Секция «Коммуникация в современном мире»

Affective computing as a source of destabilization of international psychological security

Научный руководитель – Худолей Константин Константинович

Колотаев Юрий Юрьевич
Postgraduate
Санкт-Петербургский государственный университет, Факультет международных отношений, Санкт-Петербург, Россия
E-mail: yury.kolotaev@mail.ru

Advances in the development of modern digital systems based on artificial intelligence (AI) are taking the issues of modern technology usage to a new level. One of the most promising and controversial areas of AI's application is the sphere of recognition and imitation of human emotions and feelings by computer systems. This area of AI's application is called affective computing (often also referred to as «emotional artificial intelligence» or «emotion AI»). The use of affective computing is possible in a wide area that encompasses digital or algorithmic processing of feelings expressed by a person and includes the recognition of emotional aspects of speech (spoken and written language), facial animation, nonverbal communication («body language») and other aspects of psycho-emotional monitoring. All of these types of emotion processing can be applied both for the needs of the digital space and in real life.

The potential threats posed by affective computing have a twofold nature. Emotion AI is often based on neural networks, which can inherently carry a large number of errors that affect the further use of AI. However, exogenous dangers may be a bigger threat than endogenous ones. A key role belongs here to the malicious use of artificial intelligence (MUAI), which can be applied by both private parties for personal gain or political actors to create threats in the field of international psychological security (IPS).

As for the first threat, the most obvious problems are the risks of biases and errors of affective computing, embedded (intentionally or unintentionally) in the algorithms themselves [2]. An example would be the bias towards the recreation of negative emotions and the spread of hate speech by algorithms that work for moderation on social networks. In the context of the IPS, these distortions can be deliberately exploited for manipulations in the digital space (i.e., lead to MUAI). The fundamental difficulty of recognizing the cultural or social context is equally acute and can also lead to distortions in the interpretation of people's emotions and feelings.

At the same time, the threat of MUAI to undermine the IPS using affective computing carries much greater dangers than the potential biases inherent in the algorithms themselves. "The prospect of automated detection of others' emotions adds to worries about AI's potential for pervasive, remote, and cheap monitoring and tracking at scale" [1]. In this regard, it is possible to compare open monitoring of human behaviour with a particular variation of cyber espionage, if such monitoring has a hidden and politicly motivated nature.

Even more disturbing is that emotions are a very powerful means of motivation, engagement and stimulation. They can be used as a tool to control or change the economic and political behaviour of people at the national or global levels [1]. To understand the level of danger coming from affective computing, one need only compare the capacity of simple fake news (generated often also automatically, but capable of inciting social tensions) with the potential of an algorithm that can accurately reproduce human-like behaviour or information relevant to a specific context. Such information could undermine the IPS if it affects a particularly sensitive area or topic of world politics at a particular moment.

Affective computing can be a source of instability at a less global level. For example, it could potentially reveal information that an individual or social group did not want to share (e.g., state of mind, hidden feelings or thoughts). Such processes can change our understanding of emotionality, as well as affect the experience of emotions by a person.

***

Affective computing is at the beginning of its development. This means that the timely identification of global and local threats of its use, as well as the indication of ethical aspects associated with it, are part of the priority areas in the study of the AI's potential impact on society and international relations. The ability to identify and eliminate potential threats stemming from affective computing can further reduce the risk of MUAI and increase the level of IPS.

References

1) Greene G. The Ethics of AI and Emotional Intelligence. Partnership on AI, 2020.
2) Purdy M., Zealley J., Maseli O. The Risks of Using AI to Interpret Human Emotions // Harvard Business Review. Boston, MA, 2019.