Gábor Nyáry – Cosimo Cucinelli

Cybersecurity and the Risk of Artificial Intelligence

Cyberspace is a fundamental concept in our modern world: a hotbed of escalating conflicts of power between the nations of the world, but also a preferred scene of crime, which now poses a similarly dramatic threat to members of society. We would like to provide those who wish to familiarise themselves with the problems of cybersecurity, with a brief overview and background about this strange world, which is unfolding before our eyes, but is in many ways still unknown. First we present the background against which the events of the cyber world unfold. Networking and digitalisation dramatically increase our convenience and well-being, but we have to pay a heavy price. The cyberspace where the digital economy works, where we learn, have fun, build relationships has greatly increased the vulnerability of the individual and society alike. Significant forces are loitering in this digital dimension, seeking to take advantage of these emerging weaknesses. State actors, non-state actors, groups with different motivations, and individuals with offensive intentions are all involved, threatening the online environment.

We present the wide range of these cyber actors. We also show what are the threats, different attack methods that different cyber actors operate with. We review the specific problems of cybersecurity, from intrusion detection through attribution difficulties to the topic of deterrence. We take a look at efforts that would support cyberspace security by developing a system of cyber norms. Finally, we also talk about how the latest technologies, like AI can shape cybersecurity trends.

Keywords: Artificial Intelligence, attribution, cyber actors, cyberspace, cyber threat, DDoS attack, deterrence, geopolitics, intrusion detection, malware, phishing, Tallinn Manual

Acronyms

AI	Artificial Intelligence
CCD COE	Cooperative Cyber Defence Center of Excellence
CDPF	Cyber Defence Policy Framework
CERT	Computer Emergency Response Team
DDoS	Distributed Denial of Service
DG DIGIT	Directorate-General for Informatics
EC3	Europol European Cybercrime Centre
ECCC	European Cybersecurity Competence Centre
EDA	European Defence Agency
EEAS	European External Action Service
ENISA	European Network and Information Security Agency
EU-LISA	European Union Agency for the Operational Management of Large-Scale IT Systems
GDPR	General Data Protection Regulation
GPTs	General Purpose Technologies
ICT	Information and Communication Technologies

https://doi.org/10.36250/01148_07

NLP	natural language processing
NRI	Network readiness Index
NSA	National Security Agency
PESCO	Permanent Structured Cooperation
PITM	person-in-the-middle
UN GGE	United Nations Governmental Experts Group
UN OEWG	United Nations Open-Ended Working Group

Introduction

Cyber warfare, cybercrime, cyber deterrence: these became frequently used, extremely popular expressions and concepts of both the press and also of public discourse. Cyberspace, this curious new domain that is difficult to define but is felt everywhere and by everybody, became synonymous with constant confrontation, but at least the kind of threat that is always floating there. And it is true: in our technicised societies, conflicts (whether interpersonal or interstate) are slowly filling cyberspace. Or rather: cyberspace as well. After all, one of the most noticeable phenomena of the last two or three decades is the quick proliferation and intensification of societal tensions and power clashes.

A term has appeared, or reappeared, and has become at least as popular and widely used in the world of political science and journalism as the term "cyberspace" in technical and IT discourse. The word is "geopolitics". A term with a somehow fluid content, nevertheless understood by everyone. It emphasises the importance of the environment and geographical space in the life of the states. It suggests that the undisguised, sometimes downright relentless assertion of interests has once again come to the fore in the surrounding world (Blount 2019). Then these two terms clung together and gained momentum in the form of another word combination: we are already talking about the "geopolitics of cyberspace" (Riordan 2018), and we do it for a reason.

The "cyber world" that encompasses the tangled paths, wires, computers, programs running on them, and, of course, the people who work or play with them, is by no means just the habitat of cyber fighters, hackers and cybercriminals. Conflicts, of course, often strike in this still-emerging and therefore sometimes unorganised and unregulated space. The experts of cyber diplomacy are working on solving, smoothing and regulating these. Their role will be plentiful: developments in world politics show that cyber conflicts have now become a central theme of major political rivalries instead of nuclear weapons.

The European Union (securing a leading role in regulating global digitalisation issues) has also actively participated in different multilateral forums (United Nations, Organisation for Security and Cooperation in Europe) concerning the cybersecurity domain, developing its specific tools and policy actions in the field of cyber diplomacy. The EU's basic aim is to strengthen and secure a rules-based regional (or possibly global) order in cyberspace, building also cyber resilient societies, while at the same time promoting both citizens' privacy and the freedom of the global internet. A core principle of the EU cyber diplomacy philosophy is – in the true sense of multilateralism – a "collective action",

that is, to develop policy frameworks and procedural elements to a joint response against cyberattacks which may pose a threat to the Union or its member states. The practical approach is twofold: it addresses both prevention and incident management. The major achievement of the EU cyber diplomacy strategy is the so-called cyber diplomacy toolbox which is a collection of tools ranging from classic diplomatic actions to several forms of sanctions and other coercive means. It allows a very specific, targeted and highly coordinated response to any cybersecurity threat or malicious digital action against the EU or any of its member states.

Conceptual background

When we talk about "cybersecurity", we almost always touch on economic aspects as well. And vice versa: the increasingly digital, global "cyber economy" is always a kind of security topic. National security and economy are two sides of cyberspace. The cyber domain as an economic field, both a virtual and a very real area of business interests and ambitions, is still developing today: its evolution, development directions and regulatory framework are still accompanied by questions and lively international debates (Barrinha–Renard 2020).

We say "digital economy", and with a good reason: today, this phrase has become a common term in the parlance of professionals. Since the beginning of the 2000s, a phenomenon that has accelerated at a noticeably speedy pace has shaped almost every society in the world: digitalisation. In summary, this process can be described as one in which data and networks intertwine, and permeate production processes, government and personal consumption, cross-border trade and, of course, the finances that drive the economy (Filippov et al. 2019).

However, some caution does not hurt! The term "digital economy", despite its obvious meaning, is not a uniform concept in scientific terms, it does not have an accepted definition. The International Monetary Fund, which is unavoidable in global economic statistics, declares when talking about the performance of the digital economy that there is not even a complete consensus on what we mean by the "digital sector" of the economy or what should be classified as "digital products". Although the term is utterly common in professional discourse, and so we also use the term "digital economy", let us not lose sight of the fact that it is still an evolving concept with ever-expanding meanings.

Due to the difficulty of the definition described above, the figures should be treated with caution, but the performance and pace of development of the digital economy is still remarkable (Dominioni 2019). According to benchmark calculations, the share of the digital economy in the total performance of the world economy in 2017 reached 22.5%. Traditionally, at the forefront of digitisation, the U.S., with a $5.9 trillion digital economy kicks in about 33% of the country's GDP. Experts see a particularly important role as an engine of economic growth in digital investment: this resulted in an additional 2.2% GDP growth in the U.S. by 2020 (Teoh–Mahmood 2017).

In order to take advantage of the spread of digitalisation and the growing potential of the digital economy, the widespread social acceptance and absorption of ICT technologies is also essential. The World Economic Forum's Network Readiness Index (NRI) shows the ability to exploit the potential of the digital economy. According to the 2021 report, the Netherlands, Sweden and Denmark are at the absolute top, representing the most network-ready societies. Globally, this makes Europe the leading region as to the actual potential to exploit the benefits of digitalisation (with 8 of the top 10 countries indexed). The USA, however, continues to be an up-mover, ranking 5th in 2021. Singapore is the only Asia-Pacific region country figuring among the leading 10. America, on the other hand, remains the world leader regarding future technologies, another important indicator. China, however, is moving upwards, being already the leader in some key areas like e-commerce, Artificial Intelligence, 5G and education standards (Dutta–Lanvin 2021).

The coronavirus pandemic, which – due to its global reach – has dramatically affected and wildly shaken supply chains based on a "just-in-time" logistics concept and has not slowed down the expansion of the digital economy, embodied in the technologically advanced industries. On the contrary, researchers expect a further significant expansion over the next half decade (Filippov et al. 2019).

