Forum: Youth Assembly - Action Paper 1

Issue: Promoting the ethical use of Artificial Intelligence

Student Officer: George Kokkinakis

Position: Co-Head

TOPIC INTRODUCTION

In this day and age, artificial intelligence is in everyone's lives, more prominently than ever. In 2025, according to a survey by the Statistical Office of the European Communities (eurostat), 32.7% of people aged 16-74 in the European Union (EU) used generative Artificial Intelligence (AI) tools.[1] This truly highlights how much we use AI in our day to day lives. Therefore the question is posed, why should people care about artificial intelligence ethics, what constitutes ethical use, and why should it be promoted?

Artificial intelligence is nothing new. Machine intelligence, which is the ability of a computer to think on its own, was first established as a concept in 1950 by Alan Turing, while the term "Artificial Intelligence" was coined at the Dartmouth Conference in 1956.[2] AI ethics concern humans in a more pronounced way than ever before in the last 3-5 years, most notably ever since the launch of the first large language model (LLM) to achieve widespread public adoption, ChatGPT. This was the first time where the general public actively utilised generative Artificial Intelligence, and not through a third party. That sudden exposure to AI made people curious, which led to the need for further understanding. Thus, everyone started researching and discovering that there are not only generative forms of AI, which are the ones everyone is familiar with, but others, which have been in widespread use for many years. A key issue with that is that Artificial Intelligence is largely developed without human oversight, as it is constantly evolving, making parts of it unexplainable.

A survey by the Council of European Professional Informatics Societies (CEPIS) from 2025, indicates that 84% of the EU population felt stressed by the rapid growth of AI, as well as the need for careful AI management to protect privacy and ensure transparency at work.[3] This shows that even though AI adoption rates in Europe are at approximately 33%, 84% highlight the need to uphold human dignity, showcasing the belief that AI should support, not undermine humans, something that can be safeguarded by the widely requested human agency and oversight. This is further highlighted by international initiatives, such as UNESCO's 4 core ethical principles. One of the key problems with AI is that ethical values, such as basic human dignity, are often undermined for the sake of evolution and corporate income.

---

[1] *32.7% of EU people used generative AI tools in 2025*, 16 December 2025, http://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251216-3 Accessed 14 June 2026.
[2] Scull, Sophia. "Artificial Intelligence (AI) Coined at Dartmouth." *Dartmouth*, August 1956, https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth Accessed 12 July 2026.
[3] Council of European professional informatics societies. "Europeans Support AI at Work but Call for Clear Regulations," CEPIS, 22 February 2025,

Moreover, AI can possibly be a driver for political instability. In this fragmented world, when the governance of algorithmic systems is left almost entirely to private enterprises, it accelerates global fragmentation, as set protocols are lacking, allowing tech conglomerates, which don't have the needed set of regulations, to operate with the sole purpose of advancing their technologies and increasing profitability. This is further highlighted by the fact that the terms of use in online applications, which were originally intended to ensure transparent data collection and utilisation, are constantly altered.

Diplomacy is a safeguard for human value and this year's conference theme, which aims to rediscover diplomacy, means that the international community must reclaim its responsibility from private tech enterprises. Only through coordinated, multilateral negotiations can states ensure human workers, artists and everyone's rights are protected from AI. As the Youth Assembly, the committee's focus is centered on how these frameworks will protect the upcoming generation of digital citizens, workforce entrants, and creators.

DEFINITION OF KEY-TERMS

Algorithmic Bias

"Algorithmic Bias is a feature of machine learning (ML) and artificial intelligence (AI) and it occurs when algorithms repeatedly produce systematic errors, inaccurate results, or imbalanced outcomes as a result of flawed, discriminatory, or incomplete training data."[4]

Artificial Intelligence (AI)

"Artificial intelligence (AI) is technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy."[5]

Human-in-the-Loop (HITL)

"Human-in-the-loop (HITL) refers to a system or process in which a human actively participates at some point in the AI workflow to ensure accuracy, safety, accountability or ethical decision-making."[6]

Large Language Models (LLMs)

"Large language models (LLMs) are models trained on immense amounts of data, making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks."[7]

---

[4] Greene, Jim. "Algorithmic bias | Computer Science | Research Starters." *EBSCO*, 2024, https://www.ebsco.com/research-starters/computer-science/algorithmic-bias Accessed 15 June 2026.
[5] Stryker, Cole, and Eda Kavlakoglu. "What Is Artificial Intelligence (AI)?" *IBM*, https://www.ibm.com/think/topics/artificial-intelligence Accessed 15 June 2026.
[6] Stryker, Cole. "What Is Human In The Loop (HITL)?" *IBM*, https://www.ibm.com/think/topics/human-in-the-loop Accessed 15 June 2026.
[7] Stryker, Cole. "What are Large Language Models?" *IBM*, https://www.ibm.com/think/topics/large-language-models Accessed 15 June 2026.

Lethal Autonomous Weapons Systems (LAWS)

"Lethal Autonomous Weapons Systems (LAWS) are a capability category, specifically. any weapon system incorporating autonomy in its critical functions, specifically in target selection and engagement."[8]

Machine Learning

"Machine learning is the subset of artificial intelligence (AI) focused on algorithms that can "learn" the patterns of training data and, subsequently, make accurate, logical conclusions about new data."[9]

BACKGROUND INFORMATION

Historical Background on the rise of AI

Artificial Intelligence is nothing new. It has been developing slowly ever since the 1960s. In the past few years though, artificial intelligence has taken a turn. After major breakthroughs, it is now developing more rapidly than ever before. There are a few key milestones where further attention is needed, in order to understand and encapsulate the rapid development of AI.

