UNCSTD

AGENDA ITEM:
IMPROVISATION OF THE INTERNATIONAL FRAMEWORKS FOR DATABASE QUALITY AND PRIVACY ON MACHINE LEARNING ALGORITHMS AND THEIR AUTOMATION

UNDER SECRETARY GENERAL:
DORUK ŞENTÜRK

ACADEMIC ASSISTANT:
DERIN YAVUZ

"It's all begins in the Sky"

Table of Contents

1. Letters

1.1 Letter from the Secretaries-General

Honourable participants of ÇAĞDAŞMUN'25,

As the secretary generals of ÇAĞDAŞMUN'25, taking place from November 14th to November 16th, we would like to extend our warm welcome to all participants of this prestigious conference.

Model United Nations conferences are more than just a simple event, it is a torch that shines its light upon a variety of great opportunities, guiding the youth to the future through the brightness it eradiates. It grants the opportunity to be in the minds of diplomats and decision makers allowing the participants to learn the ways of decision making and debating, at the same time giving the chance to apply the learnings in real time debates. When organised with utmost care and responsibility, one conference can shape hundreds of individuals into leaders of the future.

In the first official edition of ÇAĞDAŞMUN, our mission is to achieve what most struggle to do: committees with a wide grasp of the past, present and the future, a marvelous organisation team to be in our most perfect form and most importantly, a mission to create space for bright minds to shine the most powerful, hidden gems to come to light for the greatest jewelries and disregarded souls to prove themselves as unignorable leaders.

It all begins in the sky.

With our warmest regards,

Secretary Generals of ÇAĞDAŞMUN'25,

Mustafa Aslan and Kuzey Karlık.

1.2 Letter from the Under Secretary-General

Dear delegates, distinguished academic team and fellow participants;

I am honored to welcome you all to the first annual of ÇAĞDAŞMUN'25 conference. It's a pleasure to serve you as an Under Secretary-General of the UNCSTD committee. My name is Doruk ŞENTÜRK and I am a 12th grade student in Antalya Maya Science and Technology High School.

I'm highly believing that this guide, as it states, will help your committee roles and processes well-enough. Our general expectation for you is to study this guide sufficiently and join the committee with your knowledge of the country's AI procedure and function which you will be represented as.

Lastly, I want to thank the executive team for giving me an opportunity to serve as under secretary general in this delightful conference and my academic assistant, my sister in another mother, Derin Yavuz, for her partnership and existence. I cannot wait to meet all of you personally and I hope you can have a great time with the time you spend in the conference. You can contact me at any kind of situation or problem, do not hesitate to ask me any questions about the committee. See you soon!

Sincerely,

Doruk ŞENTÜRK
Under Secretary-General

dorukksenturk@gmail.com
+90 0506 022 234

1.3. Letter from the Academic Assistant

Dear delegates, honorable academic and executive team, and distinguished participants;

I welcome you all to to the first annual conference of ÇAĞDAŞMUN'25. I am highly honored to be your Academic Assistant and it is a great pleasure to have you as part of UNCSTD committee in order to discuss upon "improvisation of the international frameworks for database quality and privacy on machine learning algorithms and their automation." My name is Derin YAVUZ and i am a 12th grader student in Konyaaltı Bahçeşehir Anatolian High School.

We encourage each delegate to conduct thorough research on existing international regulations and initiatives related to data governance and ai ethics.
During the sessions, your active participation, and collaboration will be key. We expect every delegate to prepare themselves to any kind of debate.

I am so excited to see your great ideas, and help you with your academic journey.

Last but not least I would love to thank the Executive Team for inviting me to such a prestigious conference. And my dearest friend Doruk ŞENTÜRK for being with me in my life as my brother.

Warm regards,
Derin YAVUZ
Academic Assistance

derinyavuz2008@gmail.com

+905539030733

2. Committee Introduction

The United Nations Commission on Science and Technology for Development (UNCSTD) is a subsidiary body of the United Nations Economic and Social Council. The UNCSTD gathers annually during the United Nations Conference on Trade and Development (UNCTAD) to debate and discuss pressing issues on the use, oversight, and development of world-changing technologies Working with member states, NGOs, and other stakeholders in the global scientific community is key towards creating essential policy impact.

