Centro Studi
di Politica
Internazionale

CeSPI

AI as a Black Box for the Society

Study conducted as part of the project:

"Il balzo tecnologico e la cooperazione: quali sfide per i diritti?"

By:

Clarisa Elena Nelu

Scientific coordination by:

Prof. Michele Nicoletti

April 2024

Supported by:

Fondazione
Compagnia
di San Paolo

The implementation of the project THE TECHNOLOGICAL LEAP AND COOPERATION: WHAT CHALLENGES FOR RIGHTS?
is supported by the Compagnia di San Paolo Foundation,
as part of the "Geopolitics and Technology" 2023grant.

Table of Content

Introduction

The concept of artificial intelligence (AI) has gained influence in the last few years, especially with the new developments in the domain of different programs and systems using artificial intelligence to function. Through these changes, the impact of artificial intelligence became both positive and negative, with varying effects on society.

Artificial intelligence is reshaping economies, allowing them to make productivity gains, improve efficiency and reduce costs. It enhances people's lives and helps them make better predictions and decisions. These transformations, however, are at the beginning of their development, and there are still uncertainties on how to use them better and more effectively to tackle global challenges and promote innovation and growth rather than cause harm. It can already be seen that artificial intelligence's impacts are very present in our society; its transformational potential should be put at the service of people and organisations, but not without protection. At the same time, artificial intelligence is also fuelling anxieties and ethical concerns about its impact[1].

There are doubts about the trustworthiness of AI systems, such as that the artificial intelligence could affect the values on which the European Union is created and lead to breaches of fundamental rights, including, for instance, the freedom of assembly (Article 12), human dignity (Article 1), rights to freedom of expression (Article 11), non-discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation (Article 21), protection of personal data and private life (Article 8), or the right to an effective judicial remedy and a fair trial (Article 47), as well as consumer protection (Article 38)[2]. Another issue of AI is that fragmentation of the use of AI and the different regulations dealing with artificial intelligence are evident in artificial intelligence research activities and national AI development plans that aim to gain individual AI market leadership. No single entity has the answers to all these current challenges, and perhaps even more to appear in the future[3].

The world is divided among researchers and scholars regarding the developments of artificial intelligence. On one side is the potential of AI for good, and on the other side are the risks and negative effects that could arise while developing and using this technology. According to different scholars, artificial intelligence is changing the way people interact with technology and how the technology could have positive impacts on our lives in different fields. Luger and Sellen (2016) explained that the use of AI is helping people to execute both simple and complex tasks, from switching the light in the house just using your voice to having a car that drives itself. In other fields, Davenport and Ronanki (2018) discovered the importance of using AI tools to improve the production and delivery of services, while Szolovits (2019) talks about the improvement of healthcare services and transportation. According to Allam and Dhunny (2019), using new technologies, including AI, could help reduce climate change through the development of smart cities and using the technology more efficiently[4]. Other scholars, like Tomašev and Cornebise, suggest that building better tools and solutions while using AI could have a positive impact on achieving some of the UN' 17 Sustainable

[1] OECD, Artificial Intelligence in Society, 1st ed. Paris, 2019.
[2] Charter of Fundamental Rights of the European Union, Official Journal of the European Union, 2012.
[3] OECD, op. cit.
[4] S. F. Wamba & et al., "Are we preparing for a good AI society? A bibliometric review and research agenda", *Technological Forecasting & Social Change*, 164, 2021.

Development Goals[5]. A report issued by the UN, AI4SG, shows how the use of the new technology could improve the lives of people in developing countries while reducing the risks and discrimination. The two experts, Tomase and Cornebise, argue that to achieve sustained impact, it is important to develop long-term partnerships between application-domain experts and AI researchers and create deeply integrated collaborations that allow enough time to reach good practical outcomes[6]. According to Dulka, another potential benefit of using AI discovered by Amnesty International was the use of AI to help trained human moderators identify and analyse online abuse against women and children on social media[7].

According to Rother, healthcare is the field where people have higher hopes when talking about AI, as the latest developments show a positive impact in treating people and finding the answers to some unknown diseases. The AI community is very eager to implement more practices that could minimise suffering and help people live better and healthier lives[8]. Another field that cannot be ignored is education. According to Aslan and Zhang (2021), emerging technologies have also been transforming ways of teaching and learning in education. Expert systems can help with pedagogical planning and fully unleash the potential of learning management systems for teaching and learning while using intelligent tutors could provide customised, timely, and appropriate materials, guidance, and feedback to learn for students[9].

Two letters have been written by civil society organisations such as Access Now, AlgorithmWatch, EuroMed Rights, Human Rights Watch, etc., to ask for not trading away the fundamental rights and values of the people in the European Union. The concerns raised by those organisations are connected with discrimination against the general public, including race, gender, sexuality, religion and political views through predictive policing and biometric categorisation systems. The fear is that by using AI tools, authorities might get too much power as they could use the technology for mass biometric surveillance in public space or emotion recognition technologies like AI 'lie-detectors' in the context of migration[10]. In another letter, Turing Award winners Geoffrey Hinton and Yoshua Bengio, together with leaders of the major AI labs, including Sam Altman of OpenAI and Demis Hassabis of Google DeepMind, have signed a single-sentence statement from the Center for AI Safety that states that "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war". The possible threats that the development of AI tools could bring to our society are uncertain and impossible to predict as the technology is changing very fast without leaving enough space for policymakers to develop efficient regulations to "control" them[11].

---

[5] N. Tomašev, Cornebise, J., Hutter, F. *et al.* "AI for social good: unlocking the opportunity for positive impact", *Nat Commun* 11, 2468, 2020. https://doi.org/10.1038/s41467-020-15871-z

[6] Ibid.

[7] A. Dulka, "The Use of Artificial Intelligence in International Human Rights Law", *Standford Law School*, REV. 316, 2023. https://law.stanford.edu/wp-content/uploads/2023/08/Publish_26-STLR-316-2023_The-Use-of-Artificial-Intelligence-in-International-Human-Rights-Law8655.pdf

[8] R. Rothe, "How AI will make us live longer and healthier lives", AI Reality Bites, 2024. https://airealitybites.substack.com/p/how-ai-will-make-us-live-longer-and?r=1v54iy&utm_campaign=post&utm_medium=web&open=false

[9] K. Zhang & A. B. Aslan, "AI technologies for education: Recent research and future directions", *Computers and Education: Artificial Intelligence*, 2, 2021.

[10] Civil Society Organisations, "Letter: Open letter to EU AI Act negotiators: Do not trade away our rights!", 8 December 2023; Civil Society Organisations, "Council risks failing human rights in the AI Act", 29 November 2023.

[11] Center for AI Safety, "Statement on AI Risk: AI experts and public figures express their concern about AI risk", 2023. https://www.safe.ai/statement-on-ai-risk

It could be seen the breadth of the impact of AI on a multiplicity of areas such as work, health, education, democracy, economy, and a very high impact on a multitude of human rights. For these reasons, there is a need for cooperation among different actors involved in the international arena to guide these developments and the use of AI in a way that it is not harmful but beneficial for society. Different entities have started their work in the field of AI. However, the race to regulate artificial intelligence has tremendously increased in the last few years, especially since the EU Commission announced its plan to create an EU AI Act that regulates not only artificial intelligence but also the users of artificial intelligence. At the same time, the Council of Europe created CAI, a committee that is working on the development of an AI Convention that will supplement the AI Act regarding the protection of human rights. However, the European Union is not the only entity interested in regulating AI, as other big international players want to take part in the development of the new regulations.

At the last meeting of G7 which took place in May 2023, an idea was proposed to create an AI Code of Conduct to fill the empty spaces until the regulations are applicable. The proposal has been announced by the EU Executive Vice-President Margrethe Vestager; however, there are no details regarding the initiators of the document nor the specific content. The AI Code of Conduct was presented as a joint US-EU initiative to produce a draft set of voluntary commitments for businesses to adopt before the new regulations entered into force. The G7 could be seen as an arena where the EU and the US could collaborate to influence each other and put aside their policy differences. The Code of Conduct was released at the beginning of November 2023 under the Hiroshima AI Process, and it includes eleven voluntary guidelines for international players.

The study aims to examine the framework of artificial intelligence, the implications and developments of AI in society, and the reasons behind the urgency of creating a system to regulate artificial intelligence by the EU and the G7. The Council of Europe is working on a convention regarding AI. In 2019, the OECD produced a set of AI principles that are innovative and trustworthy, and that respect human rights and democratic values. However, these Principles are not a normative document but rather an instrument of soft-law to orient States in their activities. The European Union could have used these Principles as a starting point, instead, it decided to begin the process from scratch. The study also looks at the effect of artificial intelligence on human rights, as this is one of the biggest concerns expressed by several players, including civil society. The goal of the study is to help build a shared understanding of artificial intelligence in the present and near future and to clarify some of the uncertainties that are arising from the new binding and non-binding documents that are under development by the different actors.

Looking ahead, we see that there is a need to work together and regulate AI-related developments. However, it is also essential to better understand the concept of AI and its influence on our lives and the planet. AI technology is rapidly advancing, and its application in diverse fields is expected to proliferate in the near future. There is a need for a proper framework to keep up the speed with these faster changes and only sometimes having the regulator chasing the technology. The EU wants to avoid this scenario of chasing something that might be difficult to catch by creating a regulation based on an open framework rather than on specific/rigid rules which could already be inefficient in a few months after being realised. The urgency of regulating AI was likely coming from the desire to reduce the uncertainties that AI implies for individuals or could have been another strategy of the international players to influence when it comes to new developments, currently focusing on artificial intelligence. Either way, it is no surprise that AI has become a top priority on the EU's agenda. An answer to these doubts will be provided by saying that AI is like a black box, as the outcome seems to be entirely unexpected.

What is Artificial Intelligence?

This specific section of the study offers an overview of the existing concepts regarding artificial intelligence in order to better understand the technology and its impact on our daily lives. Some of the concepts might be already familiar as they might be part of our activities, however doubts might always arise when dealing with such a sophisticated and always-changing technology.

Definition of Artificial Intelligence

Since 1950, when Alan Turing thought of whether or not the machine could have the capacity to think, several changes have happened regarding artificial intelligence. In the last few years, breakthroughs have taken place in the field of artificial intelligence, such as machine learning, in which the machines learn from historical data and create predictions in new situations, the development of self-driving cars, the proliferation of generative AI tools like ChatGPT and Google's Bard. Currently, artificial intelligence is increasingly becoming part of everyday life. Generally speaking, artificial intelligence systems are capable of performing tasks usually connected with human behaviour, such as interpreting conversations, playing games, and identifying models. These machines are taught how to perform those tasks by using vast amounts of data and looking for similarities to shape their own decision-making. In most cases, humans will assist an AI's learning process, implementing good decisions and techniques and trying to avoid the negative ones. However, there are prototypes of AI systems that are created to learn without supervision, such as by playing a video game over and over until they finally figure out the rules and how to win[12].

As an important aspect, it is understandable that several actors tried to influence the evolution of AI in the last years, however, currently, there is no universally accepted definition by the international community to explain what AI is and what it implies when it is being used. It has been recognised by the latest AI research the fact that there is no definition for the term 'artificial intelligence'. According to Antebi (2021), developing only one accepted definition of AI is challenging for two main reasons; firstly, due to the variety and diverse approaches to research the AI field and secondly, there is a high difficulty in defining or deciding on a definition of "intelligence," because of limitations that have not yet been breached in the study of neuroscience or philosophy[13]. Therefore, the capability to evaluate these concepts about machines or apply them to machines is restricted. In 2019, the OECD released the Principles on Artificial Intelligence, which aims to increase innovation and trust in AI by advertising the responsibility of trustworthy AI while guaranteeing respect for human rights and democratic values.

