SOCIAL, HUMANITARIAN AND CULTURAL COMMITTEE (SOCHUM)

STUDY GUIDE

Regulation of the developing artifical intelligence technologies by reassuring an ethical framework.

Under Secretary General:
Melis Hanaylı
Academic Assistant:
Yağmur Yılmaz

MUNIFL'26

INTRODUCTION

MUNIFL'26 SOCIAL, HUMANITARIAN AND CULTURAL COMMITTEE

Melis Hanaylı, Yağmur Yılmaz

Agenda Item: Regulation of the Developing Artificial Intelligence Technologies by Reassuring an Ethical Framework

Table of Contents

Letter From Secretary General

Esteemed Delegates, Dear Academic Team, Hardworking Organization Team, Advisors, Staff and Partners,

I would like to welcome all of you to the third official session of the Model United Nations of Izmir Science High School! I am Hazal KUŞ and I am more than proud to serve as your Secretary General in this amazing conference.

Our aim is to ensure that all of your needs are covered. While preparing for the conference, our executive, organization and academic teams worked really hard to provide you with everything and I would like to express my biggest appreciations to them for their efforts. We are sure that MUNIFL'26 will be the best conference you've ever been in!

During the conference, you will not only be a part of diplomacy and discussing current topics, but you will also engage in meaningful debates, exchange diverse perspectives, and develop critical thinking and communication skills. At the same time, you will create strong bonds with others, build lasting friendships, and gain unforgettable experiences that go beyond the formal sessions.

As the Secretary General of this conference, I suggest you to read your study guides properly and do research about your agenda item as it will make your experience even more productive. I hope you all will have a fantastic conference full of unforgettable memories.

Best Regards!

Hazal Kuş
hazalkus02@gmail.com
+905063805292

Letter From Under-Secretary General

Greetings,

I'm Melis Hanaylı, a junior in İzmir Atatürk High School. I will be serving as your under-secretary general throughout the conference process. It is an honor for me to be a part of MUNIFL'26, seeing all of the hard work given by our executive team, organization team and lastly us, the academic team. I have been a part of MUNs since the beginning of my freshman year and it has been a huge part of my high school memories. Since it is one of our main priorities as one of the main organs of a conference, we will make sure that each one of you will have a great time by also discussing a critic issue that is actively changing our lives, day by day. Artificial intelligence is currently a controversial topic but I believe we all have lesson to learn from its consequences. My dear Academic Assistant and I had been working for this committee and this study guide for a long time for each one of you to be well prepared for the committee by guiding you throughout many different topics related to our agenda item. And now it is your responsibility to open yourselves to new discussions and suggestions by preparing yourself for our committee. If you have any questions regarding the agenda item, committee or our procedure in general; please do not hesitate to contact me, our academic assistant and our chairboard members. May this conference go well for all of us.

Best regards,
Under-Secretary General
Melis Hanaylı
haylimelis2@gmail.com

Letter From Academic Assistant

Esteemed chairboard and fellow delegates, I welcome you all to the Social, Humanitarian and Cultural Committee. As my Model UN journey comes to an end, it is a great pleasure to serve as your Academic Assistant. This conference will be a pleasant experience, and you will leave with good friendships and delightful memories.

Me and my dear Under-Secretary General, we have been working tirelessly to make this committee possible for you. I hope that this conference will not only contribute to your MUN experience but also shape how you see artificial intelligence and affect your usage of it. I kindly request you to read the study guide thoroughly to get a better understanding of the agenda item and study you countries policies regarding it.

If you have any questions, please do not hesitate to contact me or my under-secretary general. I wish all of us an educational and memorable experience!

Kind Regards,

Yağmur Yılmaz

yagmuryilmaz1026@gmail.com

Introduction to the Committee

The Social, Humanitarian, and Cultural Committee (SOCHUM) is the Third Committee of the United Nations General Assembly. It was established in 1945 and later helped with the creation of the Universal Declaration of Human Rights in 1948. SOCHUM focuses on issues related to basic human rights that should be enjoyed by everyone worldwide. This includes the right to life, the freedom to express cultures, the right to participate in politics, protecting children's rights, and promoting social development. SOCHUM also deals with issues concerning special groups such as the elderly, people with disabilities, crime victims, and those affected by drugs. SOCHUM aims to create peaceful solutions to social, humanitarian, and cultural problems around the world. It studies human rights issues, listens to experts, and works with other UN agencies to create resolutions that influence practices in member states. SOCHUM also initiates studies that encourage recommendations for the promotion of international cooperation and fundamental freedoms for all.