The UNCSTD was first established in 1992 after a successful conference in Vienna paved the way for the United Nations to consider science and technology in international policy-making. It was established to provide a forum to debate how best to accommodate developing countries in an increasingly digital world. Today, as new technologies challenge the way in which we think about health, the climate and economic opportunities, the UNCSTD is needed more than ever to promote these conversations on how the international community should put forward such innovations in the best interests of member states and the global scientific community.

3. Introduction to Agenda Item

Artificial intelligence (AI) and robotics are digital technologies that will be of major importance for the development of humanity in the near future. They are autonomous technologies minimalizing the human aspect, further optimizing our technology for a never before seen advancement of our species.

Even with the current technology, in a number of areas, AI can perform tasks much better than humans. Particularly when it comes to repetitive, detail-oriented tasks, such as analyzing large numbers of legal documents to ensure relevant fields are filled in properly, AI tools often complete jobs quickly and with relatively few errors. Because of the massive data sets it can process, AI can also give enterprises insights into their operations they might not have been aware of.

While AI tools present a range of new functionality for businesses, the use of AI also raises ethical questions because, for better or worse, an AI system will reinforce what it has already learned.

This can be problematic because machine learning algorithms, which underpin many of the most advanced AI tools, are only as smart as the data they are given in training. Because a human being selects what data is used to train an AI program, the potential for machine learning bias is inherent and must be monitored closely.

Anyone looking to use machine learning as part of real-world, in-production systems needs to factor ethics into their AI training processes and strive to avoid bias. This is especially true when using AI algorithms that are inherently unexplainable in deep learning and generative adversarial network applications.

Explainability is a potential obstacle for using AI in industries that operate under strict regulatory compliance requirements. For example, financial institutions in the United States operate under regulations that require them to explain their credit-issuing decisions. When a decision to refuse credit is made by AI programming, however, it can be difficult to explain how the decision was arrived at because the AI tools used to make such decisions operate by teasing out subtle correlations between thousands of variables. When the decision-making process cannot be explained, the program may be referred to as black box AI.

In summary, AI's ethical challenges include the following:

- Bias due to improperly trained algorithms and human prejudices.
- Elimination of jobs due to the growing capabilities of AI.
- Data privacy concerns in work spaces and social life

3.1. What is AI?

Artificial Intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. These machines and computer systems are able to perform tasks and take decisions historically only a human could do, without feeling any remorse and far more effectively.

AI is used as an umbrella term consisting of a variety of technologies such as:

- **Machine Learning:** a subfield of AI that uses algorithms trained on data sets to create self-learning models that can predict outcomes and classifying information without human intervention.
- **Deep Learning:** a method that trains computers to process information in a way that mimics human neural processes made from neural networks with three or more layers.
- **Natural Language Processing (NLP):** is a form of AI that allows computers to understand human language, whether it be written, spoken, or even scribbled. Etc.
- **Machine Vision:** This technology gives a machine the ability to see. Machine vision captures and analyses visual information using a camera, analog-to-digital conversion and digital signal processing. It is often compared to human eyesight, but machine vision is not bound by biology and can be programmed to see through walls, for example. It is used in a range of applications from signature identification to medical image analysis. Computer vision, which is focused on machine-based image processing, is often conflated with machine vision.

- **Text, image and audio generation:** Generative AI techniques, which create several types of media from text prompts, are being applied extensively across businesses to create a seemingly limitless range of content types from photorealistic art to email responses and screenplays.

3.2. What are Robots?

A robot is a machine which can be programmable by a computer, capable of carrying out a complex series of actions automatically. A robot can be guided by an external control device, or the control may be built within. Robots may be constructed to be associated with human form but most robots are task-performing machines, designed with an emphasis on stark functionality, rather than expressive aesthetics.

