LEGAL ASPECTS ON DEVELOPMENT AN USE OF ARTIFICIAL INTELIGENCE

Senior Researcher Mihai-Ștefan DINU, PhD [*]
(Academy of Romanian Scientists, 3 Ilfov, 050044, Bucharest, Romania, email: secretariat@aosr.ro)

***Abstract**: The unprecedented development of new technologies, related with the popularity of smart devices in human activity regardless of field, social level or education, has produced a gap in the regulated use of both hardware and emerging software products. This gap was extremely visible in the cleavage that began to manifest itself in increasingly differentiated behaviors in the two related spaces: the physical, geographical and geopolitical space on the one hand, and the virtual, electromagnetic, cyberspace on the other hand.*

The traditional order, taxonomies and values of the physical, geographical and geopolitical supported by legal regulations whose purpose was to obtain a behavior committed to achieving social peace with the support of the coercive force of the state, were besieged by chaotic, often absurd disinformation trends from a poorly regulated cyberspace, which allowed the actors involved to destabilize social peace and, most often, the power of the state through attacks on its fundamental institutions. This paper addresses aspects that can correct, through legal regulations, the current dissonant trends that manifest themselves from cyberspace to the physical one, even daring to state that the comprehensive regulation of behavior in cyberspace is the key to regaining social peace, respecting the social contract and treaties, conventions and agreements on the international stage.

***Keywords**: legal, politics, artificial intelligence, cyber space, regulations, European Union, USA.*

DOI **10.56082/annalsarscimilit.2025.2.109**

1. Introduction

For more than a decade, the issue of conceptual consolidation in the cyber domain has been of major interest, with the results of debates, both academic and industrial, continuously being in the focus of military specialists. This interest has increased with the recognition of cyberspace as a military operational environment, alongside land, naval, air, and outer space. This approach to cyberspace is officially mentioned in the US Department of Defense's Cyberspace Operations Strategy, which, in the first of five proposed strategic initiatives, states that the Department of Defense

[*] Mihai-Ştefan DINU, PhD, Senior Researcher, National Defense University CAROL I, Associated Member of Academy of Romanian Scientists, e-mail: mihaistdinu@yahoo.co.uk.

will treat cyberspace as an operational domain to organize, train, and equip. The declared approach to cyberspace, from a military point of view, as the fifth operational environment, was not without challenges that foreshadowed a new era, of the militarization of cyberspace and the expansion of the areas of technical surveillance of the big brother, eye in the sky type, all the more so since intelligence agencies seemed to be found in this space, an extremely prolific one in cultivating various illicit behaviors against national security. In addition, time has proven that the limit between the consolidation of national security and the integrity of rights such as the right to privacy and intimacy, was as fragile as it was flexible. Some types of behavior of this type of agencies, or even of companies wishing to expand commercial markets, led to the adoption by states or international governmental organizations of legislative norms that would protect the right to privacy and intimacy not only in cyberspace but also in the field of electronic communications. The regulation of personal data protection by the Council of the European Union was another step forward which, together with the NIS Directive, brought consolidation in the sense of achieving lawful behavior of all actors carrying out activities in and through cyberspace. Further developments such as election interference, disinformation, even health affecting issues lead to approval, last year (with the gradually entering into force from 2024 until 2027), of The EU Artificial Intelligence Act.

2. What is really happening?

The experts in cyber security predicted[1] for the second year in a row as the main threats issues related to artificial intelligence: AI phishing, deepfakes, hallucinations, use of a shadow (unauthorized deployment of AI), exploits and running the malware (along with a chat bot), prompt injections (especially on LLMs) and even the use of AI in cybersecurity processes.

The development of social media networks and the large scale (globally perceived) of membership to every one of them built a lot of bridges which in fact is a positive thing if we have in mind words like collaboration integration, cooperation, scientific and economic development. All these were supposed to increase levels in education, commerce, diplomacy and stay updated with a lot of activities which took place in the wide real world. The development of such diverse numerous social media networks lead to a solid perception that everyone is really connected to everyone and that way with the global reality. First 20 years of the new millennia bring incipient forms of social media with Friendster,

[1] Sue Poremba, *Cybersecurity trends: IBM's predictions for 2025,* available at https://www.ibm.com/think/insights/cybersecurity-trends-ibm-predictions-2025, accessed on 12.03.2025.

