Warwick Think Tank Report

>/Tech+Innovation

Technology & Innovation

December 8, 2022

>/ Who(devs our tech matters) : how {we} can make (AI = artificial Intelligence) better x utilising “regulation” + bias removal

(Inter) National Security and Artificial Intelligence in the Face of Contemporary Conflicts

(by) Marta Stankeviciute & Paweł Płonka

Our sponsors:

Institute of Economic Affairs:

Think Tank

The IEA is the UK's original free-market think tank which was founded in 1955. Their aim is to improve the understanding of the fundamental institutions of a free society by analysing the role of free markets in solving economic and social problems. Since their inception, they have worked with prominent Nobel Prize winning economists including Frederich Hayek and Milton Friedman. They have many internship opportunities for both undergraduate and postgraduate students. These include a 3-month general internship, the Epicenter Internship, and the IEA Global internship. They also have a Summer Internship aimed at undergraduate students specifically. Additionally, they are holding an essay competition where students can win a monetary prize for debating whether the current upswing in inflation is transitory or not.

Ekosgen:

Economic Development and Regeneration Consultancy

Based in Sheffield, Manchester and Glasgow, Ekosgen is a consultancy firm focused on economic and social research. From policy development to socioeconomic impact appraisal, Ekosgen works with a variety of clients in public, private and third sectors. One particularly interesting project they have undertaken is the assessment of V&A Dundee Museum on the local economy. Ekosgen has a variety of roles for interested candidates: consultants to work on the core of their projects, and a variety of specialist associates – from urban planners to market researchers, working with their consultants to engender the most appropriate solution. If bringing life back to declining towns seems to be your calling, Ekosgen is the place to be!

Briefing Note

>/Tech+Innovation

>// New line for {briefing note}

This briefing will explore Artificial Intelligence in the context of National and International Security by breaking it into four discrete intersections of AI and Security. Depending on the case, AI takes the form of both a problem and a solution (or even both). We will consider how AI-driven social media creates echo chambers, amplifies polarisation, and causes distrust in democracies; how AI can be used as a counter-terrorism tool and prevention of nuclear trade in the face of growing nuclear threats; uses of AI technologies in border & migration control and the problem of war refugees; existing problems with semi-autonomous weapon systems and growing investment into AI military technologies.

The urgency to discuss the advances, concerns, and application of AI within the domain of National and International Security comes as an epiphenomenon to witnessing the unravelling of a large-scale hybrid war in Europe - the Russo-Ukrainian War. The implications, however, extend beyond Europe as a point of concern for any nation or international alliance. Further, huge military weaponry support from the Western nations for Ukraine (and limited support from the Middle East for Russia) produces an opportunity to witness cutting-edge military technology at play.

Overview://

>/ Section 1 covers the nuances of widely-used modern social media that utilise AI algorithms to tailor the content that the users interact with. While AI algorithms ensure that social media is enjoyable and accessible to anyone, it poses security concerns through the facilitation of false news spread and the creation of echo chambers that aggravate political polarisation. These phenomena not only incidentally play part in growing radicalisation and civil unrest but are strategically employed to manipulate elections and public opinion, as well as weaponised as a hybrid warfare tool.

>/ Section 2 covers the relationship between new technologies used in border management as well as administrative migration processes and the infringement of human rights and freedoms. Although the deployment of Artificial Intelligence can facilitate and accelerate border control, the algorithms still have many flaws. One of the fundamentals is the level of bias that occurs in various systems. The models also often infringe individual's right to privacy and tend to lack accuracy. These issues are reinforced by governments' over-reliance on migration technologies and the opaque cooperation with the private sector which acquires large datasets on citizens.

>/ Section 3 covers the character of modern-day terrorism which moved almost all its preparatory actions to the cyber sphere. That regards not only the recruitment of new and radicalisation of the existing members but also financing and more complex operations. The latter concerns, for instance, nuclear terrorism and compound networks of connections related to illicit materials trafficking and nuclear proliferation. All these activities, however, can be mitigated by the deployment of new technologies and especially artificial intelligence.

>// Section 4 covers the use of semi-autonomous weapons systems in military operations, and how already existing problems could be exacerbated as investments are surging to develop and use fully autonomous AI technologies. Amongst the problems, legislative, compromisation, and use by non-state actors are of worry, along with the "black box" problem associated with deep learning neural networks utilised in fully autonomous systems.

Close [X]

>// WARNING this fact is lacking (context & has been ((independently))checked

+ our algorithm spread it {Anyway}

Privately owned Social Media (is a part of public life)

>// (Unfortunately it spreads of {Information} (true/or/false)

Social Media:

Privately owned social media platforms are a mode of public life in the information age, thus enabling the rapid spread of information such as news, irrespective of its truth value.

- By the start of July 2022, global social media users are estimated to be 59% of the world's total population - 4.7 billion people. This number has grown by more than 5% within a year (by 227 million users) and is projected to continue on increasing.[1]
- The most used social media platforms worldwide are Facebook, Youtube, WhatsApp, Instagram, WeChat and TikTok, all privately owned.[2] American technology conglomerate Meta owns 3 of these platforms - Facebook, WhatsApp, and Instagram.[3]
- A survey of 40 selected countries worldwide (sample size - around 2000 respondents per country) shows that a significant amount (a median of 55.65%) of social media users rely on it as a source of news.[4] However, the statistic varies dramatically between countries - e.g., 38% for the UK, 28% for Japan, and 82% for Kenya.
- Controlling for many factors, false news on Twitter (a trendy platform in the US) were 70% more likely to be retweeted than the truth, especially false political news.[5]

Social media operative mechanisms in part aggravate the prevalence of echo chambers and the subsequent increase in political polarisation.

- "Social media may limit the exposure to diverse perspectives and favour the formation of groups of like-minded users framing and reinforcing a shared narrative, that is, echo chambers." This, however, varies across platforms.[6]
- Exposure to opposing views does not mitigate the effects of echo chambers - experimental data from the US suggests that such exposure reinforces positional differences, and thus can increase polarisation. The effect is particularly substantial amongst conservatives (Republicans), while liberals

[1] Data Reportal, 2022, Digital 2022 July Global Statshot Report

[2] Statista, 2022, Most popular social networks worldwide

[3] Wikipedia, 2021, Meta Platforms

[4] Statista, 2022, Social media as a source of news

[5] Soroush Vosoughi, Deb Roy, and Sinan Aral, 2017, The spread of true and false news online

[6] Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. and Starnini, M., 2021, The echo chamber effect on social media, pg.1

(Democrats) “exhibited slight increases in liberal attitudes after following a conservative Twitter bot, although these effects are not statistically significant.”[7]

- “Political polarisation is on the rise not only in the United States, but also across the world”: systematic review provides evidence that “pro-attitudinal media exacerbates polarisation”, however, research on depolarization via social media is scarce.[8]

- “Opinion dynamics can be manipulated by algorithmic personalisation methods”, experimental data[9] suggests. Social media utilises Big Data and AI in algorithms that tailor the user’s exposure to social media content. The results show that filtering algorithms might exacerbate polarisation even with an equal share of opposing opinions present, but the organisation of social ties is a key factor.

- “The thriving of propaganda, disinformation, and misguided beliefs through echo chambers aggravates violence, poverty and poor health conditions.”[10]

Polarisation, misinformation & disinformation pose threats to democracy, and political regimes are prone to benefit and even weaponize them during times of war.