A recent UNESCO report describes robots as artificial beings with four characteristics:

- **Mobility**, which is important for functioning in human environments like hospitals and offices.
- **Interactivity**, made possible by sensors and actuators, which gather relevant information from the environment and enable a robot to act upon this environment.
- **Communication**, made possible by computer interfaces or voice recognition and speech synthesis systems; and
- **Autonomy**, in the sense of an ability to 'think' for themselves and make their own decisions to act upon the environment, without direct external control (UNESCO 2017: 4).

3.3 Machine Learning Algorithms

Machine learning algorithms are essentially sets of instructions that allow computers to learn from data, make predictions, and improve their performance over time without being explicitly

programmed. Machine learning algorithms are broadly categorized into three types:

- **Supervised Learning: Algorithms** learn from labeled data, where the input-output relationship is known.
- **Unsupervised Learning:** Algorithms work with unlabeled data to identify patterns or groupings.
- **Reinforcement Learning:** Algorithms learn by interacting with an environment and receiving feedback in the form of rewards or penalties.

4. Current Situation

Artificial Intelligence (AI) holds great promise for advancement but also presents significant risks across various domains. This paper examines the multifaceted dangers of AI, encompassing ethical, societal, economic, and existential concerns. Ethical dilemmas arise as AI systems make decisions affecting individuals' lives, raising questions about accountability, transparency, bias and more. Addressing these dangers requires interdisciplinary collaboration, robust regulation, ethical guidelines, and ongoing dialogue to ensure responsible development and deployment of AI technologies.

4.1. Data and Biases

In order to make AI function properly, they need to be trained with vast amounts of data. The data can take various forms, such as images, audio, text, or structured data, and each example is associated with an output label or annotation that describes what the data represents or how it should be classified. AI systems work by analysing the data for correlations and patterns, and using these patterns to make predictions about future states. In this way, a chatbot that is fed examples of text can learn to generate lifelike exchanges with people, or an image recognition tool can learn to identify and describe objects in images by reviewing millions of examples. New, rapidly improving generative AI techniques can create realistic text, images, music and other media

While high-quality AI training data is essential for building accurate, effective, and fair machine learning models, obtaining it can be challenging. Some of the most important challenges are:

- **Quality control:** Ensuring the quality of the training data can be challenging, particularly when it comes to manual labelling. Human error, inconsistency, and subjective judgments can all impact the quality of the data.

- **Lack of availability:** One of the biggest challenges in obtaining high-quality AI training data is the lack of availability. Data may be difficult or expensive to obtain, particularly for niche or sensitive domains.

- **Cost:** Another challenge in obtaining high-quality AI training data is the cost. High-quality data can be expensive to acquire, particularly if it needs to be collected or labelled manually.

- **Data labelling:** Depending on the problem being solved, obtaining high-quality AI training data may require extensive labelling efforts, which can be time-consuming and expensive.

- **Data volume:** Obtaining enough high-quality data can be a challenge, particularly when it comes to deep learning models that require enormous amounts of data to achieve high accuracy.

The arising issues are not limited to these. Another important challenge is machine learning bias. Also known as algorithm bias or Artificial Intelligence bias, refers to the tendency of algorithms to reflect human biases. It is a phenomenon that arises when an algorithm delivers systematically biased results because of erroneous assumptions of the machine learning processes. In today's climate of increasing representation and diversity, this becomes even more problematic because algorithms could be reinforcing biases. AI systems contain biases due to two reasons:

- **Cognitive biases:** These are unconscious errors in thinking that affects individuals' judgements and decisions. These biases arise from the brain's attempt to simplify processing information about the world. More than 180 human biases have been and classified defined by psychologists. Cognitive biases could seep into machine learning algorithms via either
  - ➢ designers unknowingly introducing them to the model,
  - ➢ a training data set which includes those biases.

- **Lack of complete data:** If data is not complete, it may not be representative and therefore it may include bias. For example, most psychology research studies include results from undergraduate students which are a specific group and do not represent the whole population, or imagine an AI system trained to recognize human voices but only on data from a single gender or accent. Such a system is likely to perform poorly on folks from other regions or have different accents.

