Deusto Journal of Human Rights

Revista Deusto de Derechos Humanos

ARTICLES / ARTÍCULOS

The legal rules of the European Union and the United States on artificial intelligence and human rights

Las normas jurídicas de la Unión Europea y Estados Unidos sobre inteligencia artificial y derechos humanos

Magdalena Butrymowicz
https://doi.org/10.18543/djhr.3263
E-published: June 2025

The legal rules of the European Union and the United States on artificial intelligence and human rights

Las normas jurídicas de la Unión Europea y Estados Unidos sobre inteligencia artificial y derechos humanos

Magdalena Butrymowicz ⓘ
Pontifical University of John Paul II Poland
magdalena.butrymowicz@gmail.com
ORCID: https://orcid.org/0000-0002-9920-5860

Summary: Introduction. 1. Outline of a legal problem. 2. European Union and AI legal regime. 3. United States and AI legal framework. Conclusion. References.

Abstract: In the contemporary era, artificial intelligence has become an integral component of our daily lives, permeating various facets of societal functioning with increasing prevalence. Given its pervasive presence in our lives, it is inevitable that AI exerts an influence on us, giving rise to the question of whether there is a nexus between AI and human rights. It is important to note that an algorithm created by an AI machine is devoid of feelings, emotions or prejudices. This may be indicative of a fundamental limitation of artificial intelligence as a tool in itself, constrained by the capabilities and capacities of technology, and thus incapable of perceiving and analyzing in the manner of the human mind. It is imperative that artificial intelligence is designed to be completely impartial and solely analyses the data entrusted to it for processing. Moreover, it should be capable of learning from its mistakes and being free of the emotions associated with human relationships. Conversely, it is as flawed as the human being behind its creation. In the contemporary era, there is an ongoing endeavor on the part of various nations to subjugate the domain of artificial intelligence to the provisions of a legal framework. There are three broad models for regulating AI: the first is based on regulations of the legal sector; the second is based on guidelines of inter-administrative organizations; and the third is based on legal

solutions on other similar issues. The United States and the European Union have been at the vanguard of AI regulation, with each adopting different variants. The present publication sets out to compare EU and US legislation in the context of the matrix that served to create it, and to determine whether it regulates human rights. The issue of the rights of ethnic minorities is taken as a test case for this purpose.

Keywords: human rights, artificial intelligence, ethnic minorities, private sector.

Resumen: En la era contemporánea, la inteligencia artificial se ha convertido en un componente integral de nuestra vida cotidiana, impregnando diversas facetas del funcionamiento de la sociedad con una prevalencia cada vez mayor. Dada su omnipresencia en nuestras vidas, es inevitable que la IA ejerza una influencia sobre nosotros, lo que plantea la cuestión de si existe un nexo entre la IA y los derechos humanos. Es importante señalar que un algoritmo creado por una máquina de IA carece de sentimientos, emociones o prejuicios. Esto puede ser indicativo de una limitación fundamental de la inteligencia artificial como herramienta en sí misma, constreñida por las capacidades y posibilidades de la tecnología y, por tanto, incapaz de percibir y analizar a la manera de la mente humana. Es imperativo que la inteligencia artificial esté diseñada para ser completamente imparcial y analizar únicamente los datos que se le confían para su procesamiento. Además, debe ser capaz de aprender de sus errores y estar libre de las emociones asociadas a las relaciones humanas. Por el contrario, es tan imperfecta como el ser humano que está detrás de su creación. En la era contemporánea, varias naciones se esfuerzan por someter el ámbito de la inteligencia artificial a las disposiciones de un marco jurídico. Existen tres grandes modelos de regulación de la IA: el primero se basa en normativas del sector jurídico; el segundo, en directrices de organizaciones interadministrativas; y el tercero, en soluciones jurídicas sobre otras cuestiones similares. Estados Unidos y la Unión Europea han estado a la vanguardia de la regulación de la IA, adoptando cada uno diferentes variantes. La presente publicación se propone comparar la legislación de la UE y de EE. UU. en el contexto de la matriz que sirvió para crearla, y determinar si regula los derechos humanos. Para ello, se toma como caso de prueba la cuestión de los derechos de las minorías étnicas.

Palabras clave: derechos humanos, inteligencia artificial, minorías étnicas, sector privado.