The principles focus on AI-specific issues and set a standard that is implementable and sufficiently flexible to stand the test of time in this rapidly evolving field. As part of these recommendations, the OECD proposed a definition of artificial intelligence, "an AI system is a machine-based system capable of influencing the environment by producing an output (predictions, recommendations, or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models through analysis in an automated manner (e.g., with

[12] T. Madiega, 'Artificial Intelligence Act overview', European Commission, 2023.
[13] L. Antebi, 'What is Artificial Intelligence?', *Institute for National Security Studies*, 2021.

machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy"[14]. The definition has been accepted by the countries that are part of the OECD; however, other definitions have been constructed by other entities to explain the concept of artificial intelligence.

Currently with the development of the EU AI Act, the proposed definition of AI, which is not very different from the one proposed by the OECD four years ago, is that "'artificial intelligence system' (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments"[15]. The Council of Europe proposed its definition of artificial intelligence systems in the AI Convention such as "any algorithmic system or a combination of such systems that uses computational methods derived from statistics or other mathematical techniques to carry out functions that are commonly associated with, or would otherwise require human intelligence and that either assists or replaces the judgment of human decision-makers in carrying out those functions [...]"[16]. Even with the proposed definition by the Council of Europe, the definition inside the AI Act seems stable, as in December 2023, the European Parliament has validated its negotiating position on the proposed AI Act. However, there were rumours that the European Council is thinking of narrowing down some of the definitions without meaning which of those could have been. The definition proposed by the European Council for the AI Act is "'artificial intelligence' system means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches, and produces system-generated outputs such as content, predictions, recommendations, or decisions, influencing the environments with which the AI system interacts"[17].

After the approval of the EU AI Act in December 2024, the definition of artificial intelligence systems in the AI Act aligns with internationally recognised criteria, following OECD guidelines, which define an AI system as the following "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments"[18]. The definition is an important first step in order to deal with the uncertainties posed by artificial intelligence systems and models, as the definition creates a framework in which to fit this technology. It still seems to be a broad definition; however, the regulations have to adapt to the fast developments of the technology, and perhaps a too narrow definition would not have been able to always keep up with the date.

---

[14] OECD, Artificial Intelligence in Society, 1st ed. Paris, 2019.

[15] *European Parliament of 2023 on Artificial Intelligence Act on amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).*

[16] *Council of Europe of 2023 on Committee on Artificial Intelligence (CAI) regarding the Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law.*

[17] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

[18] A. Pingen, "AI Act: Parliament and Council Reach Provisional Agreement on World's First AI Rules", *eucrim*, 2024. https://eucrim.eu/news/ai-act-parliament-and-council-reach-provisional-agreement-on-worlds-first-ai-rules/#:~:text=The%20definition%20of%20an%20AI,outputs%20such%20as%20predictions%2C%20content%2C

Symbolic Learning and Machine Learning

From the beginning, the artificial intelligence field has been divided into different domains, two of them being symbolic learning and machine learning. While symbolic learning used to dominate the sector in the beginning, machine learning has been catching up lately with its new inputs and different methods of operating. Starting with symbolic learning (symbolic AI), the system works by processing symbols that represent objectives or concepts in the world and their interactions. The primary approach of the system is to use logic-based programming, where rules and axioms are used to make inferences and deductions. Symbolic learning utilises formal languages such as logic to represent knowledge. This allows for the development of systems such as expert and decision support systems that could make predictions and inferences based on predefined data and knowledge.[19] Its principal function is to rely on symbolic representations of knowledge. In this approach, information and knowledge are mostly revealed through symbols, such as logical predicates, rules, or symbols in a formal language. Symbolic learning highlights symbolic reasoning and logic. It often applies deductive reasoning to manipulate symbols and draw outcomes. The expert systems, which are based on symbolic reasoning, are a classic example of symbolic learning. These systems use human-encoded information in the form of rules and symbols to make decisions and solve issues. Symbolic systems are often more transparent and interpretable because the knowledge is explicitly represented in a human-readable form[20].

The machine learning system is another pathway to artificial intelligence, which uses algorithms to automatically learn insights and acknowledge patterns from data, using that learning to make better decisions. The machine learning models perform specific tasks without utilising straightforward instructions but rather focusing instead on patterns and deductions, figuring out "the rules"[21]. The algorithms build mathematical models based on training data in order to make predictions. It teaches the computer how to solve the issues and to gain insights from solving those issues. This is the reason why the laptop learns automatically, without human implication: by analysing and looking for similarities in data and using feedback loops to keep track of and improve its forecasts. While humans would be overwhelmed with masses of data, machine learning thrives, and it can learn by itself and then reproduce it[22].

In other words, machine learning focuses on learning samples and representations directly from data without relying on explicit symbolic representations of knowledge. Machine learning algorithms make use of statistical techniques to identify models and make predictions or decisions based on data[23]. Deep learning, a sub-part of machine learning, has gained prominence in recent years, particularly in applications like image recognition and language processing. Deep learning models use large neural networks to learn complex patterns and make predictions independent of human input. Many machine learning models, especially

[19] D. Ernst, 'Competing in AI: Major Challenges', *Centre for International Governance Innovation*, 2020.
[20] J. Torras, 'Perspectives Symbolic AI vs. machine learning in natural language processing', *Multilingual* [online article], 2023, < https://multilingual.com/issues/may-june-2020/symbolic-ai-vs-machine-learning-in-natural-language-processing/> , accessed 29 September 2023.
[21] W. E. Agin, 'A Simple Guide to Machine Learning', *Business Law Today* , 2017, pp. 1-5.
[22] 'What is machine learning?', IBM [online journal], 2023, < https://www.ibm.com/topics/machine-learning >, accessed 29 September 2023.
[23] W. E. Agin, 'A Simple Guide to Machine Learning', *Business Law Today* , 2017, pp. 1-5.

deep neural networks, are often considered black-box models because it can be challenging to understand why they make specific decisions[24].

The two systems are different from each other due to their versatile way of using the data and then interpreting it. Symbolic learning is often used in domains where clear knowledge representation is critical, such as expert systems or rule-based reasoning systems, while machine learning, especially deep learning, has dominated in tasks where the underlying patterns are complex and not easily expressible through symbolic rules, such as image recognition and natural language understanding. Another difference among them could be the application of one of the systems. Symbolic learning could be more interpretable, making it suitable for applications where understanding the reasoning behind decisions is important, while machine learning models may be more suitable in tasks that require learning from massive amounts of data or dealing with high-dimensional and unstructured data. Symbolic learning is based on knowledge representation and reasoning, while machine learning learns samples directly from data. They have different approaches to solving issues and represent various ways of shaping and understanding intelligence. In practice, there is an increasing trend toward hybrid approaches that combine symbolic learning with machine learning models to highlight the strengths of both systems. These hybrid models aim to keep interpretability while benefiting from the data-driven abilities of machine learning.

Generative AI

Generative AI includes a category of artificial intelligence systems that generate new outcomes based on the data they have been instructed to learn. Opposite to the traditional AI systems that are formatted to recognise models and make predictions, generative AI produces new content in the form of images, text, audio, and more. Generative AI is good at tasks that involve generating new and creative content, such as text generation, image synthesis, and even music composition. It is particularly useful for tasks where the underlying patterns are complex and may need to be more easily expressed by clear rules. The models can learn from large datasets and adapt to new samples and trends. They can improve their performance over time with more data and fine-tuning. The generative AI finds applications in creative tasks like natural language generation, image generation, style transfer, and data augmentation. It could also be used in some forms of recommendation systems and data synthesis[25]. Examples of generative AI models include Generative Adversarial Networks, Variational Autoencoders, and language models like GPT-3.5. These models work by learning samples and structures from huge datasets and then using that knowledge to develop new content. For example, GANs can generate realistic images, while language models like GPT-3 can generate human-like text. Generative AI has entered into public consciousness and is increasingly present in peoples' everyday conversations worldwide[26].

In the last few years, generative artificial intelligence has taken centre stage in the public, academic, and political discussions surrounding artificial intelligence, and it is even predicted

[24] J. Torras, 'Perspectives Symbolic AI vs. machine learning in natural language processing', *Multilingual* [online article], 2023, < https://multilingual.com/issues/may-june-2020/symbolic-ai-vs-machine-learning-in-natural-language-processing/> , accessed 29 September 2023.

[25] G. Lawton, 'What is generative AI? Everything you need to know', TechTarget [online article], 2023, https://www.techtarget.com/searchenterpriseai/definition/generative-AI#:~:text=Generative%20AI%20focuses%20on%20creating,many%20types%20of%20new%20outputs., accessed 29 September 2023.

[26] A. Kumar, 'Generative AI – Concepts, Use Cases, Examples', Analytics Yogi [online article], 2023, https://vitalflux.com/generative-modeling-in-machine-learning-examples/, accessed 29 September 2023.

to create significant economic value. Growth in generative AI research, including its open-source code development, preceded the surge in investments. The generative AI produces new content in response to prompts, offering transformative potential across multiple fields such as education, entertainment, healthcare, and scientific research. Companies have begun to adopt the technology to create new business opportunities, while start-ups are competing for venture capital. While generative AI has the potential to change the industry and society in positive ways, the use of the technology also poses risks to individuals and institutions as every branch of artificial intelligence, such as potential shifts in labour markets, copyright uncertainties and risks connected with the perpetuation of societal biases and the possible for misuse in the sharing of disinformation and manipulated content[27]. Several players, such as the US, China[28] and EU member states (Netherlands)[29], acknowledge the changeable impact of generative AI and are cordially taking action to address these challenges.

In May 2023, G7 leaders addressed different subjects for discussion in the Hiroshima Artificial Intelligence Process, one of the topics being the opportunities and challenges regarding generative AI. The report presents the results of a questionnaire developed for G7 members to support stocktaking and help guide G7 discussions on common policy priorities about generative AI. In light of the advent of generative AI (e.g., ChatGPT), provisional agreement of the EU Parliament on the EU AI Act and increased ethical concerns and risks regarding Generative AI, in early July 2023, the Italian Department for Digital Transformation presented a proposal for revision of the "Strategic Plan for Artificial Intelligence 2021", including aspects regarding the use of generative AI on its territory. The proposed revision is nearing completion and will be subject to a public consultation procedure[30].

In summary, generative AI relies on machine learning and data-driven approaches to generate new content, and it can be applied across many areas of the sector. The way of operating makes it easier to interpret and understand existing data and automatically develop new content. As with any technology, there are advantages and disadvantages to using it. Some of the potential advantages of implementing generative AI are: automating the manual process of writing content, reducing the time of responding to emails, improving the response to specific technical queries, creating realistic representations of people, or simplifying the process of creating content in a particular style. Some of the disadvantages of using generative AI include the possibility of non-identification of the sources, the difficulty in identifying the bias of sources, and the difficulty in understanding how to tune in to new circumstances[31]. The model brings very promising tools for upgrading the existing technology; however, the new generative AI methods open a box of uncertainties such as accuracy, trustworthiness, bias, hallucination, and plagiarism, which might take a few years to be sorted out. The EU should realise that generative AI has the potential to change both the defensive and the

[27] OECD, 'Principles of AI', OECD.AI Policy Observatory [online], 2023, < https://oecd.ai/en/ai-principles>, accessed 29 September 2023.

[28] A. Wan, 'Why Japan is lagging behind in generative A.I. — and how it can create its own large language models', *CNBC* [online], 2023, < https://www.cnbc.com/2023/07/07/why-japan-is-lagging-behind-in-generative-ai-and-creation-of-llms.html>, accessed 12 October 2023.

[29] 'AP eist duidelijkheid techbedrijf over AI gericht op kinderen', Autoriteit Persoongegevens, 2023, < https://www.autoriteitpersoonsgegevens.nl/actueel/ap-eist-duidelijkheid-techbedrijf-over-ai-gericht-op-kinderen>, accessed 12 October 2023.

[30] OECD, 'G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI', OECDiLibrary, 2023.