- Brookings Institution in the US maintains that “one of the drivers of decreased confidence in the political system has been the explosion of misinformation deliberately aimed at disrupting the democratic process” that confuses and overwhelms voters (in the US). They use the example of Russia's cyber-tampering with the 2020 presidential election, when they successfully amplified public distrust in the electoral process (“by denigrating mail-in voting, highlighting alleged irregularities, and accusing the Democratic Party of engaging in voter fraud”). The “big lie” was reinforced by President Trump, having lasting implications on voters’ trust in election outcomes.[11]

- After the 2017 US presidential election, the concept of “alternative facts became synonymous with a willingness to persevere with a particular belief either in complete ignorance of, or with a total disregard for, reality”.[12] The term helps to characterise the “post-truth society” that is overloaded with an abundance of

---

[7] Christopher A. Bail et al., 2018, Exposure to opposing views on social media can increase political polarisation, pg. 9216

[8] Emily Kubin & Christian von Sikorski, 2021, The role of (social) media in political polarization: a systematic review, p.188

[9] Perra, N., & Rocha, L. E., 2019, Modelling opinion dynamics in the age of algorithmic personalisation, pg. 9

[10] Qureshi, I., Bhatt, B., Gupta, S., & Tiwari, A. A., 2020, Causes, Symptoms and Consequences of Social Media Induced Polarization (SMIP)

[11] Gabriel R. Sanchez, Keesha Middlemass, and Aila Rodriguez, 2022, Misinformation is eroding the public’s confidence in democracy

[12] S. I. Strong, 2016, Alternative Facts and the Post-truth Society: Meeting the Challenge

conflicting information: the truth, the lies, the gossip, and the noise, all amplified by social media.

- Systematic overloading of disinformation can be described as “censorship through noise”, Russian media analyst Vasily Gatov describes. It is non-incidental - nowadays, “Russian military theory sees information operations as integral to military operations to an unprecedented extent – Russian state media managers even received military medals for their role during the annexation of Crimea in 2014”, writes journalist Peter Pomerantsev.[13]

[13] Peter Pomerantsev, 2022, Russia’s genocidal propaganda must not be passed off as freedom of speech

Human rights and migration technology:

As cross-border movements of people surge, border control and migration management technology are remaining worryingly biased.

- Migration is a growing overarching phenomenon that leaves an imprint on various other political, social and economic issues of the world, especially the West (European Union, the United States, Canada). In Europe alone, the number of issued first residence permits to non-EU residents increased from nearly 1,5 million in 2011 to 3 million in 2021.[14]
- The number of asylum applications also escalated from ~300,000 in 2012 to over 600,000 in 2019. This is strongly dependent on geopolitical calamities - in 2015 during the peak of the migration crisis, roughly 1,250,000 applications were filled.[15]
- The above statistics show the magnitude of today's cross-border movements. Digitalization of administrative processes has been ubiquitous among a plethora of developed countries. International Organization for Migration identified issues that hinder the efficacy of automatized procedures, one of them being bias in algorithms.[16]
- The most alarming were representation bias (a wrong generalisation of an under-represented portion of the population) and historical bias (misalignment between reality and the encoded values).[17] Scientific researches prove the fallibility of AI in, for instance, facial recognition, as darker-skinned female faces are the most misclassified (with error rates up to 34.7% in comparison to lighter-skinned male faces with error rates up to 0.8%).[18]

Technology related to border surveillance and control has been shown to lack accuracy and violate individuals' right to privacy.

- As more and more refugees cross borders, the need to monitor frontiers has grown. Notwithstanding many benefits coming from the deployment of AI systems such as faster identity verification at control points and recognition of

[14] Eurostat, 2022, Residence permits - statistics on first permits issued during the year

[15] Eurostat, 2022, Asylum applicants by type of applicant, citizenship, age and sex - annual aggregated data

[16] International Organization for Migration, 2022, World Migration Report 2022

[17] International Organization for Migration, 2022, World Migration Report 2022

[18] Joy Buolamwini & Timnit Gebru, 2018, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification

individuals who might threaten national security, they also bring risks to migrants' human rights.

- Arbitrary interference with someone's "privacy, family, home or correspondence" is prohibited by international law.[19] Restrictions on personal privacy might be only introduced when there would be concerns regarding public safety, meeting the objectives of legality, necessity and proportionality.[20]

- Chatham House in its report "Refugee protection in the artificial intelligence era" underscored unlawful interference with privacy as one of the major issues facing the use of technology in border enforcement. The problem becomes more severe considering that more databases become interlinked, hence information extracted by artificial intelligence in one place can be used in another. This interoperability raises questions of "intrusive overreach against privacy standards". Moreover, the fallibility of AI tools might increase the odds of automatic pushbacks (refoulement).[21]

- Such controversial systems have been already deployed, one of the examples being the Greek "Centaur" - cameras, drones, and motion-detection algorithms used to strengthen the state's migration management. Firstly introduced at the refugee camp on Samos, the machinery consists of drone flights over the facilities, perimeter violation alarms with cameras, as well as control gates with metal detectors and x-ray devices. Despite its futuristic form, this system often has been pejoratively associated with a 'panopticon'.[22]

Bearing in mind the opacity of migration technology, governments' over-reliance on them and cooperation with the private sector are emerging as issues of concern.

- New technologies regarding migration governance give governments more leeway to introduce measures that would indirectly give fruition to their preventive policies to the detriment of refugees' human rights. Despite the assuming benevolent intentions of states, it is nonetheless possible that this sector remains mainly unregulated to make migrants more trackable. There are no institutions that could hold the stakeholders accountable.[23]

- Implementation of the newest systems often has traits of experiments with refugee camps being experimental fields. One of the projects that suited this category was Border Ctrl funded by the EU programme Horizon 2020. Its aimed to speed border control for non-EU nationals by using lie detection tests and

[19] United Nations, 1948, Universal Declaration of Human Rights

[20] United Nations, 2014, Resolution adopted by the General Assembly on 18 December 2013 (The right to privacy)

[21] Chatham House, 2022, Refugee protection in the artificial intelligence era

[22] Algorithm Watch, 2021, Greek camps for asylum seekers to introduce partly automated surveillance systems

[23] Tuba Bircan & Emre Eren Korkmaz, 2021, Big data for whose sake? Governing migration through artificial intelligence

facial recognition technology. It faced much criticism on the accounts of classifying its commercial documents as well as high inaccuracy and impression. The project raised concerns amongst many stakeholders over the militarization of the European borders.[24]

- The increasing cooperation between the government and the private sector gives rise to problems with data collection. One of the most recent prominent cases has been the $45,000,00 deal between the World Food Programme and Palantir Technology, the enterprise that contributed to the creation of many detention and deportation programmes run by American federal agencies. The data concerning 92,000,000 recipients has been shared with Palantir. Anything regarding its use remains opaque and unknown.[25]

[24]Migration Policy Institute, 2022, The Increasing Use of Artificial Intelligence in Border Zones Prompts Privacy Questions

[25] Petra Molnar, 2020, Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up

NEW SECTION
“ARTIFICIAL INTELLIGENCE”

>/ AI(in the battle) : against x
International “int” + terrorism

AI in the battle against international terrorism:

Modern-day terrorist recruitment, financing, and operation have become much more internationalized and facilitated, as activities have moved to the cyber sphere, meanwhile Artificial Intelligence is increasingly showing potential to mitigate this.

- In its heyday, there were at least 46,000, and possibly as many as 90,000, accounts that overtly supported the Islamic State on Twitter.[26] Exhibiting the physical facet of this popularity, there were 41,500 international affiliates from 80 countries in the structures of the IS in Iraq and Syria.[27]

- The Internet and social media enabled perilous groups to communicate with the masses as well as recruit online. 83% of extremist offenders were radicalised either fully online or through a combination of both online and offline sources (2015 - 2017), according to the report conducted on behalf of HM Prison and Probation Service.[28] Online radicalization increased to the detriment of offline interactions.

- AI systems that could track extremist online behaviour may efficiently identify suspicious digital activities. However, 50% of all representatives of the global law enforcement agencies stated that the AI expertise within their organisation is “rudimentary”, whilst 30% considered it “intermediary”, and only 20% - “advanced”[29], following insights covered by the United Nations Office of Counter-Terrorism survey.

