Teka Komisji Prawniczej PAN Oddział w Lublinie, vol. XIV, 2021, no. 2, pp. 503–513
https://doi.org/10.32084/tekapr.2021.14.2-36

THE EUROPEAN UNION'S SOVEREIGNTY OF ARTIFICIAL INTELLIGENCE (AI)

Dr. Agnieszka Wilk–Ilewicz

Department of Administrative Law and Public Policy Science, Faculty of Administration and Social Sciences, Warsaw University of Technology, Poland
e-mail: agnieszka.ilewicz@pw.edu.pl; https://orcid.org/0000-0002-5930-5728

Abstract. When assessing the regulatory activities of the European Union, it can of course be pointed out that the entire legislative process is running too slowly. However, I have the impression that the proposals for "controlling" AI may finally be successful. Without the implementation of a holistic solution, in which the framework of conduct / ethical rules will be imposed on the creators and the validation of AI systems by the state will be introduced, one cannot speak of any sovereignty. It seems that only such a duopoly can lead to the use of brilliant AI solutions, minimizing the risks associated with it. However, without strong state organs, this process will not be adequately secured. This article proves that sovereignty over AI seems to be a *sine qua non* condition for us to be safe in the understanding of dominating processes.

Keywords: artificial intelligence, validation of AI systems, responsibility for AI, security of AI implementation, the European Union's sovereignty

INTRODUCTION

The European Commission is working on the artificial intelligence (AI) regulations that will form the legal basis of its technological, ethical, legal and socio-economic framework. The European Parliament puts emphasis on the "human-centered" European values as well as the contribution of AI to the revival of the economy. The European approach to the Artificial Intelligence aims to promote Europe's innovation capacity in the field of artificial intelligence, while supporting the development and use of ethical and trustworthy artificial intelligence throughout the EU economy. Artificial intelligence should act on people and be a force for good in society.[1]

[1] See https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52020DC0065&from=EN [accessed: 30.04.2021].

1. TO TAME THE UNBRIDLED

For 70 years, Artificial Intelligence (AI) has been the subject of a wide ranging scientific research, in the recent years the interest in AI has intensified and its application spread across many new areas including the field of legal sciences.

Artificial intelligence is perceived as an important link at the start of the fourth technological revolution, that will introduce breakthrough changes in most areas of the economy. The undisputed weakness of the EU proposals/ directives for a legal framework across the EU countries is treating AI as a monolithic entity, while the specific examples of AI applications are considered from the perspective of the existing legal regulations. It seems that at present the axis of the analysis is focused on the reference of statutory law norms to technical applications that already exist – is it a good solution to adjust the existing legal regulations to the changing reality? [Rojszczak 2019, 1–23]. Much less attention is paid to the attempt to search for the direction of changes for the entire legal system in such a way that legal norms serve to proactively shape the industry that is just emerging. The question is whether such proactive action would not guarantee that the technology, which is evolving and probably will have a huge impact on people's lives and the functioning of entire societies, will be created from the beginning taking into account the key regulations and ethical principles underlying modern civilization? Or maybe we should implement holistic regulations, general principles defining the legal environment for the functioning of artificial intelligence, taking into account the interoperability requirements of systems, tools and services? However, whatever system we adopt (horizontal or sectoral), it seems necessary to create procedures for verification, validation and control of artificial intelligence systems based on a wide range of security and transparency standards. The issues of accountability and accountability appear to be the most serious legal issues related to AI.

Undoubtedly, the European Union has recently been striving to take control of the rapidly developing artificial intelligence. After years of lack of commitment in this regard, he seems to be moving from the observer's position to the role of the creator of duties. The activity of the European Union in the recent period has been defined more by a number of recommendations, resolutions, opinions, positions[2] than by hard legal provisions. The presented solutions

[2] Resolution of 16 February 2017 with recommendations to the Commission on civil law on robotics, resolution of 1 June 2017 on the digitization of European industry, resolution of 12 September 2018 on autonomous weapons systems, resolution on 12 February 2019 on a comprehensive European industrial policy on artificial intelligence and robotics, Commission Communication of 25 April 2018 on Artificial Intelligence for Europe (COM (2018) 0237), Commission Communication of 7 December 2018. On a Coordinated AI Plan (COM (2018)

suggest imperative actions initiated by the institutions of the European Union. They are unequivocally based on the inequality of entities, expressed in the possibility of shaping the situation of another entity by the European Union, regardless of its will, but in accordance with the subject law – hence the title of the term of sovereignty.[3]

2. MAKING MACHINES INTELLIGENT

In 1968, Marvin Minsky said, AI is "the science of making machines that would require intelligence if made by humans." Thus, all intelligent behavior belongs to the realm of AI, including playing chess, solving calculus problems, making mathematical discoveries, understanding stories, learning new concepts, interpreting visual scenes, "diagnosing disease," and analogy reasoning. As indicated, artificial intelligence is realized for at least two reasons: understanding how human intelligence works and creating useful computer programs and computers that they can perform intelligently [Rissland 1990, 1957–981].