[31] G. Lawton, 'What is generative AI? Everything you need to know', TechTarget [online article], 2023, https://www.techtarget.com/searchenterpriseai/definition/generative-AI#:~:text=Generative%20AI%20focuses%20on%20creating,many%20types%20of%20new%20outputs., accessed 29 September 2023.

offensive sides of security, from allowing security groups to better prepare for and react to threats to generating complex social engineering attacks[32].

General Purpose of AI

General purpose AI refers to AI systems that have human-like intelligence and can perform a wide range of duties, similar to how humans can adapt and learn to do various activities. The general purpose of AI is characterised by the capacity to understand, learn, and apply knowledge across different domains, tasks, and contexts. It is not limited to specific predefined tasks but can handle a broad spectrum of challenges. The system includes a wide range of possible applications, both intended and unintended by their developers. They can be used for many different tasks in various domains, often without substantial modification and fine-tuning. These systems are becoming increasingly valuable commercially due to the growing amounts of computational resources available to developers and innovative methods to use them. Current general-purpose AI systems are characterised by their scale (a lot of memory, data, and powerful hardware) as well as their capacity to transfer learning methods (applying knowledge from one task to another)[33].

These systems are sometimes called *"foundation models"* and are characterised by their widespread use as pre-trained models for other, more specialised AI systems. For example, a single general-purpose AI system for language processing could be used as the foundation for several hundred applied models (e.g., chatbots, ad generation, decision assistants, spambots, translation, etc.), some of which can then be further fine-tuned into several applications tailored to the customer. Achieving artificial general intelligence (AGI) is a long-term goal in AI research, and it seems that true artificial general intelligence has not yet been realised. Today most of the AI systems used, including popular ones like GPT-3, are not AGI but rather Narrow AI or specialised AI systems designed for specific tasks[34].

In order to eliminate any confusion that might be between the last two subchapters, general-purpose AI aims to create artificial intelligence systems with broad and versatile intelligence that can perform a wide range of tasks, mimicking human-like adaptability, while generative AI refers to AI models that create new content based on models and data they have learned. General purpose AI is a broader and more ambitious concept, while generative AI is a specific application within the wider field of artificial intelligence.

[32] R. Csernatoni, 'Generative AI Poses Challenges for Europe', Carnegie Europe [online], 2023, < https://carnegieeurope.eu/strategiceurope/90803>, accessed 9 November 2023.
[33] T. Madiega, 'General-purpose artificial intelligence', European Commission, 2023.
[34] 'General Purpose AI and the AI Act', *Future of Life Institute*, 2022.

AI in the Spotlight

The third chapter focuses on the significantly increasing influence that artificial intelligence has gained in the last few years, even a few months, and the urgency in creating regulations to impose some barriers while developing and using artificial intelligence. In these years, dominated by the development of several technologies surrounding the AI field, uncertainties and doubts about the correct use and no harm online have been raised by civil society and companies from the field. In 2019, the OECD created the Principles of AI, which highlight some recommendations for international actors on how to stay protected while using AI and measures on how to develop AI while not infringing on several human rights. In 2021, the European Commission announced the creation of a specific EU AI Act that is going to change the current scenario of developing or using AI by the people without having a clear plan or an end date for the Act. The urgency of finishing the Act seemed to appear in the first months of 2023 when the European Commission announced that the Act would be done by the end of 2023, in June, the Parliament accepted the proposed draft done by the Commission. At the same time, the Council of Europe is working on an AI Convention that is supposed to support the AI Act regarding full compliance with human rights and the rule of law. At the beginning of November 2023, the G7 released the Code of Conduct that aims to fill the gaps between the EU and the US while the EU AI Act is in progress. All these sudden movements from the big players created confusion at the international level among different stakeholders, such as companies, nations, and civil societies which were excluded from the drafting process of the global treaty on AI at the end of 2022. What could be the top priority of the big players behind all these processes and the creation of new regulations instead of improving the existing ones?

Principles of AI by the OECD – 2019

In 2019, the OECD published the final version of its recommended Principles on Artificial Intelligence. The aim of the Principles is to guide states, organisations, and individuals in creating and processing AI systems in a way that protects people's interests and guarantees that designers and operators are held responsible for their functioning. At the moment, there are 42 countries (including the 36 OECD members) that have signed up to the Principles. Part of the development process were OECD experts on AI and representatives of governments, researchers, professional organisations, and businesses such as Facebook, Google, IBM, and Microsoft. The past experiences and work carried out by the OECD's working groups and outsider entities have proved the need to shape a stable policy environment at the international level to promote trust in and adoption of AI in society[35]. Given this background and the lack of a proper regulatory framework for AI's development, the OECD Committee on Digital Economy Policy agreed to develop a draft proposal to promote a human-centric approach to safety AI that encourages research, preserves economic motives to innovate, and applies to all international actors.

The Principles were designed as an OECD legal instrument representing common ground for its member states. The document has been divided into two parts such as values-based principles and the recommendations for policymakers. The first part, values-based principles,

[35] M. Tanna, & W. Dunning, 'OECD Principles on Artificial Intelligence', Simmons+Simmons [online], 2019, < https://www.simmons-simmons.com/en/publications/ck0b9c83b7dpj0b9470kxyobl/290519-oecd-principles-on-artificial-intelligence>, accessed 29 September 2023.

includes different principles that should guide the development and use of an AI system, such as *inclusive growth, sustainable development and well-being, human-centred values and fairness, transparency and explainability, robustness, security and safety* and *accountability*. The recommendations for the policymakers focus on investing in AI research and development, fostering a digital ecosystem for AI, providing an enabling policy environment for AI, building human capacity, *preparing for labour market transition* and last but not least minor *international cooperation for trustworthy AI*[36]. The guide goes beyond the OECD's usual focus on economic issues and points out challenges such as privacy, individual and worker rights, and the reliability and safety of AI systems.

One of the issues regarding the Principles is that they are not legally binding to the member states; however, they do represent a growing international consensus on the development and governance of AI systems[37]. OECD was one of the first organisations that took the initiative to create a framework to regulate the development and use of the AI system that became, after a short period, a good AI policy practice to mitigate risks and help ensure responsible AI design followed by several international players. The situation might change, as new binding regulations are in the pipeline, and they seem to include many adjustments to the current AI framework.

EU AI Act by the European Commission – 2021/2023

The EU AI Act is expected to become the world's first comprehensive legal framework for artificial intelligence. Often referred to as the **"**Brussels Effect**"**, the EU AI Act is seen as a changing game for the regulation of artificial intelligence, as the EU General Data Protection Regulation was for the regulation of data protection a few years ago[38]. The first proposal of the AI Act was made by the European Commission in April 2021 as a political commitment by President von den Leyen, who mentioned in her political guidelines for the 2019-2024 "a Union that strives for more", and that the Commission would put forward regulation for a harmonised European approach on the human and ethical involvement of AI[39]. In December 2022, the policy was adopted by the European Council and later this year, in June 2023, after some amendments adopted by the European Parliament, the final legislative act entered into the final negotiations called "trilogues". The negotiations took place between the three involved legislative bodies, such as the European Commission, European Council (Council in the following parts) and EU Parliament, which aimed to develop a final version of the AI Act by the end of the year 2023, beginning of 2024.

The negotiations were announced to be a tough step before the deadline as the several differences between the Council, which is made by representatives of EU governments, and the Parliament, whose main priority is to secure the democratic legitimacy of the European law, need to find the middle way in the whole process[40]. Several voices have been sceptical regarding a "perfect" combination between the protection of rights in the legislative texts and

[36] OECD, Artificial Intelligence in Society, 1st ed. Paris, 2019.

[37] M. Tanna, & W. Dunning, 'OECD Principles on Artificial Intelligence', Simmons+Simmons [online], 2019, < https://www.simmons-simmons.com/en/publications/ck0b9c83b7dpj0b9470kxyobl/290519-oecd-principles-on-artificial-intelligence>, accessed 29 September 2023

[38] M. Fazlioglu, 'Contentious areas in the EU AI Act trialogues', IAPP [online], 2023, <https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues/> accessed 3 October 2023.

[39] U. von der Leyen, 'A Union that strives for more: Political Guidelines for the next European Commission 2019-2024', European Commission, 2019.

[40] M. Fazlioglu, 'Contentious areas in the EU AI Act trialogues', IAPP [online], 2023, <https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues/> accessed 3 October 2023.

the protection of the economy and social interests of the member States. It was important for every player (private and public) that the new Act will continually allow for developing new technologies without having too many barriers and regulations. In the last months, tension among the European institutions could have been noticed as the draft was out there, however, a common point on several issues was difficult to be achieved from the beginning such as the definition of AI, the list of prohibited A uses (e.g., AI use for biometric surveillance in public spaces), obligations on high-risk, foundational model and governance. The EU AI Act aims to be a Regulation which means it will be directly enforceable upon the member states once entering into force[41]. This pressure made it harder for the European institutions to make up their "minds" as the legislation is supposed to keep up with the fast changes in AI technology, as mentioned above.

The definition of AI was a starting point of the discussion as the European Commission placed it in the regulation's appendix, while the Council and Parliament thought that its place was more right in the body of the text. The Parliament also updated the definition of "AI system" under Article 3 in line with the definition developed by the OECD in 2019. This was a request, as the definition of the "AI systems" was heavily criticised due to the lack of clarity and the idea that it could have led to legal uncertainty especially for some systems that would not identify as AI systems under the draft text, while their use may have a negative influence on fundamental rights. To tackle this concern, policymakers propose to broaden the aim of the regulation to include all computational systems used in the identified high-risk domains, regardless of whether they are considered to be AI or not[42].

Other significant changes made by the Parliament were referring to users of AI systems as "deployers" and including new definitions for "affected persons", "foundation model" and "general purpose AI system". Regarding the definition of high-risk AI, the Parliament upheld a broader definition of the systems under high-risk AI, while the Council was inclined to a narrower definition of it[43]. Reasons for this issue could be many; as the Parliament had the pressure to pass the EU AI Act by the end of the year, a broader definition would allow for more room for interpretation, meaning to be accepted by a more significant majority of the European Member States. At the same time, a broader definition is more suitable for keeping up with the new advanced technology that is expected to come in the following years. On the other hand, the Council was under pressure to support the concerns raised by civil society regarding possible violations of human rights. The current definitions expose two concerns; on one hand, they could be too broad, including simple and generic terminologies that do not express too much, leaving room for interpretation; however, on the other hand, more precise definitions could endanger the law's efficiency and exclude the future developments of artificial intelligence.

The AI Act is constructed on three approaches, each of them defining a different type of AI risk system that the regulation should cover. The first approach is the unacceptable AI systems, which are prohibited completely by Article 5, such as the AI systems that deploy harmful manipulative "subliminal techniques", AI systems that exploit specific vulnerable groups, AI systems used by public authorities or on their behalf, for social scoring purposes and the last one being the "real-time" remote biometric identification systems in publicly available places for the scope of law enforcement, except in a limited number of circumstances such as:

[41] S. Lynch, 'Analysing the European Union AI Act: What Works, What Needs Improvement', Stanford Human Cantered Artificial Intelligence [online], 2023, <https://hai.stanford.edu/news/analyzing-european-union-ai-act-what-works-what-needs-improvement>, accessed 3 October 2023.

[42] Madiega, T., 'Artificial Intelligence Act overview', European Commission, 2023.

[43] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

a) "the nature of the situation giving rise to the possible use, in particular, the seriousness, probability and scale of the harm caused in the absence of the use of the system;
b) the consequences of the use of the system for the rights and freedoms of all persons concerned, in particular the seriousness, probability and scale of those consequences[44]".