- Nevertheless, the responsible deployment of this technology is encouraged by major international institutions such as United Nations Interregional Crime and Justice Research Institute and Interpol.[30]

Contemporary terrorists exploit the cyber sphere to enable illicit financing of their actions.

- Over 85% of stakeholders pointed out that anti-money-laundering and countering financing terrorism are the main benefits of the use of new technologies; AI was identified as a technology with the most potential to

[26] Berger JM and Morgan J (2015) The ISIS Twitter Census: Defining and Describing the Population of ISIS Supporters on Twitter. Washington, DC: The Brookings Institution.

[27] International Centre for the Study of Radicalisation, 2018, From Daesh to ‘Diaspora’: Tracing the Women and Minors of Islamic State

[28] HM Prison and Probation Service, 2021, Exploring the role of the Internet in radicalisation and offending of convicted extremists

[29] United Nations Office of Counter-Terrorism & United Nations Interregional Crime and Justice Research Institute, 2021, Countering Terrorism Online with Artificial Intelligence. An Overview for Law Enforcement and Counter-Terrorism Agencies in South Asia and South-East Asia

[30] International Criminal Police Organization & United Nations Interregional Crime and Justice Research Institute, 2020, Towards Responsible Artificial Intelligence Innovation

facilitate it, according to research conducted by the Financial Action Task Force.[31]

- The pilot program of the United Overseas Bank's 'Anti-Money Laundering Suite' resulted in an increase of 5% in true positives (successful predictions) and a reduction of 40% in false positives (unsuccessful predictions) in transaction monitoring.[32]

- The same program led to great enhancements in name screening; The outcomes were a 60% reduction in false positives in individual names and a 50% reduction in false positives for corporate names.[33]

The illicit trafficking of radioactive and nuclear materials as well as nuclear proliferation by non-state actors remain a fundamental threat in today's political climate.

- Terrorists would need to acquire 25kg of highly enriched uranium (HEU) to make an improvised nuclear device.[34] Currently, there is 1,255,000kg of this fissile material in the world.[35]

- Places, where nuclear weapons and materials are stockpiled, are not always properly safeguarded and there have been incidents of different groups breaking into nuclear facilities. In 2006 for example, Oleg Khinsagov was arrested along with three other Georgian accomplices with almost 80 grams of HEU.[36] In 2007, burglars infiltrated a Southern African nuclear research facility but escaped without any illicit materials.[37] As of 2010, the International Atomic Energy Agency registered 18 incidents of theft or loss of Highly Enriched Uranium or separated plutonium.[38]

- The risks of nuclear terrorism have increased, as pointed out by professor Graham Allison for National Defense University, due to 'Pakistan's growing nuclear arsenal and development of tactical nukes', 'potential for large-scale reprocessing of plutonium in China and Japan' and the 'inexorable advance of science and technology, diffusion of nuclear know-how'.[39] The Russian

---

[31] FATF, 2021, Opportunities and Challenges of New Technologies for AML/CFT

[32] Deloitte, 2018, The case for artificial intelligence in combating money laundering and terrorist financing

[33] Deloitte, 2018, The case for artificial intelligence in combating money laundering and terrorist financing

[34] Belfer Center for Science and International Affairs, Harvard Kennedy School, 2010, Nuclear Terrorism Fact Sheet

[35] International Panel on Fissile Materials, 2022, Fissile material stocks

[36] Belfer Center for Science and International Affairs, Harvard Kennedy School, 2010, Nuclear Terrorism Fact Sheet

[37] New York Times, 2007, Break-In at Nuclear Site Baffles South Africa?

[38] Belfer Center for Science and International Affairs, Harvard Kennedy School, 2010, All Stocks of Weapons-Usable Nuclear Materials Worldwide Must be Protected Against Global Terrorist Threats

[39] PRISM, 2018, Nuclear Terrorism: Did We Beat the Odds or Change Them?

invasions of Ukraine, first in 2014 and the full-scale one in 2022 have so far been fought with conventional weapons, yet the war has clear nuclear undertones. It raises “questions about the dynamics of nuclear deterrence, the future of nuclear nonproliferation, arms control and disarmament, and the international governance of nuclear energy. The ongoing war in Ukraine has profound implications for the global nuclear order.”[40]

- Machine learning implementation on list-based screening of shipments resulted in capturing 12,000 additional shipments and doubled the number of flagged shipments by known entities. The models saved the analysts 200 hours[41], according to the report prepared by C4ADS and Nuclear Threat Initiative (NTI).
- An unsupervised deep learning model that C4ADS and NTI launched analysed 4,300,000 rows of trade and identified 50 new leads for analyst review. 4 of the recognized companies that could take part in the illicit trade of nuclear materials were actioned by the U.S. government throughout research.[42]

[40] Mariana Budjeryn, 2022, Distressing a system in distress: Global nuclear order and Russia’s war against Ukraine

[41] C4ADS & NTI, 2021, Signals in the Noise: Preventing Nuclear Proliferation with Machine Learning & Publicly Available Information

[42] C4ADS & NTI, 2021, Signals in the Noise: Preventing Nuclear Proliferation with Machine Learning & Publicly Available Information

NEW SECTION
“ARTIFICIAL INTELLIGENCE”

>/ AI in:(military) : operations

AI in Military Operations:

Use of semi-autonomous weapons systems in warfare poses legislative and compromisation problems that fully autonomous device uses would likely exacerbate.

- The use of armed drones is not specifically regulated under international law. They fall under the legislation of the Geneva Convention. Their use must follow the principles of distinction (only engaging with military targets) and proportionality (no excessive collateral damage). However, "the use of both military grade and weaponized civilian drones poses challenges to implementation of these rules." While military drones can be highly precise, the accuracy "does not mean that the target was correctly identified as a military objective". Further, as drones only deliver a warhead, they can "cause excessive harm to civilians and civilian infrastructure, depending, amongst others, on the warhead used."[43]
- Between 2010 and 2020, killer drone strikes by the US in Pakistan, Afghanistan, Yemen and Somalia were estimated to have resulted in a reported 8,858 to 16,901 total individuals killed, including 910-2,200 civilians and 283-454 children.[44]
- Unmanned Aerial Vehicles, commonly used in the military for enemy surveillance and killer strikes have security vulnerabilities. "Even professional UAVs, when compromised, can be used by criminals and terrorist organizations for illegal surveillance and unmanned attacks. They may be turned off remotely, hijacked, flown away or stolen."[45]
- "Non-State actors have already begun using drones on the battlefield. The Islamic State Group (IS) has used drones prolifically." In Mosul, IS has flown 300 drone missions in one month — of which one-third were armed strike missions. Whereas a U.S. armed drone would cost $22,000, IS was using small quadcopters for $650.25 Even in the semi-autonomous mode, the small commercial systems provide intelligence, surveillance, and reconnaissance for the nonstate group."[46]

Investment into R&D of AI-powered robots and other AI systems is surging.