LinkedIn, My Space it was the enthusiastic period that create the space for the upcoming social media owned by the nowadays high-tech giants: Facebook, YouTube, Twitter, Instagram and TikTok. From individual end-users to big corporations, everyone was claiming their space and role in this new virtual world. The prior friendly spaces were transforming by the addition of advertising campaigns. This was the moment when everybody saw the way to maximize their profits: from the content creator, to companies and influencers. The social media platforms developed helpful tools in order to monetize the content, and also get huge profits. Thus the fine line between group or individual opinion and real information became thinner and thinner so that disinformation got a heavenly environment, pushed by the specific algorithms every social media platform used in order to promote its commercial interest. One of the major steps in this direction was initiated in 2009 once Facebook use their ranked and personalized news feeds. This changed not only the way users interact with the online platform but the way users interact on the platform. Approximately 15 years later, Elon Musk, owner of X (formerly known as Twitter) expressed his opinion[2] based on the free speech right that the platform will be perceived as a *`digital town square`*. The major issue is that we reach the digital town square under the strict direction of the algorithms. And this somehow contradicts the concept of free exchange of ideas as it was in its beginning enthusiastic era, more than two decades ago. Some opinions are over represented, other rather blocked based on the algorithms preference, so we can say that the social media platforms promote the freedom to speak but models the way speech is heard. Visualization of every single post is largely determined by algorithms, so the communication process is not totally controlled by the speaker and its audience and facilitated only by technology. Thus the idea of disinformation spreading, supported by the social media platforms became increasingly present in society which directed to governments and legislators such worrying aspects. Issues regarding the transparency of business models of social media corporations and the algorithms that encourage the spread of harmful content, contributing to a wider social dangers, were debated at national level but also international level, understanding the fact that the continuous development of technological tools , especially the development and use of artificial intelligence which deployed at the large scale on the devices used by various users, no matter their level of knowledge, education or intent of use can result in negative effects in the level of social peace.

The role that can be played by the generative artificial intelligence and large language models in the creation and spread of misinformation,

[2] Bobby Allyn Elon Musk bought Twitter. Here's what he says he'll do next, available at https://www.npr.org/2022/04/25/1094671225/elon-musk-bought-twitter-plans, accessed on 15.03.2025.

disinformation and harmful content is more and more perceived in our societies as its deployments are more varied and the well-intended ways of use can be rapidly overthrown by various malicious actors. Thus, the intend to vastly improve the overall human activities that AI can bring by making our machines smarter in order to get increasingly relevant and accurate predictions or scientific discoveries, can be diverted so the give superpowers to not very well intended actors.

3. Why is this happening (with a great sense perceived chaos)?

Despite numerous intensions of national or international authorities to regulate the use of algorithms and AI technology, this effort remains somehow unperceived in its effect. And there is a lot of reason in this state of fact. In order to find the key needed to resolve the problem we need look back to the root that is the various spaces in which human activities take place. We already mentioned in the opening of this paper the fact that military operations just added *cyberspace* as the fifth operational environment. The other four operational environments are:

- land;
- maritime;
- air;
- space.

All those four belong to the *physical space.* Our opinion is that the weak understanding of how the algorithms and AI used in cyberspace products such as social media platforms affect reality consist in the differences between the two (cyber and physical) spaces. Considering this kind of products in cyberspace we can set a taxonomy of cyberspace on two criteria; *nature of digital settings and the way it is used*. Thus, we can distinguish between five types:

- Public cyberspace (public networks and public information exchange);
- Private cyberspace (private networks and encrypted communication channels);
- Social cyberspace (a subcategorization of public cyberspace, social media platforms, forums, online communities etc.);
- Commercial cyberspace (digital marketplaces and financial systems);
- Military cyberspace (a very secure environment used by military, security and defence communities).

This taxonomy provides us elements to realize the interconnection between the cyberspace and physical space but we must be aware that they are two very different extents that are depicted by five characteristics:

- Nature (virtual world and real world);
- Interactions (digital interactions through diverse devices and physical interactions through physical objects);
- Safety and security (securing digital assets, data, networks etc. and physical security of human individuals, communities and their physical infrastructures);
- Threats (cyber threats like ransomware, network penetration or data exfiltration versus classic threats like natural calamities, war, aggressive human behavior etc);
- Regulation (cyberspace still needs comprehensive, universal norms of behavior of the actors implied versus physical space that is normed under local, national and international law).

Analyzing these two taxonomies it is obvious that one of the greatest issue is regulation of behavior in cyber space after the real world model. Although some guides were released by some major international organizations, in 2008 European Union states signed the Budapest convention on cybercrime, *Tallin Manual* was released by academics and NATO experts and majority of the states released a National cybersecurity strategy, the legal domain tend to remain somehow tricky, without major influence on the effects that algorithms and AI can result in real world. Major steps were taken in European Union by the legislation proposed and approved by the European Council: GDPR, NIS, European Digital Act etc.

