Hosted by The Mayflower School

SharkMUN
Model United Nations Conference

TOPIC B

Ethical AI Development

ECOSOC
Economic and Social Council

Head Chair: Belén Jacir
Deputy Chair: Emma Luksic

Economic and Social Council

Topic B: **Ethical AI Development**

Introduction

The rapid development of artificial intelligence (AI) has created multiple opportunities for the development of our world. However, these new adaptations have been raising multiple concerns about the ethics that the AI has. That's why, in these years, AI has been put to certain ethics that it has to have, so it does not generate any problems with society (UNESCO, s.f.).

Ethical AI development refers to designing and building artificial intelligence systems guided by moral principles. Prioritizing fairness, transparency, accountability, and the respect for human rights and privacy, so it ensures that AI technologies are aligned with human values (Transcend, s.f.).

With the passage of time, AI has become critical in multiple points like healthcare, public services and finances mainly, ensuring that these systems operate without bias, to not promote discrimination and maintain the user's trust.

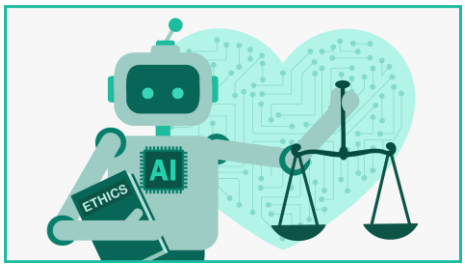
Ethical AI development aims and seeks to eliminate risks in relation to unfair treatment, privacy violations, and unfair decision making. This approach promotes responsible creation and innovation that benefits society all the way along, but at the same time safeguarding the individual rights and encouraging inclusivity among everyone.

Definition of Key Terms

- **Artificial intelligence (AI)**
  Human knowledge and intelligence processed by a machine, usually in computer systems, gives them the power to perform tasks, such as problem solving, reasoning, decision making, learning, among others.

- **Ethical AI**
  Refers to artificial intelligence that sticks to well-defined moral and ethical principles, such as individual rights, privacy, fairness, and non-discrimination. This seeks that AI principales are fair and promote positive social outcomes.

- **AI Ethics**
  It's the principales or framework that guide researchers and organizations to have a correct use of AI for the society's benefit. So it aligns with human values and it can address problems correctly, without a bias.



- **AI Bias**
  AI bias refers to AI favoring certain specific groups of people over others, leading to unfair and discriminatory outcomes. And it amplifies discrimination, stereotypes and prejudices.

- **AI Accountability**
  The state of being responsible or answerable for a system, its behavior, and its potential impacts.

- **AI Transparency**
  The openness and clarity about how AI systems work, make decisions, and learn over time. So that the AI's decision align with human values.

Background information

- **Origins of the Ethical AI Problem**

  Going back in time, the origins of artificial intelligence (AI) go back approximately to the mid 20th century, in which people like Alan Turing and John McCarthy started contemplating the machine's intelligence and morality. In 1950, Turing put to the test the famous "Turing Test," that tested the machine's intelligence on whether it could act similarly to a human or not. Then, in 1956, it was official, the term "artificial intelligence" was established, being the start of AI as something formal. The early rise of AI was mainly focused on the use of logic, reasoning, and rule-based systems, in which multiple milestones were achieved. However, as AI started being integrated into multiple systems and real-world situations, concerns were rapidly raised, and questions about their ethical implications started to emerge, since it was clear that the use of AI could have a big impact on humanity, whether positive or negative.

  The first ethical rules applied, long before modern AI existed, were from Isaac Asimov, who was a huge influence on how machine ethics were conceptualized. He established the "Three Laws of Robotics," in which they were:

  1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
  2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
  3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

  Demonstrating how important the ethic and moral principles were when talking about intelligent systems (The Origins Of Ethics In AI, s.f.).

- **Consequences of Ethical Failures in AI Nowadays**

  The failure to address and apply all these ethical frameworks and policies to AI leads to a biased AI, making AI decisions in favor of certain groups, exclusion from important services to the more vulnerable populations, and unfair treatment. All of these harms make the users distrust the AI systems and technologies, when AI is one of the most used and practical tools nowadays, that can help you with anything you need, making AI something essential in some areas. In some circumstances, the non-ethical AI has led some governments to introduce new laws, banning or restricting certain AI applications, since they don't have the morals they should have, because it disrespects human values.