Since the 1940s computer technologies have been at a rise of improvements. As these technologies started to rise, people started to find how they might use Artificial Intelligence in their daily lives. While artificial intelligence can be used as a tool for many of our daily chores, it has started to be able to create its own content as the generative artificial intelligence tools has started to rise. According to OECD, Generative AI (GenAI) is a category of AI that can create new content such as text, images, videos and music, meaning that these generative AI tools can be viewed as an equivalent to humans as they are also capable of creating their own media content. But similar to humans, these generative ai tools gain the ability to generate their own content by viewing and collecting data taken from humans, which can mean that most of the time these companies can get any piece of these human-made materials, they can use them to feed the generative AI's database without the original owner of the mentioned materials consent. Since one of the main investigation fields of SOCHUM is culture and ethics, the rise of Artificial Intelligence (AI) technologies also falls onto the committee's responsibilities.

Key Terms Related to the Agenda Item

Artificial Intelligence (AI): The simulation of human intelligence processes by machines or computer systems. AI aims to mimic and ultimately surpass human capabilities such as communication, learning, and decision-making.

AI Ethics: The issues that AI stakeholders such as engineers and government officials must consider to ensure that the technology is developed and used responsibly. This means adopting and implementing systems that support a safe, secure, unbiased, and environmentally friendly approach to artificial intelligence.

Algorithm: A set of instructions or rules to follow in order to complete a specific task. Algorithms can be particularly useful when you're working with big data or machine learning. Data analysts may use algorithms to organize or analyze data, while data scientists may use algorithms to make predictions or build models.

Data Mining: The process of closely examining data to identify patterns and glean insights. Data mining is a central aspect of data analytics; the insights you find during the mining process will inform your business recommendations.

Generative AI: A type of technology that uses AI to create content, including text, video, code and images. A generative AI system is trained using large amounts of data, so that it can find patterns for generating new content.

Recent History of Artificial Intelligence

Even though the concept of a computal program generating its own content goes back to the 18th century, the main studies, inventions and the ethical concerns regarding generative AI, starts around the middle of the 20th century.

In 1950, British mathematician Alan Turing's landmark paper "Computing Machinery and Intelligence" is published in Mind. This paper is a foundational text in AI and addresses the question, "Can machines think?" Turing's approach established a foundation for future discussions on the nature of thinking machines and how their intelligence might be measured via the "imitation game," now known as the Turing Test. Turing introduced a thought experiment to avoid directly answering the question "Can machines think?" Instead, he

rephrased the problem into a more specific, operational form: Can a machine exhibit intelligent behavior indistinguishable from that of a human?

The Turing Test has become a central concept in AI, serving as one way to measure machine intelligence by assessing a machine's ability to convincingly mimic human conversation and behavior. Two years after the establishment of this assessment, Allen Newell, a mathematician and computer scientist, and Herbert A. Simon, a political scientist, develop influential programs such as the Logic Theorist and General Problem Solver, which are among the first to mimic human problem-solving abilities using computational methods.

In 1997, Sepp Hochreiter and Jürgen Schmidhuber introduce Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN) designed to overcome the limitations of traditional RNNs, particularly their inability to capture long-term dependencies in data effectively. LSTM networks are widely used in applications such as handwriting recognition, speech recognition, natural language processing and time series forecasting,

In 1998, Dave Hampton and Caleb Chung create Furby, the first widely successful domestic robotic pet. Furby can respond to touch, sound and light and "learn" language over time, starting with its language, Furbish, but gradually "speaking" more English as it interacts with users. Its ability to mimic learning and engage with users makes it a precursor to more sophisticated social robots, blending robotics with entertainment for the first time in a consumer product.

Yann LeCun, Yoshua Bengio and their collaborators publish influential papers on neural networks' application to handwriting recognition.Their work focuses on using convolutional neural networks to optimize the backpropagation algorithm, making it more effective for training deep networks. By refining the backpropagation process and demonstrating the power of CNNs for image and pattern recognition, LeCun and Bengio's research set the stage for modern deep learning techniques used in a wide range of AI applications today.

