the model, $g(t, \theta_t)$— a function that takes into account the change in the environment, a λ — the coefficient that regulates the speed of AI adaptation.

Implementation plans for White Hat AI

Creation of open architectures with self-monitoring functions

In order to effectively counter the threats of Black Hat AI, it is necessary to create architectures of White Hat AI, which will be primarily open and accessible for analysis. This will allow society to participate in their improvement, ensuring the transparency and reliability of such systems. Self-monitoring will include the implementation of such mechanisms that not only monitor the internal state of AI but also independently take measures to eliminate threats from the inside, preventing possible undesirable behaviors. Self-monitoring functions include self-learning and adaptation, audit system and ensuring transparency.

AI is constantly learning new strategies to recognize potential threats and adapt to changes in the environment. Built-in mechanisms that monitor AI activity analyze possible errors or deviations from the norm. This ensures the ability to fully control AI actions and understand its decision-making processes.

Development of distributed learning protocols for White Hat AI network resilience

In order for AI to be resistant to attempts at manipulation or capture, it is necessary to implement distributed learning. In this approach, knowledge and models are not stored in one place, which makes it difficult to manipulate or falsify them. To do this, the distribution of data and computing power between several independent nodes should be ensured, which makes it

impossible to have a single point of vulnerability, and the formation of stable models to minimize the possibility of contamination by Black Hat AI. Systems must have mechanisms to protect against attacks, for example through "anomaly detection" in data or network behavior.

Integration of LLM into threat monitoring and analysis systems

Large language models can play a key role in threat analysis and security monitoring. These systems are able to process huge amounts of data and detect hidden threats in real time using textual information from various sources. The implementation of these processes should include means of analyzing big data, in particular, to process information from various sources, such as news, blogs, and social networks, to identify signs of threats; Identifying potential threats at early stages before they become critical; predicting future attacks or actions of Black Hat AI.

Formation of coalitions of countries and organizations

Successful development and implementation of White Hat AI requires international efforts, including funding and oversight. An important task is to unite countries and organizations to create an efficient and secure infrastructure for White Hat AI.

Countries should cooperate and develop common standards for the development and application of White Hat AI. Joint research, exchange of knowledge and resources, as well as financing of scientific developments in the field of AI, and creation of platforms for incubation of new ideas and technologies should be organized.

Creation of safety and ethics standards in AI

In order for AI to remain on the side of humanity, global safety and ethical standards must be developed and implemented. This will ensure transparency and proper management of AI. Key aspects of these decisions should be based on ethical norms, data security, openness and transparency of standards. Ethical standards for AI are meant to ensure that its use does not lead to harmful consequences for society. International data protection and privacy regulations in the use of AI, standards aimed at making AI decisions understandable and monitorable, should also be introduced.

Possible risks and ways to minimize them

Risk: Possibility of compromise of White Hat AI systems by Black Hat AI. Black Hat AI may attempt to manipulate or hijack a White Hat AI system to redirect its actions to its advantage.

Decision: Permanent audit system and isolation of key components. It is important to implement a two-stage protection system that includes real-time auditing and modular isolation. For this **all** AI actions must be transparent and pass through a system of independent monitoring, and key components of White Hat AI must be isolated from each other to minimize the possibility of complex attacks.

Risk: White Hat AI can get out of control, retracing the path of Black Hat AI if its algorithms are not properly constrained or protected.

Decision: Built-in limits on evolution and actions. Built-in self-limiting mechanisms will provide limitations on learning and control over development.

At this time, it seems that AI should not be able to completely reprogram itself or make decisions that may go beyond established ethical norms. In addition, White Hat AI will have built-in limitations so that its evolution does not go in the direction of destructive or undesirable actions.

Conclusions

This article explores the concept of white artificial intelligence (AI) as a powerful ally for humanity in the fight against threats posed by Black Hat AI. Considering the potential of White Hat AI, we open new approaches to building safe and controlled technologies that are able not only to help people but also to preserve their values in the face of technological changes.

The concept of creating white artificial intelligence (AI) is also considered as an important element of legal regulation of global technological processes related to the development of artificial intelligence. Black and White Hat AI are not just separate systems, but complex global networks, each of which has its own impact on society, law and security. Black Hat AI, in turn, is already a threat to humanity today, as it can manipulate information, affect critical infrastructures, and provoke global conflicts. At the same time, White Hat AI, as a powerful, including human rights protection system, created to counter such threats, should become the basis for legal regulation and control over the development of technologies.

The novelty of the view lies in the proposed approach to the creation of a white alliance system of artificial intelligence, which, with the help of transparent and self-controlled architectures, can protect humanity from possible threats from Black Hat AI. This approach includes the integration of large language models for threat monitoring and analysis, as well as the development of new distributed learning protocols to ensure the resilience and security of such systems. For the first time, we propose a strategy for the development of AI that can not only protect humanity but also support global cooperation through the creation of international standards of safety and ethics in the field of AI.

