GIS Odyssey Journal

2024 • Volume 4 • Number 1

https://doi.org/10.57599/gisoj.2024.4.1.66

Kamil Strzępek[1]

HUMAN RIGHTS AS A FACTOR IN THE AI ALIGNMENT

Abstract: One of the most important issues in the AI field is the alignment of AI goals with human goals. This paper examines the role of human rights in the AI alignment process. The subject of the analysis was acts adopted at the level of two European intergovernmental organisations, namely the European Union and the Council of Europe. The analysis aimed to describe the situations (their nature) in which references to human rights were made. This analysis showed that human rights may be considered a guide in European regulations related to AI. Both the AI Act and the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law are examples of third-generation human rights regulation acts. This means that the human-centric approach adopted in the context of activities within the AI lifecycle is not just a slogan. However, some studies have indicated that regulations may slow down the development of AI in Europe. There is also a question regarding the relationship between European regulations and those of other regions of the world. In particular, the provisions of the Convention clearly show the aspiration that human rights will be a factor in the alignment of goals between AI and humans worldwide.

Keywords: artificial intelligence, the AI act, the Council of Europe Framework Convention on Artificial Intelligence and Human Rights Democracy and the Rule of Law, AI alignment, human rights

Received: 18 April 2024; accepted: 06 July 2024

[1] Cardinal Stefan Wyszyński University in Warsaw, Faculty of Law and Administration, Poland, ORCID ID: https://orcid.org/0000-0001-9277-6057, email: k.strzepek@uksw.edu.pl

Introduction with analysis of the state of the problems

Artificial Intelligence (hereinafter referred to as "AI") is broadly defined as the science of teaching machines to learn humanlike capabilities (Suleyman, 2023; Boden, 2020). AI can be divided into narrow, general, and super-intelligence groups. Artificial narrow intelligence may be designed to perform singular tasks. It has been used for a long time. As an example, popular voice assistants such as "Siri" may be mentioned. Artificial general intelligence may refer to a technology with general intelligence that mimics human intelligence and/or behaviours. Artificial superintelligence may refer to a technology in which machines become self-aware and "surpass" the capacity of human intelligence and ability (Escott, 2017). Currently, there are two trends in artificial intelligence. The first is to build a universal system that could be used anywhere in the world. It would be one system in which a common idea of the world would have to be agreed upon. The second is to build small models that have a dozen or so parameters and will be dedicated to specific areas (e.g. algorithm-psychotherapist operating in a specific country). They will not be universal; they will not communicate in any language. They will consider, among other things, cultural differences present in the world (Duch, 2024). Artificial superintelligence has not yet been achieved; however, owing to rapid technological progress, it cannot be ruled out that this stage will be achieved, for example, in a year (Duch, 2024). Currently, AI (algorithms) have reached the level of development of teenagers (Dragan, 2023).

AI increases both potential values and risks (Sieja & Wach, 2023). These values indicate that technological development may facilitate access to many services, as they can be provided by AI (and not necessarily by humans). An example of this is healthcare (Gajos-Gržetić et al., 2017). AI could be used to increase the efficiency of healthcare diagnosis. For instance, "AI algorithm can be trained on a much larger set of images than a radiologist – as many as a million or more radiology images" (IBM Education, 2023). Virtual medical assistants can help to answer questions about medications. In a study conducted in San Diego, patients were asked questions about doctors and chat gpt. Of the 195 questions and answers, the evaluators preferred chatbot responses to physician responses in 78.6% (95% CI, 75.0%-81.8%) of 585 evaluations. This implies that a chatbot may generate quality and empathetic responses to patient questions (Ayers et al., 2023).

When it comes to the fears, ethical problems are primarily mentioned (Vamplew et al., 2017; Martsenko, 2022; Petrašević & Duić, 2023). One of the most frequently raised ethical issues associated with AI is bias, which may violate human rights and fundamental freedoms, resulting in discrimination (CEUR-WS.org, 2022). It is said that: "Inalienable rights and dignity become central in the context of machine-human interfaces and interactions because of the loss of control that automation and AI bring to the fore. Contemporary deep learning and predictive algorithms have unprecedented access to information, which translates, in the political realm, into the troubling consequence of unprecedented surveillance" (Montemayor, 2023).