"In addition, the use of 'real-time' remote biometric identification systems in publicly accessible spaces shall comply with necessary and proportionate safeguards and conditions in relation to the use, in particular as regards the temporal, geographic and personal limitations[45]". In the last moments of the negotiations, MEP Dragoș Tudorache revealed in an interview, that there were still some issues under Article 5 Prohibitions in the Act as the Council would want exemptions for national security regarding law enforcement, while the European Parliament advocates for a stricter approach on facial recognition technology, with a hard ban on use in public spaces. Another concern was the use of high-risk applications, where the Council advocated for more leeway for law enforcement while the Parliament was in favour of a more stringent approach[46]. Other parliamentarians have asked for the prohibition of biometrics even on devices that could have prevented a mobile phone from unlocking with fingerprints or face recognition. The proposal had not been included in the final version of the Parliament, even if it was a sensitive topic[47].

The second approach, and the most controversial one, was the high-risk AI systems, which were allowed but subject to stricter obligations and could be divided into two categories. First, it focuses on those added-in products that are already subject to third-party examination under the legislation and serve division as safety components for said products, such as medical devices or toys. The second category focuses on AI systems that are not included in other products such as biometric identification and categorisation of natural persons, management and operation of critical infrastructure, education and professional training, employment, worker supervision and access to self-employment, access to and enjoyment of necessary private services and public services and benefits, law enforcement, migration, asylum and border control governance, and administration of justice and democratic processes. These top-level categories of high-risk AI systems are under the strictest obligations under the law[48].

Article 52 of the EU AI Act offers people living in the European Union the right to know if the video they are watching is a deep fake if the person they are talking to is a chatbot or a voice assistant, and if they are under an emotion recognition examination or a biometric categorisation made by an AI system; however, it does not apply to AI systems approved by law to detect, prevent, investigate, or prosecute criminal offences. While it ultimately added

[44] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

[45] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

[46] Fleming-Jones, J., 'EU AI Act nearing agreement despite three key roadblocks – co-rapporteur', *Euronews.next* [online], 2023,<https://www.euronews.com/next/2023/10/23/eu-ai-act-nearing-agreement-despite-three-key-roadblocks-co-rapporteur>, accessed 9 November 2023.

[47] Chee, F. Y., and Mukherjee, S., 'EU lawmaker Benifei urges the bloc's countries to compromise on AI rules', Reuters [online], 2023, <https://www.reuters.com/technology/eu-lawmaker-benifei-urges-blocs-countries-compromise-ai-rules-2023-09-21/>, accessed 16 October 2023.

[48] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

the restriction on remote biometric identification systems in public by law authorities, the Parliament agreed to be used with prior judicial authorisation. It also introduced a prohibition on biometric categorisation systems that use sensitive features, such as ethnicity, citizenship, gender, race, status, religion, and political orientation[49].

At the time of the negotiations, the Parliament was pushing for a broader list of prohibited AI systems that leave some space for interpretation, including software that scrapes facial images from the web, while the Council advocated for a narrower list with more regulations. In this case, further changes to Article 5, which is the article inside the EU AI Act referring to prohibited artificial intelligence practices, are highly expected in these months during the negotiations, especially around the biometric identification ban and its inconsistencies[50].

The third approach is the low or minimal risk and it includes the rest of the concerns which do not represent the previous risks, primarily including transparency requirements that allow users to make informed decisions about them[51].

A major change that has been implemented by the Parliament at the negotiations in June was to impose a regime on foundation models as a high-risk system "there is significant uncertainty as to the way foundation models will evolve, both in terms of typology of models and in terms of self-governance. Therefore, it is essential to clarify the legal situation of providers of foundation models…"[52]. The "foundational models" are not yet real AI systems but models that are developed and 'trained' on millions of data, which can be integrated into AI systems. Such models are GPT-3 and GPT-4 which become the basis for many applications, the most notable being the ChatGPT. The Parliament introduced the "providers of a foundation model" (defined in Article 3(1)) within the scope of various obligations beyond the minimal transparency requirements[53].

These new regulations on foundation models proposed by the European Parliament within Article 28b(2) include requirements to:

- demonstrate "the identification, the reduction and mitigation of reasonably foreseeable risks to health, safety, fundamental rights, the environment and democracy and the rule of law ..." using "appropriate design, testing and analysis";
- apply certain data governance tools "to examine the suitability of the data sources and possible biases and appropriate mitigation";
- involve independent analysts, document the findings, and do "extensive testing" to reach "appropriate levels of performance, predictability, interpretability, corrigibility, safety and cybersecurity";
- develop "extensive technical documentation and intelligible instructions" that would allow compliance with Articles 16 and 28(1) by downstream providers;

[49] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

[50] M. Fazlioglu, 'Contentious areas in the EU AI Act trialogues', IAPP [online], 2023, <https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues/> accessed 3 October 2023.

[51] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

[52] *European Parliament of 2023 on Artificial Intelligence Act on amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).*

[53] Ibid.

- establish a "quality management system" that would ensure and achieve compliance with Article 28[54].

Opposite to this extending form of approaching the foundation models, the Council's approach was to request the European Commission to create tailored obligations for the parties involved after a year and a half of the AI Act's entry into force. These new changes, even if adopted in June, were still creating blockages in the discussions,[55] in the last few weeks, being contested by a coalition formed by France, Germany, and Italy.[56] Under pressure from their domestic AI companies, these three countries are advocating for 'mandatory self-regulation' for foundation models through a code of conduct, without an initial sanction regime for foundation models but rather some prescriptive obligations in the AI rulebook[57].

In order to overcome these deadlocks, the Commission suggested horizontal rules for all general-purpose AI models and codes of practices for the top-tier ones, and regarding the sanctions, the Spain presidency proposed "the fines to be settled as a percentage of the company's global annual turnover or a predetermined amount as following 6.5% for violations of the banned AI applications, 3% for violations of the AI Act's obligations, and 1.5% for the supply of incorrect information"[58]. However, the negotiations regarding these new changes were still in progress and according to the last document released in January 2024, the percentages increased from 6.5% to 7% and decreased from 15.% to 1%[59].

The last issue that had to be negotiated during the trialogues was the clarification of the governance. How will the AI Act be enforced and coordinated among different national and EU institutions? As the AI Act's structure looks similar to the GDPR, it will give various attributions to the national enforcement actors on an Artificial Intelligence Board, similar in work to the European Data Protection Board[60]. The Parliament also suggested the creation of a new EU body, the AI Office (Article 56), which would be equipped with a variety of administrative, consultative, interpretive, and enforcement-related tasks, as well as authority for managing cross-border investigations. The European AI Office will be an independent body of the Union and will have a legal personality. It will be the centre of AI expertise across the EU, playing a significant role in implementing the AI Act - especially for general-purpose AI - to foster the development and use of trustworthy AI and international cooperation. The

---

[54] *European Parliament of 2023 on Artificial Intelligence Act on amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).*

[55] Fleming-Jones, J., 'EU AI Act nearing agreement despite three key roadblocks – co-rapporteur', Euronews.next [online], 2023,<https://www.euronews.com/next/2023/10/23/eu-ai-act-nearing-agreement-despite-three-key-roadblocks-co-rapporteur>, accessed 9 November 2023.

[56] Bertuzzi, L., 'France, Germany, Italy push for 'mandatory self-regulation' for foundation models in EU's AI law', *EURACTIV* [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/news/france-germany-italy-push-for-mandatory-self-regulation-for-foundation-models-in-eus-ai-law/>, accessed 29 November, 2023.

[57] Kohn, B., 'AI Act at risk? – the regulation of foundation models and general-purpose AI', *Taylor Wessing* [online], 2023, https://www.taylorwessing.com/en/insights-and-events/insights/2023/11/ai-act-at-risk, accessed 29 November, 2023.

[58] Bertuzzi, L., 'AI Act: Spanish presidency makes last mediation attempt on foundation models', *EURACTIV* [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-spanish-presidency-makes-last-mediation-attempt-on-foundation-models/>, accessed 29 November 2023.

[59] *Council of the European Union of 2024 on Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement.*

[60] M. Fazlioglu, 'Contentious areas in the EU AI Act trialogues', IAPP [online], 2023, <https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues/> accessed 3 October 2023.

AI Office was established within the European Commission as the centre of AI expertise and forms the foundation for a single European AI governance system[61].

An important aspect of the AI Act will be not to overlap with other existing European regulations or not infringe some of the national ones. It happened in the spring of 2023 when Italy's data protection authority, the Garante, for a few weeks, it suspended ChatGPT over privacy concerns. While the Garante subsequently lifted its ban, the new regulation should be carefully analysed in its way of prohibiting or not certain systems and models[62].

The AI Act will have a broad territorial scope, applying to AI system providers in the EU market, whether operating within the EU or another country. The stakes were high as the European Parliament wanted to have a text done by the end of the year 2023, which, in fact it happened on 8th of December when the EU AI Act was adopted by the EU institutions. On the 13th of March 2024, the European Parliament finally adopted the EU Artificial Intelligence Act. The AI Act is the world's first horizontal and standalone law governing AI and a standing piece of legislation for the EU. The AI Act will probably enter into force at the end of April or early May 2024. After entry into force, the AI Act will not be immediately enforceable but will be subject to a moderate transition and implementation phase whereby it will become fully applicable 24 months after entry into force, except for: bans on prohibited practices (6 months after the entry into force), codes of practice (9 months after entry into force), general-purpose AI rules (12 months after entry into force), and obligations for high-risk systems (36 months after entry into force)[63].

Convention on Artificial Intelligence by Council of Europe – 2023

In 2021, the Committee of Ministers of the Council of Europe approved the creation of a Committee on Artificial Intelligence (CAI) assigned with developing legal tools for the development, format, and application of AI systems based on the Council' standards on human rights, the rule of law, democracy, and conducive to innovation. The purpose is to guarantee that AI systems obey human rights, respect the functioning of democracy, and observe the rule of law, regardless of whether these activities are undertaken by public or private actors. The Convention on AI has a comprehensive approach to ensure that the use of artificial intelligence does not interfere with any of the principles or fundamental rights developed in the ECHR such as the Principle of Non-discrimination (Article 3), Principle of Equality and Anti-discrimination (Article 12), Principle of Privacy and Personal Data Protection (Article 13), Principle of Accountability, Responsibility and Legal Liability (Article 14) or the Principle of Safety (Article 16)[64].

---

[61] *European Parliament of 2023 on Artificial Intelligence Act on amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).*

[62] R. Browne, 'Italy became the first Western country to ban ChatGPT. Here is what other countries are doing', CNBC [online], 2023, < https://www.cnbc.com/2023/04/04/italy-has-banned-chatgpt-heres-what-other-countries-are-doing.html> accessed 3 October 2023.

[63] W. Long et all, "EU Formally Adopts World's First AI Law", *Sidley*, March 2024. https://datamatters.sidley.com/2024/03/21/eu-formally-adopts-worlds-first-ai-law/#:~:text=On%20March%2013%2C%202024%2C%20the,of%20legislation%20for%20the%20EU.

[64] *Council of Europe of 2023 on Committee on Artificial Intelligence (CAI) regarding the Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law.*

A "Zero Draft" [Framework] Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law published in February 2023, followed by a Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law which was published in July 2023. The second draft contains provisions that have been agreed on during the first reading of the revised Zero Draft. The Council of Europe was expected to deliver a final proposal by the 15th of November 2023, however during the 4th European AI Alliance Assembly: Leading Trustworthy AI Globally, Ambassador Jānis Kārkliņš (Permanent Representative of Latvia to the Council of Europe) announced that the negotiations will be done by May 2024. In the meantime, the Convention remains open to new ideas and articles which only States could propose[65], as civil society has been excluded at the beginning of 2023 by the US representatives[66]. The AI Convention is supposed to complement the EU AI Act and other regulatory initiatives that might develop in the future. The EU AI Act focuses on the marketing of products using AI within the EU internal market, while the AI Convention aims to protect the fundamental rights of people affected by AI systems. The EU AI Act will be directly implemented upon the EU Member States, while the AI Convention is principle-based and will establish legally binding individual human rights, which will apply to the States, EU Member States and non-EU Member States that decide to ratify and sign the AI Convention in the future[67].

