AI in the Technology Sector:
Dual and Dark side of AI

Alfredo M. Ronchi
Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, 20133, Italy
Tel: +39 393 0629373, Email: alfredo.ronchi@polimi.it , ORCID 0000-0003-1230-4338

Abstract: This document considers the dark side of AI and its dual use. Starting from the sunny side of artificial intelligence and key added values provided by this technology nowadays almost embedded in any sector of daily life. As it happens often there is a dark side of this technology starting from a broad range of misuse, scams and cybercrimes operated thank to AI. The second topic, tightly related to the dark side of AI is its dual use. Nowadays one of the warfare key actors together with cyberwar attacks is artificial intelligence. AI is empowering intelligence, situational awareness, risk assessment, decision making. Intelligent and autonomous weapons, and more. All these aspects not forgetting the sister side of physical warfare that is the information and communication side deeply influenced by mainstream information, opinion formation, nudging and fake news. Last but even more relevant the economic and financial side of "confrontations".

Keywords: Digital Transformation, E-Services, Human Rights, Laws, Regulations, Privacy, Cyber Ethics, Artificial Intelligence, Machine Learning.

AI for Good - the sunny side of AI

Since the early diffusion of Chat GPT citizens started to familiarize with this technology asking to write a letter in perfect business style or a poem in old fashioned language. Artificial intelligence was immediately classified by citizens as a personal skill enhancer, which enriches results and optimizes assignments, while students found it a valuable classmate that helps them with their homework. Consequently, the idea was to upgrade from traditional performances to "super" performances. So Medical Doctors empowered by AI will become Super doctors as well as Engineers will become Super Engineers.

If we focus on the topics of safety and security, AI is now improving its role even in these fields, from both single person and groups face recognition to interconnection of data / evidence collected from different data sets or OSINT advanced investigations.

As it happens probably since the beginning of the quest for safety and security an increased level of protection often involves a reduction of our basic rights, the consensus on a golden balance between human rights and protection is a paramount. Cyber technology positively impacted citizens' safety, CCTV, mobile phones and biometrics are usually the first line of defence in case of accidents and crimes, additional cyber tools are: car emergency calls, smart phones automatic calls and messages, wearable automatic push button alarm call systems[1], satellite position tracking and more. For quite a long-time kidnapping cases were solved thanks to the localisation of the victim thanks to mobile phones tracking. Similar positive considerations can be extended in case of disasters both natural and man-made, we remember the case of the earthquake in Napal and some alerts in case of potential Tsunami in the pacific

[1] Appreciated by women and elderly people, in case of danger it is enough to press the red button and an emergency call is activated from the geopositioned device.

area. Satellite images and communication systems together with recognition drones or first aid delivery drones are key tools in similar situations. As usual, these applications represent the positive use of technologies.

Moving the focus a little toward the dark side of these technologies and features, they can be used to address unauthorised surveillance, thanks to trojan codes and position trackers, or aimed attacks to the owner of mobile phones or communication devices, but this represents a dual use of technology. The dual use can cause the reduction of human rights and surveillance of citizens as foreseen in 2019 by Shoshana Zuboff's "The Age of Surveillance Capitalism" [5 - Zuboff 2019]. This last paragraph introduces the next topic, the dark side of AI.

AI for Bad - the dark side of AI

As many skills and tools, even AI is affected by a dark side, a misuse and, we will see later, a dual use of these technologies. As described in the previous paragraph concerning some professional profiles, AI boosted creativity in the field of crime will probably generate super criminals. Many times, crimes are not perpetrated using weapons and brute force, but thanks to the development of intelligent strategies, as it happens in social engineering or other scams. AI can be used maliciously affecting cybersecurity, physical security, political security and more.

We will start considering less evident malicious use of AI. Let's consider nudging and opinion dynamic manipulation, they can deeply influence public opinion.

If we consider "creativity" as a tool to shape opinion formation, as it may happen in the field of advertisement and marketing, we deal with a complex and dynamic process mediated by interactions among individuals both offline and online. Social media have drastically changed the way opinion dynamics evolve. Social media has become a battlefield on which opinions are exchanged, often violently.