Economy, of course, is only part of a broader context. The essence of technological development is that our modern ICT tools and the Internet, are gradually interweaving our entire societies into a digital network. Our productivity, well-being and comfort increase. However, this comes at a price: the vulnerability of 21st century technology-based societies has also increased tremendously (Brangetto–Kert-Saint Aubyn 2015). At the same time, it is an opportunity, unfortunately, for the rise of cybercrime and cyber warfare.

Key concepts and problems

Terminology of cyberspace

Conceptual diversity is deep and diverse in describing the international context of cyberspace. Even the expert community is not united in naming the most important, most basic categories. We cannot undertake a systematic conceptual analysis here, but we consider it necessary to present at least some key elements in a definitive way in order to explain the phenomena.

Aside from the often theoretical debates of politics and the academia, it is clear that the fundamental expanses of geographic space, land and water are constant arenas of advocacy struggles between states from ancient times. Air, from the first third of the twentieth century, and then, from the Cold War era, has been the dimension of space joining these geopolitical arenas. As a novelty of the twentieth century, this diverse geopolitical world has been expanded to include another dimension, a brand new area of competition. Emerging cyberspace became the fifth dimension of geopolitical confrontation and advocacy.

The rapid development of informatics, computer networks and mobile technologies has given birth to this new "field of interest", so in this sense it is really a product of our time. Like so many concepts, "cyberspace" covers many interpretations, interpretive nuances (Fourkas 2004).

The term cyberspace itself is not new: it relates to a writer named William Gibson who used it for the first time in his short story published in 1982 to present a computer-generated virtual field of reality. However, it gained real popularity in 1984 through the author's next short story, *Neuromancer* (Fourkas 2004). Since then, of course, it has emerged from the imaginary matrix world of literature and is now one of the accepted terms in the fields of science. In its common professional use, it is essentially the most widely accepted synonym for the Internet (and similar technologies), the computer networks that surround the world. It is a kind of metaphor for the virtual universe of Information and Communication Technologies (ICT), which is increasingly beginning to replace the information superhighway metaphor previously prevalent for describing the Internet (and especially loved by politicians). Here it is worth highlighting a detail that cannot be ignored even in the most superficial concept of geopolitics: spaces are not something rigid, motionless, static things, but on the contrary: dynamic systems of relations in constant motion, constantly changing, interacting with their social components (Blount 2019).

The literature on the "spatial" nature of cyberspace is not uniform (Fourkas 2004). Some authors, especially experts of strategy specialising on cyber warfare, talk about the gradual (and accelerating) virtualisation of the multidimensional geopolitical environment. According to this, traditional spaces are replaced by a space that exists only in a figurative sense, without a concrete form, a kind of "spacelessness" in the context of geopolitics, with cyberspace coming to the fore as the 5th dimension of geopolitics. So much so that one of the most acclaimed contemporary representatives of strategic scientists, the American Colin Gray, defines it as a straight "counter-geographic" space, thus emphasising the elusive, plastic reality of cyberspace (Gray 2013).

On the other hand, researchers dealing with the technological aspects of the cyber complex or the social context of digitisation emphasise the spatial nature of cyberspace, that is, its very real nature, which is integrally related to physical spatial structures. In this perception, the spatiality of the concept of cyberspace appears at different levels. It is customary to peel off at least three such spatial layers of meaning in this regard. The concept first has a level of technical meaning that describes the joint technological infrastructure of a concept called cyberspace. However, the concept also has an actual geographical layer of meaning, encompassing ICT networks and the real spatial extent of their nodes. Lastly, the concept also includes a third layer of social meaning, which describes the spatial organisations of people using ICT networks. From the above, it can be seen that cyberspace, a world often called "virtual", has a very real (not just metaphorically interpretable) spatiality. It can be stated that "cyberspace" is a real spatial system: its network topology depends fundamentally on its spatial fixations, and its development is also decisively influenced by the geography of economic and technological development of the system environment.

In summary, cyberspace can be understood as a broadly common conceptual framework, as the totality of the Internet, computing devices, the software running on them, and even the users who use them and are increasingly networked (Brangetto–Kert-Saint Aubyn 2015). Its basic characteristic is that it initially appeared as a purely technical problem area, but today it has clearly changed into a domain dominated by politics, where different national and group interests, different norms and different values shape the relations. Today, not many would doubt that this cyberspace is a dimension of geopolitical advocacy, just like land, seas, air, or space. Moreover, there is a growing consensus today that cyberspace is not just one of the dimensions of conflicts of interest and advocacy, but really the defining one (Desforges 2014).

However, cyberspace is not just a new dimension of interstate conflict, where "professionals" (intelligence agents and soldiers) fight to assert their national interests (Choucri 2012). The civil sphere has also been extracting its own "cyber soldiers" for some time: by their common name, they are called "hackers" (Sigholm 2013). They, however, actually encompass very different groups of people. What they have in common is that they exploit the vulnerabilities inherent in the technical dependence of modern society. Their motivations are as diverse as the vulnerabilities they try to exploit. And their attack methods multiply by the day.

Actors of cyberspace

The "actors" of the rapidly evolving, intricate cyberspace, which is also fraught with many vulnerabilities, are those who are on the "other side" in cyber incidents and attacks. According to a strict definition, they are "states, groups, or individuals who, with malicious intent, aim to take advantage of vulnerabilities, low cyber security awareness, or technological developments to gain unauthorized access to information systems in order to access or otherwise affect victims' data, devices, systems, and networks" (Canadian Centre for Cyber Security 2020). Due to the networked nature of the Internet, intruders can launch attacks from anywhere in the world at targets anywhere else.

It is possible to classify these malicious people and groups based on different aspects; however, the motivation of the cyber actors is a particularly important characteristic in this respect (Center for Internet Security 2021). At the same time, their "expertise" and their sophistication are also important, as there are significant differences between the different categories of perpetrators in this area. Based on the intentions, incentives, i.e. motivations of the cyber actors, these persons can usually be divided into 6 main categories:

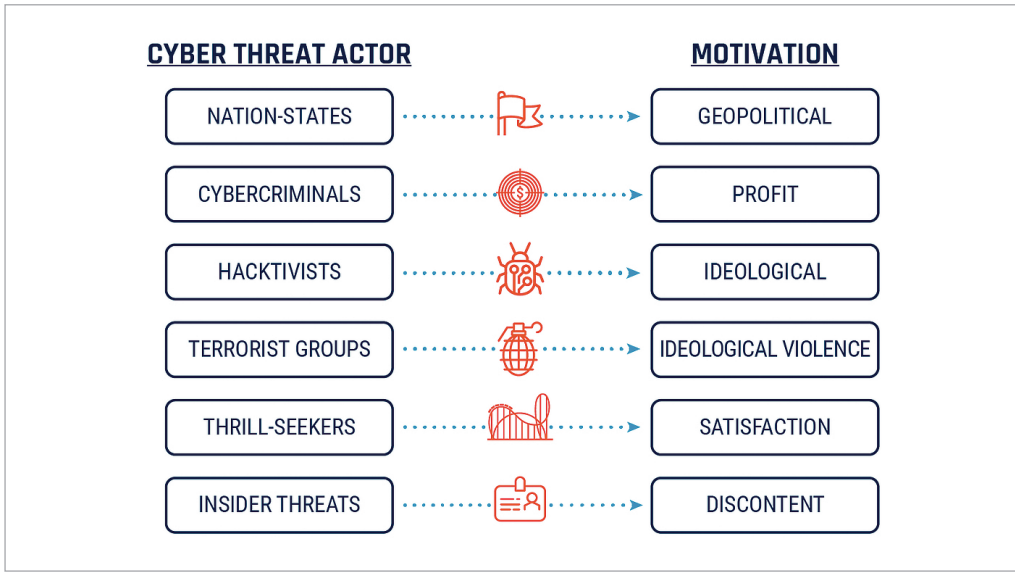

Figure 1: Threat actors and their motivation
Source: Canadian Centre for Cyber Security 2020

In the news, we often encounter hackers serving nation states. They usually work for the geopolitical interests and foreign policy goals of a sovereign state. They are either direct state actors (spies, soldiers) or other persons controlled by state organisations. They break into enemy systems to obtain or destroy data. For the most part, they are the most organised, highly educated and most sophisticated actors in cyberspace.