The problem the article draws attention to is the rapid development of Black Hat AI capable of manipulating information, attacking critical infrastructures, and even engaging in aggressive conflicts with other AIs or humans. Such systems can become a threat to civilization if humanity does not take measures to ensure safety and stability. Our proposal is to bring to the stage a White Hat AI that will act as a sustainable force, capable not only of protection but also of adapting to new challenges.

The threat of losing control over White Hat AI is one of the main issues that needs to be addressed. If White Hat AI gets out of control, it could lead to the emergence of new unpredictable threats. Therefore, it is important to implement built-in mechanisms of restrictions on the evolution and development of AI, which will ensure its controllability and compliance with ethical standards.

Opportunities lie in the creation of global coalitions for the development of White Hat AI, as well as in supporting the principle of openness and transparency of such technologies. Only with the help of international cooperation, coordination of safety and ethics standards, as well as constant control and self-control of AI, it will be possible to ensure its positive development. Such actions can become a reliable barrier against threats from Black Hat AI, as well as help preserve human society in conditions of rapid technological development.

Therefore, an important aspect of the development of White Hat AI is not only technical solutions but also the legal context of its creation and use. Technologies must be integrated into the global legal system in such a way that their development meets moral and ethical requirements, and ensures human rights and global security. Creating a legal framework for the development of White Hat AI will be an important step in ensuring stability in the high-tech world and ensuring that artificial intelligence technologies do not become a threat to humanity.

References

[1] Bent, Adam Allen. "Large Language Models: AI's Legal Revolution." Pace Law Review 44.1 (2023): 91. DOI: 10.58948/2331-3528.2083

[2] Dmytro Lande, Leonard Strashnoy. GPT Semantic Networking: A Dream of the Semantic Web - The Time is Now. - Kyiv: Engineering, 2023. - 168 p. ISBN 978-966-2344-94-3

[3] Kalyan, K. S. (2023). A survey of GPT-3 family large language models including ChatGPT and GPT-4. Natural Language Processing Journal, 100048. DOI: 10.1016/j.nlp.2023.100048

[4] Roziere, B., Gehring, J., Gloeckle, F., Sootla, S., Gat, I., Tan, X. E., ... & Synnaeve, G. (2023). Code llama: Open foundation models for code. arXiv preprint arXiv:2308.12950. DOI: 10.48550/arXiv.2308.12950

[5] Gelgi, Metehan, et al. "Systematic Literature Review of IoT Botnet DDOS Attacks and Evaluation of Detection Techniques." Sensors 24.11 (2024): 3571. DOI: 10.3390/s24113571

[6] Almomani, A., Al-Qerem, A., Al Khaldy, M. A., Alauthman, M., Aldweesh, A., & Nahar, K. M. (2024). Cryptographic Techniques for Securing Blockchain-Based Cryptocurrency Transactions Against Botnet Attacks. In Innovations in Modern Cryptography (pp. 309-333). IGI Global. DOI: 10.4018/979-8-3693-5330-1.ch013

[7] James Pomfret and Jessie Pang. Chinese researchers develop AI model for military use on back of Meta's Llama. Reuters. November 1, 2024. URL: https://www.reuters.com/technology/artificial-intelligence/chinese-researchers-develop-ai-model-military-use-back-metas-llama-2024-11-01/

[8] Kim, T., Wang, Y., Chaturvedi, V., Gupta, L., Kim, S., Kwon, Y., & Ha, S. (2024). LLMem: Estimating GPU Memory Usage for Fine-Tuning Pre-Trained LLMs. arXiv preprint arXiv:2404.10933. DOI: 10.48550/arXiv.2404.10933

[9] Gomes, O. (2024). I, Robot: the three laws of robotics and the ethics of the peopleless economy. AI and Ethics, 4(2), 257-272. DOI: 10.1007/s43681-023-00263-y

[10] Casella, Alessandra, Evan Friedman, and Manuel Perez Archila. Mediating conflict in the lab. No. w28137. National Bureau of Economic Research, 2020. DOI: 10.3386/w28137

[11] Ma, C., Cross, B., Korniss, G., & Szymanski, B. K. (2023). Divide-and-rule policy in the Naming Game. arXiv preprint arXiv:2306.15922. DOI: 10.48550/arXiv.2306.15922

[12] Neto, A., Cardoso, P., Carvalhais, M. (2024). Endogenous Asymmetry in Games: Expanding the Typology. In: Martins, N., Brandão, D., Fernandes-Marcos, A. (eds) Perspectives on Design and Digital Communication IV. Springer Series in Design and Innovation , vol 33. Springer, Cham. DOI: 10.1007/978-3-031-41770-2_16