Introduction

Ashwini (United Nation 2025) during an interactive dialogue on the occasion of the publication of her report at the 56th session of the United Nations Human Rights Council in Geneva (Switzerland), pointed out that the recent development of generative artificial intelligence (AI) and the growing use of artificial AI continue to raise serious human rights issues, including concerns about racial discrimination. She also emphasized that predictive policing can exacerbate historically excessive racial and ethnic policing in communities, because at the beginning of any AI there is a human being who is responsible for creating that particular set of technologies. Developers have, first and foremost, specific goals and intentions, creating a useful technology that they believe should prevent any kind of abuse or discrimination (Ashwini 2024a, 4-8). This means that technology is made by persons and for people and it will replicate the patterns, biases or ideology of its creators, which can result in all sorts of violations and biases. Liming Zhu et al. (2021, 15-16), wrote that if creators of the AI don't focus on the developing it responsible, it will lead to devastating effect to the humanity in general means (Zhu et al. 2021, 16-18). Joy Buolamwini (2022), one of the most widely recognized critics of facial recognition technology, has shown that systems with facial recognition cannot recognize people with very dark skin as a result of inaccurate data input at the development stage. It was conclusively demonstrated that the use of AI can lead to discrimination against the recipients or targets of a technology mainly based on their race, color, ethnicity or gender.

1. Outline of a legal problem

In light of the aforementioned considerations and the development of AI technologies, from adaptive learning platforms to administrative automation have been introduced both transformative opportunities but also ethical risks (Smith and Hill 2019, 383-387). In response to the above pointed concerns regarding the impact of AI on society, the Organization for Economic Co-operation and Development (OECD) introduced in 2019 principles on AI as a set of guidelines to promote the responsible development and use of AI. The principles provide a values-based and practical framework to assist governments and stakeholders in designing and implementing responsible AI. It is imperative that AI systems are designed in a manner that aligns with the principles of the rule of law, human rights, democratic values, and

diversity. This pattern should incorporate appropriate safeguards to ensure the establishment of a fair and just society. That is why OECD advises that AI system design and implementation adhere to core principles such as legal compliance, human rights protections, democratic governance, and cultural diversity. Systems must integrate robust safeguards to mitigate risk and ensure ethical alignment, fostering fairness and societal equity (OECD 2023, 2-8). Analyzing the above principles in the context of the search for causes of abuse caused by the improper creation or use of AI, it should be pointed out that there are two sources of the negative dimension of AI: an action of an intentional or unintentional nature by its creators or those who actually use it. Usually an unintentional action, it is the result of human prejudices, beliefs, worldviews or perceptions of a subject. Indirectly intended action is selective or biased input of previously collected data. A final reason may be the very limitation of AI as a tool limited in its capabilities and the capabilities of technology, incapable of perceiving and analyzing like the human mind. On the other hand, AI should be completely impartial and only analyses the data entrusted to it for processing, it learns from its mistakes and can be free from emotions related to a human to human relation. It is therefore the responsibility of the state to utilize the law as a mechanism for the protection of human rights, serving as a safeguard against any potential abuse by AI creators or users.

In general terms the AI is perceive as a technology programmed to analyses the world around them and take action to achieve specific goals. The AI market is 90% controlled by the private sector; only 10% is public domain (Cowger 2020, 5-9). This is an important reservation because the purpose of AI is determined by their developers and not always by their final user. As human made, the AI is generally focused on replacing people in specific activities and decision-making tool. The European Commission's High-Level Expert Group in AI described the AI as systems of software (and possibly also hardware) designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples),

machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems (European Union High-Level Expert on Artificial Intelligence 2019, 3-9). The objective of facilitating decision-making is to introduce neutral prerequisites for making decisions and objectivity in the decision-making process. However, in order to understand the real influence of the AI on the humans and the human rights, as the Karen Yeung (2019, 18) correctly pointed out, it is necessary to understand how the AI is created and developed, what is the purpose of its creation and its practical uses. In a rather important communication on AI, the European Commission described AI as a system that exhibits intelligent behavior by analyzing its environment and taking autonomous actions based on the learned environment and the data it provides. It focuses mainly on achieving specific goals. AI has two dimensions of operation: the virtual world and the real world. When the assumptions underlying the creation of AI and the purpose imposed on it contain a flaw leading to a violation of someone's rights or dignity, the law should take appropriate action. However, this is a very fine line, and general regulations will not always be able to eliminate a violation of human rights resulting from an AI (The European Commission 2018, 1-6).

As the Council of Europe has noted, a unified approach to regulating AI remains elusive. This is due to the fact that AI is a nascent, multidisciplinary field encompassing diverse scientific, theoretical, and technical domains, including mathematics, statistics, probability, neuroscience, and computer science. Consequently, it cannot be regarded exclusively as a medium for the collection or analysis of data. Common applications of AI include interpreting and analyzing results, providing behavioral suggestions, and offering emotional support to those in need. AI therefore works in a similar way to how humans analyze their environment and draw conclusions, which, according to their developers, should be reasonable and free of any bias. The legal framework can therefore be in the form of legislated laws issued by states or international organisations, or it can be introduced in the form of internal law of the particular company or other entity with legal personality that produces or uses AI.