- Various unmanned devices have been used for years, ranging from landmines to military robotics. However, the demand for incorporating AI, particularly machine learning, to control such weaponry, is growing.[47]

---

[43] Geneva Call, 2020, Humanitarian Concerns raised by the Use of Armed Drones

[44] The Bureau of Investigative Journalism, 2020, Drone warfare

[45] Vinay Chamola, Pavan Kotesh, Aayush Agarwal, Naren, Navneet Gupta, and Mohsen Guizanic, 2021, A Comprehensive Review of Unmanned Aerial Vehicle Attacks and Neutralization Techniques

[46] Sarah Kreps, The Brookings Institution, 2021, Democratizing harm: Artificial intelligence in the hands of nonstate actors

[47] International Committee of the Red Cross, 2022, What you need to know about autonomous weapons

- The applications of AI in the military are broad, spanning cybersecurity, warfare systems deployed on air, sea, land and space platforms, logistics, threat monitoring & situational awareness.[48]
- The global AI in military market size reached USD 6.5 billion in 2020 and is projected to grow by 13.4% before 2028.[49]
- Upon the Russo-Ukrainian war unfolding, NATO has launched an Innovation fund[50] that will invest 1 billion euros in start-ups and venture capital funds "developing dual-use emerging technologies" of priority to NATO. These include artificial intelligence, big-data processing, autonomy and automation - technologies critical to the security of the alliance.
- Individual militaries have ramped up their investment into military AI applications as well. Germany has earmarked just under half a billion euros for research and artificial intelligence in the military; "The Chinese military likely spends at least $1.6 billion a year on AI"; "The US Department of Defense requested $874 million for artificial intelligence for 2022, although that figure does not reflect the total of the department's AI investments, it said in a March 2022 report."[51]
- The desirable autonomy and cooperative engagement in military robotics is "not possible without integrating artificial intelligence (AI) elements such as machine vision, image recognition and natural language processing into the systems' command and control." Thus, Russia's "president Vladimir Putin identified the development of weapons with elements of AI as one of the defence ministry's five major priorities for the near future."[52]

[48] Dataconomy, 2022, Guns and Codes: the Era of AI-wars begins

[49] Emergen Research, 2020, Artificial Intelligence in Military Market

[50] NATO, 2022, NATO launches Innovation Fund

[51] MIT Technology Review, 2022, Why business is booming for military AI startups

[52] Chatham House, 2021, Advanced military technology in Russia, p. 50

YOU’RE LOOKING A
WEAPON OF THE PAST…

>/ Today with AI in:
(military) :
operations armies just
use drones to get the
Job {DONE}

Lethal Autonomous Weapons Systems (LAWS), also known as killer robots, may be finding their way into military arsenals along with human-operated drones, despite serious concerns.

- LAWS are weapons that use Artificial Intelligence (AI) to identify, select, and kill human targets without human intervention.[53]
- "A United Nations report suggested that a drone, used against militia fighters in Libya's civil war, may have selected a target autonomously.[54]" The report, however, does not comment on whether the LAWS have produced casualties, yet it is evident that it has made its military debut, raising concerns about the dangers of autonomy of such devices.
- When it comes to advanced deep learning algorithms, the black-box problem is encountered: one cannot "just look inside a deep neural network to see how it works." Its reasoning and decision-making are a result of "thousands of simulated neurons, arranged into dozens or even hundreds of intricately interconnected layers".[55] Thus, if completely autonomous weaponry is utilised, its decision-making in target choosing may not be understood, posing problems of security, accountability, and enforcement of warfare law.
- "The International Committee of the Red Cross and several NGOs had been pushing for an international treaty that would establish legally-binding new rules on the machine-operated weapons." However, the U.N. talks in 2021 have collapsed without a deal, with Russia, the USA, and India being amongst the countries expressing doubts about the need for the new LAWS treaty.[56]

[53] Future of Life Institute, 2021, Lethal Autonomous Weapons Systems

[54] The New York Times, 2021, A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says

[55] MIT Technology Review, 2017, The Dark Secret at the Heart of AI

[56] Reuters, 2021, U.N. talks adjourn without deal to regulate 'killer robots'

Policy Recommendations

>/AI+Security

//: Overview

The rapid and robust development of new technologies changed the character of the threats posed to national and international security. Combating these challenges and reinforcing the security infrastructures require the policymakers to consider using new approaches and mechanisms such as artificial intelligence, as well as legislative improvements for the highly advanced and specific applications of artificial intelligence itself.

The deployment of these measures will be examined in three areas: social media and disinformation; subsequently, migration management and border control; finally, counter-terrorism and preventing preparatory activities in the cybersphere of terrorist groups. In all these themes the common strand emerges - human rights and freedoms are vulnerable in face of implementation means that exploit artificial intelligence. This creates a huge obstacle for policymakers - to establish a balance between effectiveness and ethics, between technology and human factors, and between Artificial Intelligence causing harm and helping.

>/ Social Media:(s-o-c-i-a-l) : {media}

Social media:

While social media is synonymous with convenience, its operation aggravates existing shortcomings of human cognition and can be strategically utilised to manipulate public opinion and share disinformation, but policy and technical changes can help mitigate this.

Social media by itself is a great extension of human social life. "In a holistic sense, the popularity of social media has been driven by how user-friendly and interactive it has made modern cyberspace"[57] - arguably due to the use of complex AI algorithms that tailor the shown content to each user.

However, it can be an addictive good, where users to some extent fail to control their involvement.[58] An empirical study of 590 daily users has identified three potent factors why individuals fail to self-control their social media use: habitual checking of social media, strongly experienced online ubiquity of social media, and strong disturbances from social media notifications.[59]

While these factors are somewhat alarming at the face value, their prevalence becomes a concern when social media is redefined as not only individual social space but a mode of political and civil life as well. People rely on social media to access news; political campaigns are run on social media platforms; posts go viral; discussions take place, sometimes between (semi)anonymous agents, some of whom may be trolls.[60] In particular, there are accounts of Russia extensively using troll factories as a part of its disinformation campaign.[61]

Worryingly, the operation of social networks is prone to "limit the exposure to diverse perspectives and favour the formation of groups of like-minded users framing and reinforcing a shared narrative, that is, echo chambers, though this varies across platforms"[62] and countries (longitudinal data from Sweden indicates that media has no significant effect on polarisation in the area).[63] Thus, while an epistemic bubble where relevant opinions may be accidentally left out can be burst by exposure to facts and alternative views, echo chambers are more sturdy and self-reinforcing.[64] Hence, trolling and other disinformation campaigns may be highly destructive in their

---

[57] The British Psychological Society, 2015, Why do we 'like' social media?

[58] Trevor Hyanes, Harvard University Blog, 2018, Dopamine, Smartphones & You: A battle for your time

[59] Jie Du, Peter Kerkhof, Guido M van Koningsbruggen, 2019, Predictors of Social Media Self-Control Failure: Immediate Gratifications, Habitual Checking, Ubiquity, and Notifications

[60] Centre for Strategy & Evaluation Services, 2019, Rapid Evidence Assessment (REA): The Prevalence and Impact of Online Trolling

[61] GOV.UK, 2022, UK exposes sick Russian troll factory plaguing social media with Kremlin propaganda

[62] Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. and Starnini, M., 2021, The echo chamber effect on social media, pg.1

[63] University of Gothenburg, 2021, Few indications that the new media landscape leads to increasing polarisation in society

[64] C. Thi Nguyen, 2018, Echo Chambers and Epistemic Bubbles

reinforcement of lies and noise in efforts to manipulate public opinions, posing security threats.

The scale of political disinformation campaigns cannot be understated. The Oxford Internet Institute's survey has found organised social media manipulation campaigns in each of the 81 surveyed countries. "Governments, public relations firms and political parties are producing misinformation on an industrial scale, according to the report. It shows disinformation has become a common strategy, with more than 93% of the countries (76 out of 81) seeing disinformation deployed as part of political communication."[65]

Evidently, social media has made modern propaganda through disinformation campaigns highly attainable and widespread. The combination of human factors such as the prevalence to form epistemic bubbles and echo chambers, as well as the ubiquity of social media has created a breeding ground for extensive politically-incentivized media manipulation. While conserving the freedom of speech within social media is of critical importance, these issues cannot be ignored.