The progress of AI has made it possible to develop much more powerful nudge mechanisms thanks to the effectiveness of statistical and inferential AI systems. The impact of AI powered technology on human autonomy is huge. AI-enhanced nudges reinforce the ability to achieve the designer goals using cognitive biases, emotional impulses, and other human behavioural mechanisms both intentionally and unintentionally.

In addition, we observe the massive decrease in the level of critical thinking and the emergence of waves of information epidemics, both at the national and global level (mainstream communication, limited contraposition, censorship, and fake news). These events are many times due to a well-defined and deployed manipulation based on AI applications.

Post-truth in its heyday, with public perception, shaped more by means of addressing feelings and predetermined opinions rather than facts, with fakes, click baits, hypes and other tools introduced to form post-reality in the political and media culture [1 – Joinson, 2004].

Post-reality is changing the system of values with the "new normal" (semantic shifts, etc.), of course may appear to be politically correct. New ethics is putting personal free will and freedom of choice under question; traditional cultural regulators of social relations and processes being displaced by automated social algorithms[2]. Widespread simplified virtual mock-ups and simulacra are not only blurring the borders between the real and the digital world but also led to a mass collection of data for managing people's behaviour[3]. It has also been

[2] increasing role of algorithms and ML.

[3] evaporation of privacy, data protection.

contributing to the increasing level of conflict in both the society[4] and among states, soft concerns and more [2 – Makkuni, 2018].

Moving now to the typical malicious use of technology we can focus on cybersecurity.
Malicious use of AI in cybersecurity can include AI powered cyber-attacks, human-like Denial of Service, AI powered malware, detect attack targets based on behavioural data.
Automated social engineering attacks via impersonation of thrusted users, automated phishing. Automated vulnerability discovery will be faster and … harder to be discovered.
Deep fake of voice and video presenting celebrities as testimonials of incredible financial opportunities or your "boss" to persuade you to share critical information. Soon our face can become a trigger for the execution of malware, or even ransomware activated when target found. Recently kidnappers found in generative AI a perfect companion to persuade parents. Why do not believe that it's reality?

Impacts on safety and security

When it comes to security, we look at it from the opposite perspective, focusing not on the positive aspects but on the risks of break-ins. Experts say that burglars don't view an apartment as a closed environment but as a series of access points or security breaches. In addition, the concept of "security" it is not an absolute and permanent status, but we can identify it as a "dynamic balance" between a specific "asset" or "assets" to be secured, the specific context, sometimes for a specific time span, the range of potential threats, and more. Once more "evil is in details".

The pervasive nature of "digital" many times contributes to improve resilience, but it can be the target for attacks and generate the "perfect storm". In the "analogue" world we had different pipelines and "channels" to perform, thanks to different tools and means, our activities, in the cyber world everything depends on a single "bottleneck": cyber technology. If this pillar will fail, malfunction or be switched off our life will suffer sometimes unpredictable problems, no "cyber" and consequently no digital identity, no bank account and social security, no service delivery, no news, and connection with other "entities". Well trained AI for Bad systems can generate the "perfect storm", this is one of the components of cyberwarfare. In recent times we witnessed different effects of such "cyber weapons" blocking different tiers of critical infrastructures, from energy to transportation and more. Time ago it was the use of backdoors in industrial components and PLCs to create big damages.

This tight and tightened cyber dependence from nation state level down to citizens if misused or abused can pose serious limitations to human rights. If some global services are switched off in some areas of the planet because of sanctions or our bank account or ID is blocked it becomes very hard to survive, with the switch from traditional current currency to crypto currency [2- ECB, 2025], as some leaders are suggesting, the dependence from cyber increases. AI can easily become the judge and jury; a selective restriction of services is very easy to be applied even as an action to put some pressure on citizens due to their behaviour or activity. Back to some of the topic outlined by Shoshana Zuboff [3 – Zuboff, 2019], the risk to be manipulated or controlled becomes reality.

[4] between individuals and groups – haters, discrimination.

