"Smart" Psychological Operations in Social Media: Security Challenges in China and Germany

Darya Bazarkina[1] and Darya Matyashova[2]

[1]Department of European Integration Research, Institute of Europe of the Russian Academy of Sciences, Moscow, Russia

[2]School of International Relations, Saint Petersburg State University, Saint Petersburg, Russia

bazarkina-icspsc@yandex.ru
dasham0708@mail.ru

Abstract: Artificial intelligence (AI) is actively being incorporated into the communication process, as AI rapidly spreads and becomes cheaper for companies and other actors to use. AI has traditionally been used to run social media. It is used in the various platforms' algorithms, bots and deepfake technology, as well as for the purpose of monitoring content and targeting instruments. However, a variety of actors are now increasingly using AI technology, at times with malicious intent. For example, terrorist organizations use bots on social networks to spread their propaganda and recruit new fighters. The rise of crimes involving AI is growing at a rapid pace. The impact of this type of crime is extremely negative – mass protests which demand the restriction of the use of technology, the involvement of manipulated persons in criminal groups, the destruction of the reputation of victims of "smart" slander (sometimes leading to threats to their life and health), etc. Combating these phenomena is a task which falls to security agencies, but also civil society institutions, the academic community, legislators, politicians, and the business community, since the complex nature of the threat requires complex solutions involving the participation of all interested parties. This paper aims to find answers to the following research questions: 1) what are the current threats to the psychological security of society caused by the malicious use of AI on social networks? 2) how do malicious (primarily non-state) actors carry out psychological operations through AI on social networks? 3) what impacts (behavioral, political, etc.) do such operations have on society? 4) how can the psychological security of society be protected using existing approaches as well as innovative ones? The answer to this last question is inextricably linked to the possibilities offered by international cooperation. This paper examines the experiences of Germany and China, two leaders in the field of AI which happen to have different socio-political systems and approaches to a number of international issues. The paper concludes that by increasing international cooperation, it is possible to counter psychological operations through AI more effectively and thereby protect society's interests.

Keywords: malicious use of artificial intelligence, psychological operations, social media, Germany, China

1. Introduction

As artificial intelligence (AI) technologies spread, become cheaper, and are further incorporated into the mechanisms of social media, there are a greater number of cases of malicious use of AI (MUAI), which damage the psychological security (PS) of society. Social media have become a favorable environment for asymmetric warfare, in which even actors without mass support from citizens (for example, far-right movements) can set or change the agenda in their interests, while terrorist organizations use social media algorithms for propaganda, recruitment, and fundraising (Stalinsky and Sosnow, 2020). Malicious actors, taking advantage of the widespread use of AI, the growth of international tension, and social and economic crises, which increase the emotional perception of information by the audience, are able to conduct full-scale influence operations. In the long term, this risks serious consequences, from the total distrust of citizens to any incoming information to the destabilization of democratic institutions and the coming to power of far-right circles.

The authors, based on the definition of PSYOPs (psychological operations) as "activities designed to convey selected information and indications to a foreign audience" that "aim to influence emotions, motives, ways of thinking and ultimately the behavior of foreign governments, organizations, groups and individuals" (Miljkovic and Pesic, 2019, pp. 1081–1082), view PSYOPs primarily as a sequence of planned actions to lower the PS of a society and further the economic or political interests of the PSYOP initiator. The basis of the research methodology is a system analysis that allows assessing the relationship between economic, political, technological, and social growth factors of MUAI in PSYOPs on social media, as well as the possibility of counteracting them. Scenario analysis was partly used, which allows identifying promising threats in the area under study. The analysis made it possible to compare the experiences of Germany and China to formulate recommendations in the field of international cooperation in countering "smart" PSYOPs on social media.

An important part of the theoretical basis of the research is the three-level classification of PS threats caused by MUAI, proposed by Evgeny Pashentsev (Bazarkina and Pashentsev, 2020). The first level involves discrediting the AI itself (even without using it) or an opponent using a negative image of AI in the infosphere. The second level is associated with causing physical harm or financial damage to a person or infrastructure through AI (which inevitably causes negative emotions, anxiety, and panic, even though the main goal of such MUAI is not to influence public consciousness). The third level of MUAI threat is directly related to tasks of distorting the perception of information by the audience to influence its actions.