One of the most important innovations of the past few years were Generative Adversarial Networks (GANs). They were the key technological innovation which would eventually enable realistic-looking AI agents, deepfakes and many important advancements in an abundance of scientific fields. A concrete example is the use to generate synthetic medical images, like MRI scans for tumors, that help improve diagnostic models when real medical data are limited or difficult to obtain. While this innovation expanded human potential to an unprecedented degree, it simultaneously laid the groundwork for ethically concerning developments.

Another notable milestone in the world of AI is the creation of Sofia the robot. Developed by Hanson Robotics, she is designed to resemble a human with lifelike facial expressions and the ability to process visual data and speech, with her main purpose to interact with people in settings like healthcare, customer service, and education. She is also the first robot to be recognized with legal personhood, and she was activated in February 2016. Sofia is the first robot in the world to receive a symbolic citizenship, that of the state of Saudi Arabia. A detail that needs to be highlighted is that even though Sofia did receive a symbolic citizenship, this does not establish a universally recognized legal personhood, only to the certain states that acknowledge it. To that note, the Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies of the European

[8] Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System. *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. 10 March 2023, Geneva, Switzerland, https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_CRP.1_0.pdf. Accessed 15 June 2026.

[9] Bergmann, Dave. "What is machine learning?" *IBM*, https://www.ibm.com/think/topics/machine-learning Accessed 15 June 2026.

Parliament clarifies that “All AI-applications and machines are products. There are no technical, philosophical, or legal considerations that justify considering machines as moral or legal agents due to their intrinsic characteristics – autonomy, the ability to learn and modify itself over time, the ability to pursue a given end independent of human supervision or control”.[10]While certain states have attempted to heavily regulate humanoids, others allow little to no regulation. Sophia became a concrete milestone in AI ethics because giving a human-like robot official citizenship sparked huge debates over whether machines should ever have legal rights, how companies might trick us into thinking AI is actually conscious, and where we need to draw the line between humans and technology. Although, as observed by Sofia, no matter the regulatory frameworks that countries put into place, robots are all the more considered equal to humans.[11]

The innovation eventually led AI into what we know today, was the launch of one of the first large language models (LLMs). Open AI released GTP-1 in 2018. This kickstarted the “AI rush”, which led to the creation of all the publicly available LLMs. The milestone which really got AI in everyone’s day to day lives, and brought AI ethics to public concern was the launch of ChatGPT in 2022. According to OpenAI, the organisation that developed it, the dialogue format employed by ChatGPT allows it to answer follow-up questions, admit mistakes, challenge incorrect premises and reject inappropriate requests.[12] For the first time, a free, publicly available AI LLM could hold thoughtful conversations, solve problems, generate ideas, and even write essays, all while sounding like a real person. After ChatGPT, more tech companies followed, and launched other LLMs, dividing the market share.

Why has its rapid spread made ethical governance urgent?[13]

The ethical governance of AI is needed for the safeguard of our future, seeing as it is developing more rapidly than ever, oftentimes being utilised by enterprises, so as to cut costs. AI could and already is replacing a great amount of employees. As also seen below (Figure 1) enterprises are utilising AI to operate and evolve. It needs to be noted that in the EU, large-scale companies utilise AI more, in comparison to medium or small enterprises, consequently affecting a significant percentage of the population. Attention should be drawn to the fact that AI tools are as reliable as the data they are trained on. It is near impossible to limit the data that large scale commercially available LLMs are trained on, as they are scouting all throughout the internet for any and all information. Consequently, AI can be misleading, dangerous, it can compromise confidential information if and when leaked on the internet, as they are forever stored in a database and last but certainly not least, online privacy is jeopardized.

---

[10] Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, European Parliament. *Artificial Intelligence and Civil Liability Legal Affairs*. July 2020. Accessed 11 July 2026.
[11] Kouvaras, Nikolaos, and Andreas Pavlopoulos. *Social Robots: The case of Robot Sophia*. Department of Psychology, Panteion University, 17 May 2022, https://ejournals.epublishing.ekt.gr/index.php/homvir/article/view/30320 Accessed 11 July 2026.
[12] Open AI. “Introducing ChatGPT.” *OpenAI*, 30 November 2022, https://openai.com/index/chatgpt/ Accessed 12 July 2026.
[13] Short, Lizzie. “Ethics in AI: Why It Matters - Professional & Executive Development | Harvard DCE.” *Harvard Professional Development*, 6 March 2026, https://professional.dce.harvard.edu/blog/ethics-in-ai-why-it-matters/#Ethical-Challenges-in-AI Accessed 15 June 2026.

Furthermore, in order to ensure ethical governance, we need to understand the power of public awareness and media pressure. When the general public understands and realises the extent of the dangers of AI, they are in position to drive the attention further on these complex matters. Once leaders know where issues regarding AI ethics might exist, they can start working toward resolving them. Consequently, it is needed to urge the establishment of education on AI and more technological matters, so as to understand the true extent of the topic.