Fully autonomous artificial intelligence systems such as robots are constantly featured in various science fiction movies and books, and therefore reach the minds of the vast majority of people in the world [Naučius 2018, 113–32]. However, if creating these entities is one important task, another key goal is to determine the future legal status of fully autonomous AI entities.

Computer scientist of Stanford University, Nils Nillson, identifies the concept of artificial intelligence as "an activity devoted to making machines intelligent, and intelligence is the quality that allows an entity to function properly and be farsighted in its environment. Therefore, it is clear that AI is some kind of being made by humans and capable of performing certain tasks while being environmentally friendly" [ibid.].

0795), Commission Communication of 8 April 2019 on Building Trust in Human-Centered AI (COM (2019) 0168), Commission White Paper of 19 February 2020 on Artificial Intelligence – A European Approach to Excellence and Trust, Commission report of 19 February 2020 on the impact of Intelligence, Internet of Things and Robotics on Security and Accountability, European Parliament STOA Policy Briefing of June 2016 on Legal and Ethical Reflections on Robotics, Report of the High Level Expert Group on Artificial Intelligence of 8 April 2019 "Ethical Guidelines for Trustworthy Artificial Intelligence," Report of the High Level Expert Group on Artificial Intelligence of 8 April 2019 entitled "The Definition of Artificial Intelligence: Main Opportunities and Disciplines," Report of the High Level Expert Group on Artificial Intelligence of 26 June 2019 entitled "Policy and Investment Recommendations for Trustworthy AI," Report of the Expert Group on Liability and New Technologies – Formation of New Technologies of 21 November 2019 entitled "Responsibility for Artificial Intelligence and Other New Digital Technologies."

[3] More, for example, see Radziewicz 2005.

3. RISKS ASSOCIATED WITH AI

In the face of "making machines intelligent," the intervention of EU legislators seems necessary. In addition to consistency with existing law, it is imperative to ensure a transnational understanding of basic data economy ideas. Indeed, the digital revolution is forcing all of us, scientists and practitioners, to understand and reconsider how traditional concepts and legal principles can be adapted to new scenarios that will become science fiction [Fradera 2018, 707–12].

In the White Paper of the European Commission, Artificial Intelligence – A European approach to excellence and trust,[4] it is pointed out that as digital technology becomes an increasingly central part of every aspect of people's lives, people should be able to trust it. Credibility is therefore the basic condition for its acceptance. The document sees AI as an opportunity for Europe given […] "proven ability to create safe, reliable and sophisticated products and services, ranging from aeronautics to energy, automotive and medical equipment."

The document defines AI as a set of technologies that combine data, algorithms and computing power. Advances in computer science and increasing data availability are therefore the main drivers of the current rise in artificial intelligence. An artificial intelligence ecosystem could develop which brings the benefits of this technology to European society and the economy as a whole. It is noted that it is extremely important for European AI to be based on our values and fundamental rights, such as human dignity and the protection of privacy. And the impact of AI systems should be considered not only from an individual perspective, but also from the perspective of society as a whole. The use of artificial intelligence systems can play a significant role in achieving the SDGs and in supporting the democratic process and social rights.

As the European Commission points out, the main risks associated with the use of artificial intelligence relate to the application of principles aimed at the protection of fundamental rights (including the protection of personal data and privacy and non-discrimination), as well as security and liability issues. The use of artificial intelligence can affect the values on which the EU is founded and lead to a violation of fundamental rights, including the right to freedom of expression, freedom of assembly, human dignity, non-discrimination on the basis of sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation, as applicable in certain areas, protection of personal data and private life, or the right to an effective remedy and to a fair trial, and consumer

[4] See https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52020DC0065&from=EN [accessed: 30.04.2021].

protection. The document raises extremely important issues that these threats may result from flaws in the overall design of AI systems (including with regard to human surveillance) or the use of data without correcting possible bias.