This study draws on several groups of primary and secondary sources. The most significant groups of primary sources are legislative acts, official publications of governments and international organizations, and media materials. The bulk of the secondary sources were monographs, articles, and analytical reports on MUAI as a threat to PS, and general works on the problems of PS and PSYOPs and on the activity of and responses to aggressive economic and political actors.

2. The current threats to the PS of society caused by the MUAI on social media

The present paper studies the actions of two types of actors who carry out PSYOPs: aggressive *political* actors, such as terrorists and extremists, seeking to influence the political agenda; and *criminals* looking for profit, who use AI in social networks for malicious purposes. Both types of actors disseminate the "persuasive… information (messages), half-truths, 'misleading information' and misinformation, rumors and fake news that are distributed through the media, diplomatic channels [in the case of political PSYOPs] or the 'face-to-face' method" (Miljkovic and Pesic, 2019, p. 1084). Propaganda information "aim[s] at those psychological factors (perception, motivation, doubt, fear, stress – to psychologically shock, etc.) that, in different situations, have a decisive influence on people's behavior" (ibid). Already existing social media and mobile application tools can be widely used to disseminate such information.

When considering social networks as targets and spaces of PSYOPs, and often as their tools, it is important to keep in mind the ability of social networks to cause addiction, connected not least with targeting in the selection of content offered to the user. In 2018, some 2.6 percent of German youths aged between 12 and 17 were addicted to social media apps like WhatsApp, Instagram, and Snapchat, according to a representative study by German health insurance firm DAK (Goebel, 2018). In the PRC, where social media addiction is even more manifest (in particular, due to the greater number of young people), the government is adopting restrictions on the amount of time young people spend on social media and online games (Conklin, 2021). Under these conditions and against the backdrop of COVID-19, the threats of PSYOPs are likely to increase as the stress of the pandemic drives social media addiction (Zhao and Zhou, 2021). Malicious actors can, by their actions, both aggravate and create addiction in new users.

One of the most widely acknowledged social media threats to PS can be the malicious use of AI-based social bots, which are used to massively and rapidly spread spam, propaganda, rumors, or conspiracy theories (Gensing, 2020a). Since 2015, the so-called "Islamic State" (IS) has created thousands of Twitter bots for propaganda, fundraising, and recruitment, "as well as jamming activist communication on the platform, silencing their opponents on Twitter" (Stalinsky and Sosnow, 2020). In crises, PSYOPs using social bots can lead to particularly dire consequences, since the audience evaluates incoming information more emotionally and less critically, and leadership structures can make rash decisions.

"Deepfakes" can be distributed on social networks as part of disinformation campaigns to blackmail the "victim" for profit, with pornographic deepfakes leading among such products (Ajder et al., 2019). In a narrow sense, the process of creating deepfakes means adding one digital image or video on top of another in such a way that the added appears to be part of the original. However, without rejecting the differences in specific technologies, it is possible to use the term in a broader sense, combining it with a set of existing and future technologies for constructing pseudo-reality (Pashentsev, 2020, p. 101). These technologies are based on the ability of AI to create or modify images, video, sound, and text.

3. *Modi operandi* of the malicious actors in the AI-based PSYOPs on social networks

In the countries studied, there are varying degrees of threat from the malicious use of bots and deepfakes. The PS threats of the first and third levels appear in Germany. At the first level is the exploitation of AI by the right-wing political party "Alternative for Germany" (AfD) to attract attention during the elections: the AfD announced that it would include social bots in its election campaign strategy (Gensing, 2020a), which caused outrage across German society. Third-level threats include the PSYOPs conducted in the country, combining the use of fake accounts on social media by real people and the support of their publications by bots. Thus, during the federal election campaign of 2017, activists of the far-right movement "Reconquista Germanica" (RG) on Twitter managed to make certain hashtags, such as #reconquista and #nichtmeinekanzlerin, viral. On Discord, RG indicated supporters whose profiles should be attacked with hate speech comments. Just before the election, a growing number of AfD-supporting bot activities were discovered (Gensing, 2022). Deepfakes are recognized in the country as a possible threat to democracy, but their most famous use is in the case of fraud. In 2019, an executive in a United Kingdom-based energy company received a phone call from his boss in Germany instructing him to wire €200,000 to a Hungarian supplier within the hour. The call was a deepfake audio that "had imitated the boss's voice, tonality, punctuation, and even the German accent" (Rashid, 2021), which can be regarded as a second-level threat.