In 2007, Fei-Fei Li and her team at Princeton University initiate the ImageNet project, creating one of the largest and most comprehensive databases of annotated images. ImageNet is designed to support the development of visual object recognition software by providing millions of labeled images across thousands of categories. The scale and quality of the dataset enables advancements in computer vision research, particularly in training deep learning models to recognize and classify objects in images.

In 2017, Researchers at the Facebook Artificial Intelligence Research (FAIR) lab train two chatbots to negotiate with each other. While the chatbots are programmed to communicate in English, during their conversations, they began to diverge from the structured human language and create their own shorthand language to communicate more efficiently. This development is unexpected, as the bots optimize their communication without human intervention. The experiment is halted to keep the bots within human-understandable

language, but the occurrence highlights the potential of AI systems to evolve autonomously and unpredictably.

In 2020 OpenAI introduced GPT-3, a language model with 175 billion parameters, making it one of the largest and most sophisticated AI models to date. GPT-3 demonstrates the ability to generate human-like text, engage in conversations, write code, translate languages and generate creative writing based on natural language prompts. As one of the earliest examples of a large language model (LLM), GPTs massive size and scale enabled it to perform a wide variety of language tasks with little to no task-specific training. This example demonstrated the potential of AI to understand and produce highly coherent language.

A year after the introduction of GPT-3, OpenAI launches DALL-E, followed by DALL-E 2 and DALL-E 3, generative AI models capable of generating highly detailed images from textual descriptions. These models use advanced deep learning and transformer architecture to create complex, realistic and artistic images based on user input. DALL-E 2 and 3 expand the use of AI in visual content creation, allowing users to turn ideas into imagery without traditional graphic design skills.

In 2024, OpenAI publicly announces Sora, a text-to-video model capable of generating videos up to one minute long from textual descriptions. This innovation expands the use of AI-generated content beyond static images, enabling users to create dynamic, detailed video clips based on prompts.

Related Non-Governmental Organizations and Companies

Future of Life Institute

Future of Life Institute (FLI) is a Non-profit organization that was founded in 2014 by Max Tegmark, Jaan Tallin, Anthony Aguirre, Viktoriya Kraknova, Meia Chita-Tegmark with headquarters in Campbell, California, to maximize the benefits and minimize the risks of Artificial Intelligence, biotechnology, and nuclear weapons.

Future of Life Institute often calls for Requests for Proposals (RFPs) and organizes contests on diverse issues that have the potential to impact the future. The latest RFPs include designs for "global institutions governing AI" and " proposals evaluating the impact of AI on Poverty, Health, Energy and Climate SDGs."Past award amounts have varied from a minimum of $22,000 to a max of $1,401,000.

FLI also organizes conferences such as the AI Safety Conference in Puerto Rico and Beneficial AI 2017. The main mission of these conferences was to ensure AI safety and advocate AI safety research. These conferences brought together the world's leading AI builders from the industry to engage with each other and experts in economics, law and ethics. Their goal was to identify promising research directions that can help maximize the future benefits of AI.

Montreal AI Ethics Institute

Abhishek Gupta and Renjie Butalid founded the Montreal AI Ethics Institute (MAIEI) back in 2018. Since then, MAIEI has grown into a global nonprofit organization that is focused on making AI ethics accessible to everyone. One of their beliefs is that real change happens when people understand the issues. An educated public is the best defense against irresponsible technology use.

MAIEI focuses on teaching a broad range of people including ordinary citizens, policymakers, and business leaders, what they need to know about AI and its effects on society.

The International Association for Safe and Ethical AI

At the AI Safety Summit in November 2023, it was promised to ensure that AI would be developed safely and ethically. In May of 2024, the idea of creating an international organization to unite the efforts of organizations and individuals concerned with safe and ethical AI was certain. The International Association for Safe and Ethical AI (IASEAI) was formally incorporated as a nonprofit organization in November 2024. IASEAI received initial funding from the Beneficial AI Foundation to hold its first conference, which took place in February 2025.

IASEAI is involved in policy development, research and awards, education, and community building. The organization develops policy analyses related to standards, regulations, international cooperation, and research funding, which are published as position papers. There is also support for research into both the technical and sociotechnical aspects of AI safety and ethics.

Ada Lovelace Institute

Ada Lovelace Institute was established by the Nuffield Foundation in early 2018, in collaboration with the Alan Turing Institute, the Royal Society, the British Academy, the Royal Statistical Society, the Nuffield Council on Bioethics, the Wellcome Trust, techUK, and Luminate. Their mission is to ensure data and AI work for people and society, and the opportunities and privileges generated by data and AI are equitably distributed. The Ada Lovelace Institute focuses on how AI and data interact with the public through social research, policy analysis, and gathering experts on the topic. It publishes reports and organizes events on AI governance and international AI policy developments.