The Council of Europe is establishing a Working Group on Human Rights and Artificial Intelligence with the objective of developing a Handbook on the subject. Nevertheless, it appears that states are inclined to abstain from regulating the legal domain of AI in the context of human rights for the time being (Council of Europe,

CDDH-IA 2025). Nevertheless, the primary objective of the Working Group was to draft and submit for ratification by states an international convention known as the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. The Convention is the first legally binding international treaty designed to ensure that AI systems are developed and used in ways that uphold human rights, democratic principles, and the rule of law. Adopted by the Council of Europe on May 17, 2024, the treaty was opened for signature on September 5, 2024, in Vilnius, Lithuania (The Council of Europe 2024, 1-2). The convention establishes a comprehensive legal framework that applies to both public and private sectors involved in the AI lifecycle. It mandates that AI activities adhere to fundamental principles such as human dignity, individual autonomy, equality, non-discrimination, privacy, data protection, transparency, accountability, and reliability. Additionally, it requires parties to conduct risk and impact assessments, implement preventive and mitigating measures, and provide effective remedies and procedural safeguards for individuals affected by AI systems (The Council of Europe 2024). A thorough evaluation of the Convention reveals that it has successfully identified areas within the life cycle of artificial AI systems that have the potential to compromise human dignity and individual autonomy, human rights, democracy, and the rule of law. The authors of the Convention underscored the risk of discrimination in the digital context, particularly in the context of AI systems. However, the groups and areas deemed to be at risk of discrimination were narrowed to include only women and the disadvantaged, while other groups, such as ethnic minorities, were omitted. Additionally, the concept of an individual at risk was not adequately clarified. In summary, it is evident that the conventions predominantly pertain to the public sphere of the creation and utilization of AI, with limited relevance to the private sector that exerts significant influence over this market (The Council of Europe 2024, 2-4).

However, from the outset of its development and subsequent deployment in public domains by private entities, the field of AI has been characterized by significant controversy. Consequently, concerns regarding its potential implications for human rights have been expressed from the project's inception. Indeed, in certain domains of AI implementation, the ethical implications of potential violations of human rights have been a point of contention. These domains include mass biometric surveillance, the capture and processing of biometric data in both public and private spaces, facial recognition technology, the access to the provision of basic public services (e.g. medical services

or education), the testing of marginalized groups or, last but not least, the capture of individuals from a crowd on the street or the segregation of citizens, thus affecting the freedom of peaceful assembly and association. This problem has been noted by a number of human rights organizations, such as the European Digital Rights Association (2025), which brings together civic and human rights organizations from across Europe, and has also identified these areas as potential warning signals for potential areas of human rights violations. They also supplement the identified list with areas such as: the use of AI systems at the border or in tests on marginalized groups such as undocumented migrants, or autonomous lethal weapons and other applications that identify targets for lethal force. This situation gives rise to the conclusion that AI represents a potential risk and that a legal framework needs to be established to prevent further violations of human rights. As previously mentioned, such legislation has already been enacted and implemented, albeit primarily in the public sector and to a lesser extent in the private sector. It is imperative to acknowledge that the genesis of AI was predicated on the deceptive practices of major corporations, the very entities responsible for its creation. The second issue of some significance is whether the legal order adopted and implemented provides adequate protection of human rights. In order to achieve this objective, it is imperative to undertake a meticulous examination of the extant documentation, encompassing both state, international and private regulations, with a view to ascertaining the manner in which human rights are regulated at each stage of the process, from the conceptualization of AI to its utilization by the primary recipient.

Taking this as a point of departure and the identified reasons for potential discrimination by AI into consideration, the present author would select one indicator to test the effectiveness of private and state regulations, i.e. ethnicity. If we consider ethnicity, understood as a person's cultural-historical identity, as a cause of discrimination, it is, by and large, absent from the process of determining discriminatory risk at every stage of the creation and use of AI. The United Nations has pointed out that the term 'ethnic minority' generally refers to ethnic or racial groups in a given country in which they are in a non-dominant position vis-à-vis the dominant ethnic population: *"The term refers to a group of people in a nation State that meets one or more of the following criteria: it is numerically smaller than the rest of the population; it is not in a dominant position; it has a culture, language, religion or race that is distinct from that of the majority; and its members have a will to preserve those characteristics"* (United

Nation, Department of Economic and Social Affairs 2018, 4-8). It is evident that ethnic minorities are comparatively diminutive, thereby facilitating facile generalization. These groups are predominantly confronted with the predicament of being unrecognized, a situation compounded by the fact that they are frequently unacknowledged by their respective states. Consequently, these minority groups find themselves engaged in a persistent struggle for recognition. Moreover, in the majority of European countries, ethnic minorities remain unacknowledged, consequently hindering their capacity to influence the utilization and development of AI. It is imperative to acknowledge that ethnic minorities are predominantly perceived as part of the nation, and not as constituting potential levels of discrimination. Consequently, they are likely to become beneficiaries of any solutions imposed on them.