In China, the most famous examples of MUAI in social networks are cases of teaching chatbots anti-Chinese sentiments (presumably by ordinary users, which does not fully allow these cases to be attributed to "smart" PSYOPs), as well as first-level threats – PSYOPs that create the image of the PRC as a state that uses AI for repressive purposes (Bazarkina and Pashentsev, 2020). The third-level threat in China is the rapid spread of deepfakes, which are most often used to create pornographic content. Such content, sold through peer-to-peer networks, has attracted the attention of Chinese law enforcement not only as an illegal trade but also as a possible tool for destroying reputation (De Seta, 2021), which subsequently led to the adoption of tough legislative measures.

While conducting PSYOPs in social media, malicious non-state actors generally rely on the relevant social media's (and the Internet search engines in general) "benign" AI algorithms failures – in particular, their inability to identify hate speech in local or rare languages and, consequently, to moderate social media efficiently due to lack of data for machine learning (Heise, 2019). Systems making decisions about content further target promotion through the combination of users' interests and content popularity. There are also regular failures in tracking elements of extremist and terrorist propaganda in sounds, symbols, and censored scriptures, the meanings of which only become clear in the correct context (Weimann and Masri, 2020) due to both the sophistication of the type of tracking and to constant media communication context modifications. This approach is attractive and beneficial for malicious actors since it does not require special skills or programs and is thus cost-effective and time-saving, despite the need for creativity to "mitigate" radical rhetoric to conceal it from moderating algorithms.

This feature is exploited by different types of malicious actors, both by individual terrorists such as the 2019 shooter who relied on Twitch algorithms to widely spread the broadcast of his attack in the Halle synagogue (Jee, 2019) and by coherent groups promoting a destructive agenda, such as "Battle of the Nibelungs" using Facebook search and recommender algorithms to promote merchandise and martial art events, the latter to groom fighters for political struggle (Associated Press, 2021a); the Querdenken movement, which used Facebook to spread false information on COVID-19 vaccination and hatred to police (Associated Press, 2021b); and IS members from the Uighur minority, who produced a video in 2017 proclaiming their aggressive plans toward China and promoted it in the Chinese Internet sector (Al Jazeera, 2017).

The efficiency of the described "AI-exploitative" tactics is based on an approach that sees social media as both targets and spaces. As targets, they are exploited for entrenching radical ideas and messages into "normal" content, which becomes linked with destructive content. As spaces, social media are defined by the resources they potentially contain for malicious actors (such as people and financial support) and by fixed technical characteristics (AI algorithms) that give room to the described type of operations.

From 2016 to 2019, the transformation of social media algorithms and approaches to destructive content occurred as a reaction both to breakthroughs in AI development and the IS movement and hate crimes (Macdonald et al., 2019). These transformations require stricter rules of moderation and the usage of benign AI to prevent disruptive content from emerging and spreading. This would lead to more sophisticated operations being needed to spread disruptive content and the perfection of content promotion to circumvent moderation algorithms, as well as making the content itself more appealing and less detectable by detection systems. The given tactics of PSYOPs can include creating illusions of uncoordinated activities in social media based on fake people and automated text generation AI instruments, the generation of large amounts of radical content through the instruments of deepfakes and automated text generation, and the use of big data analytical systems to define the existing bottlenecks in social media algorithms. Their combination, in turn, is defined by the scale of a malicious actor and thus its ability to obtain sophisticated technologies, by the tactical goals of the regarded actor, by the media space, and by its target audience. Although there are no particular proven cases of deepfakes produced by terrorist groups (Semaan, 2020), the growing affordability of AI technologies as well as big data-state cooperation to fix existing algorithms flaws, limiting options for their simple exploitation, raises the probability that malicious actors will shift to new tactics based on fake content generation. These tendencies are common in the cyberspaces of both Western and non-Western regimes – in particular, we track cases of the "exploitative" malicious use of AI in two states with strikingly different regimes (Germany and China).