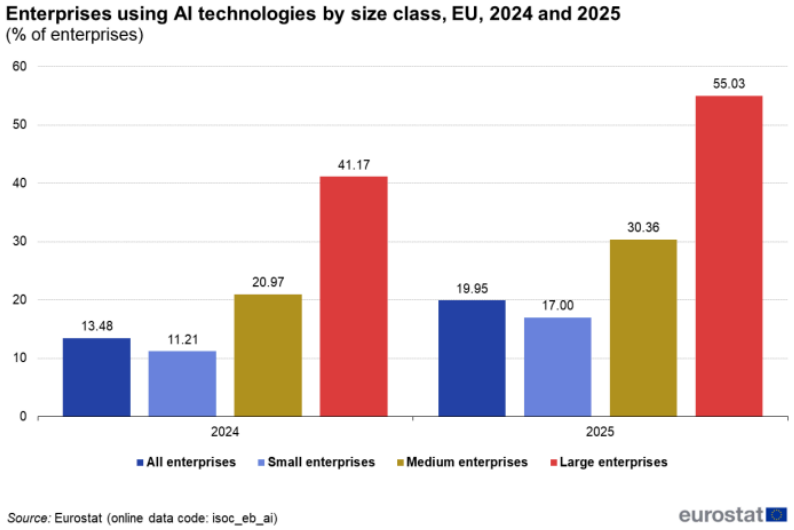

Figure 1: Enterprises using AI technologies by size class in the EU, 2024,2025.[14]

Core ethical concerns

It is crucial to understand which the core ethical concerns regarding AI are. First and foremost, the results Artificial Intelligence programmes produce reflect the data they have been trained on. This oftentimes causes great problems, especially regarding bias and discrimination. It's not rare for AI systems to be found discriminative. A case study worth investigating is Amazon's biased hiring algorithm. As stated in a Reuters article, “Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter.”[15] This was due to a long standing male domination of the tech industry. Since the AI was taught on past hiring data, and generally there were more men involved, it penalized women's applications. This case study highlights the real world impact that AI can have, which can go unnoticed, if not investigated, as it prolongs long standing discrimination and gender inequality.

Another key ethical concern regarding AI systems is the lack of online privacy, and mass surveillance. As AI systems' training rely on great amounts of data, they can often lead to possible privacy concerns, or can be utilised maliciously as mass surveillance tools. A

[14] EUROSTAT. “Enterprises Using AI Technologies by Size Class, EU, 2024, 2025.” *Eurostat*, Dec. 2025, ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises Accessed 15 June 2026.
[15] Dastin, Jeffrey. “Insight - Amazon scraps secret AI recruiting tool that showed bias against women.” *Reuters*, 11 October 2018. Accessed 12 July 2026.

case study regarding the lack of online privacy, is the one of Clearview AI. Clearview AI is a very large database, which uses AI systems to identify possible suspects, witnesses, and victims. It has over 70 billion images in their database, and is used by law enforcement worldwide.[16] The issue with Clearview is that this database was collected, without any persons legal consent or the state which they rise in agreeing upon this. This caused great security concerns, and the company was fined by multiple countries over this practice. These concerns were proven reasonable, as in February 2020, the company suffered a destructive data breach. Forbes states that, *"The data stolen in the hack included the firm's entire customer list–which will include multiple law enforcement agencies–along with information such as the number of searches they had made and how many accounts they'd set up".*[17] This case study highlights the fact that any information online can be stored and utilised by AI systems, and also can be leaked or used with bad intent.

Lastly, a significant issue is that AI systems can be used to void accountability. This is applied in many different ways, but most concerning of all is the application of Lethal Autonomous Weapon systems (LAWs), which work on their own, utilising artificial intelligence. The key problem with LAWs is that if a system commits a grave crime, states can use the fact that it was autonomous, and not human controlled, to avoid responsibility. A key example is when in March 2021, a UN Panel of Experts on Libya reported a possible use of lethal autonomous weapons systems, specifically the STM Kargu-2, which are Turkish-built autonomous Unmanned Aerial Vehicles (UAVs). They "were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true 'fire, forget and find' capability"[18]. The UN report refers to the deployment of this system in the context of the Government of National Accord Affiliated Forces (GNA-AF) with Turkish military support launching offensive campaigns against the Hafter Affiliated Forces (HAF) in what appears to be a non-international armed conflict.[19] It is oftentimes regarded as the first time when an autonomous system took human lives, though it is hard to tell with certainty, as the case is in a "gray area", highlighting the void of responsibility.

The tension between innovation and regulation

As previously established, Artificial Intelligence can be both beneficial, and detrimental, depending on the use and the development of AI systems. The key dilemma that companies, and regulatory bodies alike are facing is whether there is too much, or too little regulation regarding AI. Companies and certain states which are leading AI development, do not want to be regulated, as certain measures may hold innovation back. On the other hand, certain states and regulatory bodies push for more regulation and

---

[16] "Clearview AI - Website main page." *Clearview AI | Facial Recognition*, https://www.clearview.ai/ Accessed 12 July 2026.
[17] O'Flaherty, Kate. "Clearview AI, The Company Whose Database Has Amassed 3 Billion Photos, Hacked." *Forbes*, 26 February 2020, https://www.forbes.com/sites/kateoflahertyuk/2020/02/26/clearview-ai-the-company-whose-database-has-amassed-3-billion-photos-hacked/ Accessed 12 July 2026.
[18] United Nation Security Council. *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council*. 8 March 2021. Accessed 12 July 2026.
[19] Nasu, Hitoshi. "The Kargu-2 Autonomous Attack Drone: Legal & Ethical Dimensions - Lieber Institute West Point." *Lieber Institute - West Point*, 10 June 2021, https://lieber.westpoint.edu/kargu-2-autonomous-attack-drone-legal-ethical/ Accessed 12 July 2026.

limitations, so as to secure the public interest in data security and privacy, as well as conserve resources.