Not lowering down the meaning of the discrimination and importance of protecting minorities form the discrimination on every ground, it looks more or less like the current regulations are blurring the problems of the ethnic minorities in the connection to AI (Council of Europe/European Court of Human Rights 2021, 4-11). Discrimination is generally understood as the right not to be rejected the enjoyment of other rights, and freedom is one way to minor the ethnic minorities and they role in the society. It is evident that the anti-discrimination legislation does not guarantee respect for traditional rights associated with culture, tradition and the environment. The question will arise what it had to do with robotic lawnmower or the face recognition program? It depends on the many factors. The utilisation of facial recognition software has the potential to engender significant issues. To illustrate this point, the software has been observed to categorise individuals of Sami descent as Scandinavians of Norwegian origin, consequently classifying them as descendants of Norwegians who speak Nynorsk, despite the fact that they speak Sami language. (Eurydice 2021). In such case, the problem will be, not the skin color, but the impossibility to properly recognized ethnic background of such person. On the end, such situation can lead to discrimination, but generally it will lead to assimilation and degradation of the Sami people. Lawnmower can be also a problem when used on the traditional reindeer grazing lands, without a clear distinction, it will cut the moss on which the reindeer feed (they are tinned loose in the summer). What will lead to a decline in their population and affecting the Sami people, whose identity, traditions, culture and laws are linked to reindeer husbandry and grazing (Andersx et al. 2009, 12-23).

2. **European Union and AI legal regime**

It is evident that AI continues to be a subject of interest for international organizations and their member states. However, it should be noted that the introduction of international law (which traditionally has a reach only between signatories to an agreement or convention and does not translate into the legal system of a country), does not necessarily result in the regulation of AI within national legal orders (the mention above example of Council of Europe convention). It is therefore recommended that states that have taken the initiative at the international level should also consider taking bottom-up legislative initiatives to regulate AI within their national legal orders. In this regard, four different approaches of states can be identified. The initial approach posits that nations endeavor to institute a legal framework grounded in standards delineated by international organizations. The objective of this endeavor is to harmonize their AI legislation with internationally recognized guidelines. The result of this harmonization would be to foster cross-border collaboration and uniformity. The EU legal framework for AI is multifaceted and follows the patterns of influence of the international organization. The legislative framework of the European Union encompasses three distinct dimensions, one of which is the AI Act. This act stipulates meticulous, risk-based requirements for the operation of AI systems (European Comission 2021). The second one is Data protection laws like the General Data Protection Regulation (GDPR) that safeguard personal data (European Union 2016); and a suite of ethical guidelines and digital regulations that work together to ensure that AI is developed in a manner that is safe, transparent, and respectful of fundamental rights (European Comission 2019, 6-10). Together, these initiatives reflect the EU's commitment to fostering technological innovation while ensuring that AI systems contribute to a safe, inclusive, and rights-respecting digital society. Commission pointed out that it supports regulatory and investment approach to AI but not forgetting the risks which are associated with this new technology (European Comission 2020, 2-8).

The second approach entails the formulation of national regulations, predicated on the distinct interpretation of the pressing imperative to address the creation, development and utilization of AI. This approach is predicated on national imperatives and cultural or ethical considerations.

A third approach is the hybrid or adaptive approach, in which countries initially adopt international standards, but also introduce their own regulations adapted to the evolution of national AI challenges and

private sector expectations. Regulatory sandboxes are becoming a key tool to facilitate collaboration between the private sector and policymakers in the development of AI regulations, with the result that regulations are more aligned with market needs and societal expectations. The fundamental purpose of regulatory sandboxes is to encourage safe and ethical innovation in AI, while addressing the unique challenges posed by the rapidly evolving nature of AI. The primary benefits of regulatory sandboxes are that they serve as an effective mechanism for balancing innovation with safety and ethical concerns. By establishing a framework for real-world experimentation, they empower decision-makers to make more informed decisions and ensure that AI technologies are developed in a manner that benefits society while minimizing the potential harms associated with their existence and utilization (Blaine 2025). An illustrative example of cooperation with the private sector is the legal order of the United States. As of mid-2025, the United States does not have a comprehensive federal law specifically regulating AI, but there are various sector-specific rules, proposed legislation, executive actions, and agency guidance shaping how AI can be developed and used. At the federal level, there are only a few regulations in this area. The National Artificial Intelligence Initiative Act (US House of Representatives 2020) is part of the National Defense Authorisation Act of 2021, but it does not regulate the creation and operation of artificial intelligence. Instead, it only coordinates its use between federal agencies. It is important to acknowledge that the legislature has demonstrated a lack of awareness regarding the capabilities of AI and its potential implications for various social and economic sectors, including ethical concerns and its impact on national security and the workforce. Consequently, this act has been adopted with the objective of acquiring pertinent information from all sectors, thereby ensuring that AI functions in a manner that is both trustworthy and beneficial to all Americans. Consequently, a comprehensive research initiative was undertaken to evaluate the innovation potential at universities, non-profit research organisations, enterprises of various sizes and from diverse sectors, and within the federal administration. This initiative aimed to expedite the process (United States House of Representatives 2020).