The groups of cybercriminals are driven by a desire for personal financial gain: they pose a long-term, growing threat. In recent times, their groups, especially those specialising in extortion, have become particularly active, now posing a threat of national security proportions even in countries like the USA (Hakmeh 2017).

We would tend to underestimate the threat of ideologically motivated cybercriminals and hacktivists. Yet they are determined, action-minded fanatics whose online actions are driven by their political objectives.

Members of terrorist groups active in cyberspace tend to be the less sophisticated threat. Their goal in cyberspace is mostly to cause confusion, disruption and harassment. These illegal organisations tend to use digital spaces for organisational communication, recruitment and propaganda.

The motivation of those who attack for fun is personal, and this "aimlessness" makes them unpredictable to some degree. They are not a harmless group at all, although their professional competence and training are usually the lowest.

The last category of cyber actors is that of insiders. This should not deceive anyone: they are particularly dangerous attackers of cyberspace. They can be former or current employees, suppliers, subcontractors or possibly partners. Their strength (and danger)

lies primarily in having internal information about an organisation, a company, which is an unprecedented situational advantage.

Each group of perpetrators often specialises in one type of attack, and in fact, the cyber threat they report is one of their hallmarks (Center for Internet Security 2021).

Cyber threats

Types of attacks termed cyber threats can be just as diverse as the various cyber actor groups that watch, move and strike in cyberspace with malicious intent. Let us look at some particularly common and especially dangerous cases (Latici 2019; CSIS 2020).

Phishing and spear phishing

An attack called phishing is perhaps the most well-known malicious activity profile. It is basically a type of social engineering action that attempts to trick users into bypassing normal cybersecurity practices and giving out sensitive data, such as usernames and passwords, bank account information, or other sensitive, personal data (social security number, or any piece of information that can perhaps be used in future attacks).

The case is well-known even for non-professionals: hackers send out phishing emails that seem to originate from trusted senders such as a government office, financial institutions, or friends and co-workers. The cybercriminals try to get users to click on links in the emails that will redirect them to fraudulent websites that ask for personal information or install malware on their devices. When the target is specific (one concrete person, or organisation) the attack is called "spear phishing".

Distributed Denial of Service (DDoS)

DDoS attacks are also well-known even for civilians with no cybersecurity training. The target in this case is usually a company, or a government office, or political actor. During such incidents the cyber actors try to overwhelm the server of the target with requests, causing the temporary take-down of the organisation's website.

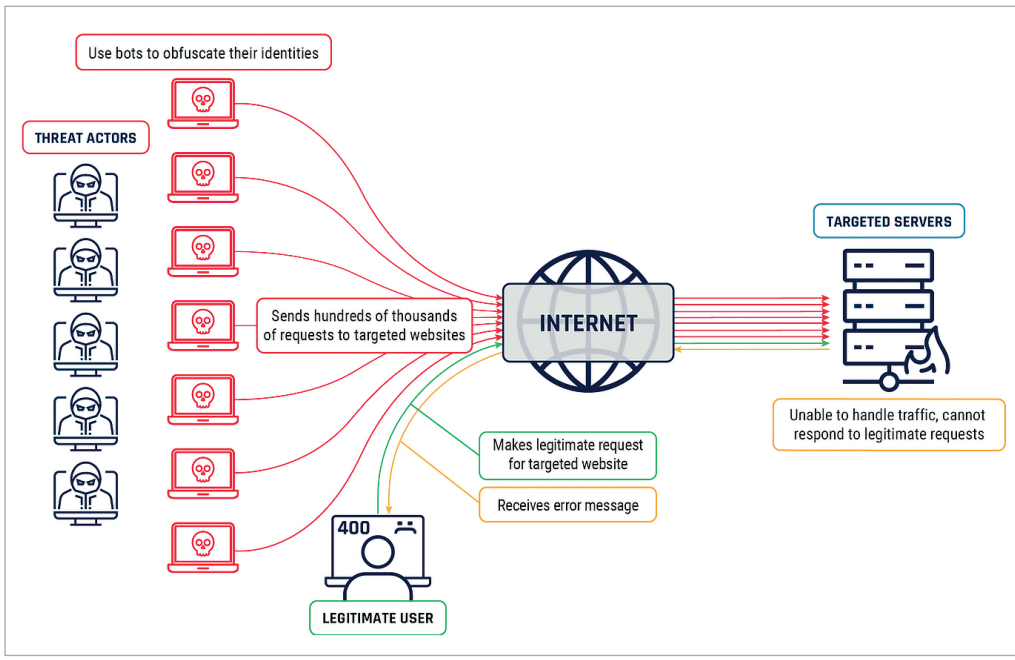

Figure 2: DDoS attack
Source: Canadian Centre for Cyber Security 2020

For being able to produce such a mass of incoming requests, the attackers use a great number of previously "hijacked" computers. Thus, requests come from hundreds or thousands of IP addresses that have probably also been compromised and tricked into continuously requesting a company's website.

Person-in-the-middle

Person-in-the-middle is also a relatively frequent type or malicious activity in cyberspace. In this case, cyber actors place themselves in the middle of a two-party communication. Once the attacker intercepts the communication, they filter and steal sensitive information and return different responses to the user.

The victim continues to believe that he is communicating, via secure connection, with a website. Sometimes the perpetrators set up fake Wi-Fi networks or install malware on users' computers or networks. Also called eavesdropping attacks, the ultimate goal of PITM attacks is to gain access to personal data (business, financial, or other).

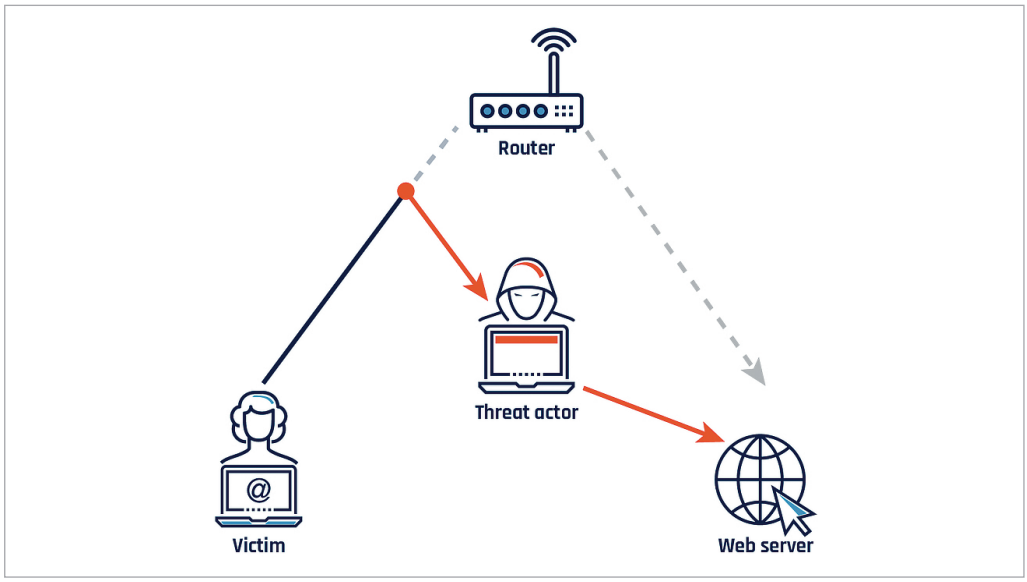

Figure 3: Person-in-the-middle attack (PITM)
Source: Canadian Centre for Cyber Security 2020

There are a very great number of other malicious activities, or cyber threats, as the development of the digitalisation of society has an impact on the development of new hacking methods, too. In short, the threat landscape is not static, on the contrary: it is evolving very rapidly, making the task of cybersecurity difficult.

However, to finish this panoramic view of cyber threats, it is worth mentioning one more category, perhaps the most common hacking tool: malicious software, in short, malware. "Malware" is, however, a generic term, denoting a very numerous family of attack tools: including trojans, backdoors, spyware, different kinds of viruses and cyber worms. One member of this group is of exceptional fame, unfortunately gaining "popularity" among cyber actors. Ransomware is also a kind of malware, and a very menacing one (Krasznay 2020).