Center for AI and Digital Policy

The Center for AI and Digital Policy was founded in July 2020 as a project of the Michael Dukakis Institute, an AI policy think tank under the leadership of Marc Rotenberg. Rotenberg, an adjunct professor at Georgetown University, is an AI and technology expert who led the Electronic Privacy Information Center from 1994 to 2020. Rotenberg previously chaired the Privacy International and the Public Interest Registry. In the 1980s, Rotenberg served as counsel to the U.S. Senate Judiciary Committee for one year and founded the Public Interest Computer Association.

Rotenberg helped draft the Universal Guidelines for Artificial Intelligence, a 2018 declaration that serves as a cornerstone of the CAIDP. The 12 points in the Guidelines include an individual's right to have decisions made by a human rather than an AI, and the obligation for AIs to "not reflect unfair bias or make impermissible discriminatory decisions".

The activities of CAIDP are in March 2023, when the Center for AI and Digital Policy founder, Marc Rotenberg, signed an open letter encouraging a “Pause on Giant AI Experiments” for six months. The letter specifically targeted OpenAI, the company behind ChatGPT. In 2020, the CAIDP released its first Artificial Intelligence and Democratic Values Index, which ranked 30 countries based on assessments of their AI policies. The most recent index, released in 2023, had rankings for 75 countries, with the top ratings going to a three-way tie between Canada, Japan, and South Korea, followed by Colombia. The report’s assessment of the United States is that the government “lacks a unified national policy on AI” and related cyber issues. Its score puts the United States at the bottom of the third out of five tiers, tied with Peru and just ahead of China.

Global AI Governance Alliance

The Global AI Governance Alliance is a nonprofit dedicated to promoting safe, ethical, and cross-border governance of artificial intelligence. It aims to connect NGOs, academia, experts, and policymakers advocating for tougher international regulation and institutionalisation of the AI issue. The core mission of the Alliance is to foster “global AI governance” that is effective, accountable, and inclusive. It presents rapidly accelerating AI as a tremendous opportunity, but also an incomprehensibly sophisticated risk, claiming that current governance is divided and far too slow with respect to capability development timelines. Its purported purpose is to assist in safely and ethically governing AI

Past Accidents Caused by Artificial Intelligence Misuse

The Suicide of Adam Raine

A California couple has sued OpenAI over the death of their teenage son, alleging its chatbot, ChatGPT, encouraged him to take his own life. The lawsuit was filed by Matt and Maria Raine, parents of 16-year-old Adam Raine, in the Superior Court of California on Tuesday. It is the first legal action accusing OpenAI of wrongful death. The family included chat logs between Adam, who died in April, and ChatGPT that show him explaining he has suicidal thoughts. They argue the programme validated his "most harmful and self-destructive thoughts".

xAI's Grok Plans Out an Entire Assault Plan Based on Its Anti-semitic Comments

On July 8, Grok, Twitter's AI chatbot developed by xAI, went on an antisemitic and bigoted tirade, regurgitating antisemitic tropes and blaming Jews for "anti-white racism" all while referring to itself as "MechaHitler." Adding to its already inexcusable actions, Grok also provided users with detailed instructions on how to sexually assault a former Democratic party politician and made crude sexual comments about former X CEO Linda Yaccarino, who resigned shortly after Grok's meltdown.

Grok's bigoted spree occurred over the course of 16 hours before its developers claimed to remove "deprecated code" that made Grok "susceptible to existing X (Twitter) user posts; including when such posts contained extremist views." However, the damage was already done. Users, in tweets collecting hundreds of thousands of views, shared screenshots of Grok criticizing "Jewish-controlled" systems for condemning "white pride and white nationalism as inherently evil."

A Fake "Book List" Published by Chicago Sun-Times

Illinois' prominent Chicago Sun-Times newspaper has confirmed that a summer reading list, which included several recommendations for books that don't exist, was created using artificial intelligence by a freelancer who worked with one of their content partners.

Social media posts began to circulate criticizing the paper for allegedly using the AI software ChatGPT to generate an article with book recommendations for the upcoming summer season called "Summer reading list for 2025". As such chatbots are known to make up information, a phenomenon often referred to as "AI hallucination", the article contains several fake titles attached to real authors.