Code of Conduct by G7 – 2023

In May 2023, during the G7 meeting, an idea was proposed to create an AI Code of Conduct to fill the empty spaces until the regulations of the EU AI Act are applicable. The AI Code of Conduct was presented as a joint US-EU initiative to produce a draft set of voluntary commitments for businesses to adopt that will focus on transparency and risk auditing. At the beginning of November 2023, the G7's leaders released the Code of Conduct, which aims "to promote safe, secure, and trustworthy AI worldwide and will provide voluntary guidance for actions by organisations developing the most advanced AI systems, including the most advanced foundation models and generative AI systems"[68].

These eleven voluntary guidelines create space for the private sector, a critical prerequisite for meaningfully implementing AI regulation and the private industry, which is expected to be able to submit feedback on the draft framework. The scope of the document is to be reviewed and updated regularly, including through ongoing inclusive multistakeholder consultations, to ensure it remains fit for purpose and responsive to this rapidly changing technology. Different jurisdictions may decide their approach to implementing these guidelines in different ways[69]. The Code of Conduct seems to emphasise the importance of human rights as the States should

---

[65] CAI - Committee on Artificial Intelligence, *Council of Europe* [online], 2023, < https://www.coe.int/en/web/artificial-intelligence/cai>, accessed 13 October 2023.

[66] L. Bertuzzi, 'US obtains exclusion of NGOs from drafting AI treaty', *EURACTIV* [online], 2023, < https://www.euractiv.com/section/digital/news/us-obtains-exclusion-of-ngos-from-drafting-ai-treaty/>, accessed 12 October 2023.

[67] H. van Kolfschooten, 'The Council of Europe's Artificial Intelligence Convention: Implications for Health and Patients', *Bill of Health: Artificial Intelligence, Biotechnology, Human Rights, International* [online], 2023, < https://blog.petrieflom.law.harvard.edu/2023/04/18/council-of-europe-artificial-intelligence-convention/#:~:text=The%20AI%20Act%20focuses%20on,people%20affected%20by%20AI%20systems.> accessed 13 October 2023.

[68] *Hiroshima Artificial Intelligence Process of 2023 on the Hiroshima Process International Code of Conduct for Advanced AI Systems.*

[69] Ibid.

comply with their obligations under international human rights law to ensure that human rights are respected and protected, while "private sector activities should be in line with international standards such as the United Nations Guiding Principles on Business and Human Rights and the OECD Guidelines for Multinational Enterprises"[70].

Before the revealing of the Code of Conduct, the main concerns were regarding its aim, as a new regulation, EU AI Act, is on its way, and a new soft law might seem insignificant at this moment and the clashes that might arise among the main parties such as the European Union and the US. The two main parties are also very different when it comes to regulations of artificial intelligence. The US prefers a vertical approach to regulation, balancing innovation and regulation, similar to the non-binding Blueprint for an AI Bill of Rights, opposite to the EU, which has taken a horizontal approach to AI regulation in its AI Act, creating four levels of risk to classify all AI systems, which aims to cover the whole risk scale[71]. Both the US and EU continue to engage with and right-size the role of the private sector as they construct a transatlantic approach to global AI governance. While the AI Code of Conduct seems to be narrow in scope, a soft law, and a nascent effort, it represents an important first step in writing the rules of the road for the development of AI systems among democracies and among two of the biggest markets worldwide.

The rapid transformations of the development of AI systems need concrete measures as both the US and EU attempt to develop long-standing regulations that will face the time. There should be no differences among the transatlantic partners. Instead, they should ensure a safe and protected space for private and public entities to carry out their activities. The Code of Conduct is no silver bullet; however, it is the first step for the transatlantic partners in shaping the global governance of AI[72]. According to the European Commission, the eleven principles adopted by the G7's leaders and the EU provide guidance for organisations in creating, sharing, and using new AI systems, such as foundation models and generative AI, to promote the safety and trustworthiness of the technology. They include commitments to reduce risks and misuse and identify vulnerabilities, encourage responsible information sharing, report incidents, and invest in cybersecurity, as well as a labelling system to enable users to identify AI-generated content. Following the outcomes of a stakeholder survey, these principles have been created under the Hiroshima Artificial Intelligence Process[73].

A point that remained unclear was the rush to create a new regulation, as the EU AI Act is very close to being approved. The respective Act is supposed to regulate every aspect of artificial intelligence and is expected to have not only an EU jurisdiction but also an intentional jurisdiction, as several other States are expected to adhere to these regulations. There was a growing discrepancy between the EU and the US in their approach to regulating artificial intelligence, as the EU has moved forward on laws around data privacy, online platforms, online e-commerce regulation, and many more, while similar legislation does not exist in the US[74]. However, it seems that the Code of Conduct brought together the two big

[70] Hiroshima Artificial Intelligence Process of 2023 on the Hiroshima Process International Code of Conduct for Advanced AI Systems.

[71] C. Nietsche, & C. Ford, 'US-EU AI Code of Conduct: First Step Towards Transatlantic Pillar of Global AI Governance?', EURACTIV [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/opinion/us-eu-ai-code-of-conduct-first-step-towards-transatlantic-pillar-of-global-ai-governance/>, accessed 3 October 2023.

[72] Ibid.

[73] European Commission of 2023 on Commission welcomes G7 leaders' agreement on Guiding Principles and a Code of Conduct on Artificial Intelligence, [press statement], < https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5379> accessed 9 November 2023.

[74] S. Lynch, 'Analysing the European Union AI Act: What Works, What Needs Improvement', *Stanford Human Cantered Artificial Intelligence* [online], 2023, <https://hai.stanford.edu/news/analyzing-european-union-ai-act-what-works-what-needs-improvement>, accessed 3 October 2023

markets in a joint document meant to help seize the benefits and address the risks and challenges brought by these new technologies[75]. For some of the G7 countries, the non-EU Member States, this Code of Conduct might have been the last chance to impose some of their preferences regarding regulating artificial intelligence before the EU AI Act enters into force.

[75] F. Y. Chee, ' Exclusive: G7 to agree AI code of conduct for companies', *Reuters* [online], 2023, <https://www.reuters.com/technology/g7-agree-ai-code-conduct-companies-g7-document-2023-10-29/>, accessed 9 November 2023.

Impact of the AI on Human Rights

General Overview

In the past few years, artificial intelligence began to have an impact on human rights, undermining some of the values and principles that are supposed to strengthen human rights rather than weaken them. AI-driven technology is becoming more influential in our lives, from smart home appliances to social media applications, and it is increasingly being used by different authorities to evaluate people's personalities or skills, allocate resources, and even make decisions that could have severe repercussions for the human rights of individuals. Finding the right balance between technological transformations and human rights protection is, therefore, an urgent matter that has to be tackled by the international community[76].

Having a look at the EU AI Act, the entire document has a specification regarding the protection of fundamental rights. However, the Act does not specify the rights that should be a priority when it comes to artificial intelligence. The Act aims to prevent a chilling effect on the rights to freedom of expression and freedom of assembly, to ensure the rights of defence and the presumption of innocence, the protection of the right to a legitimate remedy and to a fair trial, and the standard principle of good administration. The AI Act suggests that the regulations have a positive effect on the rights of a variety of particular groups, such as the workers' rights to fair and just working conditions, a high level of consumer protection, the rights of the child and the integration of persons with disabilities. As AI systems could produce adverse outcomes to the health and safety of people, in particular when these systems operate as parts of the products, people will have the right to complain and notify any abuses that might happen to them. Several times the EU AI Act mentioned "the respect of the fundamental rights of the affected persons, notably their rights to free movement, non-discrimination, protection of private life and personal data, international protection, and good administration" without giving more information on how these rights will be protected from incorrect use of artificial intelligence. There are specifications on how the national courts are entitled to apply fees and restrict the use of technologies if they do not respect the obligations under the EU AI Act, however, a clear understanding of these uncertain scenarios is still far away as the legislation is still under the process of becoming a law. Perhaps a more careful approach to how the fundamental rights would be protected should be taken in defending ways to protect any harm caused by AI technologies, not only mentioned a few times, "the right to an effective remedy and to a fair trial".

The proposal imposes some restrictions on the freedom to conduct business (Article 16) and the freedom of art and science (Article 13), in order to guarantee compliance with the most important purposes of public concerns, such as safety, consumer protection, health and the protection of other fundamental rights, when high-risk AI systems are created and used[77]. For an in-depth debate regarding these weaknesses that might be part of the EU AI Act, several organisations and active groups ask for public transparency and that the people have more power to challenge the possible harms that might arise; having better limitations on harmful and discriminatory surveillance by national security, law enforcement authorities and

[76] *Council of Europe, Commissioner of Human Rights of 2023 on thematic work regarding Artificial Intelligence and Human Rights.*

[77] *European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.*

migration controls; and removing the loopholes that could undermine the regulation[78]. As mentioned above, the EU AI Act is created in a broader sense in order to "include" every Member States' interests, however, from the civil society perspective, it seems that the key element of the Union, such as the citizens, might be the one that could be in danger in the long term if these regulations are not carefully thinking through. This could be a possible limit to the regulation such as finding the right balance between the society and the citizens.

To better understand the impact of AI on human rights, a closer look at some fields where there is an increased use of artificial intelligence, such as healthcare and education, might be a solution. There are positive human rights impacts from using artificial intelligence. Regarding the right to life, liberty, and security of a person, AI-based diagnostic systems enhance the enjoyment of the right to life by making accurate, high-quality diagnostic services more widely available. For the right to desirable work, the improved health outcomes that AI-based diagnostic systems are likely to produce will reduce the number of people who are excluded from the dignity of work for medical reasons. By detecting diseases earlier and more accurately, AI-based diagnostic systems will improve living standards and quality of life, contributing to the protection of the right to an adequate standard of living. One negative aspect could be connected with the right to privacy, as AI-based diagnostic systems require the collection of vast quantities of sensitive data relating to an individual's often immutable health characteristics, raising several privacy concerns[79].

Talking about education, there are some positive human rights impacts as well. Artificial intelligence in education research has been conducted in many countries around the world. The 40 articles reported AI in education research studies in 16 countries, with the US in first place, followed by China and then Spain and Turkey in the third position[80]. Automated grading systems may positively impact the enjoyment of economic rights, as writing is a key skill for success in the labour market, in this case fulfilling the right to an adequate standard of living. The right to education is connected with the automated grading systems that may improve educational outcomes by increasing the availability of feedback on student's writing abilities, thereby spurring performance improvements. By using artificial intelligence, professors and educators could spend more time focusing on higher-order teaching tasks and pay more attention to students rather than wasting time grading papers or preparing course materials that could be easily done by a machine[81].

With the surge of virtual reality technologies, research has started exploring the potential benefits of visualisations and virtual learning environments with AI in education. A study on an intelligent glass system also confirmed that AI technology with visualisations helped both children and adults with autism by serving as a social communication tool[82]. The role of AI in education could have several benefits, one of them being the transfer of knowledge and experiences among teachers who are not living in the same area of the globe. The advanced tools used in the education system in the North could be easily transferred to the South with the help of this new technology. There are projects in the global South that use artificial

[78] Human Rights Watch, "EU: Artificial Intelligence Regulation Should Protect People's Rights", 2023, https://www.hrw.org/news/2023/07/12/eu-artificial-intelligence-regulation-should-protect-peoples-rights

[79] F.A. Raso & H. Hilligoss, 'Artificial Intelligence & Human Rights: Opportunities & Risks', *Berkman Klein Center for Internet & Society Research Publication* Series, 2018.

[80] K. Zhang & A. B. Aslan, "AI technologies for education: Recent research and future directions", Computers and Education: Artificial Intelligence, 2, 2021.

[81] F.A. Raso & H. Hilligoss, 'Artificial Intelligence & Human Rights: Opportunities & Risks', *Berkman Klein Center for Internet & Society Research Publication* Series, 2018.