In conclusion, the EU is developing a unified, risk-based legal framework that proactively regulates AI by establishing clear standards, particularly for systems considered high-risk. This approach is closely associated with robust data protection legislation and dedication to safeguarding fundamental rights. In contrast, the US adopts a more

decentralized and reactive strategy, relying on existing sector-specific laws and voluntary guidelines to manage AI-related risks. This model is widely regarded as being more conducive to rapid innovation, although it may be perceived as lagging in terms of uniform consumer protection and accountability measures.

This thesis has deliberately eschewed discussion of the issue related to the second model of AI regulation mentioned above. This is the model in which countries analyze their own internal order and, on the basis of their existing experience, introduce AI regulations. Examples of such countries could be Australia or Switzerland. However, this analysis requires a very detailed review of the legal orders of these countries, which is beyond the scope of this publication.

As previously mentioned, the European Union has adopted a proactive, unified approach based on risk mitigation in the legal sphere to AI. This particular model of AI regulation is contingent upon the necessity of implementing specific legislation that would standardize the approach to AI and its utilization. A critical component of this analysis entails a thorough examination of the adopted model within the overarching framework of safeguarding human rights, with a particular focus on groups that are disproportionately vulnerable. The test group was chosen to be composed of ethnic minorities.

In this regard, it is necessary to refer to the three EU regulations: the Digital Markets Act, the Digital Services Act and AI Act (European Union 2022). The aforementioned EU legislation establishes a model for assessing and categorizing risks and potential impact of AI on society. They encompass several fundamental assumptions, including the harmonization of regulations, the establishment of a unified set of standards for AI systems throughout the EU, the implementation of a risk-based approach wherein AI systems are categorized according to their risk levels, and the proposal to amend and align existing EU legislative acts to address the emerging challenges and opportunities presented by AI technologies. The Digital Markets Act endeavors to establish a set of objective criteria for identifying "gatekeepers". These entities are regarded as substantial digital platforms that furnish indispensable platform services, including search engines, application stores, and messaging services. It not consider human rights as important factor and omits this issue (European Union 2022). A similar situation looks in the Digital Services Act. The subject pertains to the harmonization of the unified market, akin to the provisions outlined in the AI Act. It encompasses the definition, mitigation, and restriction of AI implementation within the EU common market. It establishes regulatory frameworks for intermediaries and online platforms,

including shopping platforms, social networks, content sharing platforms, application stores, and online travel and accommodation platforms (European Union 2022). Solid regulation but there is a noticeable lack of legal norms relating directly to human rights. It is important to acknowledge that the two cited laws, in conjunction with the aforementioned GDPR, while addressing human rights concerns, do not establish a distinct and unambiguous legal framework that would comprehensively safeguard specific groups at risk of exclusion. The prevailing sentiment, as articulated in their declarations, asserts that fundamental rights concern underpin these actions. However, these regulations conspicuously neglect to impose any form of regulatory oversight within this domain.

Nevertheless, particular emphasis must be placed on the provisions outlined in the EU AI Act, which, in concordance with its developers' stated intentions, prioritizes the imperative to safeguard both fundamental human rights during the processes of AI development, creation, and implementation. A review of the AI Act, which most comprehensively addresses respect for human rights, is warranted. Firstly, the Act accurately identifies areas of risk to human rights, including the use of biometrics, social scoring, and criminal risk assessment. Notwithstanding the implemented prohibition on the utilization of AI, states retain the option to employ AI in these domains, provided that objective factors permit it (European Union 2022). It is evident that human rights do not constitute a primary concern within the remit of this document. Consequently, the Act does not explicitly address the issue of groups at risk like 'ethnic minorities', and its fundamental principles and stipulations must therefore be interpreted in conjunction with extant EU anti-discrimination legislation, which, it should be noted, is not related to ethnic minorities but rather to ethnicity. Consequently, the Act cannot be regarded as guaranteeing that AI systems will be developed and implemented in a manner that protects ethnic minorities from prejudice and discrimination. Thus, these pieces of legislation marginalize the issue of the impact of AI on ethnic minorities, narrowing its entire impact to the problem of discrimination. In summary, it is essential to acknowledge the significance of these findings in the broader context, particularly in light of the potential implications for the study's generalizability and the field as a whole. Notwithstanding the European Union's noteworthy efforts to incorporate anti-discrimination legislation concerning AI into its legal framework, the requirement for explicit regulatory definitions of discrimination in the AI domain remains unresolved. Furthermore, contemporary legal provisions frequently

demonstrate an absence of the requisite specificity to adequately address the distinctive challenges posed by AI, particularly with respect to the safeguarding of ethnic minorities and other vulnerable groups. In order to ensure effective safeguards against algorithmic discrimination, it may be necessary for Europe to adopt more detailed, enforceable guidelines on data quality, transparency, and accountability in the context of a rapidly changing technological landscape.