The European Parliament is also speaking in a similar vein.[5] Importantly, Parliament points to the need for regulation at the level of a regulation, not a directive. It is necessary to introduce a uniform regulation throughout the European Union, due to the specific features of IS, such as: complexity, connectivity, opacity, vulnerability, the ability to change through updates, the ability to learn, autonomy, and finally the multiplicity of entities involved.

The specific characteristics of many AI technologies, including the lack of transparency, complexity, unpredictability and partially autonomous behavior, may make it difficult to verify compliance with applicable EU law and may hamper effective enforcement to protect fundamental rights. Enforcement authorities and individuals may not have the means to verify how the decision was made with AI, and therefore whether the relevant regulations were complied with. Natural and legal persons may find it difficult to effectively access justice where they may be adversely affected by such decisions.

The lack of clear safety rules on these risks can, in addition to the risks for the people concerned, create legal uncertainty for companies that sell their products using AI in the EU. Market surveillance and enforcement authorities may find themselves in a situation where they are unsure whether they can intervene because they may not be empowered to act and / or not have the appropriate technical capacity to inspect systems. Legal uncertainty can therefore lower the overall level of safety and undermine the competitiveness of European businesses. If security threats do materialize, the lack of clear requirements and features of the AI technology mentioned above makes it difficult to trace potentially problematic decisions made with the involvement of AI systems. This, in turn, can make it difficult for those who have suffered damage to obtain compensation under applicable EU and national liability rules.

Considering these threats indicated by the European Commission, it seems necessary to reduce the above-mentioned dangers, it is necessary to introduce solutions binding both the AI creators / engineers and state authorities in the process of systems admissibility / certification. It seems that only an attempt benefiting from the AI achievements.

[5] Resolution of 20 October 2020 with recommendations to the Commission on a framework for the ethical aspects of AI, robotics and related technologies (2020/2012 (INL)) Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014 (INL)) Resolution of 20 October 2020 on intellectual property rights in the field of AI technology development (2020/2015 (INI)).

4. GOOD PRACTICES

With regard to the legal / ethical framework for developers, there are already good practices in similar disciples – for example the Code of Ethics for Robotics Engineers.[6] It encourages all scientists and designers to act responsibly and to take full account of the need to respect the dignity, privacy and safety of people. Scientists conducting research in the field of robotics should adhere to the highest standards of ethics and professionalism and adhere to the following principles: benefit – robots should serve the best interest of humans, harmlessness – the principle of "no harm first," according to which robots should not harm people, autonomy – the ability to make informed, decisions about the principles of interaction with robots and justice – by fairly distributing the benefits of robotics, in particular the affordability of robots for home care and healthcare. By using the robotics experience of engineers, you can relate their principles to AI. And this is how the following catalog of rules is created: 1) any involvement in AI work should respect fundamental rights, and their design, production, dissemination and use should be in the interests of the individual and society as a whole, and with respect for the right to self-determination. Human dignity and autonomy, both physical and psychological, must always be strictly respected; 2) work on AI should follow the precautionary principle, anticipating their potential safety impact and taking appropriate precautions commensurate with the level of protection required, while promoting progress with benefits for society and the environment; 3) AI designers ensure transparency and respect for the legitimate right of access to information by all stakeholders. Integration enables all entities involved in or interested in research to participate in the decision-making process; 4) AI designers should be responsible for the social, environmental and human health impacts robotics may have now and in the future; 5) AI designers should consider and respect people's physical well-being, safety, health, and rights. Robotics engineers must promote human welfare while respecting human rights and quickly exposing factors that could threaten society or the environment; 6) reversibility, as an indispensable condition for controllability, should be a fundamental assumption in AI development; 7) the right to privacy must be strictly respected; 8) the operation of systems should always be based on a robust risk assessment process, which should be based on the principles of prudence and proportionality.

[6] See https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PL.html [accessed: 30.04.2021].

5. TRAINERS' RESPONSIBILITY

The cliche seems to be that in order to achieve the regulatory goal, every legal provision should be properly applied. With regard to AI, it is particularly important to assess whether existing / planned legislation can be adequately enforced to address the risks posed by AI systems. The European Commission believes that the legal framework could be improved by the effective application and enforcement of existing EU and national legislation, limiting the scope of existing EU legislation, modifying the functionality of AI systems, uncertainty about the division of responsibilities between different economic entities in the supply chain, and finally changes in security concept.[7] It is worth paying attention to the last-mentioned aspect. The use of artificial intelligence in products and services may pose risks that are not currently specifically addressed by EU legislation. These threats can be related to cyber threats, personal security threats, or threats from loss of connectivity, etc. These threats can occur at the time of product launch, or they arise from software updates or self-learning while using the product.