AI is used both in the public sector (e.g. automated border control e-gates and facial, fingerprint, and iris recognition systems) and in the private sector (for example, an online provider of products and services that unilaterally sets terms and conditions. If a consumer reject accepting such conditions, they will not have access to the service) (Custers, 2023; Sakowska-Baryła, 2024). Therefore it can be assumed that practical applications of AI can be found almost everywhere, namely at home (e.g. robot vacuum cleaners), in cars (including autonomous cars), schools (e.g. chatbots for enrolment and retention), offices (e.g. "cobots"), hospitals (e.g. interpretation of visual data for diagnostic purposes) in the sky (e.g. AI-powered drones) and the Internet (including the Internet of Things), and some of them are even outside our planet (e.g. robots sent to the Moon and Mars or satellites orbiting in space).

According to Zuboff, we live in the so-called surveillance capitalism era and "since surveillance capitalism targets our attention capacities, making them addicted to positional goods, we have become "entrained" by them. Unfortunately, because of this development, we are now more indifferent than ever to the balance and proper functioning of fundamental cognitive needs, most alarmingly our empathic and emotional needs" (Zuboff, 2019; Więckowski & Świerczyński, 2021).

A question arises as to how to align the goals of AI with the values that people pursue. The notion of AI alignment may be defined as the process of steering AI systems toward a person or group's intended goals, preferences, and ethical principles. An AI system is considered aligned if it strengthens its intended objective. A misaligned AI system may pursue some objectives, but not the intended ones (Russel, 2022).

Gabriel distinguished the following ways of AI alignment: "Properly aligned AI will need to take account of different forms of unethical or imprudent behaviour and incorporate design principles that prevent these outcomes. One way to do this is to build objective constraints on what artificial agents may do. More useful still would be a set of principles that situate human direction within a moral framework that is widely endorsed despite the existence of different belief systems. This requires work both in terms of the technical specification of concepts from which principles are assembled as well as identifying principles of the right kind" (Gabriel, 2020).

It is said that three main approaches in the formation of standards for the development of AI may be observed: the American one – "pragmatic", the Chinese one – "social cohesion" and the third one – "trustworthy AI", proposed in Europe and as an international standard (Szpor, 2023).

When we consider the acts being adopted in Europe at the level of intergovernmental international organisations, we will see that human rights may play a role in aligning the goals of AI with the goals of humanity. The most common concept of human rights may be characterised as a set of universal, inherent, inalienable, and inviolable human rights and freedoms independent of the state.

Material and methods

This research was conducted in 2024. The research was based on acts adopted by intergovernmental international organisations in Europe such as the European Union

(hereinafter referred to as "the EU") and the Council of Europe (hereinafter referred to as "the CoE"). The research method used in this paper was a descriptive analysis with normative theory elements (Portman, 1986) and an interpretation of recent trends in the area of AI alignment. Based on the research conducted, it was possible to answer the principal question of whether human rights play a role in the process of AI alignment?

Discussion: The AI Act

On 6 and 8 December 2023 the Council and the European Parliament of the EU came to an agreement on all political issues and successfully closed the interinstitutional negotiations on the so-called AI Act (hereinafter referred to also as "the Regulation"). It is the first act in the world, the purpose of which is a) to improve the functioning of the internal market; b) to promote the uptake of human-centric and trustworthy AI systems; c) to ensure a high level of protection of health, safety, and fundamental rights enshrined in the Charter of Fundamental Rights of the EU (including democracy, the rule of law, and environmental protection against the harmful effects of AI systems in the EU); and d) to support innovation.