4. Behavioral and political impacts of the AI-based PSYOPs on social media

The impact that AI PSYOPs have on social media can be seen both in society's attitudes toward AI adoption and in policy decisions that respond to such PSYOPs. How the real experience of using AI and its coverage in the media (including by creating fear of total surveillance) affects the mood of citizens is indicated by the results of a survey of around 8,000 people in China, India, Germany, the United Kingdom, and the United States, published in January 2022 (Tagesschau, 2022). Only half of the respondents in Germany believe that new technologies can make the world better, and about a third see AI as a threat, particularly those technologies that monitory citizens' daily lives. In China, there is somewhat greater enthusiasm for AI, with only 26% of those surveyed saying it could pose a threat. Only 46% of Germans (against almost 100% in China) consider themselves and their country well prepared for the technological age. It is advisable to correlate the dynamics of such sentiments with the information background created around AI to avoid PS threats of the first level. We can assume an indirect impact on the mood of citizens from terrorist practices and crimes using AI technologies.

The short-term political outcomes of real AI malicious use cases are connected with an intensive securitization of media communications both in the context of content and in the context of technical issues. The former aspect implies imposing new social media restrictions based on security considerations defined not only by the dominant state discourse but also the level of AI development. For instance, Germany securitized the right of vulnerable social groups to dignity through NetzDG while China has focused on the general challenges to territorial integrity and PS (Blanchard, 2015). The extensive securitization in the context of countering radical actors and their propaganda catalyzes inter-actors (for example, the KISTRA project, which pulls resources of leading German universities and Bundeskriminalamt [RWTH Aachen University, 2021] to provide technical, ethical, and legal AI-based solutions for hate crimes problem in the cyberspace [TU Berlin, 2022]) and international cooperation (for instance, China's cooperation within the iBRICS and CyberBRICS platforms [CyberBRICS, 2019]). In the R&D area, prospects are linked with the development and promotion of more perfect and sophisticated "benign" AI software (for example, Hikvision and Cloud Walk, used by China to prevent terrorists intermingling with crowds [Kaushik, 2021]). In the long term, this can lead to raising the efficiency of countering radical propaganda through combining normative state-developed decisions and business-developed best practices. Transnational (state-companies and state-experts) and international (interstate) cooperation, in turn, can give an impetus for a global AI ethics in the context of countering terrorism.

Nonetheless, in the medium-term, the prospect of political conflicts around human rights to privacy and open information access, as well as on state-business relations, are likely to increase. These conflicts can have both domestic and international dimensions, creating political polarization inside the state and "security blocking foreign policy", which would be an obstacle to forming universal AI usage rules. The cases giving the opportunities to extrapolate this trend to the near future include the critical international reaction to Chinese laws that stipulated measures on tightening Internet security management, the inspection of dangerous materials, the prevention of terrorism financing, and border controls (Refworld, 2016; BBC News, 2015) as well as German protests in reaction to the stricter regulation of social media (in particular, the requirement to remove abusive material within 24 hours or face €50m fines; Scally, 2017). Mediatization of the described conflicts can give an impetus to more aggressive radical propaganda that would rely on "returning Internet freedom" discourses. An additional threat is the "arms race" between "benign" AI (e.g., systems of moderation, tracking, and prevention that target disruptive and artificially generated content, as well as more sophisticated detectors of coordinated social harm activities) and 'malicious' AI. This threat is closer to the technical outcomes of the MUAI intensification but is bound to be catalyzed by states' normative initiatives.

The short-term behavioral outcomes of these PSYOPs are beneficial for radical movements – PSYOPs promote radical ideas and allow raising financial support through advanced crowdfunding linked with existing social media algorithms (Associated Press, 2021a) by normalizing and integrating radical content integrated through linking it by, for example, sounds and hashtags that push AI-based recommender systems to promote it. In the short term, this simplifies terrorist recruitment and propaganda (Vacca, 2021), the latter becoming possible due to content individualization by algorithms. Furthermore, the complexity of malicious operations and countering them is fruitful ground for undermining public trust in social media as well as government authorities and their ability to confront the terrorist threat. Nonetheless, the long-term prospects are ambiguous for radical groups in all the dimensions described due to the technical development of AI moderation, a rapid backlash to radical content and ideas from the moderate social groups, and the limits of radical propaganda due to the strong correlation between receptivity to propaganda and the economic situation among recruited (Cibra, 2018). These socio-behavioral and socio-economic particularities constrain the destructive influence of radical MUAI.