AI systems specifically generative AI systems also introduce several privacy concerns due to its ability to process personal data and generate potentially sensitive information. Personal data, like names, addresses, and contact details, can be inadvertently collected during interactions with AI systems. The processing of personal data by generative AI algorithms may result in unintended exposure or misuse of this information. If the training data contains sensitive data, like medical records, financial information, or other identifiers, there's a risk of unintentionally generating sensitive information that violates privacy regulations across jurisdictions and puts individuals at risk.

Misuse of data collection can be further broken down into three main issues:

a. **Data persistence**

This refers to the ongoing storage and availability of data within a system. In the context of AI and privacy, data persistence raises concerns because data, once collected and stored, can be accessed and potentially misused long after the initial purpose of collection has passed. Sensitive or personal data persisting indefinitely in databases or on the cloud is a significant privacy risk.

Some organisations have outright banned their employees from using ChatGPT, likely due to OpenAI's policy that any data input, even proprietary information, may be used to train the model.

b. **Data repurposing**

Data repurposing is the use of data for a purpose other than that for which it was originally collected. In the realm of artificial intelligence, where large datasets are often used to train and refine algorithms, data repurposing is inevitable.

This raises concerns about data privacy, particularly if individuals did not consent to the secondary use of the data, or if the new purpose deviates significantly from the original intent.

c. **Data spillovers**

Data spillover refers to the unintended or incidental exposure of data beyond its intended scope or audience. This can occur in an AI system when data, collected for specific purposes, ends up being accessible in other contexts, potentially leading to data breaches.

Data spillovers are particularly concerning in interconnected systems where data from one application or sector can inadvertently become available in another, such as through an API.

4.2. DeepFake Technologies

Deepfake technology is a process where a video or image can be manipulated by artificial intelligence (AI) to create a false representation. Individual attributes, including intricate facial features, can be studied by AI programs to generate that individual in realistic videos or images that did not take place. This technology also extends to audio, where a person's voice can essentially be harnessed and controlled using voice synthesis. Voice synthesis describes a process where algorithms can faithfully analyse, deconstruct, and then reproduce a person's voice using the correct tone, pitch, and cadence.

Deepfakes present unique challenges to national security and law enforcement. They can be exploited to incite violence through fake inflammatory statements attributed to public figures or to fabricate evidence, potentially jeopardizing key global initiatives such as climate change agreements. This technology's capability extends to the legal sector, where it can create synthetic evidence in criminal cases, influencing the outcomes and integrity of legal proceedings.

The culmination of the above process is that criminal practitioners may be presented with a piece of evidence that shows an individual doing and/or saying something that he or she did not do or say. With technological developments constantly aimed at achieving the highest levels of realism, it is not difficult to see how this can affect not only issues surrounding identity but also the commission of offences in criminal proceedings.

Consider by way of example a footage that has been submitted by a member of the public that claims to live on the same road as the defendant. This footage is grainy and not of the highest quality to begin with. It shows the defendant, wearing the exact same clothes to what he was later arrested in (a camouflage top with bright orange HOODRICH logo, Nike bottoms, and Crocs), leaving his house and walking down the street before stabbing the victim and running away. A male voice that, on the face of it, sounds very much like the defendant can be heard in the video demanding money from the victim. As the defendant is running away, he looks in the general direction of the camera, and a very unique birthmark across his face can be seen.

Setting aside issues of admissibility for the moment, it is troubling indeed that every feature of this video could have been deepfaked. That is to say, either generated from scratch by AI, or the identifying features of the defendant (clothes, physical attributes, facial features, and voice) transposed onto the actual person who committed the robbery. Deepfakes in the above context would have been created intentionally with the singular focus of framing the defendant at that scene committing the act. There is no mistake involved. Ancillary features such as his clothes and facial birthmark are carefully planted. In addition to these identifying features, the defendant can also be manipulated by AI to show him committing the act of stabbing the victim. By the same token, deepfake evidence can also support his alibi evidence elsewhere.