We are surrounded by “critical infrastructures” managed by cyber components that, in case of attacks, may create mayor or minor impact on our daily life like information services, social media, geo-positioning, home automation, smart cities, safety, and security devices, and more. It will not surprise if in few years big service platforms as GAFAM will be considered critical infrastructures[5].
Recently on the World Economic Forum 2024, Jeremy Jurgens [6], foresaw the possibility a global cyber-attack that will take us back to the stone age. Of course, even this message might be a fake news. Nowadays the key concept is “holistic security”, a “global” approach to security integrating all the different aspects and problems.

Defence and Warfare

As it happened several times in the past due to the upgrade from bronze swards to iron swards or gun power weapons instead of arrows and spears, few years ago we reached a turning point in the definition of warfare. The broad dissemination of cyber technology and its increasing role in any sector of daily life, the improvements of autonomous devices, the relevant achievements of AI and ML profoundly reshaped the warfare scenario and related defence sector[7].

We can start a short analysis subdividing by two this scenario, on one side weapons and related resources, on the other side information and communication. Let’s start considering the latter, “war” communication as one of the side component of warfare, time ago, in the 1940s key actions were disclosed to citizens through war reporters shooting images on the battlefield or recreating such actions in iconic photos, think about “The Falling Soldier” photo by Robert Capa or Iwo Jima “Raising the Flag” photo by Joe Rosenthal[8]. Later the communication about conflicts was performed by TV news with images and short video clips on Corean and Viet Nam conflicts showing napalm bombing or last overloaded helicopters leaving Saigon.
Close to the end of the XXth century TV News impressed the audience showing almost real time night images on the air forces’ bombing actions on Bagdad. On the field actions started to receive main support from cyber mediated devices, digital cockpits views, head mounted displays and actions cameras, remote UAVs operators and automatic target finder “intelligent” weapons real time videos till the impact on target. These warfare story telling implemented and mediated by “cyber” reshaped the impact on citizens generating the impression to seamlessy something happening in a different and distant world. A relevant contribution to accept and familiarise with similar news was even due to action movies and video games sometimes anticipating technologies and combat scenarios.
Cyber technology and recently more specifically AI in the shape of opinion dynamics and nudge applications together with mainstream information represent a relevant resource in case of conflicts much more than in the past. This resource, including social media, is of course subject to manipulation from the counterparts thanks to hacking, trolls and AI ghost-authors, all these to do not refer to deep fakes.
Again, a tight control of news and information in the age of information flooding is thinkable mainly thanks to AI. Consequently, we will assist to a duel AI v/s AI.

[5] The one we know as GAFAM (Google, Amazon, Facebook/Meta, Apple, Microsoft) and NATU (Netflix, Tesla, Airbnb, Uber)

[6] "Geopolitical instability makes a catastrophic cyber event likely in the next two years": WEF Managing Director Jeremy Jurgens on the Global Cybersecurity Outlook Report 2023.

[7] E.g. US Cyber Command https://www.cybercom.mil

[8] https://www.gettyimages.it/immagine/battaglia-di-iwo-jima

Leaving the side of information and communication to face the weapon and resources side, on the 2018 NATO STO meeting held in Sofia (BG) it was presented the outcome of research entitled "XXIst Century cyberwarfare[9]" outlining the radical change of warfare model reshaping completely the traditional one and another concerning "soft concerns[10]". A typical example was a PLC hacker's attack on a well-known aircraft manufacturer, which was completely unprotected from similar attacks. To provoke the participants, a scenario was imagined in which a plane could be shot down using a smartphone.

Cyber technology and related applications including autonomous unmanned vehicles together with aimed hacking attacks and network services switch off are the new "wunderwaffen". Major part of the traditional equipment is outdated, even missiles and old-fashioned ICBM belong to the past expensive approach to warfare. Relatively small drones, especially if well-equipped by AI, can impact more effectively and are harder to be neutralised by air defence systems designed to fight traditional weapons. In 2022 some contributions to the NATO STO conference already foreseen swarm of drones reaching the attack area and assigning, by shared information, specific target to each of them. Drones can fly autonomously very close to the ground invisible to the scanning devices, land and wait, properly hidden, for the appearance of a relevant target.