Mitigating them, however, poses challenges. The broad worldwide usage somewhat established the status of social media as a common good, yet all the popular platforms are privately owned by corporations or conglomerates such as Meta. The most recent reminder of this - the $44 billion dollar acquisition of Twitter by Elon Musk.[66] Musk acquired Twitter with a promise to conserve free speech, being "a vocal and longstanding critic of Twitter's moderation and suspension policies."[67] Twitter has previously banned accounts of controversial, mostly right-wing figures, amongst whom - former president of the US Donald Trump, banned for " inciting and glorifying violence surrounding the 2020 election and the Jan. 6 insurrection", and rapper Kanye West, banned for antisemitic remarks.[68] Such figures, under Musk's rule, were supposed to get 'amnesty' and be reinstated.

However, Musk's insistence on turning Twitter into "a haven of free speech" did not go unnoticed. Both the European Commission and the US Treasury have threatened Musk with grave consequences. The European Commission informed Musk that Twitter will be banned in the EU unless it abides by strict content moderation rules", delineated in the EU's new Digital Services Act, "including ditching an "arbitrary" approach to reinstating banned users, pursuing disinformation "aggressively" and agreeing to an "extensive independent audit" of the platform by next year".[69] The US Treasury Secretary Janet Yellen indicated that Washington was reviewing his purchase of the social network, voicing concerns about foreign investments that may create a national security risk.[70]

---

[65] University of Oxford, 2021, Social media manipulation by political actors an industrial scale problem - Oxford report

[66] BBC News, 2022, Elon Musk takes control of Twitter in $44bn deal

[67] Forbes, 2022, Elon Musk Is Restoring Banned Twitter Accounts—Here's Why The Most Controversial Users Were Removed And Who's Already Back

[68] ibid.

[69] Financial Times, 2022, EU and US turn up the heat on Elon Musk over Twitter

[70] ibid.

From the case study of Twitter, it is evident that public policies such as the new EU's Digital Services Act[71] may be effective in regulating the usage and operation of social media. However, the demarcation line between the private and the public still remains unclear. With various policies coming to place, a huge part of corporate accountability still relies on self-governance of questionable efficiency,[72] as well as the goodwill of the private business owners to obey.

Interestingly, while AI algorithms seem to worsen the aforementioned *status quo*, they could as well be amongst the technical changes that could improve it. "As language-processing technology develops, technology will help us identify and remove bad actors, harassment, and trolls from accredited public discourse", comments Galen Hunt, a research manager at Microsoft Research NExT.[73] "Trolling, doxxing, echo chambers, click-bait, and other problems can be solved", adds David Krager, a professor of computer science at MIT.[74] Advanced "multimodal sentimental analysis provides methods to carry out opinion analysis based on the combination of video, audio, and text[75]", significantly improving AI's ability to detect fake news and disinformation.

[71] European Commission, 2022, Digital Services Act

[72] The Guardian, 2021, The Guardian view on regulating social media: necessary but risky

[73] Pew Research Centre, 2017, The Future of Free Speech, Trolls, Anonymity and Fake News Online

[74] Ibid.

[75] Ganesh Chandrasekaran, Tu N. Nguyen, Jude Hemanth D, 2021, Multimodal sentimental analysis for social media applications: A comprehensive review

Human Rights and migration technology:

In today's climate, the line between the use of new technologies in the management of migration and their negative impact on human rights and freedoms has narrowed.

In many regions of the world, migration is becoming an increasingly central issue in both domestic and international affairs. The European Union is an exemplar of this trend - nearly 2,000,000 people immigrated to the Union in 2020.[76] These numbers skyrocketed after the outbreak of the Russo-Ukrainian War in 2022. 4.5 million people registered for Temporary Protection and 12.7 million entries into the EU from Moldova and Ukraine were registered.[77] The future remains unpredictable, but, certainly, migrations will not cease to exist. Indeed, the influx of migrants is projected to grow, having an influential impact on demography.[78] Digitalization of border management appears to be inevitable. Currently, only six Member States deploy artificial intelligence as migration technology. It is used across various stages of entering a state to, *inter alia,* identify language, detect identity fraud, assess the complexity of an application, and interact with clients. Many countries have robust plans for the development and integration of artificial intelligence with migration systems.[79] As such, the vulnerabilities of this new technology await recognition as well. AI is *not* flawless and may pose a risk to human rights. One of the key reasons in this regard is the bias exhibited by many algorithms. Research conducted by the National Institute of Standards and Technology (U.S. Department of Commerce) tested 189 facial recognition algorithms with the result of most of them being biased and leaning towards more false positives.[80] On this account, the main problem does not lie in the use of technology but with the training datasets and the programme itself.[81]

New migration management technologies may also lead to an increase in privacy infringements. Crossing a border requires providing biometric data which is very often later collected and stored without the consent of individuals. Moreover, owing to the interoperability of the systems, data became more fluid and information is based on other sets. A centralised collection of immense amounts of information on millions of people is also an attractive target for hackers.[82] Another dire facet of using new

[76] European Commission, 2021, Statistics on migration to Europe

[77] European Commission, 2022, Migration management: Welcoming refugees from Ukraine

[78] European Commission, 2019, Demographic Scenarios for the EU - Migration, Population and Education

[79] European Migration Network & Organisation for Economic Co-operation and Development, 2022, The use of digitalisation and artificial intelligence in migration management

[80] National Institute of Standards and Technology, 2019, Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects

[81] European Parliamentary Research Service, 2021, Artificial intelligence at EU borders. Overview of applications and key issues

[82] Lucia Nalbandian, 2022, An eye for an 'I:' a critical assessment of artificial intelligence tools in migration and asylum management

technologies in border control is monitoring and detecting the flows of migrants. Autonomous border surveillance systems can improve present governance of the resources on the frontiers, although AI is also able to make profound mistakes resulting in innocent people losing lives. Such measures have been used, for instance, in the United States. The concept of a "smart border" comprises surveillance towers and drones using AI to 'patrol' the terrain. The project thereby contributed to doubling migrant deaths as well as reorientated migrant routes to more perilous areas, endangering even more lives.[83] Similar initiatives are planned to be established in the EU. For instance, ROBORDER aims to use unmanned mobile robots to provide "early identification of criminal activities". FOLDOUT is another project with the same goal of "through-foliage" surveillance. Nonetheless, the programmes still require to be tested and refined to the point they no longer imperil migrants' human rights. Even though border surveillance technologies are prone to elevate the likelihood of violation of human rights, they were often classified as "low risk" in regulations like the AI Act proposal.[84]

Many controversies also arise on the grounds of transparency of process-relevant information, the cooperation between the private and the public sectors, and the over-reliance on technology. All these aspects are connected as transparency might be aggravated even more when external suppliers produce the necessary components. How a certain technology operates is usually opaque and there is little way to hold the authorities accountable for that. Sometimes it is even impossible; UNHCR provides scant information on the workings of the biometric identity management system (BISM) and has no citizenry to answer to.[85] The systems might be also used maliciously to steer policies toward more anti-immigration courses to fly over the radar of non-governmental organisations and other stakeholders. This was done, for instance, in the United States where the Risk Classification Assessment, a computational tool used to determine an immigrant's danger to society, was modified in an aforementioned way. The system automatically recommended to the Immigration and Customs Enforcement officials either "release" or "detain", however, in 2017 the "release" option was removed. The triple increase in the number of immigrants with no criminal history who got detained was caused in part by these minor unnoticeable changes.[86]

---

[83] Petra Molnar, 2020, Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up

[84] Statewatch, 2022, A clear and present danger Missing safeguards on migration and asylum in the EU's AI Act

[85]Lucia Nalbandian, 2022, An eye for an 'I:' a critical assessment of artificial intelligence tools in migration and asylum management

[86] Reuters Investigates, 2018, Trump's catch-and-detain policy snares many who call the U.S. home

>/ transnational : terrorism : {AND} Counter(terrorism)

Transnational terrorism and counter-terrorism:

The changing character of modern-day terrorism (which uses new means to finance its actions and threaten society) might be alleviated by deploying innovative technologies such as artificial intelligence systems to counter suspicious activities.