MAJOR COUNTRIES AND ORGANISATIONS INVOLVED

China

China stands as a major global developer of artificial intelligence, utilizing a heavily state-led approach to drive technological advancement. However, its implementation has raised significant international surveillance concerns, and the nation frequently resists external oversight regarding its domestic AI practices. China has their own state backed LLM, Deepseek, which has 96,88 million monthly active users, ranking it 4th on the market.[20]

The United States of America

The United States leads global private sector AI innovation, driven by tech giants and massive venture capital. Historically, the U.S. favored little to no regulation to avoid stifling innovation, though it has recently shifted toward oversight through new executive orders aimed at managing AI risks. The US is the leader in AI development, as a great share of the AI market is occupied by American companies, such as OpenAI or Google.

The United Arab Emirates

The United Arab Emirates (UAE) represents an ambitious and unique middle-ground case in the global landscape. Demonstrating its commitment to a dedicated AI strategy, the UAE became the first country in the world to appoint an official Minister of AI, a position that later expanded to cover the Digital Economy and Remote Work Applications.To achieve this, the minister's initiatives center on building a supportive AI ecosystem, providing essential data and regulatory infrastructure to make the nation a premier testing ground for innovations, and strengthening priority sectors.[21]

Rwanda

Rwanda serves as a prominent example of how African Less Economically Developed Countries (LEDCs) interact with the technology. In this context, AI is primarily deployed to solve critical development challenges. Rwanda is one of the pioneering states in AI governance, since as of June 11th 2026, they established a National Artificial Intelligence Agency, to oversee tech adoption, drive digital transformation, and promote ethical standards across essential sectors like healthcare and agriculture.[22]

---

[20] Backlinko Team. "DeepSeek AI Usage Stats." *Backlinko*, 16 January 2026, https://backlinko.com/deepseek-stats Accessed 12 July 2026.
[21] "Omar Sultan Al Olama | World Economic Forum." *The World Economic Forum*, https://www.weforum.org/people/omar-sultan-al-olama/ Accessed 12 July 2026.
[22] Mbego, Steve. "Rwanda Establishes National Artificial Intelligence Agency." *CIO Africa*, 11 June 2026, https://cioafrica.co/rwanda-establishes-national-artificial-intelligence-agency/ Accessed 12 July 2026.

UNESCO

UNESCO produced the first truly global AI ethics recommendation in 2021. This influential framework is anchored by four core ethical principles designed to protect human rights, dignity, and environmental sustainability in the age of AI. Applicable across member states, this landmark standard emphasizes transparency, fairness, and human oversight, establishing vital policy action areas for governance, gender equality, and data protection to translate moral values into concrete global practices.[23]

The Organisation for Economic Co-operation and Development (OECD)

The Organisation for Economic Co-operation and Development (OECD) developed early, foundational AI principles. These principles have since been widely adopted by member governments to shape their national regulatory frameworks.Serving as the first intergovernmental standard on artificial intelligence, these guidelines prioritize trustworthy, human-centric AI by establishing key standards for transparency, safety, fairness, and accountability, providing a flexible blueprint that helps policymakers globally manage AI risks while promoting responsible innovation.[24]

The Partnership on AI (PAI)

The Partnership on AI (PAI), founded in 2016, represents a major attempt at private sector self-governance. It is a multi-stakeholder non-profit established by tech giants, including Google, Amazon, Meta, Microsoft, and Apple, alongside civil society organizations and academic institutions. PAI's primary goals are to study and formulate best practices on AI ethics, while actively bringing together industry leaders, researchers, and public advocacy groups to ensure AI is developed responsibly.[25]

TIMELINE OF EVENTS

DATE	DESCRIPTION OF EVENT
October 1950	The concept of Machine Thinking is introduced by Alan Turing's seminal paper Computing Machinery and Intelligence.[26]

[23] UNESCO. "Recommendation on the Ethics of Artificial Intelligence." *Unesco*, 26 September 2024, https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence Accessed 15 June 2026.
[24] OECD. "AI Principles." OECD, Organisation for Economic Co-operation and Development, 2019, revised 2024, https://www.oecd.org/en/topics/ai-principles.html Accessed 14 June 2026.
[25] PAI. *Partnership on AI - Home - Partnership on AI*, https://partnershiponai.org/ Accessed 12 July 2026.
[26] Turing, Alan. "COMPUTING MACHINERY AND INTELLIGENCE." October 1950, https://courses.cs.umbc.edu/471/papers/turing.pdf Accessed 12 July 2026.