Consequently, one can see deepfake evidence being of critical importance at trial. As of now, the current regimes addressing deepfake include the Online Safety Bill, The Fraud Act, Audiovisual Media Services Regulation, and The Data Protection Act. However, these frameworks focus on combatting misinformation, privacy concerns, and using deepfakes to commit fraud. There is no awareness yet on how deepfake evidence can be admitted as evidence at trial to affect the outcome.

It is submitted that it is only a matter of time before deepfake evidence begins to appear in criminal proceedings. There is a growing need for the courts and practitioners to start thinking about establishing rules and procedures to determine the authenticity of evidence where there is suspicion of deepfake involvement. Delaying this acknowledgment, awareness, and action will, in my view, make it more likely that individuals may be wrongly convicted by deepfaked evidence.

The commercial sector is equally at risk. Deepfakes can facilitate corporate sabotage by spreading false information about companies. They enable sophisticated social engineering attacks, tricking employees into making costly financial decisions. Similarly, the financial industry is vulnerable, with deepfakes potentially manipulating stock markets and compromising banking security.

The personal impacts of deepfakes are profound and disturbing. They enable new forms of cyberbullying, attacking individuals' reputations through manufactured content.

The political landscape is not immune to the influences of deepfakes. They have been utilised to impersonate political figures, manipulate public sentiment, and even incite chaos or conflicts, as seen during Slovakia's contested parliamentary elections and a doctored TV interview with a US Senator circulated on social media. These examples highlight the technology's potential to disrupt political processes and sway public opinion on an international scale.

Notable incidents further demonstrate the extensive reach of deepfakes. One time Sachin Tendulkar's deepfake video circulated in which he was promoting an online gaming app. In 2023, the crypto sector saw a surge in advanced scams using AI and deepfakes. A notable case involved a fake video of MicroStrategy CEO Michael Saylor offering to "double money instantly," leading viewers to send Bitcoin to fraudsters. This highlighted Deepfakes' role in complex financial fraud. Similarly, back in 2019, a video of Facebook CEO Mark Zuckerberg speaking about controlling the world's population surfaced, illustrating the potential for public opinion manipulation and widespread confusion.

These scenarios collectively highlight the diverse and severe implications of this technology, emphasizing the critical need for heightened awareness and the development of effective strategies to counteract these emerging threats.

4.3. Abusive use of AI

AI-powered cyber attacks use machine learning to analyze a human or machine target and find techniques most likely to help compromise an organization. This could be generating an email based on your people's social media profile or using small bits of information to predict the most likely vulnerabilities in a target system and launch an attack. These attacks can be highly targeted and can bypass traditional cybersecurity solutions that are not equipped to detect them. Yesterday's advice - like paying attention to mis-spelling and poor grammar in an email that helped detect human-oriented attacks, or high amounts of bad traffic in a machine-generated scan - are techniques of the past.

One of the key risks of AI-powered cyber attacks is their ability to learn and adapt to new defenses. Traditional cybersecurity solutions often rely on known patterns and signatures to detect and block attacks. However, AI-powered attacks can learn from these defenses and find new ways to bypass them.

Another risk of AI-powered cyber attacks is their potential to cause widespread damage. These attacks can target critical infrastructure, such as power grids and transportation systems, and disrupt entire economies. They can also steal sensitive data, such as financial information and intellectual property, which can have long-lasting consequences for organizations and individuals.

In 2017, the WannaCry ransomware attack, which affected organizations across the globe, was launched using an automated tool that scanned for vulnerabilities in networks and systems. The automated tool made it possible for the attack to be launched on a massive scale, with little human intervention required. The attack highlighted the potential for automated hacking tools to cause widespread damage.

The risks of AI powered cyber attacks are not limited to digital space. Hackers all around the world can harness the power of AI to cause physical damage. Stuart Madnick, an MIT professor of engineering systems and co-founder of Cybersecurity at MIT Sloan (CAMS), said that he and his team have simulated cyberattacks in the lab, resulting in explosions. They were able to hack into computer-controlled motors with pumps and make them incinerate. Attacks that cause temperature gauges to malfunction, pressure values to jam, and circuits to be circumvented can also cause blasts in lab settings. Such an outcome, Madnick said, would do far more than simply taking a system offline for a while, as a typical cyberattack does.