It is worth recalling that in the context of determining responsibility for autonomous systems, the basic concept was expressed in the judgment of Greenman against Yuba Power Prod. Inc., in which the court stated that "a defect may appear in the mind of designers as well as at the hands of a worker" (case of 1963). It is clear that there is a risk of artificial intelligence being "taken over" and it is important to maintain control over the system. Opponents call the takeover argument a "paranoid anthropocentric argument," and oppose it by saying that because robotic technology can pose a threat to humans, the only solution is not to manufacture robots [Adriano 2015, 370].

As an example, Tesla requires buyers to sign a contract that obliges them to keep their hands on the steering wheel at all times, even when the autopilot is engaged [Kowert 2017, 181–204]. It seems that once the ultimately responsible parties have been identified, their responsibilities should generally be proportionate to the level of instructions given to the robot and its degree of autonomy. Thus, the more learning or autonomy a given robot has, and the longer the robot has been "trained," the more responsibility should rest with the trainer. Importantly, when looking for a person who is actually responsible for the harmful behavior of the robot, you should not confuse the skills resulting from the robot's "training" with skills that depend strictly on the robot's ability to learn independently. At least at this stage, the responsibility must lie with the person, not the job.

[7] Read more cited above The White Book.

6. TRUSTWORTHY AI

A key issue for the future detailed regulatory framework for AI is to define its scope of application. In the opinion of the European Commission, the requirements for high-risk AI applications may consist of the following key functions: data, data and documentation storage, information to be provided, robustness and accuracy, human supervision, detailed requirements for some specific AI applications, such as these used for remote biometric identification. To ensure legal certainty, these requirements will be clarified to provide a clear benchmark for all actors who need to comply with them.

Considering the complexity and opacity of many AI systems and the associated difficulties that may exist in order to effectively check compliance and enforce the applicable rules, it is urged to meet the record keeping requirements related to algorithm programming, the data used for training high-risk AI systems, and in some cases, to store the data itself. These requirements generally allow potentially problematic actions or decisions of AI systems to be traced and verified. This should not only facilitate supervision and enforcement. It may also increase the incentive for economic operators to consider the need to comply with these rules at an early stage.

There is no doubt that it seems necessary to introduce an appropriate regulatory framework. These could identify the exact set of data used to train and test AI systems, including a description of the main characteristics and how to select the data set. A necessary condition seems to be the presentation of documentation on programming and methodology of training processes and techniques used to build a given AI system. In the process of validating such a system, it must be proven that the safety has been guaranteed and that any bias that could lead to prohibited discrimination has been excluded. Records, documentation and, where applicable, data sets would need to be kept for a limited, reasonable period to ensure the effective enforcement of the relevant provisions. Data from these files should be available on request of the relevant administrative authorities. At the same time, it is necessary to ensure the protection of confidential information (e.g. business secrets).

For the creators of such systems, it seems necessary to provide clear information about the possibilities and limitations of the AI system. They need to know that they are required to clearly state the purpose for which the systems are intended and the conditions under which they can be expected to function as intended, and last but not least, the level of accuracy expected in achieving the stated objective. This information is especially important for system implementers, but may also be relevant to competent authorities and stakeholders.

Citizens should be clearly informed when they are interacting with an AI system and not with humans. It is important that the information provided is

objective, concise and easily understood. The way information is to be communicated should be context specific.

AI systems need to be technically robust and accurate to be trustworthy. This means that such systems have to be developed responsibly and with due diligence ex ante, taking due account of the risks they may generate. Their development and operation must ensure the reliable operation of AI systems as intended. It is therefore necessary to ensure that AI systems are robust and accurate, or at least correctly reflect their level of accuracy, at all stages of the life cycle. Furthermore, they ensure reproducible results and deal with errors or inconsistencies at all stages. It is also necessary to make AI systems resilient to both blatant attacks and more subtle attempts to manipulate data or algorithms.