[82] K. Zhang & A. B. Aslan, "AI technologies for education: Recent research and future directions", Computers and Education: Artificial Intelligence, 2, 2021.

intelligence and virtual tutors to reach as many teachers as possible with brief and focused courses to help them face critical day-to-day issues of teaching[83].

Besides the impact on individual rights, AI could also affect the social rights, not only of workers in different sectors but also of citizens, such as equality of access to AI, equality of treatment, and non-discrimination by AI. Access to the internet is not granted in all parts of the world, with many regions having no access to the internet or a network in general. A term that has been defined by the Antonio Casilli is "micro-workers", who is a special type of digital laboratory that is required for training and programming artificial intelligence. These people are everywhere around the world, however the large number is located in the low-wage countries, as the big companies want to reduce their labour costs. It is, therefore, no surprise that the companies that produce AI systems are based in the global North, while the workers needed to keep the machines going are mostly placed in the global South[84].

According to the research done by Caselli, the situation is different in Europe and the rest of the world. If, in some European countries such as France and Germany, these micro-workers might get a relatively good salary as they have an above-average level of education, the situation is changing for the people coming from the country of the global South, such as Africa or South America. According to his studies, Venezuela and Madagascar were two very important examples, as Venezuela is the source of micro-work for both Spanish-speaking and English-speaking countries, while Madagascar is for France. Even if people could double their salary by doing this kind of work, it does not mean that micro-work is well paid for. Above all, it means that the living conditions in many countries are so low that even a small amount of money makes a difference. However, many AI companies are taking advantage of this situation of people from the Global South, as they might need better alternatives to improve their living conditions[85]. In terms of their qualification, the majority of people doing these tasks are well qualified compared to the population average. However, they earn little compared to the skills they have. Their salary could be higher, according to their qualifications, if they would come from the global North. This type of work also comes with risks, such as health issues due to exposure to a lot of violent and/or pornographic content. The main task of the click workers is to train algorithms in such a way that they recognise, filter, and sort out or block that type of online content. Changing their situation, improving their working conditions, and recognising their rights are quite tricky for two reasons. Firstly, it is due to the large distances involved, which makes effective self-organisation to protect their interests and rights more challenging, and secondly, it is due to the public invisibility of these micro-workers who are working from home. In most situations, nobody knows about their work and the struggle that they might encounter while doing this type of job, as they might be unable to talk about that due to the confidentiality clauses with the private companies[86].

Another challenge that AI might bring to human rights is the impact on the prohibition of discrimination and the right to equal treatment. As exemplified earlier, the involvement of AI could lead to different treatment of the workers, but it could also enshrine discriminatory or unacceptable biases. As an example, it could be Amazon's recruitment AI process, which favoured men over women because it was trained on profiles of successful Amazon

[83] Teachers Outreach, "Project: The global teachers initiative", 2023. https://teachersoutreach.org/

[84] Antonio A. Caselli, " "The More AI Progresses, the Greater the Need for Human Labor", *Extensive Interview in the German Journal Soziopolis*, 24 January 2024, https://www.casilli.fr/2024/01/24/the-more-ai-progresses-the-greater-the-need-for-human-labor-extensive-interview-in-the-german-journal-soziopolis-24-jan-2024/.

[85] Ibid.

[86] Antonio A. Caselli, " "The More AI Progresses, the Greater the Need for Human Labor", *Extensive Interview in the German Journal Soziopolis*, 24 January 2024, https://www.casilli.fr/2024/01/24/the-more-ai-progresses-the-greater-the-need-for-human-labor-extensive-interview-in-the-german-journal-soziopolis-24-jan-2024/.

employees, which happened to be only men. The AI system not only filtered out women but also looked at the characteristics of successful employees, such as typical wording and phrasing, and filtered out CVs that did not show these characteristics[87].

Looking at the current AI system, where the systems check for correlations and keywords to match, all types of unacceptable biases could arise in the field of employment even if the companies want to ensure that the system is not jeopardised, just a look at the actual data in used show that this might not be true. The system uses an algorithm that uses data from past examples, which is already biased as in some companies, "successful candidate" might mean different than in other companies, and it might already leave out some underrepresented communities, such as persons with disabilities if successful records of these communities are not part of the algorithm. The biases could come from several backgrounds, such as race, gender, skin colour, or the choice of language used. To prevent the biases from happening, private companies should implement more transparency and provide information about how the AI systems operate. They should have specific requirements in the system that could stop the biases from happening and identify their negative impacts on the rights of people with different backgrounds. An inclusion of these communities in the building process of these technologies could be a further step to prevent the biases made by the AI systems[88].

Staying in the same field, AI can bring benefits when used for heavy, dirty, and exhausting work that could put people in danger or it is too boring to be done. However, the use of AI in the working field, such as monitoring the workers, distributing work without human intervention or assessing worker performance regarding hiring and firing situations, could bring infringe to the workers' human rights, such as the right to conditions of work, safe and healthy working conditions, dignity at work as well as the right to organise (Art. 2 and 3 of European Social Charter (ESC))[89]. The involvement of the AI in such decisions, could reduce the qualifications of the workers, as more and more decisions that were made by humans are taken over by the AI systems. This could not only lead to a less skilled workforce but also increases the risk of systemic mistakes, where only a few humans are capable of working with AI systems and responding to situations where these systems fail[90]. For the moment, it is unknown how AI is going to affect the job market and if it is going to have a positive or a negative impact on people's right to work, however, a gap between vulnerable labour forces and required skills could lead to technological unemployment in the future if the countries do not properly regulate the market.

Another field where AI might interfere with some fundamental human rights is regarding liberty and security, fair trial, and no punishment without law, which are very important articles (5, 6, 7) in the European Convention on Human Rights. The fact that AI can maintain or increase existing biases is particularly essential when used in law enforcement and the judiciary. In situations where physical freedom or personal security is at risk, such as predictive policing, recidivism risk determination and sentencing, the right to liberty and security combined with the right to a fair trial are becoming vulnerable[91].

---

[87] Jerneja Turin, "Artificial intelligence and its impact on the human rights of persons with disabilities", *ENNHRI*, 3 December 2023, https://ennhri.org/news-and-blog/artificial-intelligence-and-its-impact-on-the-human-rights-of-persons-with-disabilities/.

[88] Jerneja Turin, "Artificial intelligence and its impact on the human rights of persons with disabilities", *ENNHRI*, 3 December 2023, https://ennhri.org/news-and-blog/artificial-intelligence-and-its-impact-on-the-human-rights-of-persons-with-disabilities/.

[89] Catelijne Muller, "The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law", *ALLAI*, 20 March 2020.

[90] Ibid.

[91] Ibid.

The use of AI and other new technologies can have significant consequences on people's lives, including both the younger and older generations. It may lead to a digital divide between the ages, referring to the gap between people who face difficulties accessing, affording, or using digital tools. These disadvantages reduce their quality of life and increase insecurities and gaps between people. The older generation is likely to be the most affected as they have been living in a different society than the one we know today. Most of them do not have access to the internet and are unfamiliar with the fast advancements in technology. As we move facilities like healthcare, payments, and communication into digitalisation, their accessibility may be affected. During the pandemic, the gap between people with digital access and those without was evident, as those without access faced several risks, such as depression, loneliness, or anxiety, when they had to stay inside their houses and were unable to visit family or friends[92]. As AI tools progressively redefine our workplaces, activities, communication, and everyday lives, there is a risk that AI could deepen, rather than narrow, the digital divide. We should be careful not to forget the older generation in this fast-paced technological advancement and consider the positive impact that the use of AI could have on their lives if society chooses to improve their quality of life, such as better medical treatments and communication with the ones who are far away from them.

As can be seen, the relationship between artificial intelligence and human rights could be very complex, and without appropriate regulations, technology could have a harmful effect on several human rights. In order to make use of the positive potential use of AI and prevent the harmful consequences that could arise, societies need to ensure that the advantages of artificial intelligence outweigh the risks in the long term. To do that, there is a need for regulations and limitations. The EU AI Act takes responsibility for ensuring that human rights are protected while society embraces the advantages of AI transformations. However, there are still concerns about how the AI Act will guarantee the balance between the protection of human rights and the use of AI technology, as it could have been seen from the examples that overstepping and infringing on some human rights might happen without the intention behind it.

Concerns by the Civil Society Organisations

Since January 2023, when the US abdicated for the exclusion of civil society from participating in the drafting of the AI treaty by the Council of Europe[93], in the last months, several civil society organisations such as Algorithm Watch, Statewatch, European Digital Rights, European Civic Forum and many more, have gathered together in an attempt to make their voices heard regarding the protection of human rights and the rule of law through different letters addressed to the EU institutions. As AI systems are more and more used by law enforcement, migration control and national security authorities, the EU AI Act is a significant opportunity to prevent harm, protect human rights and issue legal measures for authorities to use artificial intelligence within the boundaries of respecting the rule of law.

Civil society organisations are worried about the use of AI systems that could increase discriminatory checks and mass surveillance, and in order to prevent that from happening,

[92] Sarah Bentley and Claire Naughtin, "The 'digital divide' is already hurting people's quality of life. Will AI make it better or worse?", *CSIRO*, March 2024. https://www.csiro.au/en/news/all/articles/2024/march/digital-divide-ai

[93] L. Bertuzzi, 'US obtains exclusion of NGOs from drafting AI treaty', *EURACTIV* [online], 2023, < https://www.euractiv.com/section/digital/news/us-obtains-exclusion-of-ngos-from-drafting-ai-treaty/>, accessed 12 October 2023.

they were calling for some measures that the AI Act should include, such as legal limits prohibiting AI for uses that show an unacceptable risk for fundamental rights, public transparency and oversight when police, migration and national security agencies use 'high-risk' AI, and the insurance that the AI Act properly regulates the uses of AI in policing, migration and national security that pose risk to human rights[94]. In the last drafts proposed by the EU Parliament, in the section regarding high-risk AI, some loopholes would undermine the rule of law, as the Parliament does not want to provide only a specific list regarding the systems under the high-risk AI section but also a section called "additional layer". Such a structure would allow AI developers, without responsibility or qualifications, to decide whether or not their systems pose a 'significant' risk to justify their "good behaviour" under the AI Act. This freedom of self-deliberation could undermine the entire purpose of the AI Act, shifting to self-regulation, creating numerous challenges for enforcement and harmonisation, and encouraging the developers to categorise their own AI systems[95].

In December 2023, the final version of the EU AI Act was adopted. However, the text raised concerns from different civil society organisations that do not see that the framework of the Act is fully protecting fundamental human rights. According to the final proposal, some AI systems are banned from the EU on the grounds that they pose an unacceptable risk to basic rights and freedoms. They include emotion-recognition tools in the workplace or educational institutions, as well as the biometric recognition of sensitive data such as sexual orientation and some cases of predictive policing. The use of real-time remote biometric recognition, or facial identification, in public places is prohibited, except for use by law authorities to counter terrorism and to search for victims or perpetrators of serious crimes[96]. The fear that persists in society is that everything might look good on paper; however, going in-depth into the content might not seem as good as it appears, as the exemption created in the law might lead to abuses in the future.

At the beginning of the negotiations, the ban on emotion recognition was adopted in four contexts such as education, workplace, law enforcement and migration. However, under pressure from different Member States, the prohibition on using it in law enforcement and migration contexts was removed from the final text, demonstrating that the AI Act might deem migrant people and already-marginalised people less worthy of protection. The deal might also allow exceptions to the ban on emotion recognition for medical or safety purposes, leading to a gaping loophole that could prove extremely dangerous. There is the so-called "aggression detection" system that could identify images of black men as more aggressive than white men; if such a system were used in schools or workplaces, it could lead to racist surveillance of black students or workers, for example[97].