August 1956	Dartmouth Conference coins the term Artificial Intelligence.[27]
January 1966	The first chatbot was created at MIT, named ELIZA.
May 11, 1997	IBM's Deep Blue defeats world chess champion Garry Kasparov.[28]
2014	Amazon begins developing an AI hiring tool, later found biased.[29]
September 2016	Partnership on AI (PAI) is founded by major tech companies.
May 25, 2018	The EU introduced GDPR, the first major data protection law.
April 8, 2019	The EU released its Ethics Guidelines for Trustworthy AI.
May 22, 2019	The Organisation for Economic Co-operation and Development (OECD) publishes its AI principles, adopted by 47 countries.
November 23, 2021	UNESCO adopts the first global recommendation on the ethics of AI.
November 30, 2022	ChatGPT launches, bringing AI ethics into mainstream public debate overnight.
October 30, 2023	The US Executive Order on AI safety was signed by President Joe Biden.
March 13, 2024	The EU AI Act was formally passed by the European Parliament, becoming the world's first comprehensive AI law.
March 21, 2024	The UN adopts the first resolution on safe and trustworthy AI.
May 21, 2024	The EU AI act is formally adopted by the European Council.
August 1st, 2024	The EU AI Act officially enters into force.
February 2026	Studies expose deep-seated representational bias in Gemini Flash 2.5 Image (NanoBanana).[30]

[27] Scull, Sophia. "Artificial Intelligence (AI) Coined at Dartmouth." *Dartmouth*, August 1956, https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth. Accessed 12 July 2026.
[28] IBM. "Deep Blue." *IBM*, https://www.ibm.com/history/deep-blue. Accessed 12 July 2026.
[29] Dastin, Jeffrey. "Insight - Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters*, 11 October 2018 Accessed 12 July 2026.
[30] Balestri, Roberto. "Neutral Prompts, Non-Neutral People: Quantifying Gender and Skin-Tone Bias in Gemini Flash 2.5 Image and GPT Image 1.5." Department of the Arts, Universita di Bologna, Italy, 12 February 2026, https://arxiv.org/pdf/2602.12133 Accessed 14 June 2026.

RELEVANT UN RESOLUTIONS, TREATIES AND EVENTS

UNGA Resolution A/RES/78/311:

The United Nations General Assembly resolution on enhancing international cooperation on capacity-building of artificial intelligence, is a monumental milestone on the international regulation of AI systems. It is one of the first UNGA resolutions which is dedicated specifically on artificial intelligence. Overall, it calls on member states to develop safe, secure and trustworthy AI systems. It also emphasizes that AI governance must respect all human rights.

HRC Resolution 47/23 - New and emerging digital technologies and human rights (2021)

The Human Rights Council's resolution on new and emerging digital technologies and human rights flagged AI as a potential threat to human rights, privacy and non discrimination, making it a core legal ground on AI development.By establishing that international human rights law applies fully to digital technologies, this landmark resolution urges states and tech companies to conduct strict human rights due diligence to prevent algorithmic bias, invasive surveillance, and discrimination.

UNESCO Recommendation on the Ethics of AI (2021)[31]

The recommendation on the Ethics of AI is one of the most important regulatory frameworks, currently in place today. It is not a binding treaty which is adopted by all 193 UNESCO member states. This marks it as the broadest international consensus document on AI ethics in existence. It covers Human Rights, transparency, accountability, privacy, and the environment.

OECD AI Principles (2019) [32]

The AI principles by the Organisation for Economic Co-operation and Development, adopted by 47 countries, set the first intergovernmental standards on trustworthy AI. Many national AI policies such as the ones implemented within the European Union, are built on these principles.Focusing on human-centric innovation, these guidelines establish key criteria for safety, transparency, fairness, and accountability, providing a flexible blueprint for governments to manage technological risks while encouraging responsible growth.

Council of Europe Framework on Convention on AI (2024)[33]

---

[31]UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization, 21 November 2021. Accessed 14 June 2026.
[32] OECD. "AI principles." *OECD*, Organisation for Economic Co-operation and Development, 2019, revised 2024, https://www.oecd.org/en/topics/ai-principles.html Accessed 14 June 2026.
[33] Council of Europe. "The Framework Convention on Artificial Intelligence - Artificial Intelligence." *The Council of Europe*, 5 September 2024,

The first legally binding international treaty on AI, opened for signature in 2024. Required signatories to ensure AI systems respect Human rights, democracy, and the rule of law. Applicable to both public and private sectors, this treaty establishes enforceable legal standards for risk management, transparency, and accountability, creating a unified framework to protect democratic processes and individual freedoms from AI-driven harms.

World Summit on the Information Society (WSIS) (2003 - today)[34]

The WSIS is a UN- led forum that has increasingly addressed AI governance, digital inclusion and ethical technology use. It is an ongoing, multilateral space where AI ethics are constantly debated. By connecting governments, private tech leaders, and civil society, the summit works to bridge the global digital divide while establishing global consensus on how emerging technologies can be deployed responsibly to support sustainable development.

PREVIOUS ATTEMPTS TO SOLVE THE ISSUE

Global digital compact (2024)[35]

The global digital compact was agreed at the UN summit of the future, in September 2024. It is a landmark commitment by member states to govern digital technologies, such as AI, in an inclusive and rights-respecting way. The GDC successfully established five core objectives that 193 member states officially agreed to pursue. It successfully forced topics like algorithmic accountability, bridging the digital divide in developing countries, and data protection into a singular, high-level international framework. It initiated concrete structures within the UN system. It established the Independent International Scientific Panel on AI so as to achieve a shared understanding of AI capabilities and risk, the Global Dialogue on AI Governance, facilitating ongoing intergovernmental debates and the creation of the UN Office for Emerging and Digital Technologies (ODET) to oversee its long-term goals. It is worth mentioning that it is widely criticised, as it is a non binding treaty, meaning that it cannot force any of the proposed solutions, as a great deal of the treaty was compromised, from pressure from MEDCs.

