kaspersky

KASPERSKY SUBMISSION TO THE SUPPORTING RESPONSIBLE AI: DISCUSSION PAPER

(AUSTRALIA, JULY 2023)

Kaspersky, a leading international cybersecurity company, has a long and successful record of utilizing artificial intelligence (AI) technologies to prevent and combat cyberthreats – particularly, by using machine learning (ML) systems. The following submission is based on Kaspersky's profound expertise in AI as well as its extensive experience of engaging in UN-led initiatives on information and communications technologies, and reflects company's views on the main issues highlighted in the Supporting Responsible AI discussion paper.

Generally, when assessing potential AI-related risks, it is important to proceed from the fact that AI systems are not harmful and dangerous by design. It is the malicious use of AI technologies that poses risks – not artificial intelligence itself.

Classification of AI Systems

Kaspersky's stance is that, due to the heterogeneity of AI systems, regulation of their use needs to vary. First of all, it should be proportional to the level of risks associated with specific types of AI systems. This approach receives substantial support among the expert community and is currently under consideration by some jurisdictions, including the EU (the Artificial Intelligence Act), Brazil (Draft AI Law of December 2022), and Canada (the Artificial Intelligence and Data Act).

In this regard, Kaspersky proposes that the following classification of AI systems, which incorporate the elements of a risk-based approach, be taken into consideration when developing future regulations:

- AI systems with potential "intolerable" risks should be restricted (systems that may pose a threat to basic values and human rights);
- AI systems with "high-risk" should be regulated by a set of mandatory requirements (for example, AI systems utilized in critical infrastructure);
- AI systems with "limited or minimal risk" should be subject to minimal obligations (for example, recommendation services).

We believe that future standards and regulations for AI systems should reflect a balance between ensuring safety and security (including proper protection of human rights and basic values) and avoiding the creation of artificial impediments to the development of AI technologies.

Regulation of AI Use in Public and Private Sectors

In addition to this, Kaspersky also suggests that different approaches to public and private sector use of AI technologies be applied, with stricter regulations being introduced for public entities.

The potential failure of an ICT system in the public sector can have a more detrimental impact as it usually involves working with more sensitive information and dealing with more critical processes. Moreover, by adhering to stricter standards, the public sector can provide private entities with an example regarding the responsible and transparent use of AI technologies.

kaspersky

Ethical and Regulatory Principles

As AI/ML systems not only positively impact different areas of life, but also create a number of risks, the responsible, safe, and secure use of these technologies needs to be based on a set of both ethical and regulatory principles.

We believe that, in general, the following voluntary ethical principles should be implemented industry-wide (including cybersecurity vendors) to lay the groundwork for the responsible use of AI.

- All AI/ML systems need to be transparent both in operational and decision-making processes;
- Companies that use AI/ML need to implement advanced testing methods (ideally with the involvement of external expert testing organizations) to validate decisions made by ML algorithms, and use only trusted ML training data;
- Human control, including real-time access to all necessary data and processes for experts, needs to be a key element when implementing AI/ML systems;
- Personal information in training data sets used for AI/ML systems needs to be both minimized and anonymized wherever possible. Use of personal data in AI should comply with existing regulatory practices regarding the use of personal data in general;
- AI/ML systems must not be used for any negative or unlawful purposes.
- Developers of AI/ML systems should be actively engaged in exchanging best practices in the use of AI systems.

At the same time, regulation of the AI domain should not be limited to ethical principles only – especially when it comes to AI systems used in decision-making processes involving substantial levels of responsibility and risk. Taking into consideration the critical importance of their safety and security, oversight of these systems should be implemented through mandatory regulation tools. We believe that, in order to be efficient and effective, future regulation should match the following principles:

- Regulation should not create artificial obstacles for AI/ML developers, but provide additional incentives for companies engaged in developing and implementing AI/ML;
- Regulation should, in particular, cover multi-purpose AI systems trained on extensive amounts of open-source data and AI systems utilized in decision-making processes involving substantial levels of responsibility and risk;
- Regulation should reflect the various requirements for AI systems used in different domains, including cybersecurity;
- Regulation should be adaptive and stay abreast of both technology and market developments.

kaspersky

- Experts with relevant experience from the corporate sector and academia should be actively invited to take part in developing and discussing future regulation;
- Countries should actively cooperate in developing national AI regulatory frameworks in order to maximize international harmonization of AI regulation.

Transparency of AI System Utilization

Kaspersky's position is that transparency should serve as an underlying principle in the use of all types of AI systems by both public and private entities. While transparency is important across the whole AI lifecycle, it is most essential in the following processes:

- **Use of training datasets**

  Training an AI system largely depends on the data utilized for this purpose. It is the quality of the datasets used that greatly determines the efficiency of AI training and, eventually, the quality of its performance. In this regard, specific tools should be implemented in order to ensure and validate the quality of a training dataset, including in terms of bias, copyright, 'hostile data', etc.

  As there is a large amount of data available on the internet that was uploaded before the advent of AI technology, and the question of whether a person was ready to share her/his data to be used for AI systems development/training never arose. In order to respect the rights of individuals to their own data a special mechanism should be introduced requiring explicit consent from the user allowing the data they created to be utilized for AI training purposes. Moreover, a creator/owner of data should be also entitled to explicitly prohibit the use of his/her data for AI training purposes by, for example, adding a hash tag to indicate that the data can't be used for AI training.

  Additionally, the training of large AI systems or those that belong to a particular area, needs to be implemented only with the permission of a regulator.

- **Verification of results**

  The transparency of AI systems should be underpinned by the ability to prove the validity of their output. In particular, verification tools could be extremely useful for detecting and countering AI hallucinations[1].

In conclusion, Kaspersky believes that AI/ML domain regulation should be flexible and vary depending on the level of risk associated with specific types of AI systems, prioritize the protection of individual rights and personal data, and feature human control as an obligatory element of its development and use of artificial intelligence.

About Kaspersky

[1] AI hallucination – a response that is confident while also counterfactual.

kaspersky

Kaspersky is a global cybersecurity company founded in 1997. Kaspersky's deep threat intelligence and security expertise is constantly being transformed into innovative security solutions and services to protect businesses, critical infrastructure, governments and consumers around the globe. The company's comprehensive security portfolio includes leading endpoint protection and a number of specialized security solutions and services to fight sophisticated and evolving digital threats. Over 400 million users are protected by Kaspersky technologies and we help 220,000 corporate clients protect what matters to them most. Learn more at www.kaspersky.com.

For further information regarding this paper, please reach out to Yuliya Shlychkova (Yuliya.Shlychkova@kaspersky.com), Public Affairs Director, Igor Kumagin (Igor.Kumagin@kaspersky.com), Senior Project Manager, and Andrey Ochepovsky (Andrey.Ochepovsky@kaspersky.com), Public Affairs Manager.