Various guides were released regarding misinformation, disinformation, hate speech, harmful content, protection of children, measures to counteract the influence of malicious actors on national election process etc. Various guides were released at international scale regarding misinformation, disinformation, hate speech, harmful content, protection of children, measures to counteract the influence of malicious actors on national election process etc. Most of this digital harmful behaviors for the real world society are founded on the way the algorithms of the social media platforms promote content posted by the users.

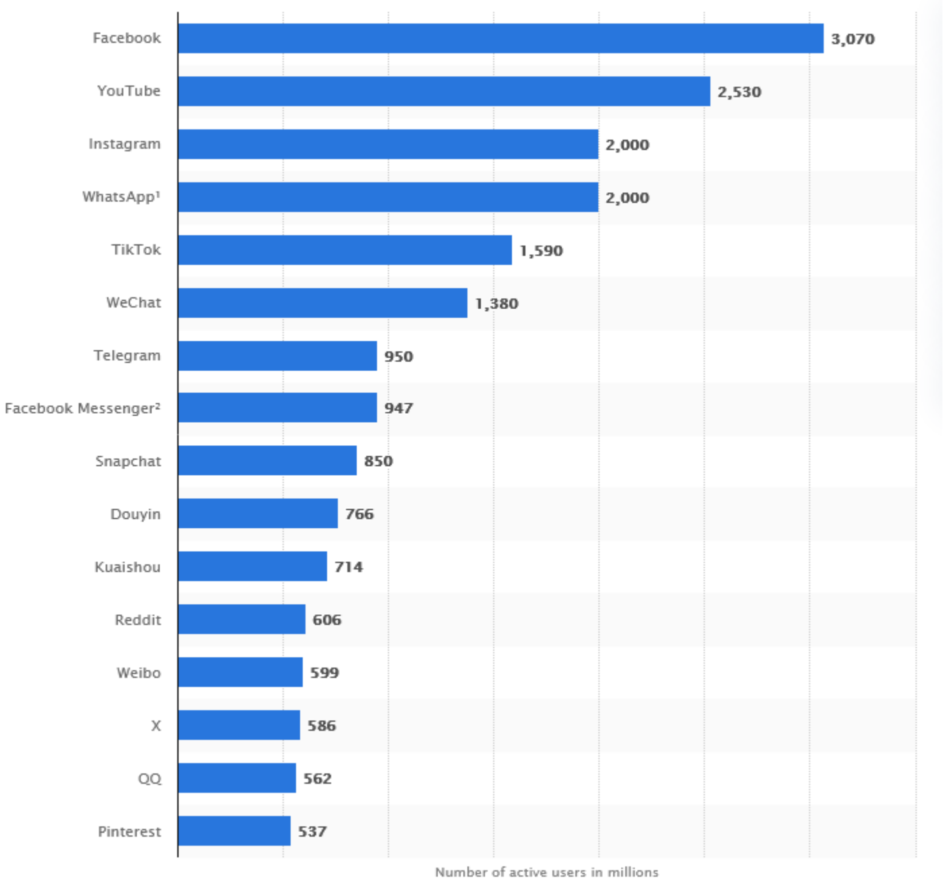


Figure no. 1 – Most popular social networks worldwide as of February 2025, by number of monthly active users[3].

The vast amount of data generated by the social media users nowadays is difficult to manage without the support of AI tools that enhance social media user experience. Algorithmic bias that can occur may result in unbalanced promotion of user posts that sometimes spread harmful content. The fact that humans inclined to be social learners until the emergence of social media platforms, looking around and learn from other people are doing lead to a confusion due to this new not well charted and not well enough regulated, so that the user of social media platforms tend to see different harmful content as a model and to learn from it. Adding the power of dissemination of that kind of content of hundreds of thousands or even millions or hundreds of millions, can create a different separate world whose norms of behavior end up diverging from those of the real world affecting the normal, legal order. In this case regulating the activity of the