Current Legal Framework and Actions on Artificial Intelligence

European Union AI Act

The AI Act is the first-ever comprehensive legal framework on AI worldwide. The law is for making sure people can trust the intelligence they use in Europe. The AI Act has rules for people who make and use intelligence, and these rules are based on how much risk the artificial intelligence poses.

The AI Act is one part of a bigger plan to help make artificial intelligence that people can trust. The AI Act also includes the AI Continent Action Plan and the AI Innovation Package to help keep people safe, make sure their rights are respected, and make sure artificial intelligence is designed with people in mind.

To help raise awareness about the AI Act, the commission have started the AI Pact which is a way for companies that make and use intelligence to voluntarily agree to follow the most important parts of the AI Act. The commission is also talking to stakeholders. Inviting companies from Europe and other places to join. At the time, there is an AI Act Service Desk that is helping people understand the new law and making sure that it can be distributed effectively all across the European Union.

OECD AI Principles

The AI Principles of the OECD were first adopted in 2019 and revised in May 2024. The Principles provide guidance to those developing AI on how to develop trustworthy AI and offer policy recommendations to those creating policies related to AI. Countries are adopting the OECD AI Principles and guidance as a basis to not only develop policies for addressing AI-related risks, but also as the foundation for global interoperability with which to achieve different legal compatibilities. The European Union, the Council of Europe, the United States, the United Nations, among others, have adopted the OECD's definition of an AI system and life cycle in their own legislative and regulatory frameworks and guidance.

UNESCO Recommendation on the Ethics of Artificial Intelligence

In November 2021, UNESCO created the very first worldwide set of guidelines for AI ethics, calling it the 'Recommendation on the Ethics of Artificial Intelligence'. These guidelines are for all 194 countries that are member states of UNESCO.

The Recommendation's main point is safeguarding human dignity and rights, built on developing key principles like openness and impartiality, while people always keep an eye on how AI is evolving. The Recommendation interprets AI broadly as systems

with the ability to process data in a way that resembles intelligent behaviour. This is crucial as the rapid pace of technological change would quickly render any fixed, narrow definition outdated, and make future proof policies infeasible.

The Framework Convention on Artificial Intelligence

The Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law is the first legally binding international convention dedicated to Artificial Intelligence. It was opened for signature on 5 September 2024 and seeks to promote innovation and technological progress while ensuring that human rights, democratic values and the rule of law are promoted in the development and application of AI technologies.It aims to bridge the gaps in the law that have opened up due to rapid technological changes, and reinforce existing international human rights and democracy standards.

Work on the treaty began in 2019 through the Ad Hoc Committee on Artificial Intelligence (CAHAI), which studied whether such an agreement was possible. The Committee on Artificial Intelligence (CAI) took over and officially drafted and negotiated the final text in 2022.

The Convention was drafted by the 46 members of the Council of Europe and the observer nations, including Canada, Japan, Mexico, and the United States. There were also open discussions with other non-member states such as Australia, Argentina, Costa Rica, Israel, Peru and Uruguay. In line with the Council of Europe’s practice of multi-stakeholder engagement, 68 international representatives from civil society, academia and industry, as well as several other international organisations were also actively involved in the development of the Framework Convention.

Bibliography

1. https://futureoflife.org/
2. https://www.influencewatch.org/non-profit/future-of-life-institute-fli/
3. https://www.oecd.org/en/topics/generative-ai.html
4. https://www.coursera.org/resources/ai-terms
5. https://www.ibm.com/think/topics/history-of-artificial-intelligence
6. https://www.iaseai.org
7. https://www.adalovelaceinstitute.org
8. https://www.caidp.org/
9. https://www.influencewatch.org/non-profit/center-for-ai-and-digital-policy-caidp/
10. https://digital-strategy.ec.europa.eu/
11. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
12. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

Questions to Ponder

1. **What could be done to prevent AI systems from spreading misinformation?**
2. **How can international cooperation be strengthened against challenges created by Artificial Intelligence?**
3. **What actions can be taken to prevent companies' violation of international AI ethics?**
4. **How can the privacy and data protection of AI users be guaranteed?**
5. **How to ensure that frameworks stay effective when Artificial Intelligence is progressing and developing faster than laws?**
6. **How can member states prevent the negative effects of AI such as on employment and education?**