The beginning of the year 2021 was characterised by an alarming growth in the use of this special kind of malware. The truth is that this trend can go back a long way. The global coronavirus pandemic, however, caused a sharp increase in the number of ransomware attacks. The average amount of ransom paid by affected companies is also steadily rising.

Ransomware is basically a malicious software that, in many cases, restricts access to a computer or a device and its data by encrypting its content. The computer is effectively locked, and the users cannot get access to their datasets. The cybercriminals demand that a ransom be paid, usually via a cryptocurrency such as bitcoin, in order for the victim to regain access to systems and information.

Is anybody there? The problem of intrusion detection

Cyber defence, namely the response to conflicts in cyberspace also raises a number of novel problems, mainly due to special digital technologies, which are not supported by the usual conflict management principles and practices developed during the Cold War era. Right at the beginning, there is the difficulty of intrusion detection (Champion 2020).

In general, opponents manoeuvring in cyberspace (whether they are state actors working for foreign policy purposes or even ordinary criminals) are trying to hide in the darkness of online spaces. Although there are exceptions (since in a blackmail virus attack, it is indeed important for the victim to detect the action), it is not as easy to detect attacks – attempts and intrusions – as we might think. Previously, according to a common concept, protection was focused on the boundary line of the system, called the perimeter (Schaffer 2021). But the truth is that a sizable portion of the attacks are internal actions (insider threats), so they already arise within the perimeter. At the same time, significant cyberattacks in recent times (especially the so-called SolarWinds action) have drawn attention to the fact that sophisticated attackers of our time can penetrate external defences easily; and perimeter-focused protection is incapable of detecting attackers who have already entered the system as a result of such breakthroughs. The owners of the attacked systems and networks are often unable to reveal intruders for months until it is too late to prevent damage (Krasznay 2020).

External perimeter protection devices (firewalls, secure web gateways, antivirus solutions) are no longer able to reliably intercept external intruders; as stated by U.S. government agencies investigating SolarWinds hacking, perfect protection just does not exist. Intrusion detection is also essential to complement external protection: these are procedures (and tools) that continuously analyse the entire ecosystem of a system to be protected. They are constantly looking for traces of malicious activities that could compromise the network. Detection is actually based on a threat intelligence activity that continuously analyses (legitimate) user behaviour and continuously compares it with signatures that were captured from previous attacks.

Who did it? The problem of attribution

In relation to cyberspace conflicts, attribution is undoubtedly one of the most difficult tasks (Assumpção 2020). The activity (in which the answer to this basic, simple-looking question of who "staged" the malicious event in cyberspace is sought) is essentially a process of investigation and analysis, in which cybersecurity professionals gather probative information and set event schedules, and from all this they laboriously reconstruct the attack history and profile. The purpose of such an activity, also called forensics, is to establish in a demonstrable way who was the perpetrator of a cyber action or who may be the actual person responsible for an offensive action.

Some experts argue that, while attribution is indeed a difficult task, it is by no means impossible. There are always many clues left after an offensive action from which

the image of the perpetrator can be compiled with meticulous work. However, this seems to be contradicted by the fact that even great powers with almost unlimited resources and vast knowledge capacities present evidence only in an exceptionally rare case when naming an alleged perpetrator of an attack. In the vast majority of cases, attributions contain a number of hypothetical elements, and the strength of the evidence is a matter of point of view, so it is far from being legally clear, or corroborative (Tsagourias–Farrell 2020).

Nevertheless, the meticulous tracing of the perpetrator is a mandatory activity for cyber defence professionals after every cyber incident, even if the work involved takes months or even years. It is worth noting, analysts are not just investigating the perpetrators of successful actions (resulting in actual damages): in fact, you can learn just as much from the details of unsuccessful, possibly aborted, offensive actions as you can from successful ones.

Generally speaking, a forensic investigation is considering different "attribution layers" (Dévai 2020). The first level, or layer is trying to uncover and understand the technical parameters of an attack. This is the tactical layer. The next step is to understand the attack's high-level architecture and the attacker's profile. This is the operational layer. And lastly, the strategic goal of the investigation is to discover who is responsible for the attack. Finally, communicating the results of a lengthy forensic investigation is also an important task of the attribution process.

The difficulty of identifying the perpetrators, the process of attribution stems primarily from the fact that the perpetrators try very carefully and sometimes in very sophisticated ways to remove all traces of their intrusion (Assumpção 2020). One of the tricks, by no means an unusual practice, is to use others (mostly third parties) to cover your own operations. This process, known in Anglo-Saxon professional circles as a "false-flag" operation, is also frequently used in the cyber action sphere. Using the attack methods, procedures, or even (previously compromised, or illegally obtained) offensive software of hackers from other countries, they disguise themselves as someone else. In one such case that later became known, for example, in the fall of 2019, it was Russian hackers who "captured" the identities of others, an Iranian cyber group, to gain access to the networks of government and economic actors in a dozen countries. Incidentally, according to U.S. experts (as seems to be confirmed by Snowden leaks and information made public through WikiLeaks), the U.S. signal detection and cyber action agency, the NSA, also prefers (and has a high level of technical expertise) to use such methods to cover its actions.

How to discourage them? The problem of deterrence

In the era of the Cold War, the concept of nuclear powers holding each other in checkmate situation was the so-called principle of deterrence. The point was that both sides knew it was not worth attacking the other because the challenged party would make the attacker pay a heavy price for that act, thus making "victory" meaningless. A brutally simple principle − and it worked! The great tragedy of the cyber age is that this scheme, which

guaranteed security for a long time, is essentially not applicable in the realm of digital devices and networks. Where an attack can go unnoticed or can be easily denied if it is revealed, it is not possible to know for sure who the addressee of such a threat of retaliation should be. Therefore, one of the best defences, discouraging any potential attackers seems hardly viable in the new cyberspace environment.

From the above, a sequentially cumulative series of cyber defence problems can already be seen. In the previous subsection, we have shown that establishing the attribution, the identity of the perpetrator (that is, in a clear, proven way) is one of the most difficult tasks. The problem is exacerbated by the fact that a not very elegant but effective way of curbing conflicts, the deterrence of potential perpetrators, is practically based on the possibility of attribution. Deterrence theory, developed in the period and situation of the classical rivalry of the great powers of the nuclear age, assumes the existence of three important elements for maintaining the equilibrium (deterring potential attackers): attribution, credible signalling, deterrence strategies (Krasznay 2020).

For the deterrence to work effectively, the perpetrator (or potential perpetrator) had to be clearly identified. This should be followed by a credible signalling of the attacked party's determination to retaliate. Signalling is possible in a comprehensive way, as a general warning. However, this has only a limited persuasive, deterrent power. Tailored signalling can send a much more focused, and powerful alert. However, it is a condition that the perpetrator is known, without which tailored signalling can be a particularly risky move. The source of an additional problem is, when signalling intent, the more specific the threat, the more plausible. However, the signalling power does not want to reveal too much about his own capabilities. And this is a problem, since credible deterrence is based on two pillars: a credible will and appropriate assets to retaliate (Schaffer 2021).

Analysing the classical superpower rivalries of the Cold War era, another important detail emerges: the principle of reciprocity had to be associated with the operation of deterrence, that is, the preservation of a sensitive equilibrium. In the case of opposing great powers, it would not have been sufficient for power A to have indicated to power B that it would retaliate if necessary. The changing of signals between the U.S. and the Soviet Union actually sounded like: "If you attack – I will retaliate. But I also know that you would do the same, in case I would attack you. Therefore, you just should not attack and also need not attack." And peace, however hot, reigned all those dramatic years.

What should be the rules? The problem of legal frameworks

A very special characteristic of cyberspace, with reference to conflict and attacks, is the limited jurisdiction of the legal framework governing the behaviour of different actors. However, contrary to popular belief, cyberspace is neither some kind of "digital Wild West". Rather, the case is that, due to its special nature, the operation of cyberspace raises many novel technological and legal dilemmas. At the same time, there are norms here, well-thought-out, clear rules for conflicts in online spaces, for behaviours to be followed or forbidden (Stadnik 2017).