- **Current Situation Towards Ethical AI Development**

  In response to these complicated challenges, governments, industries, companies, countries and agencies have all worked towards striking a biased AI and making it as similar as possible to be aligned with human values. For example global organisations like UNESCO have established and recommended normative frameworks like the 2021 Recommendation on the Ethics of Artificial Intelligence, that had as a base principles like, fairness, human rights, transparency and accountability. By this, UNESCO'S recommendation has been taken as an example to follow by many countries, making that many technology companies publish their own ethical AI principles and frameworks and created multiple programs to constantly monitor the AI so it doesn't operate without any bias.

  Despite all of these advances, there are still challenges along the way. The ethical frameworks and policies apply differently depending on the country, that's why it has been complicated to establish norms that can be applied universally. Either way, lots of industries and governments are constantly working with each other over seas to get to an AI that's beneficial for everyone

  That's why ethical AI development is a must and it's rapidly evolving, especially centered around discipline on the creation of intelligence systems that align with human-values so it can be accepted socially and not cause any problems. Since AI has been integrated into areas such as healthcare, finances, among many others, concerns about bias, discrimination and unfairness have been a problem. Developers must constantly identify and eliminate those biases, through extensive research, data practices and fairness-aware systems. Transparency is very important, as the decisions from the AI's often are complex and involve tedious processes behind them, causing difficult understanding for users or regulators. Making AI decisions explainable and easy, secures the users trust and enables accountability. Principles such as fairness, transparency, accountability, and the respect for human rights and privacy are the key and well recognized in the process of having an ethical AI system for everyone.

Major Parties Involved

- **Kingdom of the Netherlands**

Netherlands is known for being a leader in AI Ethics, since they have a national strategy that involves ethical principles such as human-rights, inclusivity and transparency. The government does multiple collaborations with various entities to promote and have an AI that creates a system where social values and innovation are the protagonists. Its policies align very closely to the European Union (EU) regulations, as well as the EU AI Act (European Commission, 2024).

This progress is seen in education, research and multiple interactions with the public, to be able to build trust and know what it's better and have a better understanding of AI (UNESCO, 2023). The Netherlands has been cited as one of the principal models for responsible and ethical AI development in Europe, due to its very well approach to the topic and being balanced with innovation and the regulations applied. Plus being the number one country in 2024 for the most responsible and ethical AI country.

- **Germany**

Germany has been widely recognized as one of the strongest countries with ethical AI development, since it aligns with the EU AI Act on which it "Focuses on developing trustworthy AI technologies to enhance Europe's competitiveness while safeguarding and advancing our democratic values. It aims to bring the benefits of AI to various sectors such as healthcare, education, industry, and environmental sustainability." (European Commission, 2025)

The German government has been hugely invested in ethical AI, since they do massive research on it, promoting collaborations with academia and industries. Since Germany's approach is very human-centered, it aligns with human-wellbeing and promotes regulatory experiments, including regular sandboxes (AI Regulatory Sandbox Approaches: EU Member State Overview, 2025). And they do broad public dialogue, so it shows the commitment to apply a safe AI for their community.

- **Canada**

Canada has been an explorer in ethical AI, it has been recognized internationally for its commitment and responsibility on responsible and ethical AI. Canada is well known for their Pan-Canadian AI Strategy, made in 2017, in which it was centered on privacy protection, transparency, fairness and inclusivity. Canada's participation is reflected in the frameworks such as Montreal Declaration for Responsible AI and the Toronto Declaration on which mainly it promotes equality and fairness in AI systems.

Canada has proposed the Artificial Intelligence and Data Act (AIDA), that aims to regulate the high impact AI's systems, with obligatory requirements for fairness.

- **United Kingdom**

The United Kingdom is well known for its strong institutional framework and their policy leadership. Their government made the "Centre of Data Ethics and Innovation" to provide guidance in transparency, policy, deep AI research and more. The United Kingdom strategy for AI, seeks ethical research and the public's participation, so it ensures that AI has the public's interest and respect for human rights, no discrimination and fairness. The UK is in constant participation in international AI governance efforts, for example the UK AI Safety Summit on which it led to the Bletchley Declaration and promoting a more human AI with international collaboration. Plus the UK invests in ethical AI testing, by initiatives like the AI Safety Institute, empathizing the big importance of trust and social benefits in AI.