5. Ways to protect the PS of society: existing approaches and innovations

Countermeasures against “smart” PSYOPs on social media are taken both in the “human” (legislative, political, and educational), and technical dimensions. Such measures can also be complex, representing well-coordinated counter-PSYOPs aimed at counteracting criminal manipulations. Thus, proposals for the use of social bots are useful, such as online services that allow military personnel and law enforcement officers to control virtual personalities that can interact on social media “to counter violent extremist and enemy propaganda” (Paganini, 2013). Of course, in situations of international tension, such technologies could be used by states against each other, but terrorism remains a common threat that creates the need for international cooperation.

In China, attempts are being made to use citizens’ addiction to social media (and the AI built into them) for positive purposes; technology companies are developing online training programs, using people’s desire to receive the encouragement of subscribers (“likes”) to complete tasks and pass checkpoints (Liu, 2018). It seems, however, that such practices risk abuse of the addictions of ordinary users by business entities. The measures taken by the Chinese government to limit the time spent by children and adolescents on social networks and online games can be aimed at minimizing them. Both in China and EU countries, including Germany, the removal of malicious content has been established, usually carried out by the administrators of social networks themselves at the request of law enforcement officers. In particular, this work is carried out in the EU in accordance with the Code of Practice on Disinformation, signed by major online platforms. When discussing technical measures to identify and remove content in Germany, proposals have been made to start countering PSYOPs with the help of AI on social networks, choosing as a starting point the content of the messaging, not the technology with which it was created (Gensing, 2020b), to overcome imperfections of bot identification tools. China has become an innovator in AI regulation, including in technologies such as deepfakes (国家互联网信息办公室 文化和旅游部 国家广播电视总局, 2019). In 2019, China announced new rules governing video and audio content on the Internet, including a ban on the publication and distribution of “fake news” created using AI and virtual reality. Any use of technologies like deepfakes should be clearly marked and clearly visible to Internet users, and failure to comply with these rules can be considered a criminal offense (Pashentsev and Bazarkina, 2022). Significantly, it is not the technology that is prohibited, but deliberate misleading an audience with it.

The element most vulnerable to PSYOPs through AI in the information and cybersecurity system remains the human, making relevant the education of citizens not only in the technical sphere but also in the political or psychological sphere. In particular, it is important to familiarize the audience with the goals of criminals, including terrorist organizations; the political forces that manipulate public consciousness; and the tactics and techniques used by both criminals who seek only profit and politically motivated actors. The cooperation of political institutions and cross-border security structures in the accumulation of expertise and the development of political and legislative solutions remains a relevant policy measure at the international level. This can be done, for example, by Interpol and the Centre for AI and Robotics at the UN Interregional Crime and Justice Research Institute (UNICRI and Interpol, 2019, p. 5). Today more than ever, civil society structures should be involved in such an exchange. It is also necessary to develop international scientific cooperation in both technical and social sciences, including interdisciplinary research projects aimed at countering high-tech terrorism.

6. Conclusion

The analysis shows that in both countries under study, MUAI threats of the first (discrediting AI itself or someone with the help of its negative image) or third (using AI directly in PSYOPs) levels are predominantly manifested. There is reason to believe that these threats will increase in the near future due to a number of factors:

- The growing dependence of the audience on social media against the background of the pandemic;
- International tension that impedes the development of effective international norms for regulating AI and mechanisms to combat MUAI, as well as creates conditions for the growth of social instability, which strengthens aggressive political and criminal actors;
- The level of technology development in which MUAI PSYOPs actors are ahead of structures that oppose MUAI (for example, due to insufficient data for training AI monitoring in the languages of countries where AI is not so highly developed).