As a general remark, it should be noted that the Regulation shall apply (in general) 24 months following its entry into force, that is, in 2026. Tite I (General provisions) and Title II (Prohibited AI practices) shall apply 6 months following the entry into force of the Regulation, i.e. in 2024. Title III Chapter 4 (Notifying Authorities and Notified Bodies), Title VI (Governance), Title VIII(a) (General Purpose AI Models), Title X (Confidentiality and Penalties) shall apply 12 months following the entry into force of the Regulation, i.e. in 2025. Article 6(1) and the corresponding obligations in the Regulation shall apply 36 months following the enactment of the Regulation. Article 6(1) of the Regulation concerns the so-called high-risk systems used as a safety component of a product covered by EU harmonisation legislation (Debunking the EU AI Act, 2024).

The legal basis of the AI Act is Article 114 of the Treaty on the Functioning of the EU. Indeed, AI systems can be easily deployed in many sectors of the economy and society, including across borders, and are distributed throughout the EU. The EU has adopted a human-centric approach to regulating AI systems. The human rights approach should be taken into account not only by public authorities but also by private entities acting as a) providers who place on the market or put into service AI systems or place on the market general-purpose AI models in the EU, irrespective of whether those providers are established or who are located within the EU or in a third country; b) deployers of AI systems that have their place of establishment or who are located within the EU; c) providers and deployers of AI systems that have their place of establishment or who are located in a third country; d) importers and distributors of AI systems; e) product manufacturers who place on the market or put into service an AI system together with their product and under their own name or trademark; f) authorised representatives of providers, which are not established in the EU; and g) affected persons that are located in the EU.

It is worth noting that the terms "provider" and "deployer" refer to any natural or legal person, public authority or any other body using an AI system. Whereas terms

"authorised representative", "importer" or "distributor" refer to any natural or legal person in the supply chain. It is worth emphasising that the obligations arising from the Regulation involve not only "public authorities" but also "private entities".

There are a few exceptions to the application of the Regulation. For instance, the Regulation shall not apply to areas outside the scope of EU law, and in any event, shall not affect the competencies of the Member States concerning national security. Moreover, the Regulation shall not apply to AI systems if and insofar placed on the market, put into service, or used with or without modification of such systems exclusively for military, defence, or national security purposes, regardless of the type of entity carrying out those activities.

The AI Act includes several practices prohibited in publicly accessible spaces (Szpor, 2023). Many refer to the potential violations of values protected by human rights. For instance, it is prohibited to place on the market, put into service, or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's behaviour by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision that the person would not have otherwise taken in a manner that causes or is likely to cause that person, another person, or group of persons significant harm.

The above practices can be considered in the context of a violation of the right to respect a person's private life. If the right to respect private life includes the right to conscious self-determination, then the above practice undoubtedly has a negative impact on human will.

Another example of a prohibited practice is to place on the market or put into service for this specific purpose, or use biometric categorisation systems that categorise individual natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life, or sexual orientation. However, this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data, or categorisation of biometric data in the area of law enforcement.

The above practices can result not only in violation of the right to respect for a person's private life but also, for example, the right to non-discrimination, freedom of thought, conscience and religion, freedom of expression, freedom of assembly, and association.

The Regulation also indicates the circumstances in which each use for the purpose of law enforcement of a "real-time" remote biometric identification system in publicly accessible spaces shall be subject to a prior authorisation granted by a judicial authority or an independent administrative authority whose decision is binding of the Member State in which the use is to take place, issued upon a reasoned request, and in accordance with the detailed rules of national law.

The AI Act classifies AI systems. There can be distinguished: a) prohibited AI systems, namely, the already mentioned subliminal techniques beyond a person's consciousness to distort a person's behaviour in a manner that causes harm; b) high-risk AI systems, that

is, AI systems that may pose a risk of harm to people's health, safety, and fundamental rights; c) general-purpose AI models and generative AI, for example, large language models; d) limited-risk AI systems, that is, systems with specific transparency requirements (e.g. chatbots); and e) unregulated AI systems, such as video games.