Human surveillance helps to ensure that the AI system does not undermine human autonomy or cause other undesirable effects. The goal of a credible, ethical and human-centered AI can only be achieved by ensuring that people are properly engaged with regard to high-risk AI applications. Ultimately, it is the human who should make the decision / finally approve the operation of the AI application. In the opinion of the European Commission, human surveillance may be that the output from an AI system does not become effective unless it has been previously validated and approved by a human. Or it may be that the output from the AI system becomes effective immediately, but human intervention is ensured later. We must absolutely agree with this approach. Monitoring the AI system during operation and the possibility of real-time intervention and deactivation should be obligatorily entered as the guiding principle of the SI operation, and it is at the design stage that such operational restrictions should be imposed on the AI system.

Any such obligation should be directed to the creator who is best prepared to deal with the potential risk. For example, while AI developers may be best equipped to deal with risks arising from the development phase, their ability to control risk during the use phase may be more limited. In this case, the implementer should be subject to appropriate obligations. This is without prejudice to the question of whether, in order to ensure accountability to end-users or other parties suffering a loss and to ensure effective access to justice, that party should be liable for any damage caused. Under EU product liability law, liability for defective products is assigned to the manufacturer, without prejudice to national legislation that may also allow recovery from other parties.

The Commission also considers it extremely important that the requirements apply to all relevant economic operators delivering AI-enabled products or services in the EU, whether or not they are based in the EU. Otherwise, the aforementioned objectives of legislative intervention could not be fully achieved.

We should also agree with this postulate of the Commission. Conformity assessments would be mandatory for all affected economic operators, irrespective of their place of establishment. To reduce the burden on entrepreneurs, a support structure could be envisaged, including through digital innovation hubs. In addition, standards and dedicated online tools can facilitate compliance.

Any prior conformity assessment should be without prejudice to compliance monitoring and ex post enforcement by national competent authorities. Ex-post controls should be made possible by properly documenting the relevant AI request and, where appropriate, allowing such applications to be tested by third parties such as competent authorities. This can be especially important where there are threats to fundamental rights which depend on the context. Such compliance monitoring should be part of the continued market surveillance system.

As shown by bad experiences related to, for example, sanitary services, it seems necessary to increase the capacity of administrative bodies in the Member States of the European Union in the field of testing and certification of IS. In this context, it is necessary to support the competent national authorities to enable them to fulfill their mandate when AI is used. The entire process indicated above, without qualified and efficient state authorities, will not function, and the individual will not be adequately protected.

CONCLUSION

When assessing the regulatory activities of the European Union, it can of course be pointed out that the entire legislative process is running too slowly. However, I have the impression that the proposals for "controlling" AI may finally be successful. Without the implementation of a holistic solution, in which the framework of conduct / ethical rules will be imposed on the creators and the validation of AI systems by the state will be introduced, one cannot speak of any sovereignty. It seems that only such a duopoly can lead to the use of brilliant AI solutions, minimizing the risks associated with it. However, without strong state organs, this process will not be adequately secured. This title sovereignty over AI seems to be a *sine qua non* condition for us to be safe in the understanding of dominating processes.

REFERENCES

Adriano Elvia A.Q. 2015. "The Natural Person, Legal Entity or Judicial Person and Judicial Personality." *Penn State Journal of Law & International Affairs* 4, no. 1:356–91.

Fradera, Francesca. 2018. "Report from the conference entitled Digital revolution: data protection, artificial intelligence, smart products, Blockchain technology and virtual currencies. Challenges to law in practice." *European Review of Private Law* 26, no. 5:707–12.

Kowert, Weston. 2017. "The foreseeability of human-artificial intelligence interactions." *Texas Law Review* 96, no. 1:181–204.

Naučius, Mindaugas. 2018. "Ar visiškai autonomiškiems dirbtinio intelekto subjektams turi būti suteikiamas teisinis subjektiškumas?" *Teisės apžvalga. Law review* 1 (17):113–32.

Rissland, Edwina L. 1990. "AI and law – a springboard to the model of legal justification." *The Yale Law Journal* 99, no. 8:1957–981.

Radziewicz, Piotr. 2005. "Administracyjnoprawne pojęcie władztwa publicznego." *Kwartalnik Prawa Publicznego* 4:121–44.

Rojszczak, Marcin. 2019. "Prawne aspekty systemów sztucznej inteligencji – zarys problemu." In *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, edited by Kinga Flaga–Gieruszyńska, Jacek Gołaczyński, and Dariusz Szostek, 1–23. Warsaw: C.H. Beck.