According to civil society organisations, the creation of the EU AI Act should empower people and public interest actors to take positive actions and understand when the AI systems might cause damages and violations of fundamental rights. In order to do that, it is essential that the EU AI Act provides a structure based on accountability, transparency, and accessibility. The EU AI Act should include a total ban on real-time and post-remote biometric identification in publicly accessible spaces, by all actors, without exception; a

[94] 'Police AI needs strict limits and controls, warn civil society organisations', *Statewatch* [online], 2023, < https://www.statewatch.org/news/2023/september/police-ai-needs-strict-limits-and-controls-warn-civil-society-organisations/>, accessed 4 October 2023.

[95] A civil society statement on fundamental rights in the EU Artificial Intelligence Act, 'EU Trialogues: The AI Act must protect people's rights', 2023.

[96] J. Gill, "What does the EU's AI act mean for human rights?", *Thomas Reuters Foundation*, December 2023, https://www.context.news/ai/what-does-the-eus-ai-act-mean-for-human-rights

[97] D Leufer, C. Rodelli and Fanny, "Human rights protections…with exceptions: what's (not) in the EU's AI Act deal", *Access Now*, December 2023, https://www.accessnow.org/whats-not-in-the-eu-ai-act-deal/

prohibition of all forms of predictive and profiling systems in law enforcement and criminal justice; prohibitions on AI in migration contexts to make individual risk assessments and profiles based on personal and sensitive data; a ban on the use of emotion recognition systems to infer people's emotions and mental states; remove exceptions and loopholes for law enforcement and migration control introduced by the Council; and secure public transparency as to what, when and how public actors share high-risk AI in fields of law enforcement and migration control, avoiding any exemption to the obligation to record high-risk operations into the EU AI database[98].

In a last attempt to express their concerns, civil society organisations required that the AI Act should protect the rule of law without any exceptions and the introduction of mandatory fundamental rights impact assessments (FRIAs) that have been discussed at some point by the Parliament. The principle of the rule of law is part of Article 2 TEU, and it is connected with the protection of democracy and fundamental rights. Therefore, the nexus between the development and deployment of AI systems and the rule of law should be created. The AI FRIAs might be considered in the case of an emerging trend toward the creation of standards and certifications for correct AI governance measures. It should be an obligation for all high-risk AI systems to make sure that they are produced in a way that does not infringe on the principles of justice, fairness, and accountability. The FRIAs aim to develop a structure to evaluate and avoid possible fundamental rights violations and make sure that AI systems are functioning for the benefit of the people[99].

The goal of these civil society organisations is to protect the human rights of the citizens, especially in the process of creating new regulations that might infringe on those rights. In the case of the EU AI Act, the right to participation has been taken away from these organisations, and the final draft has been approved, but they are still advocating for an EU AI Act that includes the proper measures that will not violate people's rights. According to some civil societies, the AI Act is not a bulletproof regulation. From its earliest days, it has had a lack of inclination toward industry and law enforcement, and despite the efforts of some actors involved in the negotiations, for the moment, it seems that it is inclined to protect the authorities and private companies far more than the people[100].

New Digital Rights

Due to the rapid transformation of technology, perhaps there is a need for "new" digital rights that could be added as part of the fundamental rights inside the Charter. According to Custers, the specific rights are not new, as some of them have been implemented. However, most of them are novel concepts only in the literature. The digital rights that he is suggesting are the following: the right to be offline (to disconnect at some point), the right to internet access (parts of the European Union without internet access, in remote areas), the right not to know (a life without being assaulted by information), the right to change your mind, the right to start over with a clean digital footprint, the right to expiry dates for data and the right to know the value of your data[101].

---

[98] A civil society statement on fundamental rights in the EU Artificial Intelligence Act, 'EU Trialogues: The AI Act must protect people's rights', 2023.

[99] Civil Liberties Union for Europe, European Civic Forum, European Center for Not-for-profit law, 'The AI Act Must Protect the Rule of Law', Statewatch, 2023.

[100] D Leufer, C. Rodelli and Fanny, "Human rights protections…with exceptions: what's (not) in the EU's AI Act deal", *Access Now*, December 2023, https://www.accessnow.org/whats-not-in-the-eu-ai-act-deal/

[101] B. Custers, 'New digital rights: Imagining additional fundamental rights for the digital era', *ScienceDirect* [online], volume 44, 2022,

These rights might not seem highly important at first glance; however, people spend most of their time online, perhaps without realising the importance of being protected while navigating or using different applications. The purpose of the EU AI Act is to introduce regulations to ensure that citizens are protected online. However, unintended violations could happen at any point. Having explicit rights that could guarantee our safety online while using different technologies might be an idea to think about for the future.

<https://www.sciencedirect.com/science/article/pii/S0267364921001096?ref=pdf_download&fr=RR-2&rr=810cbcce9b4ea32d#section-cited-by>, accessed 4 October 2023.

Conclusion

To conclude, the AI system seems to be like a "black box" for society that cannot be 100% predicted as the new technologies appear on a conveyor belt and the inputs and outputs are visible to the public, but the steps a system takes to get from one input to other output are not very transparent. The system is continuously changing, bringing new challenges and insecurities to the international legal framework as the laws have to keep up with these developments. The latest AI Act, proposed by the European Union, is expected to become the world's first comprehensive legal framework for artificial intelligence that will change the behaviour of how AI is used and developed. The AI Act aims to bring positive changes into society, however, according to civil society, some of the proposals seem not to align with other fundamental principles such as human rights or the rule of law. Once the AI Act enters into force, the international dynamics among the different players in regard to artificial intelligence will change, as the aim of the Act is to have a worldwide influence.

By the first quarter of 2024, the Council of Europe is expected to have a final draft of the AI Convention, which will supplement the EU AI Act regarding protecting human rights and the rule of law. It could be that the European Commission's ambition is to internationalise the EU AI Act through the Council of Europe so that it will have the same influence as the GDPR. For these reasons, it seems that other non-European big players felt left out, currently trying to influence the situation through the G7's Code of Conduct, a voluntary document that it seems not very appropriate at this point in time, as its effect is not as powerful as the other Acts. On top of the urgency of developing these legal papers, civil society organisations are concerned about how these new transformations will affect the fundamental rights of the users and the general framework of using/developing artificial intelligence.

The regulations cannot be too narrow as they will leave out some of the new technologies coming along. However, a document that is too broad will not change the actual system either, as it might be too self-regulated. A balance has to be found between controlling artificial intelligence and protecting human rights in an environment that respects all the parties involved. The stakes are high, as on one side, there are the fundamental rights that must be protected, and on the other side, there are the different interests involved from the Member States, leaving the burden on the European Union as it is the one which started the "game", however not the only international player that wants to regulate the use of artificial intelligence.

The speed at which AI is developing has "forced" lawmakers to think outside the box and begin the development of new regulations that could keep up with these new technologies. The EU chose a 'top-down' approach to regulate AI on the European continent, while the UK is taking a sector-based approach. In November 2023, 28 countries, including the United States and China, signed a declaration to support transparency and accountability from developers of AI technology to mitigate potential harms. Before the meeting, in October 2023, President Biden issued an executive order called the Blueprint for an AI Bill of Rights, requiring developers of AI systems which pose a risk to national security, the economy or public health to share the outcomes of safety tests before they are released to the public[102]. Opposite to the EU AI Act, the Blueprint is a non-binding legislation and includes five principles that are in place to reduce the potential risks from AI systems. On the East side of

[102] J. Gill, "What does the EU's AI act mean for human rights?", *Thomas Reuters Foundation*, December 2023, https://www.context.news/ai/what-does-the-eus-ai-act-mean-for-human-rights

the European Union, China has established a series of intern measures to strengthen security requirements for AI products since 2017, when it established the "Next Generation Artificial Intelligence Development Plan". In October 2023, the Chinese government also released a list of security requirements for services using generative AI, such as OpenAI's ChatGPT. Brazil and Japan are other countries that, in the last months, have drawn attention to the development of AI, with Brazil working on its first law to regulate AI and Japan being one of the G7 that was leading in the creation of the Code of Conduct in 2023[103].

As can be seen, the "negotiations" between the different international and regional players are still taking place, even if they might not be as official as before. AI is not just a matter of European Union businesses but an international one that needs to be carefully addressed as it has diverse consequences in different fields where technology is used. The EU AI Act has been adopted by the European Member States without any way to return. However, the law is still in the ongoing process, leaving room for civil society to advocate and get involved in the topic, as the people's fundamental rights are a key element to be at risk in this game. The Convention on Human Rights by the Council of Europe is still an important document that is supposed to bring a balance between the use of future technologies while respecting and protecting fundamental rights without any risks to the people. However, a broader discussion might be an effective step, involving civil society in a debate regarding the importance of the protection of fundamental rights and how AI should be used for specific needs, not in every project that might happen in the future. There is still a lack of understanding from the international community about artificial intelligence, what it means, and how it can be used appropriately without generating harm. These specific gaps could be filled out by establishing discussions and debates on the ongoing topic, involving players using a 'down top' approach, where everybody could bring their contribution to the debate.

One last aspect that cannot be ignored when using AI is the potential risk that AI could highlight the inequalities between developed and developing countries. AI technologies often require significant resources, including infrastructure, talent, and funding, which can be more readily available in wealthier countries. Additionally, the data needed to train AI systems may be more abundant in specific regions or industries, leading to disparities in the development and deployment of AI solutions. Another criterion is that the availability of increasingly advanced computers conditions the use of AI and, thus, the availability of technology and raw materials that are increasingly centralised in the hands of a few countries, most of the time, the developed ones. However, an aspect that needs to be mentioned is that AI has the potential to address some challenges faced by poorer countries, such as improving healthcare, agriculture, and education. Efforts to democratise AI and make it more accessible to all countries, as well as initiatives to ensure that the benefits of AI are shared equitably, will be crucial in mitigating the potential gap between rich and poor nations. Understanding the potential impact of AI on global inequalities is an essential aspect of studying the technology.

The following months seem uncertain regarding this topic, as many documents still need to be released and even approved by the States, and there are still questions if this new regulation is going to have a "Brussels effect", speeding around the world, or due to its "rigid" construction, it will just have a European jurisdiction. At the end of the day, the crucial part of the debate is that the fundamental rights are not being traded off in exchange for profits and other interests involved in the game, as one of the primary responsibilities of the Member States should be the protection of the fundamental rights in any circumstances and that the future technologies will have a positive impact on the society, rather than a negative one.

[103] B. Kohn & F.U. Pieper, "AI Regulation around the World", *Taylar Wessing* [Online], May 2023. https://www.taylorwessing.com/en/interface/2023/ai---are-we-getting-the-balance-between-regulation-and-innovation-right/ai-regulation-around-the-world

Bibliography

A civil society statement on fundamental rights in the EU Artificial Intelligence Act, 'EU Trilogues: The AI Act must protect people's rights', 2023.

Agin, W. E., 'A Simple Guide to Machine Learning', *Business Law Today* , 2017, pp. 1-5.

Anon., 'AP eist duidelijkheid techbedrijf over AI gericht op kinderen', Autoriteit Persoongegevens, 2023, < https://www.autoriteitpersoonsgegevens.nl/actueel/ap-eist-duidelijkheid-techbedrijf-over-ai-gericht-op-kinderen>, accessed 12 October 2023.

Anon., 'General Purpose AI and the AI Act', *Future of Life Institute*, 2022.

Anon., 'Police AI needs strict limits and controls, warn civil society organisations', *Statewatch* [online], 2023, < https://www.statewatch.org/news/2023/september/police-ai-needs-strict-limits-and-controls-warn-civil-society-organisations/>, accessed 4 October 2023.

Anon,, 'What is machine learning?', *IBM* [online journal], 2023, < https://www.ibm.com/topics/machine-learning >, accessed 29 September 2023.

Antebi, L., 'What is Artificial Intelligence?', Institute for National Security Studies, 2021.