The advent of the Internet and the Atomic age has advanced our society across many dimensions. However, it has also made our world more perilous. Present terrorism differs from its past forms by exploiting the above measures. The operations have become more complex as well as covert. The spadework (propaganda, recruitment, financing, execution, etc.) has moved to the cyber sphere, whilst the building of extensive networks for the trading of illicit materials has increased the feasibility of obtaining weapons of mass destruction. Notwithstanding the decreasing numbers of successful attacks and arrests, the menace of terrorism is continually up-to-date, with the likelihood of an escalation in the next years. Considering recent and current global calamities such as the cost of living crisis and the pandemic of Covid-19, the law enforcement agencies such as Europol underscore the potential for the rise of politically-driven offences.[87] Recently, this category of terrorism has undergone an insurgence in the West - politically motivated terrorism overtook religiously motivated terrorism which declined by 82% in 2021.[88] Worldwide, the epicentre of terrorism has moved to Africa, especially to the region of Sahel. Another pivot lies in the Russo-Ukrainian war which is also expected to negatively influence the spread of violence, mainly concerning cyberattacks.[89] The intensifying terrorism in Africa and Eastern Europe poses risk to the whole international community.[90]

Owing to the globalization of terrorism, many of its preparatory activities can now be found in the online environment. In 2014, Simon Wiesenthal Centre identified more than 30,000 websites, forums, and social media users that promote hate and terrorism worldwide.[91] Social media has become almost a ubiquitous method to disseminate propaganda and radicalise individuals. Participating in internet communities and different social platforms accounted for the "radicalization and mobilisation processes of 88.23% of the lone actors and 76% of individuals who were members of extremist groups or radical cliques (i.e. non-lone actors)".[92] The promotion of extremist rhetoric and violence also resonates with international and domestic clandestine recruitment. Through such means, the capabilities of various organisations are extensive; consider the 41,500 international affiliates from 80 countries who were part of the Islamic State

[87] Europol, 2021, European Union Terrorism Situation and Trend report 2022 (TE-SAT)

[88] Institute for Strategic Dialogue, 2022, Global Terrorism Index 2022: Key findings in 6 Charts

[89] Institute for Strategic Dialogue, 2022, Global Terrorism Index 2022: Key findings in 6 Charts

[90] United Nations News, 2022, Terrorism intensifying across Africa, exploiting instability and conflict

[91] Simon Wiesenthal Center, 2014, District Attorney Vance and Rabbi Abraham Cooper Announce the Simon Wiesenthal Center's Report on Digital Terrorism and Hate

[92] The National Consortium for the Study of Terrorism and Responses to Terrorism, 2018, The Use of Social Media by United States Extremists

in Iraq and Syria.[93] Terrorists can therefore captivate anyone's mind in the world, especially minors. By spreading propaganda via email, chat rooms, e-groups, message boards, and even cartoons, music videos and computer games, they can reach different social groups (usually marginalised and vulnerable), whose ideological beliefs are reinforced or attract young Internet users. The first set of people is often allured by content tailored to them through so-called "targeted advertising". The second set is usually subject to a technique of "grooming", which means "learning about the individual's interests in order to tailor the approach and build up a relationship of trust".[94] After this initial stage of sparking interest, potential accomplices are recruited and trained via password-protected websites and other online platforms.[95] Further instrumentality depends on the profile of an organisation, however, preparing for an attack has been the most common use. It is worth noting that extreme-right-wing offenders are over four times more likely to search for this information compared to their Islamist counterparts.[96] Preparation might include activities such as seeking instructions on making improvised explosive devices, collecting data on potential targets using publicly available information, and others. The final stage is the execution. It has not been uncommon for terrorists to Livestream attacks, and then inspire copycat acts. This was the case with, for instance, the Christchurch mosque shootings in New Zealand in 2019 during which the perpetrator live-streamed his deeds on Facebook, motivating similar assaults in El Paso, Poway, Baerum, Oslo, and Halle.[97]

In order to pursue their goals, terrorist organizations seek different measures to gather necessary funds. Simultaneously, they must conceal their sources to avoid exposure. The Internet developed to be the perfect space for these actions, as it acts to minimise the risk of capture as well as to increase the efficiency of collecting money. Smooth financing of terrorism is simplified by the use of services such as PayPal or donation platforms. Fraudulent measures are also deployed. The most significant reported case recently was the money laundering of profits from stolen credit cards by Younis Tsouli (UK). The fact that approximately 1,600,000 pounds of illicit funds were transferred to support terrorist groups by using 1,400 credit cards shows the magnitude of an individual contribution.[98] The issue is therefore intertwined with a broader problem of money laundering. UNODC estimated that 2-5% of global GDP goes through a laundering cycle per annum.[99] Major financial institutions continue to struggle with combating the concealment of illegal funds. Traditional technological approaches seem to be unsuccessful as many results are incorrect - some organisations' analyses

---

[93] International Centre for the Study of Radicalisation, 2018, From Daesh to 'Diaspora': Tracing the Women and Minors of Islamic State

[94] United Nations Office on Drugs and Crime, 2017, Handbook on Children Recruited and Exploited by Terrorist and Violent Extremist Groups: The Role of the Justice System

[95] United Nations Office on Drugs and Crime, 2012, The use of the Internet for terrorist purposes

[96] Paul Gill, Emily Corner, Maura Conway, Amy Thornton, Mia Bloom, John Horgan, 2017, Terrorist Use of the Internet by the Numbers

[97] The Institute for Economics & Peace, 2022, Global Terrorism Index 2022

[98] United Nations Office on Drugs and Crime, 2012, The use of the Internet for terrorist purposes

[99] United Nations Office on Drugs and Crime, Money Laundering, Proceeds of Crime and the Financing of Terrorism

produced 95% of false positives.[100] At the same time, terrorists find more and more adroit ways to manage their clandestine accounts. Financial technologies such as cryptocurrencies gain traction owing to their anonymity. Currently, the most popular cryptocurrency among extremists is Bitcoin, however, they also use less common ones such as Monero. One of the practices used to increase the operational security of organisations is creating new wallets for every crypto transaction which was introduced, for instance, by Hamas.[101] Another malicious scheme that may become more common in the future is the use of audio deepfakes to not only raise funds but also obtain sensitive and secret information.[102] The need for cooperation and coordination of proper anti-money laundering and counter-financing terrorism mechanisms is stressed by the key stakeholders, e.g. the World Bank.[103]

The Internet is not the only measure that is used to elevate terror. Other more ghastly and deathly means are chemical, biological, radiological, and nuclear devices which attract the attention of militant groups. Nuclear terrorism brings the most apocalyptic visions. Weapons of such kind can be acquired by three different pathways: transfer, leakage, and indigenous production. The first two require buying or stealing a bomb - the third, much more dangerous, concerns manufacturing nuclear devices without the assistance of a state.[104] Non-state actors right now have the capabilities to acquire all necessary materials to create at least a 'crude' nuclear device according to national security agencies.[105] Terrorists would need to acquire 25kg of highly enriched uranium to make an improvised nuclear device.[106] Currently, there is 1,255,000kg of this fissile material in the world.[107] Data from the IAEA Incident and Trafficking Database show that 14% of all incidents regarding the possession or trafficking of radioactive materials involved nuclear material between 1993-2021.[108] The international trade in illicit nuclear goods appears to be the main issue. Trade networks specialising in nuclear-related materials existed in the past. The most prominent one was the network established and coordinated by Pakistani scientist Abdul Qadeer Khan. His network facilitated Pakistan to become a nuclear state as well as supplied Iran, Libya, and North Korea which then sold missile technology to Syria and Myanmar. The network was unravelled in 2004, however, its remnants and components most likely exist to this day.[109]