According to the provisions of the Regulation, enforcement activities will be performed by the EU and national-level competent authorities. At the EU level, there will be the "European AI Office", and "European Artificial Intelligence Board", whose task will be, among other things, to advise and assist the Commission and the Member States in order to facilitate the consistent and effective application of the Regulation. An "Advisory forum" shall be established to advise and provide technical expertise to the Board and the Commission to contribute to their tasks under the Regulation. Moreover, the Commission shall, by means of an implementing act, make provisions on the establishment of a scientific panel of independent experts (the "scientific panel") intended to support enforcement activities under the Regulation. Additionally, each Member State shall establish or designate at least one notifying authority and at least one market surveillance authority for the purpose of the Regulation as national competent authorities. These national competent authorities shall exercise their powers independently, impartially, and without bias so as to safeguard the principles of objectivity of their activities and tasks and to ensure the application and implementation of the Regulation.

Any affected person, that is, any natural or legal person, shall have the right to lodge a complaint with the relevant market surveillance authority having grounds to consider that there has been an infringement of the provisions of the Regulation. Non-compliance with the prohibition of AI practices is subject to administrative fines.

It follows from the above analysis that a key issue is the issue of "inferring" by AI systems. The AI system uses data to infer how to generate results, such as predictions or recommendations that may impact the physical or virtual environment. At the same time, depending on the circumstances regarding its specific application, use, and level of technological development, AI systems may generate risks and harm public interests and fundamental rights that are protected by EU law.

Discussion: The European Ethical Charter on the use of AI in judicial systems

The European Commission for the Efficiency of Justice (CEPEJ) of the CoE adopted in December 2018 the first European text to establish ethical principles related to the use of AI in judicial systems (CoE CEPEJ, 2018).

The European Ethical Charter on the Use of AI in Judicial Systems and Their Environment provides a framework of principles that can guide policymakers, legislators, and justice professionals when they grapple with the rapid development of AI in national judicial processes. Five principles are proposed: 1. Principle of respect for fundamental rights; 2. Principle of non-discrimination; 3. Principle of quality and security; 4. Principle of transparency, impartiality, and fairness; and 5. Principle "under user control".

The first of these principles clearly states the need for compliance with "the fundamental rights guaranteed by the European Convention on Human Rights and the Convention on the Protection of Personal Data" (CoE CEPEJ, 2018).

As regards the second principle, it faces one of the biggest problems when using AI systems, namely the "bias" problem. It is worth remembering that the most frequently cited risk associated with AI systems in the literature is bias (CEUR-WS.org, 2022).

The principle of quality and security refers to, among other things, the need for designers of AI models to draw widely on the expertise of the relevant justice system professionals (judges, prosecutors, lawyers, etc.) and researchers/lecturers in the fields of law and social sciences (e.g. for example, economists, sociologists and philosophers) (CoE CEPEJ, 2018; Dobák, 2021). The AI models created must also be able to be stored and executed in secure environments to ensure system integrity and intangibility (CoE CEPEJ, 2018).

The principle of transparency, impartiality, and fairness refers to a fair balance between "the intellectual property of certain processing methods and the need for transparency (access to the design process), impartiality (absence of bias), fairness, and intellectual integrity (prioritising the interests of justice) when tools are used that may have legal consequences or may significantly affect people's lives" (CoE CEPEJ, 2018).

The principle "under user control" may be considered in two ways. The first aspect "refers" to professionals in the judicial system who should be able to review judicial decisions and the data used to produce a result and continue not to be necessarily bound by the light of specific features of that particular case (CoE CEPEJ, 2018). The second aspect "refers" to the user who should be aware of whether or not the solutions offered by AI systems are final and binding, of the different options available, and that she/he has the right to legal aid and the right to access a court (CoE CEPEJ, 2018).

Discussion: The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (hereinafter referred to also as "the Convention") was adopted on 17 May 2024 by the Committee of Ministers of the Council of Europe and will be opened for signature on the occasion of the Conference of Ministers of Justice in Vilnius (Lithuania) on 5 September 2024.