3. **United States and artificial intelligence legal framework**

Government of United States addresses the issue of human rights in correlation with AI, in a guidance document called the Risk Management Profile for Artificial Intelligence and Human Rights (U.S. Departament of State, Bureau of Cyberspace and Digital Policy 2024), which is a practical guide for organizations -including governments, the private sector and civil society- to design, develop, deploy, use and manage AI in a manner consistent with respect for international human rights. The guide looks in detail at potential situations where human rights may be violated when using AI. According to the authors' findings, the use of AI typically leads to unintentional human rights violations based on ethnicity. This confirms that without internal regulation by states, the use of AI is fraught with significant risks of human rights violations. Especially as the very methodology of the international human rights system has not kept pace with technological developments. In my view, it is not just the process of using AI itself in all areas where it is used but the purpose for which AI is to be used and the actual intentions of AI developers and users. The Report similarly points out as above, that human rights violations can occur at any stage, i.e.: initial planning and design, data collection and analysis, and subsequent use and processing. The authors of the Report therefore propose the introduction of a 4-step process for assessing the impact of AI on human rights. The model is therefore based on: 1) Governance (setting up institutional structures and processes). 2) Mapping (understanding the context and identifying risks). 3) Measurement (assessing and monitoring risks and impacts), and 4) Management (prioritizing, preventing and responding to incidents) (U.S. Departament of State, Bureau of Cyberspace and Digital Policy 2024). The model, the authors claim, can be applied across all applications, stakeholders and sectors, and across the lifecycle of AI. While the model may not be revealed, it introduces

some systematization into the legal realm of AI. It is certainly not a substitute for regulation but is a fairly good start to creating it. It should also be stressed that these guidelines are aimed at governments and the private sector, which is well ahead of the public sector in this respect (Council of Europe 2025).

However, these guidelines are not biding documents and an analysis of the U.S. regulatory framework for AI in the context of safeguarding human rights, particularly with regard to ethnic minorities, reveals notable parallels between this approach and the European Union's model. It is important to emphasize that the U.S. model is characterized by fragmentation, reactivity and the prevalence of political and economic pressures over a unified legal philosophy or coherent regulatory framework. Evidence suggests that a similar pattern is also exhibited by the private sector, which has a significant impact on the U.S. legislature. In this study, the three companies selected for analysis are IBM, Microsoft and Samsung. These companies have been chosen due to their strong leadership in the field of AI, as well as their involvement in international AI regulation. In the IBM Everyday Ethics for Artificial Intelligence there is only one statement relating to ethnic minorities: "*diverse teams help to represent a wider variation of experiences to minimize bias. Embrace team members of different ages, ethnicities, genders, educational disciplines, and cultural perspectives*" (IBM 2025). In the case of Microsoft, there is no specific legal regulation. Microsoft has scattered the rules of conduct and ethics in several different toolkits and guidelines (Microsoft 2025a), from which it is difficult to identify one specific rule of internal law. From the general description of the rules of conduct, it can only be read that AI should be structured in such way as to avoid biases. A positive regulation in Guidelines for Human-AI Interaction relating to the process of creating AI is the Guideline 5: "*Keep in mind that social and cultural norms vary across groups and cultures. For example, an informal tone may be perceived as friendly in the United States and impolite in more formal cultures*" (Microsoft 2025b). It sensitizes the creator that matching relevant social norms may requires linking AI with the social and cultural context of its recipient (the final user), and consulting its expectations with them is required. Nevertheless, a more thorough examination of the norm reveals that it is overly general in nature and aligns with the language utilized. The other Guideline number 6 Mitigate social biases is not even a rule but a reminder to plan for identifying, testing, and mitigating fairness harms (Microsoft 2025c). This leads to the assumption that social biases can be present in any AI product, as a

rule. Microsoft's obligation is only to try to mitigate them, and not to construct an AI tool in such a way that it does not lead to social biases at all. The last one company, Samsung, approached the issue of AI similarly to IBM, introducing a group of ethical AI principles: fairness, transparency and accountability:

> Fairness: We will apply the values of equality and diversity in AI throughout its entire life cycle. We will not encourage or propagate negative or unfair bias. We will endeavor to provide easy access to all users. Transparency: Users will be aware that they are interacting with AI. AI will be explainable for users to understand its decision or recommendation to the extent technol ogically feasible. The process of collecting or utilizing personal data will be transparent. Accountability: We will apply the principles of social and ethical responsibility to AI. AI will be adequately protected and have security measures to prevent data breach and cyberattacks. We will work to benefit society and promote corporate citizenship through the AI system (Samsung 2025).

As evidenced by the precedent set by IBM and Microsoft, this provision is a soft clause that pertains to the issue of respect for humans. It does not guarantee or protect any human rights. It is challenging to extrapolate any rights from this provision due to its extensive scope.