The latter trend may change, but effective countermeasures are needed to accelerate this process, possible only based on international cooperation in the field of combating PS threats caused by MUAI. In this context, it is necessary to further study the transformation of the phenomenon of crime, including terrorism, during the "fourth industrial revolution"; the socio-psychological, economic, and political conditions in which AI is developed; and the psychological mechanisms that aggressive actors use to incorporate MUAI in their PSYOPs.

Fundings

The reported study was funded by RFBR, under number N 21-514-92001 "Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia."

References

Ajder, H., Patrini, G., Cavalli, F. and Cullen, L. (2019) *The State of Deepfakes: Landscape, Threats, and Impact*, Deeptrace, Amsterdam.

Al Jazeera. (2017) "ISIL video threatens China with 'rivers of bloodshed'" [online], https://www.aljazeera.com/news/2017/3/1/isil-video-threatens-china-with-rivers-of-bloodshed.

Associated Press. (2021a) "Neo-Nazis are still on Facebook. And they're making money" [online], New York Post, https://nypost.com/2021/09/27/neo-nazis-are-still-on-facebook-and-theyre-making-money/.

Associated Press. (2021b) "Facebook bans German accounts under new 'social harm' policy" [online], Euronews, https://www.euronews.com/next/2021/09/17/facebook-bans-german-accounts-under-new-social-harm-policy.

Bazarkina, D., and Pashentsev, E. (2020) "Malicious Use of Artificial Intelligence", *Russia in Global Affairs*, Vol. 18(4), pp154-177, doi: 10.31278/1810-6374-2020-18-4-154-177.

BBC News. (2015) "China passes controversial new anti-terror laws" [online], https://www.bbc.com/news/world-asia-china-35188137.

Blanchard, B. (2015) "China passes controversial counter-terrorism law" [online], Reuters, https://www.reuters.com/article/us-china-security-idUSKBN0UA07220151228.

Cibra, V. (2018) *Social Media and Terrorist Organizations: Observing Success of Recruitment through Social Media*, University of Central Florida, Spring Term.

Conklin, A. (2021) "China moves to limit social media, gaming as teens get more addicted and mental health is impacted" [online], Fox Business, https://www.foxbusiness.com/technology/china-social-media-limits-minors.

CyberBRICS. (2019) "CyberBRICS: Building the Next Generation Internet, STEP by Step" [online], https://cyberbrics.info/cyberbrics-building-the-next-generation-internet-step-by-step/.

De Seta, G. (2021) "Huanlian, or changing faces: Deepfakes on Chinese digital media platforms", *Convergence: The International Journal of Research into New Media Technologies*, Vol. 27(4), pp935-953, doi: 10.1177/13548565211030185.

Gensing, P. (2020a) "Bots im Wahlkampf: Interaktiv, aber nicht intelligent" [online], Tagesschau, https://www.tagesschau.de/faktenfinder/inland/social-bots-101.html.

Gensing, P. (2020b) "Das Problem mit den Social Bots" [online], Tagesschau, https://www.tagesschau.de/faktenfinder/social-bots-111.html.

Gensing, P. (2022) "Wie Trolle im Wahlkampf manipulierten" [online], Tagesschau, https://www.tagesschau.de/faktenfinder/inland/manipulation-wahlkampf-101.html.

Goebel, N. (2018) "100,000 German teenagers addicted to social media, study finds" [online], Deutsche Welle, https://www.dw.com/en/100000-german-teenagers-addicted-to-social-media-study-finds/a-42783211.

Heise. (2019) "Overblocking in Südkorea" [online], https://www.heise.de/newsticker/meldung/IGF-Kuenstliche-Intelligenz-versagt-oft-im-Kampf-gegen-Hass-und-Terror-4597914.html?seite=2.

Jee, C. (2019) "Germany's synagogue shooting was live-streamed on Twitch—but almost no one saw it" [online], MIT Technology Review, https://www.technologyreview.com/2019/10/10/132662/germanys-synagogue-shooting-was-live-streamed-on-twitch-but-almost-no-one-saw-it/.

Kaushik, N. (2021) "Artificial Intelligence in counter-terrorism" [online], The Pioneer, https://www.dailypioneer.com/2021/columnists/artificial-intelligence-in-counter-terrorism.html.