Partnership on AI (2016)

The partnership on AI was the tech industry's attempt to self govern voluntarily. It was founded by large tech companies, including Google, Microsoft and Apple, and it produces best practises guidelines on fairness, transparency and accountability. Before PAI, tech companies were developing AI in complete silos with no industry-wide ethical standards. PAI successfully established a set of tenets that guided the industry on safety, transparency, and fairness. It created highly regarded, practical frameworks, such as the Synthetic Media Framework which is one of the earliest global guidelines on how to responsibly create and

---

https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence Accessed 14 June 2026.

[34]Miao, Fengchun, and Wayne Holmes. “World Summit on the Information Society (WSIS).” *UNESCO*, https://www.unesco.org/en/wsis Accessed 14 June 2026.

[35] United Nations. *A/79/L.2 - Global Digital Compact*. 22 September 2024, Summit of the Future, New York, https://www.un.org/digital-emerging-technologies/global-digital-compact/intergovernmental-process

disclose AI-generated content (like deepfakes) to prevent misinformation, and the Responsible Deployment Guidelines which standardizes how companies should test AI models for bias and security vulnerabilities before releasing them to the public. Furthermore, it forced multi-billion-dollar tech giants to sit at the same table with civil rights groups (like the ACLU) and academic researchers. This ensured that corporate AI development was constantly being challenged by independent watchdogs who prioritized human rights over corporate profit. It is worth mentioning that critics have frequently accused PAI of being a tool for "ethics washing", by allowing big tech companies to point to their membership as proof they are being responsible, while continuing to prioritize profit and market dominance. Because the tech giants provide the vast majority of funding and hold immense sway over the industry, smaller civil society members are often sidelined or outmaneuvered. As Michael Sandel, a professor of Government at Harvard University stated at the Harvard Gazette, ”The problem is these big tech companies are neither self-regulating, nor subject to adequate government regulation. I think there needs to be more of both”, highlighting that PAI is inadequate, and that government regulation lacks.

The United States' Sectoral and Executive Approach

Historically, the United States has favored private-sector AI innovation with little to no overarching federal regulation. By avoiding heavy, restrictive federal laws, the U.S. approach allowed its private sector (companies like OpenAI, Google, Meta, and Anthropic) to move incredibly fast. Tech companies didn't have to wait for government pre-clearance to launch new technologies, keeping the U.S. at the absolute forefront of global AI innovation. It has relied on soft-law measures and executive actions, culminating in the US Executive Order on AI Safety signed by President Joe Biden on October 30, 2023.[36] On the downside, this approach relies heavily on voluntary compliance, creating a fragmented regulatory environment across different states while lacking strict, legally binding federal enforcement.

China's State-Led and Target-Driven Approach

China has implemented a highly centralized, state-led approach to AI development. Its regulatory strategy focuses on specific, targeted laws, including regulations targeting recommendation algorithms and generative AI tools, which are heavily tied to national security, social stability, and state surveillance systems. China's domestic policies strongly resist external oversight or binding international frameworks, prioritizing state control over universal human rights principles. A concrete example is the Algorithmic Recommendation Provisions (2022),[37] which targets the code behind content delivery feeds and forces platforms to give users the right to opt-out of algorithmic profiling. It further mandates that algorithms promote "positive energy," and establishes a formal Algorithm Registration/Filing System where tech companies must pull back the curtain on their code for state auditors.

---

[36] Harris, Laurie, and Chris Jaikaran. *Highlights of the 2023 Executive Order on Artificial Intelligence for Congress*. 3 April 2024, https://www.congress.gov/crs-product/R47843 . Accessed 15 June 2026.
[37] Costigan, Johanna. “Translation: Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022.” *DigiChina - Stanford University*, 1 March 2022, https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/ Accessed 15 June 2026.

Pre-AI Act Framework

Before passing comprehensive legislation, the European Union relied on existing data protection laws, introducing the General Data Protection Regulation (GDPR) on May 25, 2018, to handle initial data privacy anxieties. This was supplemented by the release of the "Ethics Guidelines for Trustworthy AI" on April 8, 2019. The Guidelines put forward a set of 7 key requirements that AI systems should meet in order to be deemed trustworthy. A specific assessment list aims to help verify the application of each of the key requirements.[38] A key limitation is that the GDPR was not originally designed to handle complex deep learning or generative AI models, and the 2019 guidelines were completely voluntary, lacking concrete penalties or enforcement mechanisms until the formal passage of the EU AI Act.

EU AI Act

The European Union's AI Act is the world's first comprehensive, legal framework designed to regulate artificial intelligence based on risk. It abandons purely voluntary guidelines in favor of a strict tier-based system, and it was formally adopted in 2024. Banning AI applications deemed an unacceptable threat to fundamental rights, such as real-time biometric surveillance in public spaces and social scoring, while enforcing rigorous data governance, transparency, and human oversight requirements on high-risk AI systems such as those used in critical infrastructure, hiring, and law enforcement. A standout strength of the Act is its extraterritorial reach, often termed the "Brussels Effect", which forces global tech companies to comply with European standards if their AI models process data or operate within the EU market. However, the legislation has faced significant criticism from European tech advocates who argue that its heavy compliance burdens and steep penalties, ranging up to millions of euros or a percentage of global turnover, will negatively affect domestic startup innovation and put Europe at a competitive disadvantage against other countries, such as the U.S. and China.