As much as cyber is deeply rooted in our life as much as the warfare is completely reshaped in electronic countermeasures, unmanned vehicles, automatic weapons, and more but this is only the direct set of tools. As outlined before, direct tools represent only part of the warfare scenario, aimed communication, nudging, opinion manipulation, and of course a broad range of hacking activities aimed to impact any kind of service are a fundamental complement to depict the whole cyberwarfare. Once again, the dark side of AI can perform a well-structured set of aimed attacks potentially generating a perfect storm. On the opposite side real time trained AI cyber defence systems can intercept and deactivate such attacks and perform counter actions.
The tight dependence of our daily life from digital resources represents a critical weakness even accessible by hacking citizens. Hackers can interfere both with national and the private services, and commercial ones. To do not mention the direct will to switch them off by the providers if belonging to the counterpart. These actions can generate more impacting effects, simply considering actions impacting critical infrastructure, supply value chains or delivery facilities.

This relevant evolution of weapons and side components changed the relation between competing armies. Many times, the completely unbalanced military power forced some party to switch from traditional warfare to guerrilla or even terrorism. This is probably the extreme unbalance between costs and effects; one single person or explosive device can create a physical or psychologic impact far bigger than an expensive middle range missile or an ICBM.

This approach paved the way to a very complicated and hard to control situation leveraging on an immediate feeling of danger at citizen's level. In such a case, there is not an enemy country or official government, a complete reshape of defence measure become a must. A well-

[9] Ronchi A. (2018) - XXIst Century cyberwarfare, proceedings NATO SCO, International Journal of Information Security, ISSN 1615-5262 New York Springer-Verlag

[10] Ronchi A. (2021) - Soft but still concerns, proceedings International Conference on Homeland Security - Emerging Trends, Challenging Aspects, Gaziantep, Turkey, ISBN 978-605-67664-7-3, Hasan Kalyoncu Üniversitesi Yayınları

structured network of intelligence and information services, accurate borders control are the starting point. International level security information sharing, CCTV and biometric data analytics suspect individual shadowing, open-source intelligence together with big data analytics can be considered some of the available tools[11].
In addition, cyber technology and AI offered some potential contribution thanks to unmanned land vehicles and robots or unmanned aircrafts even having tiny dimensions.

XXth century weapons and means characterised by high technology and related high costs changed the structure of supply chains and financial side of conflicts leading to and unpair ratio between the economic value of the offense weapon and the cost of the counter action. As it happens if you launch a high-tech earth / air missile to take down a handmade rocket or self-build drone. At the beginning of the XXIst AI started to offer a key contribution to reduce the number of soldiers on the field in favour of newly developed technological devices.
These considerations on the different economic weight on the two actors take us to fucus on the evolution of nation state structure that pushed some countries to outsource defence materials providers to private companies while other countries kept this industrial sector as national patrimony. This flourished the creation of new private companies structured as defence sector providers and the issue by governments of call for bids. Private companies will act following business rules, sometimes competing other times in a kind of monopoly while state owned companies have usually limited competition and act in tight cooperation with government. Private enterprises are publicly traded, and many times shares are hold by investments funds, this involves they must generate profits. Consequently, they by default pursue profitable projects and add a revenue to the direct cost of each product. This, of course, it is not happening in case of state-owned companies.

Why dealing with AI and more in general with cyberwarfare we consider the economic aspect? Because the new scenario imposes the need to dramatically reduce the activation time of countermeasures and related costs. This paves the way to investments in research and development of low-cost measures to fight against cyber threats and neutralisation of low cost physical weapons paying a fraction of their costs. The idea that flourished in the 1980s, Strategic Defence Initiative (SDI) known as Ronald Regan's Star Wars foresaw, in addition to nuclear space weapons, the use of nuclear beam weapons and high energy laser weapons to destroy enemy's weapons with a minimum cost.