NATO – Tallinn Manuals

After previous professional attempts, in 2007, in the wake of the crippling cyberattack on Estonia, the search for and elaboration of universal norms applicable in cyberspace conflicts gained momentum. The then-established NATO headquarters for research on cyber warfare, CCD COE (NATO Cooperative Cyber Defence Center of Excellence), began examining the rules applicable to cyber warfare with the help of international experts. Their efforts led to the birth of the Tallinn Manual (2013), which focused mainly on the use of force (ius ad bellum) and the validity of international humanitarian law (ius in bello).

To complement the recommendations, the Tallinn Manual 2.0, published in 2017, focused on topics not previously covered: it sought primarily to find applicable principles and rules for conflicts and actions "below the stimulus threshold". A common feature of both Manuals is that they contain only recommendations to be followed and not mandatory legislation. Their rules are not legally binding (Vihul 2013).

Experts are already working on the compilation of the Tallinn Manual 3.0, which will integrate responses to the increasingly sophisticated yet dangerous cyber actions of recent years into a single system.

United Nations – UNGGE and OEWG

The United Nations (UN) also plays an important role in the international regulation of cyberspace (Digital Watch 2021). The so-called United Nations Governmental Expert Groups (UN GGE) of the world organisation have been working – since 2021 – with the international expert community to develop voluntary standards for reducing cyber-attack threats and to establish responsible state behaviour in cyberspace (Ruhl 2020). The Group, set up by the UN on the basis of a proposal put forward by Russia much earlier, opened a working platform for rival powers where they could try to establish some sort of common ground, especially in the area of much-needed confidence building. Representatives of the great powers were seated in the work organisation, and power rivalries ultimately left their mark on the group's activities. Nevertheless, their joint working material (report) completed in the summer of 2021 is a major step in strengthening international confidence. The World Organization has launched another initiative to map the problems of cyberspace and to develop the normative systems to be followed. Partly based on the experience of the UN GGE group, learning from its difficulties (the organisation could summarise its work in the form of a standard report after a very lengthy process), otherwise again on Russian initiative, in 2017 the UN decided to set up an Open-Ended Working Group (UN OEWG). The main virtue of a format open to all Member States is that it involved all UN Member States in developing common sets of rules.

Bilateral constructions

The common weakness of the above (regional or nationwide) multilateral norm-setting activities is that they set important principles and rules, but they are not binding even for states that recognise the regulation (i.e. adhere to their final documents). Rulemaking efforts that encompass only two (maybe three) major powers (therefore bilateral in nature) may promise more success, because they can result in contract-like legally binding rules. An important feature of them is that they focus on regulating only a small number of important issues out of the diverse, complex problem areas of cyberspace.

The act of launching such a bilateral international regulation effort could be the case of the summit of the presidents of the two great powers, the U.S. and Russia, in the summer of 2021. Negotiations have begun between leading politicians in the two states over what should be the minimum basis for responsible state conduct in cyberspace, for a kind of "peaceful cyber coexistence". The starting point for the discussions is to jointly define the range of civil and state critical infrastructures that are essential to the functioning of modern societies and which the parties will refrain from attacking.

Challenges and answers – The European Union as a global norm-setter

The European Union has long positioned itself – with great success – as a powerful norm-setter of the often frontier-like cyber domain (Dévai 2020; Ruhl 2020). The unique opportunities, as well as major challenges posed by the rapid digital transformation clearly have not escaped the attention of the continent's decision-makers.

At the end of 2020, the European Commission presented the most ambitious reform package for the European digital space to date. The two pieces of legislation presented by the panel − the Digital Services Act and the Preliminary Version of the Digital Market Act − aim at not less than to energise the whole range of digital services and online market places in the Union by creating a long-term and coherent regulatory environment. A couple of days later, to this already impressive package was added another important element with the European Commission's presentation of the Union's new cybersecurity strategy, which aims to ensure that the European digital space spurred is not only economically fruitful, but at the same time remains a free and safe medium (Modern Diplomacy 2020).

The economy comes first, but closely followed by security issues

The new European cybersecurity strategy (2020), presented by the EU Commission and the High Representative of the Union for Foreign Affairs and Security Policy is a remarkable policy document, guiding future cyber defence efforts not only on the European level (European Commission 2020b).

Cybersecurity has been one of the Union's top priorities for some time. During the coronavirus crisis, cyberattacks against healthcare institutions (research sites, manufacturing plants and hospitals) multiplied, demonstrating the importance of protecting infrastructure.

Under the EU's new strategy document, cybersecurity would be integrated into all elements of the supply chain, and EU activities and resources would be even more closely connected across the four cybersecurity communities – internal market, law enforcement, diplomacy and defence. The new strategy builds on the Communication on Planning for Europe's Digital Future and the EU Strategy for the Security Union, as well as on a number of pieces of legislation and initiatives to strengthen the EU's cybersecurity capabilities and increase Europe's resilience to cyberattacks. In this respect, of particular importance are the cybersecurity strategies adopted in 2013 (revised in 2017), as well as the Commission's *European Security Strategy 2015–2020.* In the field of legal regulation, the Cybersecurity Directive (EU Network and Information Security Directive 2016/1148 (the NIS Directive), which entered into force in 2016, was a pioneering initiative: it resulted in a uniformly high level of security of network and information systems across the EU.

The Union has developed a comprehensive, systems-based international cyberspace policy since the 2013 EU Cybersecurity Strategy (Dévai 2020). Through bilateral, regional and international cooperation with its partners, it has promoted the creation of a global, open, stable and secure cyberspace, guided by the EU's core values and based on the rule of law. The EU has also supported third countries in enhancing their resilience to cyberattacks and in tackling cybercrime more effectively, and has contributed to international security and stability in cyberspace through the 2017 EU Cyber Diplomacy Toolkit. As a memorable recent move, it applied for the first time the cybercrime sanction system introduced in 2019, listing 8 individuals and 4 organisations. This is "naming and shaming" first and foremost, utilising the power of international public opinion as a strong deterrent force. Besides, the economic and personal consequences of these sanctions are also to be felt. The Union has also made significant progress in cyber defence cooperation, including on cyber defence capabilities, mainly in the context of the Cyber Defence Policy Framework (CDPF) and through the work of the Permanent Structured Cooperation (PESCO) and the European Defence Agency (Modern Diplomacy 2020).

The European Union has long recognised the need to guarantee the resilience of critical infrastructures, that is, the all-important social infrastructure that provides services that are essential for the smooth functioning of the internal market and for the daily life and livelihood of European citizens. It therefore established a European Program for Critical Infrastructure Protection in 2006 and adopted the European Critical Infrastructure Directive in 2008 for the energy and transport sectors. These measures were complemented in later years by various sectoral and cross-sectoral measures on specific aspects such as examining resilience to the effects of climate change, strengthening civil protection or the resilience of foreign direct investment.

Actions for strengthening cybersecurity through Europe

In principle, the newly adopted European cybersecurity strategy aims to preserve the global and open internet, while ensuring that, in addition to security, European values and fundamental rights for all are also protected. Besides, the document also sets out concrete proposals for action in three areas (European Commission 2020b):

Resilience, technological sovereignty and leadership

As part of the review of the aforementioned cybersecurity directive launched in February 2020, the Commission is proposing a reform of the rules on the security of network and information systems. The aim is to increase the resilience of the critical public and private sectors (hospitals, energy networks, railways, but also data centres, administrations, research laboratories and the manufacture of critical medical devices and medicines, and other critical infrastructures and services) to cyberattacks. It proposes strengthening the role of digital innovation centres and stepping up efforts to train and develop the workforce in order to establish the Union's technological sovereignty and leadership.

Operational capacity building: Facilitating prevention, deterrence and response

As a key element, the European Commission is preparing to set up a new joint cybersecurity unit in the Member States. The aim is to significantly increase the capacity and effectiveness of cyberattacks prevention, deterrence and incident response through cooperation. It is also a priority to strengthen the cyber diplomatic toolbox, in particular to respond effectively to attacks on critical infrastructures, supply chains and democratic institutions. The EU will also seek to further strengthen cyber defence cooperation and develop state-of-the-art effective defence capabilities among EU Member States.