POSSIBLE SOLUTIONS

Pushing for Global Oversight and Institutional Architecture

A possible way to resolve this issue could be establishing an International Artificial Intelligence Governance Agency, closely modeled after the International Atomic Energy Agency (IAEA).[39] It would be tasked with tracking frontier machine learning models and LLMs, conducting security checks, such as pre-deployment safety evaluations. It will also ensure that private-sector tech enterprises conform strictly to universal human rights standards. Mirroring China's domestic algorithmic filing system on an international level, the agency will maintain a unified registry where sovereign developers operating under a member state will submit general-purpose LLMs and high-risk algorithmic models for structural auditing.

---

[38] European Commision. "Ethics guidelines for trustworthy AI." *Shaping Europe's digital future*, 8 April 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai Accessed 15 June 2026.
[39] International Atomic Energy Agency. *IAEA - Governance*. How the IAEA operates from within. IAEA, https://www.iaea.org/about/governance Accessed 15 June 2026.

Establishing "Human-in-the-loop" workplace legislation

Another important need that needs to be preserved is ensuring that humans are threatened by AI as little as possible. This could be done by urging member states to pass legislation ensuring that high-stakes choices cannot be made solely on AI algorithms. A qualified human must always review and make the final decision to prevent cold, algorithmic errors from disrupting people's livelihoods. This applies to cases such as military usages utilising Lethal Autonomous Weapons (LAWS), where humans should always have the final say on each target, to ensure full accountability, medical usages, where human intuition and emotion is required, as the algorithm which receives pure data does not account for the emotional capacity, mental health or overall impact on the patients on the medical personnel, and last but not least, handling human resources in businesses, where an AI may hire or fire a someone, purely based on their resume, disregarding their overall capacity,

Implementing Accountability and a Human-Centric Approach

Accountability is of high importance in sectors where AI is used as a free-for-all pass. It can be ensured up to a degree by calling upon member states to utilise legislation to protect Human rights and values threatened by AI, such as people's right to privacy and autonomy. Approach the matter legally, so as to protect the people from losing their workplace to AI. Discourage AI usage in artistic environments, and enforce strict, transparent, intellectual property safeguards, to ensure that the humanity that grants art its value is secured. Last but not least, urge for the limitation of AI in the military. Artificial intelligence cannot be held accountable, meaning that any military utilising AI in uses such as LAWs oftentimes voids accountability.

Implementing AI literacy frameworks in education

The negative impact that AI is having in literacy can be monitored by urging member states to implement AI literacy, data privacy and algorithmic awareness into primary and secondary education. These frameworks could focus on critical digital education. It will train the upcoming generation to identify algorithmic bias, detect AI-generated synthetic media, including deepfakes , and understand how their personal data is harvested for model training. The main goal is to empower the youth to actively question automated systems rather than blindly trusting them. By funding public education on AI limitations, states can build societal resilience against AI-driven misinformation and prepare future workforce entrants to navigate AI-integrated workplace environments safely.

Regulating the use of AI in military operations

Create a unified, international body, which will oversee the production and use of AI enhanced military equipment; Establish clear legal definitions on LAWS and other autonomous military systems, to safeguard full legal accountability; Urge for the limitation of AI in the military. Artificial intelligence cannot be held accountable, meaning that any military utilising AI in uses such as LAWs is frequently not held accountable upon their actions.

BIBLIOGRAPHY

General Bibliography

32.7% of EU people used generative AI tools in 2025, 16 December 2025, http://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251216-3 . Accessed 15 June 2026.

Backlinko Team. "DeepSeek AI Usage Stats." *Backlinko*, 16 January 2026, https://backlinko.com/deepseek-stats. Accessed 12 July 2026.

Balestri, Roberto. *Neutral Prompts, Non-Neutral People: Quantifying Gender and Skin-Tone Bias in Gemini Flash 2.5 Image and GPT Image 1.5.* Department of the Arts, Universita di Bologna, Italy, 12 February 2026, https://arxiv.org/pdf/2602 .12133. Accessed 14 June 2026.

Bergmann, Dave. "What is machine learning." *IBM*, https://www.ibm.com/think/topics/machine-learning . Accessed 15 June 2026.

"Clearview AI - Website main page." *Clearview AI | Facial Recognition*, https://www.clearview.ai/ . Accessed 12 July 2026.

Costigan, Johanna. "Translation: Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022." *DigiChina - Stanford University*, 1 March 2022, https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/ . Accessed 15 June 2026.

Council of Europe. ""The Framework Convention on Artificial Intelligence - Artificial Intelligence." The Council of Europe, 5 September 2024, https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence . Accessed June 14 2026.

Council of European professional informatics societies. *Europeans Support AI at Work but Call for Clear Regulations*. CEPIS, 22 February 2025, https://cepis.org/new-eurobarometer-report-europeans-support-ai-at-work-but-call-for-clear-regulations/ .

Dastin, Jeffrey. “Insight - Amazon scraps secret AI recruiting tool that showed bias against women.” *Reuters*, 11 October 2018. Accessed 12 July 2026.

European Commision. “Ethics guidelines for trustworthy AI.” *Shaping Europe's digital future*, 8 April 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai . Accessed 15 June 2026.