- **United States of America**

The United States seeks the responsible use of AI, that's why they go under 5 important principles which were established when it comes to AI ethics: responsibility, equitable, traceable, reliable and governable. They have some key initiatives such as the National AI Research and Development Strategic Plan and the AI Bill of Rights and both of them focus on bias mitigation, fairness and privacy. The US plays a big role in the international AI governance discussions, so they can come with innovation and creativity to help develop ethical AI development policies.

RANKINGS	COUNTRY	REGION	INDEX SCORE	PILLAR SCORE			DIMENSION SCORE		
				Government frameworks	Government actions	Non-state actors	Human rights and AI	Responsible AI capacities	Responsible AI governance
1	Netherlands	Europe	86.16	74.33	95.46	91.23	78.74	88.59	91.12
2	Germany	Europe	82.77	72.69	93.00	82.48	80.30	64.03	90.94
3	Ireland	Europe	74.98	81.71	74.17	63.16	84.11	57.60	73.68
4	United Kingdom	Europe	73.12	60.66	80.90	82.48	67.59	66.54	79.62
5	USA	North America	72.81	62.41	79.19	80.87	65.37	84.34	74.75
6	Estonia	Europe	67.61	65.58	81.52	43.86	65.85	58.54	72.01
7	Italy	Europe	61.8	59.99	55.96	77.10	66.34	38.21	66.13
8	France	Europe	57.62	56.23	58.92	57.81	39.84	55.17	72.27
9	Canada	North America	57.39	37.80	80.84	49.68	55.37	57.06	59.07
10	Australia	Asia and Oceania	56.22	34.69	72.44	66.82	71.31	31.62	52.67

Source: Global Index on Responsible AI Report 2024

Previous Attempts to Solve the Issue

Over the past few years, international organizations, industries and governments have done multiple discussions to implement efficient and realistic solutions aimed at AI technologies that are developed and used with responsibility and have an ethical criterion behind them. These solutions have seeked mainly to focus on safeguarding transparency, privacy, fairness and respect for human values, for the benefit of society.

1. **GDPR (General Protection Data Regulation)**

   Introduced by the European Union, the GDPR introduces a strict set of guidelines on how the personal data should be gathered, processed and saved. Its main goal is to give control to the users of their personal data, so it's not used against them. This means that in AI, the organisations have to ensure that there is transparency on how the users data is used. So they have to include systems where the users can protect their data from unwanted and unauthorized access to it. Transparency is the key and an obligation in AI decisions (University of San Diego, s.f.)

2. **IBM'S AI Ethics Board**

   IBM has created an AI ethics board, that in 2022 the OPTR set in motion the Privacy and AI Management Systems (PIMS), so it helps IMB to manage ML models with trustworthiness, promote accountability, transparency and facilitate obedience with privacy and AI regulations (IBM, s.f.) So by this, IBM has positioned itself as a big leader in a trustworthy AI, ensuring bias mitigation principales and trust with the user.

3. **Microsoft's Responsible AI Standard**

   Microsoft has dedicated itself to develop AI systems that function with trust, transparency and reliability. That's why Microsoft identified 6 principles that AI should have to ensure a better development of it, which are; fairness, inclusiveness, transparency, accountability, reliability and safety, and privacy and security. All of these principles are completely ethical and align with our human values, to ensure that the AI has a reasonable and responsible use behind it so it doesn't operate without any bias. (Responsible AI Principles And Approach | Microsoft AI, s. f.)

Possible Solutions

Ethical AI Development itself is a problem that has been solved by many countries, but there are numerous ways in which these solutions can be improved, or the addition to new ones for better policies and frameworks that align perfectly with human values, without causing any harm in society

1. **Organization and governance strategies**
    a) Customizing AI Frameworks to the necessities and rules of the country
    b) A functional and cooperative work with other countries
    c) Ethics Training programs to ensure a well done AI with ethic values

2. **Technical/Development Solutions**
    a) Data privacy improvement
    b) Constant auditing and monitoring AI
    c) Bias detection and mitigation tools
    d) Ethical impact assessments

3. **Good AI use**
    a) Transparency in the interactions between AI and the person
    b) AI use for social initiatives or events with good intentions