---

[100] Deloitte, 2018, The case for artificial intelligence in combating money laundering and terrorist financing

[101] Royal United Services Institute for Defence and Security Studies, 2020, New Technologies but Old Methods in Terrorism Financing

[102] United Nations Office of Counter-Terrorism & United Nations Interregional Crime and Justice Research Institute, 2021, Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes

[103] World Bank Group, 2022, Preventing Money Laundering and Terrorist Financing

[104] Wilson Center, Robert S. Litwak, 2016, Deterring Nuclear Terrorism

[105] The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction, 2005, Report to the President of the United States March 31, 2005

[106] Belfer Center for Science and International Affairs, Harvard Kennedy School, 2010, Nuclear Terrorism Fact Sheet

[107] International Panel on Fissile Materials, 2022, Fissile material stocks

[108] IAEA, 2022, IAEA Incident and Trafficking Database (ITDB) 2022 Factsheet

[109] Andrew Futter, 2015, The Politics of Nuclear Weapons

>/ the Internet : and all its {many} spaces created + the perfect outlet for terrorists to gather //: funds
>/ While also ((concealing)) their : identities

NEW LINE :/>
They can do all this :While evading the >RISK< of capture

Welcome to the wild west of the Internet: where terrorism is >>transnational

Policy Recommendations
>/AI+Security

>// Overview

The intersection of AI and Security is a particularly salient one, urged by recent global unrest and seemingly endless technological advancements. The unstable nature of the contemporary world and the rapid progress of Artificial Intelligence and associated technologies urge policymakers to adjust just as rapidly. Such adjustments are non-trivial and multifaceted, establishing the need for multi-agent international involvement. We thus propose 3 areas of policy recommendations:

>// Action 1 - regulation of social media is critical to ensure national and international security where social networks uphold the freedom of expression but are harsh to disinformation campaigns, trolls, and misinformation.

//> Action 2 - The adequate international regulatory framework for developing and deployment of AI to counter the spread of terrorist activities online (propaganda, radicalisation, financing) must be established.

>//Action 3 - A lot more restraining regulations must be introduced regarding new technologies dedicated to the management of migration and border control; impartial and independent bodies should watch the use of these technologies.

>/AI+Security

Regulation of secure and accountable online environments should employ legislation, technical implementations and education.

Regulation of secure and accountable online environments should employ legislation, technical implementations and education.

Building upon The Brookings Institution Senior Fellow Mark MacCarthy's argument, the first and foremost step in social media and regulation is transparency.[110] While (by themselves) transparency legislations are not self-enforcing, "a dedicated regulatory agency must define and implement them through rulemaking and must have full enforcement powers, including the ability to fine and issue injunctions."[111] Without transparency as the first building block, "no other regulatory measures will be effective".[112]

In this regard, the new EU Digital Services Act serves as an example of a legislative international basis for transparency. Whether it will be effective in combating disinformation, fake news and polarisation is, however, contingent on the regulatory mechanisms in place. Nonetheless, in the discussed case of Musk's acquisition of Twitter, stark warnings by the EU Commission have seemingly worked in upholding the social media platform to the EU block's standards, but its continuous reinforcement remains to be seen.

Nonetheless, one principal problem that has been encountered in the past is the efficiency of fines brought upon online platforms. While fines seem to be a straightforward measure to incentivize private corporations to obey, the case of GDPR violation fines suggests that the amounts may sometimes be insufficient to ensure the cooperation of the private platforms, or that, perhaps, they are not effective altogether. An important feature of corporate fining is that the fine "primarily affects shareholders, not necessarily the individuals who committed or even benefited from the crime." While the company's stock falls after incurring the fine, it bounces back, usually - "shareholders might not demand an appropriate reduction in activity levels, nor the right amount of firm-wide monitoring, to avoid future instances of crime".[113]

Amongst the largest fines receivers, 3 out of the top 5 agents are Meta-owned companies: in September 2021, WhatsApp was fined USD 223m; in December 2021 - Facebook received a fine of USD 60m; in September 2022 - a hefty fine of USD 402m for Instagram.[114] While Meta intends to appeal the Instagram fine and asserts that they have already fixed the underlying issues for which the fine was issued[115], the whole Meta conglomerate has accrued yet another hefty fine of USD 275m in

[110] Mark MacCarthy, The Brookings Institution, 2022, Transparency is essential for effective social media regulation

[111] ibid.

[112] ibid.

[113] Dorothy S. Lund and Natasha Sarin, 2020, The Cost of Doing Business

[114] Silicon Republic, 2022, How the €405m Instagram fine compares to other GDPR penalties

[115] ibid.

November 2022.[116] Some consider "whether such fines actually influence corporate behaviour or if some companies simply see them as an added cost of doing business".[117]

Thus, not only international legislative basis for transparency must be ensured, but effective enforcement mechanisms as well, which is not a trivial task. Following the EU Commission's warnings to Twitter, an appropriate enforcement mechanism should be multifaceted, e.g., a combination of a monetary punishment, the suspension of activities, and independent auditing in the future to ensure that proper changes have been made.

The second course of action involves global investments into technical implementations improving "the ability of AI models to recognise contextual nuance in social media discourse and adapt more rapidly to recognise novel pieces of disinformation".[118] Provided that through appropriate systems of reinforcement, private social media corporations are keen on avoiding violations, this creates a window of opportunity for policymakers, scientists, and corporate agents to collaborate on these advancements, with all allocating resources to these developments and implementations, as well as involving independent technology developers.[119] To date, investment in such technologies is scarce.[120]

Third, the alliances must as well consistently invest in the education of the public, improving media literacy and critical thinking. Efforts are being made to include these topics in the schools' curriculum. The European Commission has released guidelines for teachers and educators for tackling these topics in the classroom as a deliverable part of the Digital Education Action Plan.[121] In the UK, "elderly, disabled and other vulnerable people will get better support to stay safe online and avoid being misled by disinformation" utilising the funding boost from UNESCO.[122] However, "educational interventions have significant limitations: chiefly, they require individuals who are motivated to seek and voluntarily engage them"[123], but these interventions could aid the overall effectiveness of proposed policies.

Additionally, in devising these policies, a set of fundamental norms must be followed to ensure unified and consistent application. Such principles would for example include "enabling fair and equal access; avoiding obvious falsehoods; offering and engaging with reason; supporting epistemic respite (allowing "time-outs" for individuals to

---

[116] BBC News, 2022, Facebook: Meta fined €265m by Irish Data Protection Commission

[117] The Guardian, 2022, Meta fined €265m over data protection breach that hit more than 500m users

[118] Linda Slapakova, Rand Corporation, 2021, Towards an AI-Based Counter-Disinformation Framework

[119] ibid.

[120] The German Marshall Fund, 2022, AI Startups and the Fight Against Mis/Disinformation Online

[121] The European Commission, 2022, Guidelines for teachers and educators on tackling disinformation and promoting digital literacy through education and training

[122] GOV.UK, 2022, Help for vulnerable people to spot disinformation and boost online safety

[123] M. V Bronstein, S. Vinogradov, 2021, Education alone is insufficient to combat online medical misinformation

process information)".[124] The aforementioned principles were devised as a part of the "Norms for the New Public Sphere" philosophy project, showcasing the need to involve not only policymakers, political scientists, and Technology sector agents, but philosophers and social scientists to ensure the consistency and logic of the foundational regulatory bases.

Finally, the scale of such policy changes must be accentuated. To date, "regulation of the digital environment is fragmented", thus measures contained in one area may lack the efficiency of enacting sufficient change.[125] Thus, the EU's Digital Services Act may act as a great example of legislation that has the potential to be the framework for further policies and legislations to be devised that would span areas outside of the European Union.