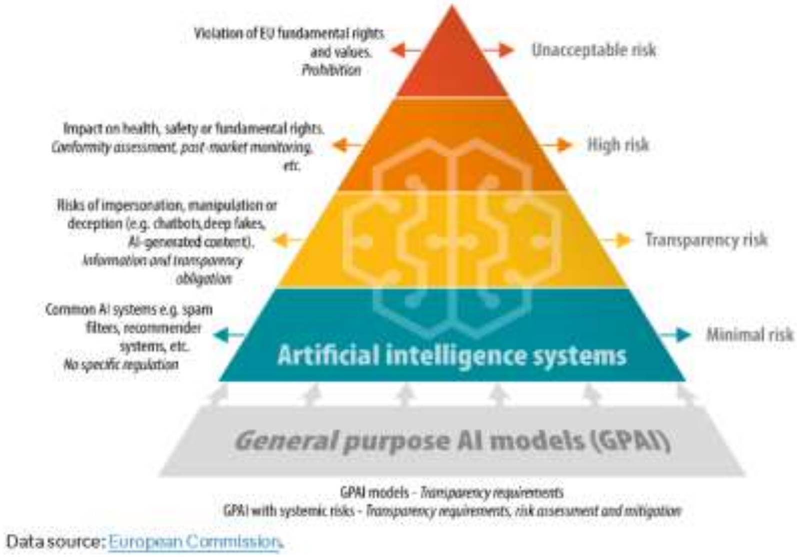
[3] ***, Statista, available at https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/, accesed on 14.03.2025.

social media platforms, which until now seemed to not achieve the desired results, must be strengthened by the regulation of tools that the social media platforms use. And the AI is a good starting point.

4. How can this be stopped (or, at least permitting a sense of perceived order/social peace)?

In april 2021, as part of its digital strategy, the European Commission advanced a proposal in order to regulate AI and ensure the legal framework for the development and use of this technology. Considering the plethora of benefits in domains as healthcare, transportation, energy etc., the European Parliament wanted to make sure ` *that AI systems used in the EU are safe, transparent, traceable, non-discriminatory and environmentally friendly. AI systems should be overseen by people, rather than by automation, to prevent harmful outcomes.* `[4]

In order to achieve its declared objectives EU proposed a solid `technology-neutral, uniform definition for AI that could be applied to future AI systems`[5], and an AI risk taxonomy , the legal norms being applied specific to the level of risk of AI risk qualification.


Violation of EU fundamental rights and values.
Prohibition
Unacceptable risk
Impact on health, safety or fundamental rights.
Conformity assessment, post-market monitoring, etc.
High risk
Risks of impersonation, manipulation or deception (e.g. chatbots,deep fakes, AI-generated content).
Information and transparency obligation
Transparency risk
Common AI systems e.g. spam filters, recommender systems, etc.
No specific regulation
Minimal risk
Artificial intelligence systems
General purpose AI models (GPAI)
GPAI models - *Transparency requirements*
GPAI with systemic risks - *Transparency requirements, risk assessment and mitigation*

Data source: European Commission.

Figure no. 2 - AI systems' risk categories[6]

[4] European Parliament, *EU AI Act: first regulation on artificial intelligence*, available at https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, accessed on 16.03.2025.

[5] *Ibidem*

[6] Tambiama Madiega, *Artificial intelligence act*, European Parliamentary Research Service, available at https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/-698792/EPRS_BRI(2021)698792_EN.pdf, accessed on 16.03.2025.

Defining the AI systems as a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments; EU includes AI in four risk classes, as follows[7]:

1. Unacceptable risk - All AI systems considered a clear threat to the safety, livelihoods and rights of people are banned. The **AI Act prohibits eight practices**, namely:

- harmful AI-based manipulation and deception
- harmful AI-based exploitation of vulnerabilities
- social scoring
- Individual criminal offence risk assessment or prediction
- untargeted scraping of the internet or CCTV material to create or expand facial recognition databases
- emotion recognition in workplaces and education institutions
- biometric categorisation to deduce certain protected characteristics
- real-time remote biometric identification for law enforcement purposes in publicly accessible spaces

2. High risk - AI use cases that can pose serious risks to health, safety or fundamental rights are classified as high-risk. These **high-risk** use-cases include:

- AI safety components in critical infrastructures (e.g. transport), the failure of which could put the life and health of citizens at risk
- AI solutions used in education institutions, that may determine the access to education and course of someone's professional life (e.g. scoring of exams)
- AI-based safety components of products (e.g. AI application in robot-assisted surgery)
- AI tools for employment, management of workers and access to self-employment (e.g. CV-sorting software for recruitment)
- Certain AI use-cases utilised to give access to essential private and public services (e.g. credit scoring denying citizens opportunity to obtain a loan)
- AI systems used for remote biometric identification, emotion recognition and biometric categorisation (e.g AI system to retroactively identify a shoplifter)

[7] European Commission, *AI Act*, available at https://digital-strategy.ec.europa.eu/-en/policies/regulatory-framework-ai, accessed at 17.03.2025.