Bentley Sarah and Naughtin Claire, "The 'digital divide' is already hurting people's quality of life. Will AI make it better or worse?", *CSIRO*, March 2024. https://www.csiro.au/en/news/all/articles/2024/march/digital-divide-ai

Bertuzzi, L., 'US obtains exclusion of NGOs from drafting AI treaty', *EURACTIV* [online], 2023, < https://www.euractiv.com/section/digital/news/us-obtains-exclusion-of-ngos-from-drafting-ai-treaty/>, accessed 12 October 2023.

Bertuzzi, L., 'France, Germany, Italy push for 'mandatory self-regulation' for foundation models in EU's AI law', *EURACTIV* [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/news/france-germany-italy-push-for-mandatory-self-regulation-for-foundation-models-in-eus-ai-law/>, accessed 29 November, 2023.

Bertuzzi, L., 'AI Act: Spanish presidency makes last mediation attempt on foundation models', *EURACTIV* [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-spanish-presidency-makes-last-mediation-attempt-on-foundation-models/>, accessed 29 November 2023.

Browne, R., 'Italy became the first Western country to ban ChatGPT. Here is what other countries are doing', *CNBC* [online], 2023, < https://www.cnbc.com/2023/04/04/italy-has-banned-chatgpt-heres-what-other-countries-are-doing.html> accessed 3 October 2023.

CAI - Committee on Artificial Intelligence, *Council of Europe* [online], 2023, < https://www.coe.int/en/web/artificial-intelligence/cai>, accessed 13 October 2023.

Caselli Antonio A., "The More AI Progresses, the Greater the Need for Human Labor", *Extensive Interview in the German Journal Soziopolis*, 24 January 2024, https://www.casilli.fr/2024/01/24/the-more-ai-progresses-the-greater-the-need-for-human-labor-extensive-interview-in-the-german-journal-soziopolis-24-jan-2024/.

Center for AI Safety, "Statement on AI Risk: AI experts and public figures express their concern about AI risk", 2023. https://www.safe.ai/statement-on-ai-risk

Charter of Fundamental Rights of the European Union, Official Journal of the European Union, 2012.

Chee, F. Y., 'Exclusive: G7 to agree AI code of conduct for companies', *Reuters* [online], 2023, <https://www.reuters.com/technology/g7-agree-ai-code-conduct-companies-g7-document-2023-10-29/>, accessed 9 November 2023.

Chee, F. Y., and Mukherjee, S., 'EU lawmaker Benifei urges the bloc's countries to compromise on AI rules', *Reuters* [online], 2023, <https://www.reuters.com/technology/eu-lawmaker-benifei-urges-blocs-countries-compromise-ai-rules-2023-09-21/>, accessed 16 October 2023.

Civil Society Organisations, "Letter: Open letter to EU AI Act negotiators: Do not trade away our rights!", 8 December 2023; Civil Society Organisations, "Council risks failing human rights in the AI Act", 29 November 2023

Civil Liberties Union for Europe, European Civic Forum, European Center for Not-for-profit law, 'The AI Act Must Protect the Rule of Law', *Statewatch,* 2023.

Council of Europe, Commissioner of Human Rights of 2023 on thematic work regarding Artificial Intelligence and Human Rights.

Council of Europe of 2023 on Committee on Artificial Intelligence (CAI) regarding the Revised Zero Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law.

Council of the European Union of 2024 on Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement.

Csernatoni, R., 'Generative AI Poses Challenges for Europe', *Carnegie Europe* [online], 2023, < https://carnegieeurope.eu/strategiceurope/90803>, accessed 9 November 2023.

Custers, B., 'New digital rights: Imagining additional fundamental rights for the digital era', *ScienceDirect* [online], volume 44, 2022, <https://www.sciencedirect.com/science/article/pii/S0267364921001096?ref=pdf_download&fr=RR-2&rr=810cbcce9b4ea32d#section-cited-by>, accessed 4 October 2023.

Dulka A., "The Use of Artificial Intelligence in International Human Rights Law", *Standford Law School*, REV. 316, 2023, <https://law.stanford.edu/wp-content/uploads/2023/08/Publish_26-STLR-316-2023_The-Use-of-Artificial-Intelligence-in-International-Human-Rights-Law8655.pdf>, accessed 24 February 2024.

European Commission of 2021 on proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (artificial intelligence act) and amending certain union legislative acts.

European Commission of 2023 on Commission welcomes G7 leaders' agreement on Guiding Principles and a Code of Conduct on Artificial Intelligence, [press statement], < https://ec.europa.eu/commission/presscorner/detail/en/ip_23_5379> accessed 9 November 2023.

European Parliament of 2023 on Artificial Intelligence Act on amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).

Ernst, D., 'Competing in AI: Major Challenges', *Centre for International Governance Innovation,* 2020.

Fazlioglu, M., 'Contentious areas in the EU AI Act trialogues', *IAPP* [online], 2023, <https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues/> accessed 3 October 2023.

Fleming-Jones, J., 'EU AI Act nearing agreement despite three key roadblocks – co-rapporteur', *Euronews.next* [online], 2023,<https://www.euronews.com/next/2023/10/23/eu-ai-act-nearing-agreement-despite-three-key-roadblocks-co-rapporteur>, accessed 9 November 2023.

Future of Life Institute, "Policymaking in the Pause", 2023.

Gill J., "What does the EU's AI act mean for human rights?", *Thomas Reuters Foundation*, December 2023, https://www.context.news/ai/what-does-the-eus-ai-act-mean-for-human-rights

Hiroshima Artificial Intelligence Process of 2023 on the Hiroshima Process International Code of Conduct for Advanced AI Systems.

Human Rights Watch, "EU: Artificial Intelligence Regulation Should Protect People's Rights", 2023, <https://www.hrw.org/news/2023/07/12/eu-artificial-intelligence-regulation-should-protect-peoples-rights>, accessed 25 February 2024.

Kohn, B., 'AI Act at risk? – the regulation of foundation models and general-purpose AI', *Taylor Wessing* [online], 2023, https://www.taylorwessing.com/en/insights-and-events/insights/2023/11/ai-act-at-risk, accessed 29 November, 2023.

Kohn, B. & Pieper F.U., "AI Regulation around the World", *Taylar Wessing* [online], May 2023. https://www.taylorwessing.com/en/interface/2023/ai---are-we-getting-the-balance-between-regulation-and-innovation-right/ai-regulation-around-the-world

Kumar, A., 'Generative AI – Concepts, Use Cases, Examples', *Analytics Yogi* [online], 2023, https://vitalflux.com/generative-modeling-in-machine-learning-examples/, accessed 29 September 2023.

Lawton, G., 'What is generative AI? Everything you need to know', *TechTarget* [online], 2023, https://www.techtarget.com/searchenterpriseai/definition/generative-AI#:~:text=Generative%20AI%20focuses%20on%20creating,many%20types%20of%20new%20outputs., accessed 29 September 2023.

Leufer D., Rodelli C. and Fanny, "Human rights protections…with exceptions: what's (not) in the EU's AI Act deal", *Access Now*, December 2023, https://www.accessnow.org/whats-not-in-the-eu-ai-act-deal/

Long W et all, "EU Formally Adopts World's First AI Law", *Sidley*, March 2024. https://datamatters.sidley.com/2024/03/21/eu-formally-adopts-worlds-first-ai-law/#:~:text=On%20March%2013%2C%202024%2C%20the,of%20legislation%20for%20the%20EU.

Lynch, S., 'Analysing the European Union AI Act: What Works, What Needs Improvement', *Stanford Human Cantered Artificial Intelligence* [online], 2023, < https://hai.stanford.edu/news/analyzing-european-union-ai-act-what-works-what-needs-improvement>, accessed 3 October 2023.

Madiega, T., 'Artificial Intelligence Act overview', *European Commission*, 2023.

Madiega, T., 'General-purpose artificial intelligence', *European Commission*, 2023.

Muller Catelijne, "The Impact of Artificial Intelligence on Human Rights, Democracy and the Rule of Law", *ALLAI*, 20 March 2020.

Nietsche, C., & Ford, C., 'US-EU AI Code of Conduct: First Step Towards Transatlantic Pillar of Global AI Governance?', *EURACTIV* [online], 2023, <https://www.euractiv.com/section/artificial-intelligence/opinion/us-eu-ai-code-of-conduct-first-step-towards-transatlantic-pillar-of-global-ai-governance/> accessed 3 October 2023.

OECD, *Artificial Intelligence in Society*, 1st ed. Paris, 2019.

OECD, 'G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI', *OECDiLibrary*, 2023.

OECD, 'Principles of AI', *OECD.AI Policy Observatory* [online], 2023, <https://oecd.ai/en/ai-principles>, accessed 29 September 2023.

Pingen P., "AI Act: Parliament and Council Reach Provisional Agreement on World's First AI Rules", *eucrim*, 2024, <https://eucrim.eu/news/ai-act-parliament-and-council-reach-provisional-agreement-on-worlds-first-ai-rules/#:~:text=The%20definition%20of%20an%20AI,outputs%20such%20as%20predictions%2C%20content%2C>, accessed 24 February 2024.

Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

Raso, F.A., & Hilligoss, H., 'Artificial Intelligence & Human Rights: Opportunities & Risks', *Berkman Klein Center for Internet & Society Research Publication Series*, 2018.

Rothe R., "How AI will make us live longer and healthier lives", AI Reality Bites, 2024, <https://airealitybites.substack.com/p/how-ai-will-make-us-live-longer-and?r=1v54iy&utm_campaign=post&utm_medium=web&open=false>, accessed 24 February 2024.

Tanna, M., & Dunning, W., 'OECD Principles on Artificial Intelligence', *Simmons+Simmons* [online], 2019, < https://www.simmons-simmons.com/en/publications/ck0b9c83b7dpj0b9470kxyobl/290519-oecd-principles-on-artificial-intelligence>, accessed 29 September 2023.

Teachers Outreach, "Project: The global teachers initiative", 2023. https://teachersoutreach.org/

Tomašev N., Cornebise, J., Hutter, F. *et al.* "AI for social good: unlocking the opportunity for positive impact", *Nat Commun* 11, 2468, 2020. https://doi.org/10.1038/s41467-020-15871-z

Torras, J., 'Perspectives Symbolic AI vs. machine learning in natural language processing', *Multilingual* [online], 2023, <https://multilingual.com/issues/may-june-2020/symbolic-ai-vs-machine-learning-in-natural-language-processing/> , accessed 29 September 2023.

Turin Jerneja, "Artificial intelligence and its impact on the human rights of persons with disabilities", *ENNHRI*, 3 December 2023, https://ennhri.org/news-and-blog/artificial-intelligence-and-its-impact-on-the-human-rights-of-persons-with-disabilities/.

van Kolfschooten, H., 'The Council of Europe's Artificial Intelligence Convention: Implications for Health and Patients', *Bill of Health: Artificial Intelligence, Biotechnology, Human Rights, International* [online], 2023, <https://blog.petrieflom.law.harvard.edu/2023/04/18/council-of-europe-artificial-intelligence-convention/#:~:text=The%20AI%20Act%20focuses%20on,people%20affected%20by%20AI%20systems.>, accessed 13 October 2023.

von der Leyen, U., ‘A Union that strives for more: Political Guidelines for the next European Commission 2019-2024’, European Commission, 2019.

Wamba S. F. & et al., “Are we preparing for a good AI society? A bibliometric review and research agenda”, *Technological Forecasting & Social Change,* 164, 2021.

Wan, A., ‘Why Japan is lagging behind in generative A.I. — and how it can create its own large language models’, *CNBC* [online], 2023, < https://www.cnbc.com/2023/07/07/why-japan-is-lagging-behind-in-generative-ai-and-creation-of-llms.html>, accessed 12 October 2023.

Zhang K. & Aslan A. B., “AI technologies for education: Recent research and future directions”, *Computers and Education: Artificial Intelligence,* 2, 2021.