Supporting the development and operation of global and open cyberspace

The EU's top foreign policy priority is the rule-based world order and the representation and protection of human rights and fundamental freedoms. This concept must be reflected in all the Union's cyberspace policies. In line with this, the new cybersecurity strategy considers guaranteeing the international security of cyberspace to be a key objective. The Union intends to promote international norms and standards that reflect these core EU values by working with its international partners in the UN and other relevant fora. The EU will further strengthen the EU cyber diplomacy toolbox and step up its cyber capacity building efforts in third countries through the development of a comprehensive EU agenda. As a vital institutional development, the European Union intends to set up a global EU cyber diplomacy network to promote its cyberspace ideas internationally.

The long-term expectation of the new cybersecurity strategy is to enable the European Union to increase its leadership in international norms and standards for cyberspace and to strengthen cooperation with its partners worldwide to develop a global, open and secure cyber domain. In the light of the above measures, we can conclude that the EU clearly aims for a role of international norm-setter (normative power) in the all important cyber domain.

Institutional framework

Due to the comprehensive nature of cybersecurity, practically all EU institutions, bodies and agencies are involved in the preparation and implementation of cybersecurity policy (the Directorate-General CONNECT). The Directorate-General for Informatics (DG DIGIT) provides digital services for departments of the European Commission and other EU institutions. DIGIT hosts CERT-EU (Computer Emergency Response Team). The European Network and Information Security Agency (ENISA) was established in 2004. It helps Member States, EU institutions and all other stakeholders in their cyber policies. The European Cybersecurity Industrial, Technology and Research Competence Centre (Cybersecurity Competence Centre, ECCC) was established in 2021 in order to improve the coordination of research and innovation in cybersecurity. The Europol European Cybercrime Centre (EC3) was set up in 2013 to protect European citizens and businesses from cyber threats and support governments against cybercrime. The European Union Agency for the Operational Management of Large-Scale IT Systems (EU-LISA), was established in 2011. This agency is responsible for the operational management of large-scale IT systems in the area of justice, security and freedom. It helps the implementation of the asylum, border management and migration policies of the EU. The European External Action Service (EEAS) has a central role in the field of cyber diplomacy, strategic communication and the policies concerning cyber defence. In this field it closely cooperates with the European Defence Agency (EDA) (Molnár 2020).

Artificial Intelligence in cyberspace: Defensive and offensive roles

"Whoever leads in Artificial Intelligence, will rule the world." The quote from Russian President Vladimir Putin three years ago could have been chosen with a calm heart as the motto of the subchapter (Meyer 2017).

Innovations, identified by researchers as the engine of great social transformation, are the so-called "general purpose technologies" (GPTs), which of course might be more accurately called technologies of "comprehensive scope". After the steam engine, then electricity and informatics, Artificial Intelligence is now coming on the back of the fourth wave. And if it is true that cyberspace is a new, defining dimension of geopolitical advocacy, just as crime is now a prominent field of it, then Artificial Intelligence will be the most important piece on this all-powerful social playing field (Gill 2020).

What is Artificial Intelligence?

Over the last years, the new disruptive technology generally known as Artificial Intelligence noticeably imposes itself, holding a promise that is very hard to be delivered, namely, to drastically transform citizens' life as well as to improve people's quality style of life. In fact, it is not a coincidence that the EU is striving towards becoming the world-leading region for developing and deploying ethical and secure AI (European Commission 2018b).

"Artificial Intelligence" does not refer to a single technology, but rather to a typically interdisciplinary field of research in computer science and the technologies and applications developed in connection with them. The focus of AI is the simulation of intelligent activities (processes) characteristic of humans by computer systems (Taulli 2019).

These activities include, above all, the ability of human learning (i.e. the ability to obtain information and the rules necessary for the use of that information); they also encompass the ability to reason humanly (i.e. to be able to draw conclusions based on rules) and, as a particularly important feature, the ability to self-correct.

Despite the wide debate and research, there is no agreed definition of AI. As usual, this is not an easy task. However, in this course we can consider AI as "…systems that display intelligent behaviour by analysing their environment and taking action – with some degree of autonomy – to achieve specific goals…". Such definition, explained within the European Commission Coordinated Plan on AI (European Commission 2018a), appears so flexible as necessary to comprise the whole set of different shapes the AI may assume.

Artificial Intelligence is closely related to the concept of big data. Big data covers a large amount of data that is extremely varied, complex and changes rapidly. These masses of information can no longer be managed with traditional tools (e.g. database managers). It is the AI technologies that help to process them. At the same time, it can be said that machine learning procedures, which are one of the most important technologies of AI, cannot be imagined without a significant amount of data (mostly classified as big data) for training algorithms (Taulli 2019).

AI technologies are commonly categorised either as Artificial Intelligence with general ability, also known as "strong AI", or as Artificial Intelligence with narrow application, also known as "weak AI". "Weak AI" essentially covers Artificial Intelligence technology designed and trained to perform a single target task. A typical example is a chat robot used in public relations systems. One promising field for "powerful" AI applications is cyber defence. Unfortunately, however, both opposing parties would be happy to count on this new weapon.

Potential risks and benefits of the new technologies

Even though most part of the general public have their personal ideas and blurry vision of AI, all of them sooner or later in their life have already experienced and wholeheartedly agree about the benefits and the advantages each AI application reserves for its users.

Given these premises, in any case, it is not safe to assume AI systems have, *a priori,* an undeniable capability for ethical reasoning, if anything, quite the contrary (Dignum et al. 2018). In this regard, an environment of trust and accountability around the development and use of AI is needed, for both citizens and companies (European Commission 2018b).

It means an AI ecosystem of excellence and trust, according to the principle of "ethics and security by design", by means of a common set of actions comprised in the Coordinated Plan on Artificial Intelligence (European Commission 2018a).

In fact, the lack of trust is the major risk to face, due to, for instance, the current uncertainty about:

- the allocation of responsibilities related to material or non-material damages AI could impose
- the opaqueness of AI decision-making processes

Despite the great divergence between Member States' legal frameworks related to AI, in any case any "trustworthy AI" firstly should comply with the law, secondly should fulfil ethical principles, and thirdly should be robust against cyber and hybrid threats (European Commission 2019).

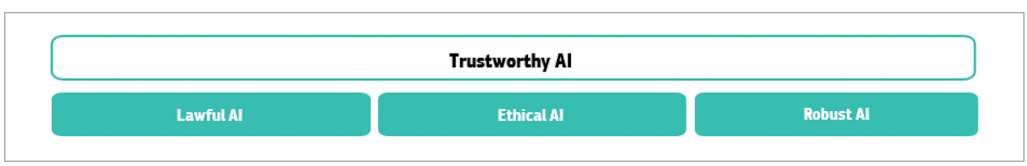

Figure 4: Trustworthy AI three pillars
Source: AI HLEG 2019

These are the three pillars at the base of *Guidelines for Trustworthy AI* drafted by the AI high-level expert group set up by the European Commission and in charge of drafting AI ethics guidelines as well as preparing the basis for a human-centric AI (European Commission 2021).

Thus, research should aim at consolidating AI decision-making process no longer:

- opaque
- bias-based
- not compliant with the privacy principles
- vulnerable to cybercriminal attacks

In the light of the above, there is the urgency of a trustworthy AI to be enhanced by and to be embodied in the following requirements (AI HLEG 2019), whenever AI solutions are exploited:

- human agency and oversight, implementing a human governance mechanism intervention
- technical robustness and safety, granting resilience and increasing the risk tolerance threshold
- Privacy and Data Governance, without prejudices against the grateful and remarkable achievement awarded by virtue of the GDPR
- transparency, facing the matter of the explainability of the algorithmic decision-making process
- societal and environmental well-being
- diversity, non-discrimination and fairness, avoiding that training schemes for AI inherit biases from their programmer and improving a fully compliant AI
- societal and environmental wellbeing
- accountability, confirming the innovative approach adopted within the EU

Potential risks

Technology itself is neutral, contrary to its use by humans: it could express either ethical or unethical aptitudes, posing new and challenging high risks to the fundamental rights fully recognised in the constitutional traditions of the Member States.