The United States administration, under the auspices of President John Biden, developed and publicly disclosed the White House Blueprint for an AI Bill of Rights. This initiative follows an assessment of the prevailing legal framework for regulating the AI market in U.S., and it is intended to establish a set of guidelines to safeguard the rights of individuals and ensure the responsible development of AI technologies. As demonstrated previously, the efficacy of private sector regulation in ensuring adequate human rights protection has been demonstrated to be inadequate. The Blueprint for an AI Bill of Rights is a set of five principles and associated practices to help guide the design, use, and deployment of automated systems to protect the rights of the American public in the age of artificial intelligence. The list includes the following provisions: the implementation of secure and effective systems to prevent discrimination, the implementation of algorithmic protection of personal data privacy, the obligation to notify and explain, the mandate to implement so-called human alternatives and fallback solutions (U.S. White House 2023). This Act is characterized by its lack of binding force, which is a key issue that must be considered.

Therefore, it can be concluded that neither federal regulations nor state or private sector regulations directly address the protection of human rights in the U.S. AI legal framework. It is noteworthy that no state regulations were identified that pertain to human rights and AI. From a critical legal perspective, the regulation of AI in the U.S. is reactive, underdeveloped, and structurally biased in comparison to the domestic law of private actors. It is recommended that a federal legal framework for AI be established, with mandatory audits introduced and explicit lines of responsibility and accountability delineated. Absent legal action in this regard, there is a risk of perpetuating the injustice of opaque, unregulated automated AI systems.

Conclusion

To summarise, the preceding analysis demonstrates that the two legislative models differ not only in terms of the sources of law, but also in terms of the conceptual approach to the legal model for the control of AI use and production. It is evident that both E.U. regulations and U.S. law are inadequate in terms of safeguarding human rights in the domain of AI creation and utilization. Undoubtedly, a certain intermediate solution would be the introduction of guidelines by both legal systems and the partial regulation of this issue in the adopted legislation. Nevertheless, the fundamental issue is to ascertain the interplay between human rights and AI. This remains a salient consideration in the broader context of societal advancement and evolution. To summaries the analysis of all the legal norms cited above, it is important to note that they only contain general standards or objectives that should apply to the creation of AI. There is no mention of the need to regulate the market in specific areas of its functioning. There are no explicit standards or principles that address the specific risks that may arise in connection with the production and use of AI. The degree of generality is such that it is impossible to properly predict the specific behavior of a given entrepreneur in terms of the ethical creation and use of AI. The problem is their generality and lack of specifics. Despite the critical approach to the regulations in question, it should be considered positive that states see the need to adapt AI to the social and cultural environment of its users and recipients and to create some basic regulation. The objective is to ensure the protection of the general public.

The EU legal framework has been criticized for its apparent inadequacy in adequately protecting human rights. This is because it

relies solely on a risk assessment and mitigation procedure. In contrast, the U.S. has moved towards the implementation of non-binding guidelines. While the guidelines and guidance comprehensively address the protection of societal rights, they are non-binding on the private sector. In terms of data privacy and automated decisions, an analysis of the EU reveals that while the GDPR aims to protect personal data, it does not fully address the risks associated with automated decision-making processes or the potential bias of algorithms. In terms of transparency and accountability, there is a paucity of comprehensive regulation to ensure that AI systems operate transparently and that those affected by automated decisions have clear avenues of redress. Finally, the question of fundamental rights is of particular concern. The prevailing legal framework may not adequately safeguard fundamental rights such as the right to due process, freedom of expression, and non-discrimination in the context of AI systems. The potential for circumventing restrictions is inherent in EU law, provided that the state deems it justifiable. It is evident from the examples provided that as AI technologies become increasingly prevalent, there is an imperative for the development of more robust and targeted legislation in order to ensure the adequate protection of human rights.

References

Andersx, Oskal, Johan Mathis Turi, and Svein D. Mathiesen. 2009. *Reindeer herding, traditional knowledge and adaptationto Climate Change And Loss Of Grazing Land.* Alta: Arctic Council.

Ashwini, K.P. 2024a. *Bias from the past leads to bias in the future.* Accessed june 16, 2025: https://www.ohchr.org/en/stories/2024/07/racism-and-ai-bias-past-leads-bias-future (accessed February 23, 2025).

Ashwini, K.P. 2024b. *United Nation Human Rights Council.* July 30. Accessed June 16, 2025. https://www.ohchr.org/en/stories/2024/07/racism-and-ai-bias-past-leads-bias-future.

Blaine, Tristan. 2025. «Guide to AI Laws and Regulations in the U.S.» *LawSoup.* Accessed February 6, 2025. https://lawsoup.org/legal-guides/ai-laws-in-the-us-artificial-intelligence-regulations.

Buolamwini, Joy. 2022. *How I'm fighting bias in algorithms.* Accessed May 15, 2022: https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?language=en.

Council of Europe. 2025. «Human Rights and artificial intelligence.» *Human Rights Intergovernmental Cooperation.* Accessed January 1, 2025: https://www.coe.int/en/web/human-rights-intergovernmental-cooperation/intelligence-artificielle.