In the Preamble to the Convention, we can read as follows: "Recognising the value of fostering co-operation between the Parties to this Convention and of extending such co-operation to other States that share the same values". In a sense, it reaffirms the aspiration that human rights will be a factor in the alignment of goals between AI and humans not only in Europe but also worldwide. It is worth noting that States such as Argentina, Australia, Canada, Costa Rica, the Holy See, Israel, Japan, Mexico, Peru, the United States, Uruguay, and the European Union participated in the negotiations of the Convention (Explanatory Report, 2024). The global aspirations of the Convention may also be evidenced by the references in the Preamble to the 1989 United Nations Convention on the Rights of the Child and the 2006 United Nations Convention on Persons with Disabilities.

Article 1 of the Convention indicates its basic purpose, which is: "to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with

human rights, democracy and the rule of law." "In this sense, the Framework Convention is aligned with the applicable human rights protection systems and mechanisms of each Party, including their international law obligations and other international commitments and their applicable domestic law" (Explanatory Report, 2024). It is worth noting that the Convention is not intended to create new human rights or undermine the scope and content of the existing applicable protections, but rather to facilitate the effective implementation of the applicable human rights obligations of each Party in the context of the new challenges raised by AI (Explanatory Report, 2024). The latest goal of the Convention is to be achieved by the provisions of Chapters II-VI. It can therefore be said that the Convention was an attempt to implement human rights in the context of third-generation human rights. The purpose of the Convention is not so much to create new human rights, but to implement the observance of existing human rights "by AI". Regarding the scope of the application of the Convention, it is worth noting the exclusions provided in Article 3 of the Convention (e.g. national security interests).

Article 6 of the Convention sets forth general common principles that each Party shall implement with regard to AI in a manner appropriate to its domestic legal system and other obligations of the Convention. For instance, each Party should adopt or maintain measures to respect human dignity and individual autonomy in relation to activities within the AI lifecycle. They are drafted with a high level of generality (Explanatory Report, 2024; Wiśniewski, 2023). The intention is to be an overarching requirement that can be flexibly applied in a variety of rapidly changing contexts (Explanatory Report, 2024). There is also an obligation to adopt or maintain measures tailored to specific risks with respect to activities within the AI lifecycle (Article 8 of the Convention). In this context, the AI Act may be viewed as complementary to the Convention. The AI Act may also be viewed as a complementary act to Article 9 of the Convention which deals with the adoption or maintenance of measures tailored to ensure accountability and responsibility for adverse impacts on human rights resulting from activities within the AI lifecycle.

The Convention also specifies procedural guarantees. Each Party shall ensure that procedural guarantees, safeguards, and rights prescribed in applicable international and domestic human rights laws remain available and effective in the context of AI (Article 15 of the Convention). However, this does not mean the obligation to create new procedural guarantees.

It is worth noting that the obligations prescribed in the Convention are obligations of results and not obligations of means, which means that the applicable provisions of law in the States Parties may already meet the obligations indicated in the Convention (at least to a certain extent).

Conclusions

Human rights have been classified traditionally in terms of the notion of three "generations" of human rights. The first generation encompasses civil and political human rights (late 18th and 19th centuries); the second generation encompasses economic, social, and cultural human rights (20th century); and the third generation encompasses human rights of solidarity (recent years) (Masferrer, 2023). The first generation is based

on the issue of freedom, the second on equality, and the third on solidarity (fraternity). It is within the third generation of human rights that, among other things, we refer to, the use of advances in science and technology, although some have assigned them to the fourth generation (Masferrer, 2023). It can be assumed that the third generation of human rights is complementary to those of the previous two generations. It can also be assumed that third-generation human rights come from the values of the first and second generations of rights but concern a "different regulatory space" (Milczarek, 2024).

The above assumptions seem to be confirmed by the content of the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. The Convention does not create new human rights in its content, but its aim is to ensure that activities within the lifecycle of AI are fully consistent with "existing" human rights, democracy and the rule of law. It is also worth noting that the obligations prescribed in the Convention are obligations of results and not obligations of means, which means that the applicable provisions of law in the States Parties may already meet the obligations indicated in the Convention (at least to a certain extent).