References

- Transcend. (n.d.). *AI Ethics: Why it matters and how to practice it*. Transcend. https://transcend.io/blog/ai-ethics
- UNESCO. (n.d.). *Recommendation on the ethics of artificial intelligence*. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
- Coursera. (2023, March 29). *AI ethics: What it is and why it matters*. https://www.coursera.org/articles/ai-ethics
- SmartDev. (2023, August 15). *A comprehensive guide to ethical AI development: Best practices, challenges, and the future*. https://smartdev.com/a-comprehensive-guide-to-ethical-ai-development-best-practices-challenges-and-the-future/
- Rapid Innovation. (2023, June 6). *Ethical AI development guide*. https://www.rapidinnovation.io/post/ethical-ai-development-guide
- SAP. (s.f.). *What is AI bias?* SAP. https://www.sap.com/resources/what-is-ai-bias#:~:text=Artificial%20intelligence%20bias%2C%20or%20AI,discrimination%2C%20prejudice%2C%20and%20stereotyping.
- Holistic AI. (s.f.). *What is ethical AI*. https://www.holisticai.com/blog/what-is-ethical-ai
- Pure Storage. (2023, marzo 30). *What is ethical AI?* https://blog.purestorage.com/purely-educational/what-is-ethical-ai/
- Neurons Inc. (s.f.). *Ethical AI*. https://www.neuronsinc.com/glossary/ethical-ai
- University of Helsinki. (s.f.). *What is accountability?* In *Ethics of AI*. https://ethics-of-ai.mooc.fi/chapter-3/2-what-is-accountability
- Shelf.io. (s.f.). *Ethical AI uncovered: 10 fundamental pillars of AI transparency*. https://shelf.io/blog/ethical-ai-uncovered-10-fundamental-pillars-of-ai-transparency/#:~:text=At%20its%20core%2C%20AI%20transparency,ethical%20standards%20and%20societal%20values
- BytePlus. (s.f.). *The origins of ethics in AI: A historical overview*. https://www.byteplus.com/en/topic/403397?title=the-origins-of-ethics-in-ai-a-historical-overview
- McCormack, L. (n.d.). *Top 3 countries leading responsible AI: Ireland leads the way*. LinkedIn. https://www.linkedin.com/pulse/top-3-countries-leading-responsible-ai-ireland-louise-mccormack-kw6ae
- Rejolut. (n.d.). *13 top AI countries*. https://rejolut.com/blog/13-top-ai-countries/
- Preiksa, U. (n.d.). *Global AI power rankings: Top 35 countries in 2025*. LinkedIn. https://www.linkedin.com/pulse/global-ai-power-rankings-top-35-countries-2025-preiksa-ueenf
- Pandey, P. (n.d.). *Who is leading the world in artificial intelligence in 2025?* Dev.to. https://dev.to/priyankapandey/who-is-leading-the-world-in-artificial-intelligence-in-2025-e32
- Business Elites Africa. (2024, April 10). *Top 10 countries leading in artificial intelligence in 2025*. https://businesselitesafrica.com/top-10-countries-leading-in-artificial-intelligence-in-2025/?v=161747ec4dc9
- ResearchGate. (n.d.). *Top ten countries leading in AI ethics development and research according to Global AI Index*. https://www.researchgate.net/figure/Top-Ten-Countries-Leading-in-AI-Ethics-Development-and-Research-according-to-Global-AI_fig1_381992270
- Zawya. (2024, February 15). *2025: Top 10 countries leading in AI*. https://www.zawya.com/en/business/technology-and-telecom/2025-top-10-countries-leading-in-ai-euhy3ijr