Thus, although our society could not back-pedal on the advantages of the use of AI, however, nowadays all citizens are subject to a set of numerous decisions based solely on automated, complex algorithms, and must face the twofold nature of those solutions.

This wide use of AI applications raises many concerns, and deserves an outstanding focus on more than just one perspective, namely in terms of:

- Cybersecurity vulnerability: according to the AI asset taxonomy carried out by the ENISA (ENISA 2020), the key assets are the data and the processes, namely the set of operations performed on the data, the models which the AI resorts to, the actors involved, the Environment/Tools, which include the Machine learning platform and the monitoring tools, and, finally, the artefacts, such as the data and metadata schemata (ENISA 2020). Once defined the intervention boundary, it is possible to straight proceed through the threats modelling activity, with two paramount aims: firstly, identifying, secondly, prioritising threats, in order to implement the appropriate countermeasures.
- Opaqueness of AI decision-making processes: the principle of "explainability" is not new within the European Union legal framework. On the contrary, the right to explanation, already enshrined in Article 22 of the GDPR, should be transposed and clearly stated even with regard to AI applications. Without information about the logic criteria lead an AI system to a certain prejudicial decision, the latter cannot be duly contested, thus, citizens will have no shields to defend themselves

from prejudicial legal effects produced by such decision born from a "black box" (AI HLEG 2019). Irrespective of the type, the race, the class of the data subject, algorithmic transparency should ensure, for instance, people, either who were denied jobs or who stumbled into a rise of health insurance premiums considering time and nature of food consumption habits collected by a mobile-application (Art 29 2014 WP), to be able to grasp the reasons of such decisions. Although explaining why profiling and automated processing of personal data lead to a certain direction it is not an easy task, however, explaining the evaluation method of certain personal aspects about a natural person is the key point for "Contestability by design" (Almada 2019).

- Privacy and data protection: with new advancements in technologies, huge amounts of data could be collected, analysed and stored. It poses relevant issues concerning the lawful treatment of personal data, not only in respect of surveillance of civilians by governments (e.g. predictive policing algorithms) (Rodrigues 2020), but also regarding the misuse of anonymisation techniques: the risk is that personal data may be produced from non-personal data by pinpointing the relations between certain anonymised datasets and additional data, e.g. harvested by web scraping.
- Allocation of responsibilities: in case of incidents and material or non-material damages provoked by AI: a common liability rules framework is deemed as an urgency, for instance, among others, to clearly regulate autonomous driving solutions, in case of malfunctioning of sensors detecting and/or avoiding potential collisions as well as recognising the traffic signs.
- Discrimination: there is a significant surge in current AI research efforts avoiding training schemes for AI inherit biases from their programmer. In fact, a fully compliant AI should not present any affection of data sets' historic bias, resulting into discriminations against certain vulnerable members, groups and social classes, in terms of unequal opportunities for access to education and employment (Rodrigues 2019).

All the risks listed could be summarised in just one major risk: the lack of trust. The latter must be faced by investments in research, training programs and awareness campaigns, as well as by providing common legal framework, certifications and standards to resort to, as with both data protection regulation and cybersecurity regulation (European Commission 2020a).

AI, potent weapon in the armoury of cyber actors

There is broad consensus that cyberspace is a dimension where both the attacking and the defending party will soon seek to operate using Artificial Intelligence-enabled systems (Crane 2021). Many people believe that this is (also) an area where those who want to maintain order and security are at a disadvantage. This is because machine learning can now bypass and break down cyber defence systems so quickly that conventional

protection tools cannot be kept up. Of course, machine learning is also used by cyber defence professionals. AI is used, for example, to identify threatening online patterns of behaviour.

However, experts are still optimistic: while it is clear that integrating Artificial Intelligence technologies and devices into existing cybersecurity systems is not an easy task, the expected benefits can still do much to strengthen the rapidly deteriorating cybersecurity environment. Key cybersecurity areas (functions) where AI applications can significantly increase defence effectiveness: AI can be used to create more accurate, biometric based login techniques; it can also be used for detecting threats and malicious activities using predictive analytics. It can also serve the cyber defenders by enhancing learning and analysis through natural language processing (NLP), one of the major areas of AI development. Artificial Intelligence can also support traditional cybersecurity functions by securing conditional authentication and access (Crane 2021).

On the other hand, the unparalleled capabilities of Artificial Intelligence technologies, as mentioned before, provide more than just a new set of tools for defence professionals. Attackers (whether ordinary cybercriminals or public service intelligence agents, cyber soldiers) can launch attacks that are much more sophisticated than they are today, using AI tools. Basically, Artificial Intelligence technologies allow attackers to produce much more complex, and more adaptive, malicious software. This means two things: on the one hand, attackers can adapt better and faster to new means and procedures of defence. On the other hand, the cost and time to develop complex offensive software are dramatically reduced. Thus, AI can also make a significant contribution to the proliferation of cyber weapons.

Conclusions

Compared to the former conditions of the original Cold War era, where the normality of the two poles gave the world some stability, in this new kind of "21st century Cold War", peculiar, difficult-to-follow logics prevail both in real geopolitical spaces and in the cyber dimension. The coronavirus epidemic also revealed a harsher and more ruthless cyber world. Based on the chronology of the cyber incidents of recent decades, and especially the cyberspace rivalries of the competing superpowers, the following characteristics of the cyber world as a geopolitical "battlefield" seem to emerge. The proliferation of IT tools and procedures suitable for cyber warfare seems difficult to stop. As a result, cyber warfare capabilities will spread rapidly among medium-sized powers, but even among less significant power actors. At the same time, however, the great powers will continue to dominate cyberspace (as well), as only countries with a strong technological background will still be able to carry out complex attacks. Finally, as a particularly disturbing development, experts consider it conceivable that cyberattacks, with the escalation of strikes-counter-attacks, could degenerate into real ("kinetic") damage (Domingo 2016: 166).

The study of the civil sphere of cyber conflicts suggests similarly threatening perspectives. A clear trend is the increase in the number and intensity of cyberattacks, as the damage caused becomes more and more serious. Within cybercrime, extortionist attacks (mainly ransomware) are clearly taking over. Moreover, they now target the infrastructures that are essential for the functioning of modern societies, so it has been suggested that they should be treated in the same way as terrorist acts. In cyberspace, common crime is beginning to become a threat to national security.

References

AI HLEG (2019): *Ethics Guidelines for Trustworthy AI.* High Level Group on Artificial Intelligence. Online: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60651

Almada, M. (2019): *Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems.* In 17th International Conference on Artificial Intelligence and Law (ICAIL 2019).

Assumpção, C. (2020): The Problem of Cyber Attribution Between States. *E-International Relations,* 06 May 2020. Online: www.e-ir.info/2020/05/06/the-problem-of-cyber-attribution-between-states/

Barrinha, A. – Renard, T. (2020): Power and Diplomacy in the Post-Liberal Cyberspace. *International Affairs,* 96(3), 749–766. Online: https://doi.org/10.1093/ia/iiz274

Blount, P. (2019): *Reprogramming the World. Cyberspace and the Geography of Global Order.* Bristol: E-International Relations Publishing.

Brangetto, P. – Kert-Saint Aubyn, M. (2015): *Economic Aspects of National Cyber Security Strategies.* Project Report. Tallin: CCDCOE.

Canadian Centre for Cyber Security (2020): *An Introduction to the Cyber Threat Environment.* Online: https://cyber.gc.ca/en/guidance/introduction-cyber-threat-environment

CCDCOE (2013): *The Tallinn Manual.* Online: https://ccdcoe.org/research/tallinn-manual/

Center for Internet Security (2021): *Election Security Spotlight – Cyber Threat Actors.* Online: www.cisecurity.org/spotlight/cybersecurity-spotlight-cyber-threat-actors/

Champion, A. (2020): *Cyberattack Detection Challenges and How to Meet Them.* Online: www.f-secure.com/en/consulting/our-thinking/challenges-of-cyber-attack-detection

Choucri, N. (2012): *Cyberpolitics in International Relations.* Cambridge, Mass.: MIT Press.