"If you cause a power plant to stop from a typical cyberattack, it will be back up and online pretty quickly, but if hackers cause it to explode or burn down, you are not back online a day or two later; it will be weeks and months because a lot of the parts in these specialized systems are custom made. People don't realize downtimes can be substantial," Madnick said.

He added that the technology, now boosted by AI, exists to wreak havoc on physical systems.

4.4. Inequality in AI Distribution

As of October 2021, 44 countries were reported to have their own national AI strategic plans, showing their willingness to forge ahead in the global AI race. These include emerging economies like China and India, which are leading the way in building national AI plans within the developing world.

Oxford Insights, a consultancy firm that advises organisations and governments on matters relating to digital transformation, has ranked the preparedness of 160 countries across the world when it comes to using AI in public services. The US ranks first in their 2021 Government AI Readiness Index, followed by Singapore and the UK.

Notably, the lowest-scoring regions in this index include much of the developing world, such as sub-Saharan Africa, the Caribbean and Latin America, as well as some central and south Asian countries.

The developed world has an inevitable edge in making rapid progress in the AI revolution. With greater economic capacity, these wealthier countries are naturally best positioned to make large investments in the research and development needed for creating modern AI models.

In contrast, developing countries often have more urgent priorities, such as education, sanitation, healthcare and feeding the population, which override any significant investment in digital transformation. In this climate, AI could widen the digital divide that already exists between developed and developing countries.

4.5. Carbon Emissions of AI

Today data centers run 24/7 and most derive their energy from fossil fuels, although there are increasing efforts to use renewable energy resources. Because of the energy the world's data centers consume, they account for 2.5 to 3.7 percent of global greenhouse gas emissions, exceeding even those of the aviation industry.

Most of a data center's energy is used to operate processors and chips. Like other computer systems, AI systems process information using zeros and ones. Every time a bit—the smallest amount of data computers can process—changes its state between one and zero, it consumes a small amount of electricity and generates heat. Because servers must be kept cool to function, around 40 percent of the electricity data centers use goes towards massive air conditioners. Without them, servers would overheat and fail.

In 2021, global data center electricity use was about 0.9 to 1.3 percent of global electricity demand. One study estimated it could increase to 1.86 percent by 2030. As the capabilities and complexity of AI models rapidly increase over the next few years, their processing and energy consumption needs will too. One research company predicted that by 2028, there will be a four-fold improvement in computing performance, and a 50-fold increase in processing workloads due to increased use, more demanding queries, and more sophisticated models with many more parameters. It's estimated that the energy consumption of data centers on the European continent will grow 28 percent by 2030.

With AI already being integrated into search engines like Bing and Bard, more computing power is needed to train and run models. Experts say this could increase the computing power needed—as well as the energy used—by up to five times per research. Moreover, AI models need to be continually retrained to keep up to date with current information.