The "solidarity" of the third generation of human rights manifests itself in at least two aspects: 1) the scope of entities covered by the obligations arising from these rights is wide; 2) due to the subject of regulation (e.g. natural environment, peace, freedom and security of communication technologies), they require cooperation between States.

The AI Act is an example of a third-generation human rights regulation act. The addressees of obligations specified in the AI Act are not only public authorities but also any natural or legal person. By its very nature, it is also an act of cooperation between States (the EU Member States).

The Convention and the AI Act (which may be viewed as a complementary act to the Convention) adopt the so-called human-centric approach. On the one hand, from a human point of view, the adoption of the so-called human-centric approach in the regulations in question is undoubtedly laudable. On the other hand, representatives of companies in the IT sector indicate that the regulations in question may slow down the development of new technologies (including AI) in Europe (AI supremacy, n.d.).

The regulations in question cover companies with headquarters outside of Europe that operate in the European market. Reciprocity is one of the most important principles in international law. The question arises whether restrictions related to the activities of "foreign" companies in Europe will result in restrictions imposed on the activities of European companies in other regions of the world. The question also arises as to whether other regions of the world adopt a regulatory model similar to that in Europe. In particular, the provisions of the Convention clearly show the aspiration that human rights will influence the alignment of goals between AI and humans worldwide.

References

AI supremacy: The artificial intelligence battle between China, USA and Europe | DW Documentary. (n.d.). https://www.youtube.com/watch?v=Le122vas9aM [access: 17.04.2024].

ALEXA STAGE – Prof. Włodzisław Duch o tym czy nadal warto uczyć się języków obcych (*Prof. Włodzisław Duch about whether it is still worth learning foreign languages*) (2024). https://www.youtube.com/watch?v=61AOh1WJlzU [access: 17.04.2024].

Ayers J.W., Poliak A., Dredze M., Leas E.C., Zhu Z., Kelley J.B., Faix D.J., Goodman A.M., Longhurst C.A., Hogarth M., Smith D.M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Internal Medicine, 183(6), pp. 589–596. https://doi.org/10.1001/jamainternmed.2023.1838 [access: 17.04.2024].

Boden M.A. (2018). Artificial intelligence: A very short introduction. Oxford University Press, Oxford.

CEUR-WS.org/Vol-3399 – Ongoing Research, Practitioners, Posters, Workshops, and Projects of the International Conference EGOV-CeDEM-ePart 2022. Ceur-Ws.org. https://ceur-ws.org/Vol-3399/ [access: 17.04.2024].

Coe CEPEJ European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment (2018). European Commission for the Efficiency of Justice (CEPEJ). https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment [access: 17.04.2024].

Committee on Artificial Intelligence (CAI) (2024). Explanatory Report. https://www.coe.int/en/web/artificial-intelligence/cai [access: 17.04.2024].

Committee on Artificial Intelligence (CAI) (2024). The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. https://www.coe.int/en/web/artificial-intelligence/cai [access: 17.04.2024].

Council of the European Union, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (the AI Act) and amending certain Union legislative acts – Analysis of the final compromise text with a view to agreement (2024). Interinstitutional File: 2021/0106(COD).

Custers B. (2023). The Right to Break the Law? Perfect Enforcement of the Law Using Technology Impedes the Development of Legal Systems. Ethics and Information Technology 25(58). https://doi.org/10.1007/s10676-023-09737-3 [access: 16.07.2024].

Debunking the EU AI Act: an overview of the new legal framework (2024). https://www.youtube.com/watch?v=iO0nfxxeYjE [access: 17.04.2024].

Dobák I. (2021). Thoughts on the evolution of national security in cyberspace. Security and Defence Quarterly, 1(33), pp. 75–85.

Dragan A. (2023). O sztucznej inteligencji. Tempo zmian jest niewiarygodne (*About artificial intelligence. The pace of change is incredible*) (2023). https://www.youtube.com/watch?v=-p0kRrioFWA [access: 17.04.2024].