- Ng, J. (n.d.). *A global perspective on AI ethics frameworks*. LinkedIn. https://www.linkedin.com/pulse/global-perspective-ai-ethics-frameworks-jean-ng--1tdec
- UNESCO. (n.d.). *The Netherlands: Championing AI ethics in Europe*. https://www.unesco.org/ethics-ai/en/netherlands#:~:text=Dutch%20policy%20actively%20integrates%20considerations,those%20that%20are%20most%20marginalized
- European Commission. (n.d.). *The European approach to artificial intelligence*. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence
- OECD. (2024). *GSG background note* [PDF]. https://www.oecd.org/content/dam/oecd/en/about/programmes/strategic-foresight/GSG%20Background%20Note_GSG(2024)1en.pdf
- UNESCO. (2023). *Recommendation on the ethics of artificial intelligence* [PDF]. https://unesdoc.unesco.org/ark:/48223/pf0000393240
- UNESCO. (n.d.). *Ethics of artificial intelligence*. https://www.unesco.org/ethics-ai/en
- Digital Watch Observatory. (2024, February 8). *UNESCO and the Netherlands launch European AI supervision initiative*. https://dig.watch/updates/unesco-and-the-netherlands-launch-european-ai-supervision-initiative
- European Commission. (n.d.). *The European approach to artificial intelligence*. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence
- Artificial Intelligence Act. (2023, June 15). *AI regulatory sandbox approaches – EU Member State overview*. https://artificialintelligenceact.eu/ai-regulatory-sandbox-approaches-eu-member-state-overview/
- Deutschland.de. (2023, February 20). *German Ethics Council: Risks of artificial intelligence*. https://www.deutschland.de/en/topic/culture/german-ethics-council-risks-of-artificial-intelligence
- OECD. (2023). *OECD artificial intelligence review of Germany*. https://www.oecd.org/en/publications/oecd-artificial-intelligence-review-of-germany_609808d6-en.html
- Canada.ca. (2025, March). *Canada moves toward safe and responsible artificial intelligence*. https://www.canada.ca/en/innovation-science-economic-development/news/2025/03/canada-moves-toward-safe-and-responsible-artificial-intelligence.html
- Government of Canada. (n.d.). *Innovation, Science and Economic Development Canada (ISED)*. https://ised-isde.canada.ca/site/ised/en
- Oxford Insights. (n.d.). *Spotlight series: Canada*. https://oxfordinsights.com/insights/spotlight-series-canada/
- University Affairs. (2024, January 9). *Ensuring a responsible and ethical development of artificial intelligence*. https://universityaffairs.ca/news/ensuring-a-responsible-and-ethical-development-of-artificial-intelligence/
- Centre for Data Ethics and Innovation. (n.d.). *About us*. UK Government. https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation
- Department of Defense. (2020, February 24). *DoD adopts ethical principles for artificial intelligence*. https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/
- Government of the United Kingdom. (n.d.). *AI Safety Institute*. https://www.gov.uk/government/organisations/ai-safety-institute
- Government of the United Kingdom. (n.d.). *AI Safety Summit 2023*. https://www.gov.uk/government/topical-events/ai-safety-summit-2023/about

- Government of the United Kingdom. (n.d.). *Understanding artificial intelligence ethics and safety*. https://www.gov.uk/data-ethics-guidance/understanding-artificial-intelligence-ethics-and-safety
- Montreal Declaration. (n.d.). *Montreal declaration for responsible AI*. https://montrealdeclaration-responsibleai.com/
- National Coordination Office for Networking and Information Technology Research and Development (NITRD). (2023). *National artificial intelligence research and development strategic plan: 2023 update* [PDF]. https://www.nitrd.gov/pubs/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf
- U.S. Department of State. (n.d.). *Artificial intelligence*. https://2021-2025.state.gov/artificial-intelligence/
- IBM. (n.d.). *IBM Office of Privacy and Responsible Technology*. https://www.ibm.com/case-studies/ibm-office-of-privacy-and-responsible-technology
- Microsoft. (n.d.). *Responsible AI principles and approach*. https://www.microsoft.com/en-us/ai/principles-and-approach
- University of San Diego. (n.d.). *Ethics in artificial intelligence*. https://onlinedegrees.sandiego.edu/ethics-in-ai/
- Convin.ai. (n.d.). *Responsible AI: Building ethical AI models in the age of automation*. https://convin.ai/blog/responsible-ai
- LinkedIn. (n.d.). Saikarthik, A. K. *Ethical AI in action: Real-world case studies*. https://www.linkedin.com/pulse/ethical-ai-action-real-world-case-studies-saikarthik-ak-xwjrc
- NCube. (2023, October 17). *AI software development ethics: A guide to building responsible AI solutions*. https://ncube.com/ai-software-development-ethics-guide-to-building-responsible-ai-solutions
- SATORI Project. (n.d.). *Section 5: Ethical impact assessment*. https://satoriproject.eu/framework/section-5-ethical-impact-assessment/#:~:text=The%20ethical%20impact%20evaluation%20stage%20is%20aimed%20at%20assessing%20the,value%20conflicts%20that%20may%20arise