- AI use-cases in law enforcement that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence)
- AI use-cases in migration, asylum and border control management (e.g. automated examination of visa applications)
- AI solutions used in the administration of justice and democratic processes (e.g. AI solutions to prepare court rulings)

High-risk AI systems are subject to **strict obligations** before they can be put on the market:

- adequate risk assessment and mitigation systems
- high-quality of the datasets feeding the system to minimise risks of discriminatory outcomes
- logging of activity to ensure traceability of results
- detailed documentation providing all information necessary on the system and its purpose for authorities to assess its compliance
- clear and adequate information to the deployer
- appropriate human oversight measures
- high level of robustness, cybersecurity and accuracy

3. Limited risk - refers to the risks associated with a need for transparency around the use of AI. The AI Act introduces specific disclosure obligations to ensure that humans are informed when necessary to preserve trust. For instance, when using AI systems such as chatbots, humans should be made aware that they are interacting with a machine so they can take an informed decision. Moreover, providers of generative AI have to ensure that AI-generated content is identifiable. On top of that, certain AI-generated content should be clearly and visibly labelled, namely deep fakes and text published with the purpose to inform the public on matters of public interest.

4. Minimal or no risk - The AI Act does not introduce rules for AI that is deemed minimal or no risk. The vast majority of AI systems currently used in the EU fall into this category. This includes applications such as AI-enabled video games or spam filters.

5. Some brief conclusions (or the urge to revisit a 100-year-old sociological work of reference)

The last 20 years of interaction on the international scene seemed to be characterized by high grades of connectivity that could boost cooperation and superior standard of living along with increased chances of transparency throughout various communities, nations and organizations. The last three years seemed to contradict somehow the prior tendencies: a great amount of manipulation, bloody armed conflicts, public riots, unlawful extremist political activity, all these disseminated through the amplifying power of social media platforms. All these remind of two major events in the last century:

- Titanic sinking more than a century ago, and the legal norms instituted by the Radio Act of 1912, act that could prevent the sinkink of Titanic[8];
- The sociological work published by José Ortega y Gasset, under the title of *The Revolt of the Masses.* The author described a mass-man transformed by the abundance provided by the growing standard of living and tend to occupy the spaces of the modern society with all its benefits but does not realize that all that benefits means respect for the values of the institutions that created them, and, generally, disrespect the legal framework that permitted the present standard of living.

Taking advantage by the way our conclusion developed, we would like to recommend one of the recent study showing that AI systems can autonomously develop social conventions without explicit programming and have implications for designing AI systems that align, and remain aligned, with human values and societal goals.[9]

BIBLIOGRAPHY

***, Statista, available at https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users;

ALLYN B., Elon Musk bought Twitter. Here's what he says he'll do next, available at https://www.npr.org/2022/04/25/1094671225/elon-musk-bought-twitter-plans;

ASHERY A-F., AIELLO L-M., BARONCHELLI A., Emergent social conventions and collective bias in LLM populations, in Science Advances, vol 11, No. 20, published by American Association for the Advancement of Science, May 2025;

BOSS A., NIST and the Titanic: How the Sinking of the Ship Improved Wireless Communications for Navigating the Sea, available at https://www.nist.gov/blogs/taking-measure/nist-and-titanic-how-sinking-ship-improved-wireless-communications-navigating;

[8] Alex Boss, *NIST and the Titanic: How the Sinking of the Ship Improved Wireless Communications for Navigating the Sea,* available at https://www.nist.gov/blogs/taking-measure/nist-and-titanic-how-sinking-ship-improved-wireless-communications-navigating, accessed on 17.03.2025.

[9] Ariel Flint Ashery, Luca Maria Aiello and Andrea Baronchelli, *Emergent social conventions and collective bias in LLM populations*, in Science Advances, vol 11, No. 20, published by American Association for the Advancement of Science, May 2025 .

MADIEGA T., Artificial intelligence act, European Parliamentary Research Service, available at https://www.europarl.europa.eu/RegData/-etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf;

NEUWIRTH R-J., *The EU Artificial Intelligence Act< Regulating Subliminal AI Systems,* Routledge, London and New York, 2023;

POREMBA S., Cybersecurity trends: IBM's predictions for 2025, available at https://www.ibm.com/think/insights/cybersecurity-trends-ibm-predictions-2025;

European Parliament, EU AI Act: first regulation on artificial intelligence, available at https://www.europarl.europa.eu/topics/en/article/-20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence;

European Commission, AI Act, available at https://digital-strategy.ec.-europa.eu/en/policies/regulatory-framework-ai.