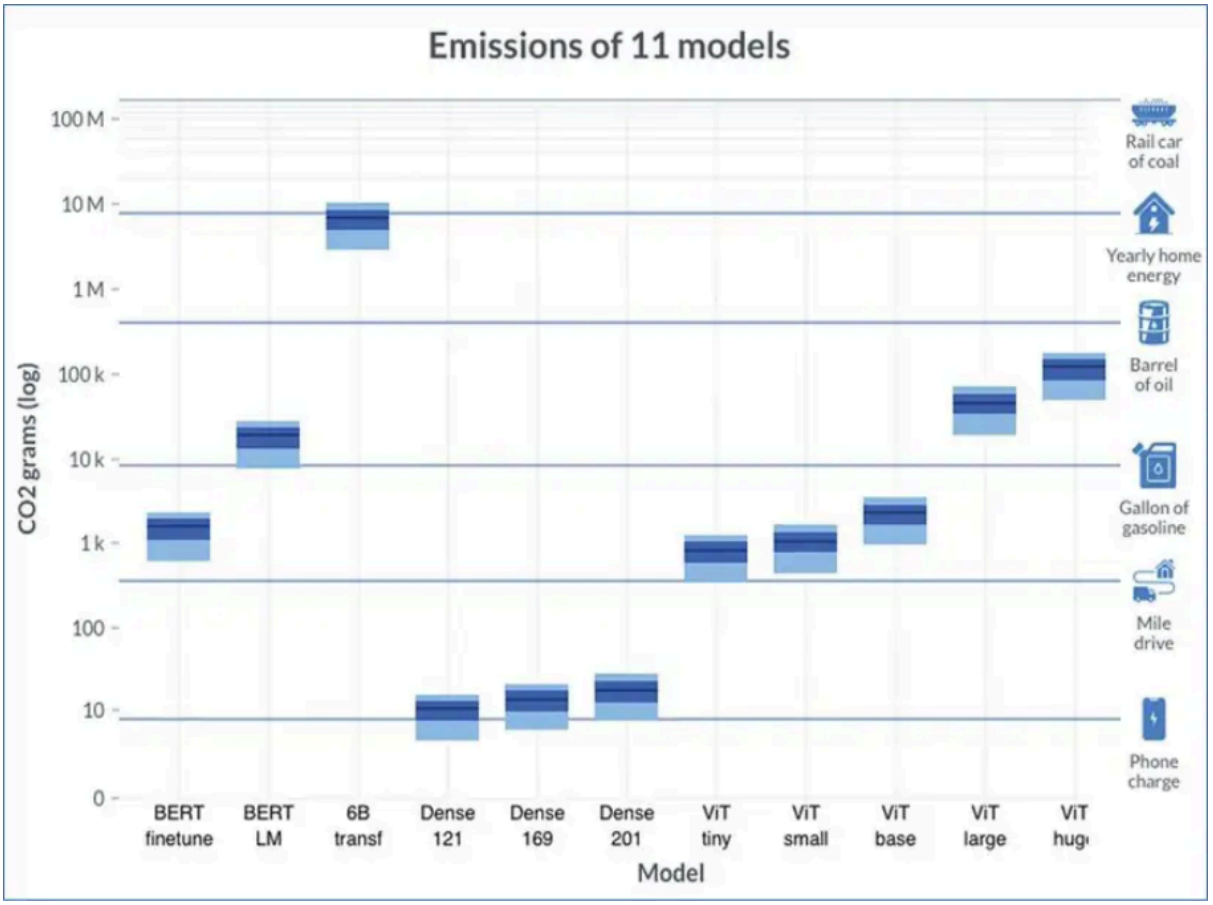
In 2019, University of Massachusetts Amherst researchers trained several large language models and found that training a single AI model can emit over 626.000 pounds of CO2, equivalent to the emissions of five cars over their lifetimes.

A more recent study reported that training GPT-3 with 175 billion parameters consumed 1287 MWh of electricity, and resulted in carbon emissions of 502 metric tons of carbon, equivalent to driving 112 gasoline powered cars for a year.

Once models are deployed, inference—the mode where the AI makes predictions about new data and responds to queries—may consume even more energy than training. Google estimated that of the energy used in AI for training and inference, 60 percent goes towards inference, and 40 percent for training. GPT-3's daily carbon footprint has been estimated to be equivalent to 50 pounds of CO2 or 8.4 tons of CO2 in a year.

Inference energy consumption is high because while training is usually done multiple times to keep models current and optimized, inference is used many many times to serve millions of users. Two months after its launch, ChatGPT had 100 million active users. Instead of employing existing web searches that rely on smaller AI models, many people are eager to use AI for everything, but a single request in can consume 100 times more energy than one Google search, according to one tech expert.

Northeastern University and MIT researchers estimated that inference consumes more energy than training, but there is still debate over which mode is the greater energy consumer. What is certain, though, is that as OpenAI, Google, Microsoft, and the Chinese search company Baidu compete to create larger, more sophisticated models, and as more people use them, their carbon footprints will grow. This could potentially make decarbonizing our societies much more difficult.