Escott E. (2017). What Are The 3 Types Of AI? A Guide To Narrow, General, And Super Artificial Intelligence. Codebots. https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible [access: 17.04.2024].

Gabriel I. (2020). Artificial Intelligence, Values, and Alignment. Minds and Machines, 30. https://doi.org/10.1007/s11023-020-09539-2 [access: 17.04.2024].

Gajos-Gržetić M., Rahmonov O., Stach S. (2017). Geospatial technologies in biology and medicine: analysis of GIS repository. In: Geographic Information Systems Conference and Exhibition "GIS Odyssey 2017". Trento-Vattaro.

IBM Education (2023). The benefits of AI in healthcare. IBM Blog; IBM. https://www.ibm.com/blog/the-benefits-of-ai-in-healthcare/ [access: 17.04.2024].

Martsenko N. (2022). Artificial intelligence and human rights: a scientific review of impacts and interactions. Studia Prawnoustrojowe, (58), pp. 315–331. https://doi.org/10.31648/sp.8245 [access: 16.07.2024].

Masferrer A. (2023). The Making of Dignity and Human Rights in the Western Tradition. Springer, New York City.

Milczarek E. (2024). Cyfrowe prawa człowieka – czy potrzebujemy ich kodyfikacji? (*Digital human rights – do we need their codification?*). In: J. Jaskiernia, K. Spryszak (ed.), Efektywność krajowych i międzynarodowych systemów ochrony praw człowieka drugiej generacji (*The effectiveness of second-generation national and international human rights protection systems*). Adam Marszałek, Toruń.

Montemayor C. (2023). The Prospect of a Humanitarian Artificial Intelligence. Bloomsbury Publishing, London.

Petrašević T., Duić D. (2023). Data Protection and Cybersecurity: Case-Law of Two European Courts. EU and Comparative Law Issues and Challenges Series (ECLIC 7 – Special Issue), pp. 94–118.

Portman C.P. (1986). Jurisprudence: A Descriptive and Normative Analysis of Law. Michigan Law Review, 84(4/5), 1041. https://doi.org/10.2307/1288875 [access: 17.04.2024].

Russel S., Norvig P. (2021). Artificial intelligence: a Modern approach. (4th ed.). Prentice Hall, Hoboken.

Sakowska-Baryła M. (2024). Sztuczna inteligencja w sektorze publicznym (*Artificial Intelligence in the public sector*). In: B. Fischer, A. Pązik, M. Świerczyński (ed.), Prawo sztucznej inteligencji (*Artificial Intelligence law*). Wolters Kluwer Polska.

Sieja M., Wach K. (2023). Revolutionary artificial intelligence or rogue technology? The promises and pitfalls of ChatGPT. 9(4), pp. 101–115. https://doi.org/10.15678/ier.2023.0904.07 [access: 17.04.2024].

Suleyman M. (2023). The Coming Wave. The Bodley Head, London.

Szpor G. (2023). European legal framework for the use of artificial intelligence in publicly accessible space. GIS Odyssey Journal, vol. 3, no. 2, pp. 25–36. https://doi.org/10.57599/gisoj.2023.3.2.25 [access: 17.04.2024].

Vamplew P., Dazeley R., Foale C., Firmin S., Mummery J. (2017). Human-aligned artificial intelligence is a multiobjective problem. Ethics and Information Technology, 20(1), pp. 27–40. https://doi.org/10.1007/s10676-017-9440-6 [access: 17.04.2024].

Więckowski Z., Świerczyński M. (2021). Sztuczna inteligencja w prawie międzynarodowym. Rekomendacje wybranych rozwiązań (*Artificial intelligence in international law*. *Recommendations for selected solutions*). Difin.

Wiśniewski A. (2023). Sztuczna inteligencja i prawa człowieka w kontekście prawa międzynarodowego (*Artificial intelligence and human rights in the context of*

international law). Prawo i Więź, 4(47), pp. 29–52. https://doi.org/10.36128/PRIW.VI47.785 [access: 16.07.2024].

Zuboff S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. NY: Public Affairs, New York.