Crane, C. (2021): *Artificial Intelligence in Cyber Security. The Saviour or Enemy of Your Business?* Online: www.thesslstore.com/blog/artificial-intelligence-in-cyber-security-the-savior-or-enemy-of-your-business/

CSIS (2020): *Significant Cyber Incidents.* Online: www.csis.org/programs/technology-policy-program/significant-cyber-incidents

Desforges, A. (2014): Representations of Cyberspace: A Geopolitical Tool. *Hérodote,* 152–153(1–2), 67–81.

Dévai, D. (2020): The International Cyberspace Policy of the European Union. In Török, B. (ed.): *Információ és kiberbiztonság.* Budapest: Ludovika Egyetemi Kiadó. 469–485.

Digital Watch (2021): *UN GGE and OEWG.* Online: https://dig.watch/processes/un-gge

Dignum, V. – Baldoni, M. – Baroglio, C. – Caon, M. – Chatila, R. – Dennis, L. – Génova, G. – Haim, G. – Kliess, M. S. – Lopez-Sanchez, M. – Micalizio, R. – Pavón, J. – Slavkovik, M. – Smakman, M. – van Steenbergen, M. – Tedeschi, S. – van der Torre, L. – Villata, S. – de Wildt, T. (2018): *Ethics by Design: Necessity or Curse?* In 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18), 2–3 February 2018, New Orleans, LA: ACM. 60–66. Online: https://doi.org/10.1145/3278721.3278745

Domingo, F. C. (2016): Conquering a New Domain: Explaining Great Power Competition in Cyberspace. *Comparative Strategy,* 35(2), 154–168. Online: https://doi.org/10.1080/01495933.2016.1176467

Dominioni, S. (2019): *Digital Economic Powers and Digital Political Rulers.* Online: www.ispionline.it/en/pubblicazione/digital-economic-powers-and-digital-political-rulers-24187

Dutta, S. – Lanvin, B. (2021): *Network Readiness Index 2021.* Online: https://networkreadinessindex.org/

ENISA (2020): *Artificial Intelligence Cybersecurity Challenges.* Online: www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges

European Commission (2018a): *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Coordinated Plan on Artificial Intelligence.* Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0795

European Commission (2018b): *Staff Working Document. Artificial Intelligence for Europe.* Brussels: European Commission.

European Commission (2019): *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Building Trust in Human-Centric Artificial Intelligence.* Brussels: European Commission. Online: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52019DC0168

European Commission (2020a): *White Paper on Artificial Intelligence. A European Approach to Excellence and Trust.* Online: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission (2020b): *The EU's Cybersecurity Strategy for the Digital Decade.* Online: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=72164

European Commission (2021): *Shaping Europe's Digital Future.* Online: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai#:~:text=The%20European%20Commission%20appointed%20a

European Parliament (2017): *Resolution of 14 March 2017 on Fundamental Rights Implications of Big Data: Privacy, Data Protection, Non-Discrimination, Security and Law Enforcement [2016/2225(INI)].* Online: www.europarl.europa.eu/doceo/document/TA-8-2017-0076_EN.html

Filippov, V. – Chursin, A. – Ragulina, J. – Popkova, E. G. (2019): *The Cyber Economy. Opportunities and Challenges for Artificial Intelligence in the Digital Workplace.* Cham: Springer Nature.

Foreign Policy (2020): *Global Data Governance. Part One: Emerging Data Governance Practices.* Online: https://foreignpolicy.com/2020/05/13/data-governance-privacy-internet-regulation-localization-global-technology-power-map/?utm_source=PostU

Fourkas, V. (2004): What Is 'Cyberspace'? *Media Development,* 3, 6–7.

Gill, I. (2020): *Whoever Leads in Artificial Intelligence in 2030 Will Rule the World Until 2100.* Online: www.brookings.edu/blog/future-development/2020/01/17/whoever-leads-in-artificial-intelligence-in-2030-will-rule-the-world-until-2100/?fbclid=IwAR3UJSDfUG4aFGXpHRDdch76HQu9ZgDIOQXGZo2t8iVjn-0HIQPNPnD42MM

Gray, C. S. (2013): *Making Strategic Sense of Cyber Power: Why the Sky Is Not Falling.* Carlisle Barracks SSI: US Army War College Press.

Hakmeh, J. (2017): *Cybercrime and the Digital Economy in the GCC Countries.* London: Chatham House.

IMF (2018): *Measuring the Digital Economy.* Washington: International Monetary Fund.

Krasznay, Cs. (2020): Case Study: The NotPetya Campaign. In Török, B. (ed.): *Információ és kiberbiztonság.* Budapest: Ludovika Egyetemi Kiadó. 485–501.

Latici, T. (2019): *Cyber: How Big Is the Threat?* Online: www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_ATA(2019)637980

Meyer, D. (2017): Vladimir Putin Says Whoever Leads in Artificial Intelligence Will Rule the World. *Fortune,* 4 September 2017. Online: https://fortune.com/2017/09/04/ai-artificial-intelligence-putin-rule-world/

Modern Diplomacy (2020): *Europe Fit for the Digital Age: Commission Proposes New Rules for Digital Platforms.* Online: https://moderndiplomacy.eu/2020/12/16/europe-fit-for-the-digital-age-commission-proposes-new-rules-for-digital-platforms/

Molnár, A. (2020): European Union – Cybersecurity. In Török, B. (ed.): *Információ és kiberbiztonság.* Budapest: Ludovika Egyetemi Kiadó. 437–457.

Riordan, S. (2018): The Geopolitics of Cyberspace: A Diplomatic Perspective. *Brill Research Perspectives in Diplomacy and Foreign Policy,* 3(3), 1–84. Online: https://doi.org/10.1163/24056006-12340011

Rodrigues, R. (2020): Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities. *Journal of Responsible Technology,* 4. Online: https://doi.org/10.1016/j.jrt.2020.100005

Ruhl, Ch. (2020): *Cyberspace and Geopolitics. Assessing Global Cybersecurity Norm Processes at Crossroads.* Washington: Carnegie Endowment.

Schaffer, A. (2021): The Cybersecurity 202: Legal Scholars Are Working on New Rules for International Hacking Conflicts. *The Washington Post,* 21 June 2021. Online: www.washingtonpost.com/politics/2021/06/21/cybersecurity-202-legal-scholars-are-working-new-rules-international-hacking-conflicts/

Sigholm, J. (2013): Non-state Actors in Cyberspace Operations. *Journal of Military Studies,* 4(1), 1–37. Online: https://doi.org/10.1515/jms-2016-0184

Stadnik, I. (2017): What Is an International Cyber Regime and How We Can Achieve It? *Masaryk University Journal of Law and Technology,* 11(1), 129–154.

Stempel, J. – Finkle, J. (2017): Yahoo Says All Three Billion Accounts Hacked in 2013 Data Theft. *Reuters,* 03 October 2017. Online: www.reuters.com/article/us-yahoo-cyber/yahoo-says-all-three-billion-accounts-hacked-in-2013-data-theft-idUSKCN1C82O1

Taulli, T. (2019): *Artificial Intelligence Basics. A Non-Technical Introduction.* Monrovia: Apress.

Teoh, C. – Mahmood, A. (2017): National Cyber Security Strategies for Digital Economy. *Journal of Theoretical and Applied Information Technology,* 95(23), 6510–6522. Online: https://doi.org/10.1109/ICRIIS.2017.8002519

Tsagourias, N. – Farrell, M. (2020): Cyber Attribution: Technical and Legal Approaches and Challenges. *European Journal of International Law,* 31(3), 941–967. Online: https://doi.org/10.1093/ejil/chaa057

Vihul, J. (2013): *The Tallin Manual on the International Law Applicable to Cyber Warfare.* Online: www.ejiltalk.org/the-tallinn-manual-on-the-international-law-applicable-to-cyber-warfare/