Liu, C. (2018) "How tech firms use China's addiction to 'likes' to teach English" [online], South China Morning Post, https://www.scmp.com/week-asia/business/article/2136105/how-tech-firms-are-using-chinas-social-media-addiction-teach.

Macdonald, S., Correia, S., and Watkin, A. (2019) "Regulating terrorist content on social media: automation and the rule of law", *International Journal of Law in Context*, Vol. 15(2), pp183–197, doi: 10.1017/s1744552319000119.

Miljkovic, M., and Pesic, A. (2019) "Informational and Psychological Aspects of Security Threats in Contemporary Environment", *TEME*, Vol. XLIII(4), pp1079–1094, doi: 10.22190/teme191015064p.

Paganini, P. (2013) "PsyOps and Socialbots" [online], Infosec, https://resources.infosecinstitute.com/topic/psyops-and-socialbots/.

Pashentsev, E. (2020) Malicious Use of Deepfakes and Political Stability, in: Matos, F. (ed.) *Proceedings of the 3rd European Conference on the Impact of Artificial Intelligence and Robotics*, Academic Conferences International Limited, Reading, pp100-107.

Pashentsev, E., and Bazarkina, D. (2022) "The malicious use of artificial intelligence against government and political institutions in the psychological area", in: Bielicki, D. (ed.) *Regulating artificial intelligence in industry*, Routledge, Abingdon and New York, pp37-52.

Rashid, R. (2021) "Improved Technology for Deepfakes Highlights a Supply Chain Problem" [online], IEEE Spectrum, https://spectrum.ieee.org/deepfakes-supply-chain.

Refworld. (2016) "Country Reports on Terrorism 2015 – China (Hong Kong and Macau)" [online], https://www.refworld.org/docid/57518dcb2c.html.

RWTH Aachen University. (2021) "KISTRA – Use of Artificial Intelligence for Early Detection of Crimes" [online], https://www.comm.rwth-aachen.de/cms/COMM/Forschung/Projekte/~jeohm/KISTRA/lidx/1/.

Scally, D. (2017) "German social media law sparks protest" [online], The Irish Times, https://www.irishtimes.com/news/world/europe/german-social-media-law-sparks-protest-1.3159300.

Semaan, N. (2020) "Democratising Deepfakes" [online], Konrad Adenauer Stiftung, https://www.kas.de/en/web/auslandsinformationen/artikel/detail/-/content/democratising-deepfakes.

Stalinsky, S. and Sosnow, R. (2020) "Jihadi Use of Bots on the Encrypted Messaging Platform Telegram" [online], MEMRI, https://www.memri.org/reports/jihadi-use-bots-encrypted-messaging-platform-telegram.

Tagesschau. (2022) "Deutsche neigen zu Technologie-Skepsis" [online], https://www.tagesschau.de/wirtschaft/technologie/technikbegeisterung-deutschland-101.html.

TU Berlin. (2022) "Using AI to Detect Hate Crimes in the Net" [online], https://www.tu.berlin/en/about/profile/press-releases-news/2020/november/hate-crimes-in-the-net/.

UNICRI and Interpol. (2019) *Artificial Intelligence and Robotics for Law Enforcement*, UNICRI and Interpol, Torino – Lyon.

Vacca, J. (ed.) (2021) *Online Terrorist Propaganda, Recruitment, and Radicalization*, CRC Press, Boca Raton.

Weimann, G., and Masri, N. (2020) "Research Note: Spreading Hate on TikTok", *Studies in Conflict and Terrorism*, pp1-14, doi: 10.1080/1057610x.2020.1780027.

Zhao, N., and Zhou, G. (2021) "COVID-19 Stress and Addictive Social Media Use (SMU): Mediating Role of Active Use and Social Media Flow", *Frontiers in Psychiatry*, Vol. 12, doi: 10.3389/fpsyt.2021.635546.

国家互联网信息办公室 文化和旅游部 国家广播电视总局 (State Internet Information Office Ministry of Culture and Tourism State Administration of Radio and Television). (2019) "关于印发《网络音视频信息服务管理规定》的通知 (Notice on Issuing the 'Regulations on the Administration of Network Audio and Video Information Services)" [online], http://www.law-lib.com/law/law_view.asp?id=671676.