5. AI Policy of Some Critic Countries

5.1. Russia

In October 2019, Russia adopted its National strategy for the development of artificial intelligence by 2030, mostly coordinated between state-owned commercial companies and the Russian Government. In September 2017, when Putin stated that artificial intelligence masters would “rule the world”, Russia announced that it was entering a global race to develop artificial intelligence technologies. Previously, though, despite the aforementioned obstacles to artificial intelligence development, Russian government organizations and companies had begun to adopt artificial intelligence for their own use. For example, in April 2016, Sberbank, a state-owned bank, created a venture capital fund focused on investing financially in startups.

5.2. People’s Republic of China

China has enforced a new law to regulate generative artificial intelligence (AI), a move made to manage increasing international apprehension regarding privacy invasion, violation of intellectual property rights and the circulation of false information. While some analysts contend that this legislation indicates a relaxed approach towards AI, there is an agreement that China's focus on AI governance remains strong, balancing innovation with control in managing industry setbacks and economic turbulence, with the regulations serving an essential role in securing international regulatory leadership and reducing risks posed by AI.

5.3) Japan

Japan's approach to AI regulation focuses on fostering innovation, ensuring public trust, and addressing ethical considerations. Unlike some jurisdictions that have implemented sector-specific regulations for AI, Japan has not established assessment or audit requirements specific to AI. This relatively hands-off approach indicates a regulatory environment that encourages developing and applying AI technologies across various sectors without imposing stringent sector-specific mandates.

5.4. United Kingdom

The UK Government has adopted a cross-sector and outcome-based framework for regulating AI, underpinned by five core principles. These are safety, security and robustness, appropriate transparency and explainability, fairness, accountability and governance, and contestability and redress.

5.5. United States of America

This initiative identified five key lines of effort, including increasing AI research investment, unleashing Federal AI computing and data resources, setting AI technical standards, building America's AI workforce, and engaging with international allies.

5.6. Italy

The importance of AI for the National Security of a country has been growing steadily in the last five years. Hence, Italy is fully committed to investing in AI applications that ensure the security of its citizens. This includes individual and national cybersecurity, where AI has been contributing to the development of new-generation detection and resolution software.

5.7. Spain

The new plan for the digitalisation of the public sector in 2021-2025 further stimulates the use of AI in the public administration by presenting policy initiatives to foster digitalisation and automation in the public sector to finally improve the quality of public policies and service to citizens.

6. Questions to be Addressed (QTBA)

1) How can we prevent possible data leaks and biases in AI?

2) What are some ways to optimize and control the AI training data in areas such as but not limited to volume, storage of data, quality and availability?

3) Which developments can be done with youth generations and interested officials for their education grade?

4) What measures can be taken in order to prevent the harmful usage of illegal technologies such as Deepfake?

5) What can be done to mitigate the digital, physical and physiological effects of and be protected against cyber attacks powered by AI?

6) What are the ways to equally distribute AI technologies across the globe to prevent the increasement of inequalities?

7) What can we do to reduce carbon emissions and energy use of AI technologies to make them more sustainable for human life?

7. Bibliography

https://news-climate-columbia-edu.translate.goog/2023/06/09/ais-growing-carbon-footprint/?_x_tr_sl=en&_x_tr_tl=tr&_x_tr_hl=tr&_x_tr_pto=sc

https://www.supermicro.com/en/article/ai-training-5-tips-reduce-environmental-impact

https://theconversation.com/developing-countries-are-being-left-behind-in-the-ai-race-and-thats-a-problem-for-all-of-us-180218

https://miguelgfierro.com/blog/2018/10-ethical-issues-of-artificial-intelligence-and-robotics/

https://philarchive.org/archive/MLLEOA-4v2

https://www.datagrail.io/blog/data-privacy/generative-ai-privacy-issues/

https://www.reuters.com/legal/legalindustry/privacy-paradox-with-ai-2023-10-31/

https://transcend.io/blog/ai-and-privacy

https://www.jpe.ox.ac.uk/papers/first-steps-towards-an-ethics-of-robots-and-artificial-intelligence/

https://plato.stanford.edu/entries/ethics-ai/

https://www.geeksforgeeks.org/machine-learning/machine-learning-algorithms