[124] Norms for the New Public Sphere, 2022, Shaping Democracy in the Digital Age, pg. 5-6

[125] House of Lords, 2019, Regulating in a digital world, pg. 3

AI should be embraced as a means of countering terrorism, but appropriate national and international frameworks must be developed first.

>/AI+Security

AI should be embraced as a means of countering terrorism, but appropriate national and international frameworks *must* be developed first.

Given how modern-day terrorism leverages the internet and other new technologies, there is a need for policymakers around the world to establish national and international frameworks that address these issues. Concerning the character of international security and technological advancements, the dialogue should involve experts representing both perspectives. Voices from all stakeholders should be taken into account, inter alia, Big Tech, other minor information technology companies, civil society organisations and human rights advocates.[126] Indeed, as shown by the Financial Action Task Force, a favourable regulatory framework or incentives are, by and large, the most important preconditions in the adoption and use of new technologies which also applies to algorithms countering-terrorism.[127] The emphasis should be put on developing both domestic and international technology that would use artificial intelligence in detecting suspicious actions online. Cooperation between the government and the private sector in deploying models advancing machine learning is not unprecedented with the Home Office and ASI Data Science working in accord as an example. Universal algorithms, like the one developed by the two above-mentioned subjects, can be shared with smaller companies leading to positive outcomes in detecting terrorism.[128]

Social Media platforms should be regulated too, although the balance between the freedom of expression and safety must be maintained. Companies should comply with the mechanism of removing content that may be considered illegal or that competent authorities recognize as such. An analogous model *was* introduced in the EU, however, it was not flawless. Before implementing the aforementioned system of checking online content, the authorities should also bear in mind key features such as support for smaller platforms to tackle the threat of terrorist exploitations; measurement of activities on platforms; proper guidelines and limitations for competent authorities so they could not abuse their powers and restrain freedom of expression politically.[129] These issues do not occur just in the EU but also appear in regulations in other countries such as the United Kingdom.[130] Hence, policymakers should be especially aware of these aspects.

---

[126] United Nations Office of Counter-Terrorism & United Nations Interregional Crime and Justice Research Institute, 2021, Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes

[127] FATF, 2021, Opportunities and Challenges of New Technologies for AML/CFT

[128] Home Office, 2018, New technology revealed to help fight terrorist content online

[129] Tech against technology, 2021, Tech Against Terrorism response to the EU's terrorist content online regulation

[130] Independent Reviewer of Terrorism Legislation, 2022, Missing Pieces: Terrorism Legislation and the Online Safety Bill

Regarding anti-money laundering strategies, first and foremost, adequate policies must revolve around (inter)national cooperation as well as coordination between the authoritative financial supervisors and the private sector. In the guidelines and requirements prepared by the regulator in accordance with the major private stakeholders, the ideas of implementation of artificial intelligence mechanisms should be evident. The recognition of these models and the further incentivisation of their deployment by the regulators will be beneficial for the combat against terrorist financing. Advantages of these measures include enhancement of adverse media screening, monitoring of patterns and anomalies in transnational activities as well as transaction thresholds.[131] On account of the new infrastructure, institutions must conduct even more meticulous individual risk assessments. Nonetheless, the overall governance over the financial institutions cannot be too restraining as they need space for agility and adjusting innovative measures to changing patterns. They must, however, follow key principles of fairness, ethics, accountability and transparency (Monetary Authority of Singapore).[132] The number of legislations regulating this issue to a degree has recently emerged in G7 countries (the UK, Germany, France, the US) which may indicate that the trend in amending existing public policy is on the rise.

Another element of public policy of counter-terrorism where artificial intelligence can be of high use is the prevention of radioactive and nuclear materials trafficking. So far, the most cutting-edge research and analysis on machine learning is in the detection of new illicit proliferation networks. This was carried out by the Nuclear Threat Initiative and the American institutes Center for Advanced Defense Studies. Their recommendations are worth sharing explicitly:

- Integration of publicly available information into existing monitoring, verification and export control regimes;
- Deployment of adequate analytical tools such as machine learning infrastructure to conceptualise models on complex datasets from different sources;
- Broadening partnerships between analysts to enable better information and knowledge flow as well as data sources;
- 'Use of entity-level trade data as well as other diverse sets of the publicly available information in future international non-proliferation initiatives'.[133]

[131] Comply Advantage, 2022, AML and AI: How AI is Changing the AML Landscape

[132] Monetary Authority of Singapore, 2018, Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector

[133] C4ADS & NTI, 2021, Signals in the Noise: Preventing Nuclear Proliferation with Machine Learning & Publicly Available Information

>/AI+Security

Proper regulation of AI in migration technology must be introduced, including the establishment of an impartial independent adjudicator.

Proper regulation of AI in migration technology must be introduced, including the establishment of an impartial independent adjudicator.

According to Petra Molar, a world-class expert in migration law, no appropriate regulations exist right now that would establish guidelines and control the deployment of artificial intelligence in migration management.[134] The debate on the governance of the sector, therefore, should especially focus on the negative outcomes of the new technologies and what causes them. Currently, bias remains one of the fundamental issues regarding migration technology. Reduction of the chance of a bias occurrence in a system may be enabled by broadening datasets the algorithm learns from and constantly rectifying the code. These actions should be codified and in case of non-compliance adequately sanctioned. Moreover, bearing in mind the fallibility of the systems, automated decisions made by them must be at least reviewed by an impartial entity until full reliability is established. These models must also comply with internationally protected human rights frameworks and be constantly under the scrutiny of authoritative supervisors.

Another crucial part of regulating migration technology involves effective oversight and proper recognition of harmful models. Firstly, new technologies developed either under the aegis of national or supranational institutions, like Europol or Frontex, or under the auspices of private entities should be overseen by an independent body. Creating a civic institution that would check if certain stakeholders do not perpetuate existing discriminatory practices such as non-refoulment or non-entree is pivotal in monitoring the malicious use of artificial intelligence.

Secondly, any present regulation regarding migration technology must be overarching in respect of assessment. This issue has been recently visible, especially in the case of the European Artificial Intelligence Act. The piece of legislation does not establish a sufficiently wide regulatory framework that would restrain the usage of highly detrimental mechanisms. The initial proposal of the regulation did not assess as 'unacceptable uses' the systems like AI-based individual risk assessment and profiling systems in the migration context that draws on personal and sensitive data; AI polygraphs in the migration context; predictive analytic systems when used to interdict, curtail and prevent migration; and remote biometric identification. Many should-be 'high-risk' technologies have been absent in the legislation as well.[135] The act, despite its faults, is the first regional attempt to regulate such a broad array of technology that has a direct, although opaque, impact on people's rights and freedom at the same time. Other governments and regions should follow European regulatory footsteps

[134] Petra Molnar, 2020, Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up

[135] Platform for International Cooperation for Undocumented Migrants, 2022, Regulating migration tech: How the EU's AI Act can better protect people on the move

which, even though are not flawless, contribute highly to the debate regarding how migration technology should be governed.[136]

[136] Robert Bosch Stiftung, 2022, The EU's AI Act and its Human Rights Impacts on People Crossing Borders

Conclusions:

Our policy recommendations collectively push for the development of international frameworks and regulations to preserve security and mitigate threats. Multi-agent coordination is highly important in each and every case we cover, as well as striking a balance between threat management and upholding rights to life, free speech, and privacy. While the recommendations may be ambitious due to their large scale of implementation and complexity, we believe that appropriate policy measures that at the foundational level encompass the principles of fairness, transparency, accountability, progress, and reason are the cardinal in tackling the multifaceted issues that are affecting national and international security.

Artificial Intelligence is a prominent tool which's utmost capabilities and applications are in the process of discovery, yet the policymakers, along with the public must be able to adjust as rapidly as the technology is advancing. Significant effort must be dedicated to harnessing and exploring AI while ensuring its applications' compliance with ethics and human rights.