Due to the recent innovation trend, a swarm of drones controlled by interconnected AI "pilots", yesterday painting in our sky wonderful shapes dynamically changing every minute, can become an intelligent swarm of killer weapons targeting specific objects and infrastructures.
Relevant industrial production exhibits in recent times are full of any kind of unmanned vehicles (land, sea, air) and electronic countermeasures. Small trucks are on exhibit as take-off and landing areas for drones. Cyberwarfare and cybersecurity are intertwined. After different decades of research and development in the field of anthropomorphic and zoomorphic robots now dogs shaped, human shaped robots and more are on stage both with and without weapons. Are we getting toward a "Terminator" scenario populated by robots and cyborg?
One of the key problems in the past was information sharing and interoperability of information systems, this even in the field of communication, wired, wireless, different bandwidth and protocols. Cyber technology nowadays offers AI based software applications to support

[11] Ronchi A. (2018), TAS: Trust Assessment System, International Journal of Information Security, ISSN:1615-5262 vol. vol.39, and TAS: Risk Analysis & Clustered Sensors, IST-Africa 2018, ISBN:978-1-905824-59-5 vol. 2018

decision making in real-time collecting and integrating information from different sources: networks, satellites, drones, land recognisance, and more. These systems privately connected empower any action.

One of the relevant actors in the field of AI based decision support systems is Palantir[12] a US company founded in 2003 in Denver, with the support of different agencies. Palantir emphasizes that it is **not a data company**, but rather a provider of **data integration and privacy-preserving tools**. Their customers often make **critical or life-saving decisions**, so Palantir ensures:

- **Purpose-based access controls** to restrict data usage to specific, justified purposes.
- **Beyond-anonymization techniques** to minimize re-identification risks.
- A **framework combining powerful data tools with robust privacy controls**.
- **Transparency** for users into data origins, preparation steps, and processing pipelines.
- **Privacy protection by design**, including features for **data deletion** to meet regulatory and customer requirements.

Considering this complex scenario nowadays filtering direct human responsibilities through autonomous weapons or cyber-mediated "videogames" push Ethics principles far from decision makers, actors and the battlefield. Even what it was considered unforeseeable simply fifteen years ago, like a nuclear conflict, it is now a potential option on the table. The full respect of human rights and ethics must guide the research and development in the civil use of AI and cyber technologies in general, the same must guide the developments in the defence sector. Tight competition generating a never-ending sequence of measures and countermeasures must comply to general ethics principles and human rights.

To conclude

Starting from the actual "fluid" context as the result of a mix of different crisis the evolution of the impacts on different sectors is shortly depicted. The impact on society due to the acceleration of the digital transition during the pandemic. The desire of decision makers to go digital, sometimes forgetting some wise principles. The goal to digitise as much as possible reaching a cyber-based society relaying on "digital", this pillar is quite fragile, potentially subject to attacks and suitable for top-down discrimination. The Internet distributes all-over the world "homogenised" content that can jeopardise cultural identities. Additional potential drawbacks due to lives spent in cyber-bubbles, cyber-mediation of human relations, citizens experience the world thanks to a cyber device mediated approach, mainstream influence on opinion dynamics and nudging. Cyber-loneliness, one of the foreseeable risks is a kind of addiction to this "parallel life" training users to shift from real to meta-life blurring the border between them.

If we focus on defence, it is more than evident that this sector is under deep transformation or better, revolution. Traditional approach and methodologies become suddenly outdated since cyber technology took a prominent position in this field. Apart from the relevant economic and production gap due to state owned hardware and software providers and private ones, the use of UAVs and intelligent weapons characterised by limited cost compared with traditional weapons tilted the table.

Ethics plays a key role on both scenarios both security and defence. The challenges for the upcoming years are the ways to sustain the human's role and the inviolable right to freedom and personal privacy in an era of unlimited information gathering. Once again, the need to find a proper balance between humanities and technologies is omnipresent. Social sciences and

[12] https://www.palantir.com

humanities must establish a tight cooperation in designing or co-creation of cyber technologies always keeping humans in the focus.

References

[1.] Joinson Adam (2004) Internet Behaviour and the Design of Virtual Methods, http://www.joinson.com/home/pubs/02%20Virtual%20methods021-34.pdf

[2.] Ranjit Makkuni, (2018), Betrayed IT Revolution, https://www.ranjitmakkuni.world/betrayal-of-the-it-revolution

[3.] European Central Bank (2025) - https://www.ecb.europa.eu/press/pr/date/2025/html/ecb.pr250716~463e72bbcb.en.html

[4.] Zuboff Shoshana (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, ISBN 9781781256855, Profile Books