Council of Europe, CDDH-IA. 2025. *Human Rights and artificial intelligence.* Accessed January 15, 2025: https://www.coe.int/en/web.

Council of Europe/European Court of Human Rights. 2021. *Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention.* Strasbourg: Council of Europe.

Cowger, Alfred R. 2020. *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms.* Lanham: Rowman & Littlefield.

European Comission. 2021. *Shaping Europe's digital future - AI Pact.* Accessed December 20, 2024: https://digital-strategy.ec.europa.eu/en/policies/ai-pact.

European Comission. 2020. *On Artificial Intelligence - A European approach to excellence and trust.* White paper, Brussels: European Comission.

European Comission. 2019. «Shaping Europe's digital future.» *Ethics guidelines for trustworthy AI.* April 8. Accessed February 20, 2025: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

European Union. 2024a. «Regulation 2024/1689 laying down harmonised rules on artificial intelligence.» *Artificial Intelligence Act.* Brussel: Official Journal of the European Union. Accessed June 13, 2024: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.

European Union High-Level Expert on Artificial Intelligence. 2019. *A definition of AI: Main capabilities and scientific disciplines.* Brussels. Accessed June 14, 2025: file:///C:/Users/VT97/Downloads/ai_hleg_ai_definition_final_DF06F793-EA01-3573-16D2ACD625E2BDB0_56341.pdf.

European Union. 2016. «Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.» *Official Journal of the European Union.* Accessed December 20, 2024: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

Eurydice. 2021. *Norway, Population: demographic situation, languages and religions.* Accessed May 21, 2022: https://national-policies.eacea.ec.europa.eu/youthwiki/chapters/norway/overview.

IBM. 2025. *Everyday Ethics for Artificial Intelligence.* Accessed May 15, 2025: https://www.ibm.com/artificial-intelligence/ai-ethics.

Microsoft. 2025a. *Responsible AI resources.* Accessed May 15, 2025: https://www.microsoft.com/en-us/ai/responsible-ai-resources.

Microsoft. 2025b. *Guideline 5 Match relevant social norms*. Accessed May 15, 2025: https://www.microsoft.com/en-us/haxtoolkit/guideline/match-relevant-social-norms.

Microsoft. 2025c. *Guideline 6, Mitigate social biases*. Accessed may 15, 2025: https://www.microsoft.com/en-us/haxtoolkit/guideline/mitigate-social-biases/.

OECD. 2023. «AI Policy Observatory.» Accessed May 15, 2022: https://oecd.ai/en/dashboards/ai-principles/P6.

Samsung. 2025. *AI Ethics.* Accessed May 15, 2025: https://www.samsung.com/latin_en/sustainability/digital-responsibility/ai-ethics/.

Smith, Karen and Anna Hill. 2019. *Defining the Nature of Blended Learning through Its Depiction in Current Research.* Milton Park: Higher Education Research & Development.

The Council of Europe. 2024. *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.* Vilnus. Accessed September 5, 2024: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

The European Commission. 2018. *Artificial Intelligence for Europe.* Accessed March 27, 2025: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN.

The European Digital Rights Association. 2025. *EDRI.* 2025. Accessed May 15, 2025: https://edri.org/.

The European Parliament and of the Council. 2022. «Regulation 2022/1925 on contestable and fair markets in the digital sector.» *Digital Markets Act.* Brussel: EUR-Lex, September 14.

The European Parliament and of the Council. 2022. «Regulation 2022/2065 of the of on a Single Market for Digital Services.» *Single Market for Digital Services.* Official Journal of the European Union, October 19.

U.S. White House. 2023. *White House Blueprint for an AI Bill of Rights.* Accessed February 14, 2025: https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

U.S. Departament of State, Bureau of Cyberspace and Digital Policy. 2024. «Risk Management Profile for Artificial Intelligence and Human Rights.» July 25. Accessed December 20, 2024: https://2021-2025.state.gov/risk-management-profile-for-ai-and-human-rights.

United Nation. 2025. *Ms. Ashwini K.P. Special Rapporteur on contemporary forms of racism.* Accessed June 12, 2025: https://www.ohchr.org/en/special-procedures/sr-racism/ms-ashwini-kp.

United Nation, Department of Economic and Social Affairs. 2018. *The Report on the World Social Situation 2018, Promoting Inclusion Through Social Protection*. DOI: https://doi.org/10.18356/5ef37a49-en.

United States House of Representatives. 2020. *H.R.6216 - National Artificial Intelligence Initiative Act of 2020.* Washington: Congress, March 12.

Yeung, Karen. 2019. *A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework. Council of Europe study DGI(2019)05.* Council of Europe.

Zhu, Liming , Xiwei Xu, Guido Governatori, and Jo Whittle. 2021. «AI and Ethics-Operationalizing Responsible AI.» In *Humanity Driven AI*, edited by Jianlong Zhou and Fang Chen, 15-16. New York: Springer International Publishing.