Greene, Jim. “Algorithmic bias | Computer Science | Research Starters.” *EBSCO*, 2024, https://www.ebsco.com/research-starters/computer-science/algorithmic-bias . Accessed 15 June 2026.

Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System. *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. 10 March 2023, Geneva, Switzerland, https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_CRP.1_0.pdf . Accessed 15 June 2026.

Harris, Laurie, and Chris Jaikaran. *Highlights of the 2023 Executive Order on Artificial Intelligence for Congress*. 3 April 2024, https://www.congress.gov/crs-product/R47843 . Accessed 15 June 2026.

IBM. “Deep Blue.” *IBM*, https://www.ibm.com/history/deep-blue. Accessed 12 July 2026.

International Atomic Energy Agency. *IAEA - Governance*. How the IAEA operates from within. IAEA, https://www.iaea.org/about/governance . Accessed 15 June 2026.

Kouvaras, Nikolaos, and Andreas Pavlopoulos. *Social Robots: The case of Robot Sophia*. Department of Psychology, Panteion University, 17 May 2022, https://ejournals.epublishing.ekt.gr/index.php/homvir/article/view/30320 . Accessed 11 July 2026.

Mbego, Steve. "Rwanda Establishes National Artificial Intelligence Agency." *CIO Africa*, 11 June 2026, https://cioafrica.co/rwanda-establishes-national-artificial-intelligence-agency/ . Accessed 12 July 2026.

Miao, Fengchun, and Wayne Holmes. "World Summit on the Information Society (WSIS)." *UNESCO*, https://www.unesco.org/en/wsis . Accessed 14 June 2026.

Nasu, Hitoshi. "The Kargu-2 Autonomous Attack Drone: Legal & Ethical Dimensions - Lieber Institute West Point." *Lieber Institute - West Point*, 10 June 2021, https://lieber.westpoint.edu/kargu-2-autonomous-attack-drone-legal-ethical/ . Accessed 12 July 2026.

OECD. "AI Principles." OECD, Organisation for Economic Co-operation and Development, 2019, revised 2024, https://www.oecd.org/en/topics/ai-principles.html . Accessed 14 June 2026.

O'Flaherty, Kate. "Clearview AI, The Company Whose Database Has Amassed 3 Billion Photos, Hacked." *Forbes*, 26 February 2020, https://www.forbes.com/sites/kateoflahertyuk/2020/02/26/clearview-ai-the-company-whose-database-has-amassed-3-billion-photos-hacked/ . Accessed 12 July 2026.

"Omar Sultan Al Olama | World Economic Forum." *The World Economic Forum*, https://www.weforum.org/people/omar-sultan-al-olama/ . Accessed 12 July 2026.

Open AI. “Introducing ChatGPT.” *OpenAI*, 30 November 2022, https://openai.com/index/chatgpt/ . Accessed 12 July 2026.

PAI. *Partnership on AI - Home - Partnership on AI*, https://partnershiponai.org/ . Accessed 12 July 2026.

Pazzanese, Christina. “Ethical concerns mount as AI takes bigger decision-making role.” *Harvard Gazette*, 26 October 2020, https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/ . Accessed 15 June 2026.

Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, European Parliament. *Artificial Intelligence and Civil Liability Legal Affairs*. July 2020. Accessed 11 July 2026.

Scull, Sophia. “Artificial Intelligence (AI) Coined at Dartmouth.” *Dartmouth*, August 1956, https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth . Accessed 12 July 2026.

Short, Lizzie. “Ethics in AI: Why It Matters - Professional & Executive Development | Harvard DCE.” *Harvard Professional Development*, 6 March 2026, https://professional.dce.harvard.edu/blog/ethics-in-ai-why-it-matters/#Ethical-Challenges-in-AI . Accessed 15 June 2026.

Stryker, Cole. “What are Large Language Models.” *IBM*, https://www.ibm.com/think/topics/large-language-models . Accessed 15 June 2026.

Stryker, Cole. “What Is Human In The Loop (HITL)?” *IBM*, https://www.ibm.com/think/topics/human-in-the-loop . Accessed 15 June 2026.

Stryker, Cole, and Eda Kavlakoglu. “What Is Artificial Intelligence (AI)?” *IBM*, https://www.ibm.com/think/topics/artificial-intelligence . Accessed 15 June 2026.

Turing, Alan. “COMPUTING MACHINERY AND INTELLIGENCE.” October 1950, https://courses.cs.umbc.edu/471/papers/turing.pdf . Accessed 12 July 2026.

Turner, Georgia. “10 AI milestones of the last 10 years.” *The Royal Institution*, R.I., 18 December 2023, https://www.rigb.org/explore-science/explore/blog/10-ai-milestones-last-10-years . Accessed 15 June 2026.

UNESCO. “Recommendation on the Ethics of Artificial Intelligence.” *Unesco*, 26 September 2024, https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence . Accessed 15 June 2026.

United Nations. *A/79/L.2 - Global Digital Compact*. 22 September 2024, Summit of the Future, New York, https://www.un.org/digital-emerging-technologies/global-digital-compact/intergovernmental-process .

United Nation Security Council. *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council*. 8 March 2021. Accessed 12 July 2026.

Media Bibliography

Figure 1 - EUROSTAT. “Enterprises Using AI Technologies by Size Class, EU, 2024, 2025.” *Eurostat*, Dec. 2025, ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises . Accessed 15 June 2026.