**Forum:** Social, Humanitarian & Cultural Committee

**Issue:** Reviewing the Global Threat of Artificial Intelligence

**Student Officer:** Jesse-Krishna Johnston

**Position:** Co-chair

---

# PERSONAL INTRODUCTION

Dear Delegates of QEGSMUN 2024,

My name is Jesse-Krishna Johnston, a year 13 student at Queen Elizabeth's Grammar School. I welcome you to the Social, Humanitarian & Cultural Committee during the 3rd annual QEGSMUN conference.

This study guide will serve as an introduction to the topic of "Reviewing the Global Threat of Artificial Intelligence", a topic which is a passion of mine, reflected in my Extended Project Qualification discussing the Ethics of Artificial Intelligence in Modern Society as well as my future degree course, providing me with a helpful insight into the topic. Delegates should note that this guide will not cover all the information that they will require, and therefore are advised to indulge in further study, especially as differing delegations may have different viewpoints on the discussed topic, so take your assigned country into consideration.

The matter at hand is a developing problem, meaning that delegates should try to refer to more recent sources of information as artificial intelligence has grown in popularity within recent years. There may have been more development between the time of the conference and the creation of this study guide so continue to keep in mind recent developments.

I hope this study guide provides you with a useful insight into the topic which will be discussed across the 18th and 19th of October. If you have any questions on the information of this guide or about how to prepare for the conference feel free to contact me at 3389@queenelizabeths.kent.sch.uk.

Yours Sincerely,
Jesse-Krishna Johnston, Co-Chair of the Social, Humanitarian & Cultural Committee

# TOPIC INTRODUCTION

Within recent years the topic of artificial intelligence (AI) has become a growing controversy, with differing groups either in support or against the development. Since the recent introduction of generative AI platforms such as ChatGPT and Copilot, growing fears for a variety of different sectors, including things such as education, employment and warfare.

Artificial intelligence's integration into society has been met with fears of mass unemployment due to the ability of the systems being able to mimic human intelligence and efficiency. The introduction of artificial intelligence's possible impact on production has caused many large corporations, such as Google and Meta, to invest substantial amounts into the development of AI. Google Deepmind CEO Demis Hassabis, expects Google alone to spend more than $100 billion on the development of artificial intelligence. ``

AI has also stirred further worries surrounding the spread of false information. New generative AI tools have led to an influx of very convincing fake news, 'deep fake' videos, and other misleading content - all of which could undermine things such as democratic processes, and therefore begin to destabilise society as a whole. Other security risks include the threat to cybersecurity - AI can now be used to enhance cyberattacks, with AI-driven hacking tools now automating attacks meaning they are both more effective and harder to detect.

AI has also begun to be used to enhance current weapons in use, or even develop new autonomous weapons. This could lead to completely different forms of warfare to what we currently see, which would be much more destructive and harder to control than current ones. Ethically this is also a big issue, questions surrounding the decision making of such systems poses several moral complications. AI systems have been known to perpetuate and exaggerate existing biases in society, potentially leading to unfair treatment and discrimination.

All of this highlights and exaggerates the necessity of developing robust regulatory frameworks and ethical guidelines to mitigate the risks associated with AI. It's a complex issue that requires global cooperation and careful consideration.

# KEY TERMS

**Artificial Intelligence (AI):** any application or machine which has the ability to mimic human levels of intelligence.

**Machine Learning (ML):** A form of training artificial intelligence which teaches the AI model to develop decisions based on correlations between training data, this is performed through many types of algorithms.

**Algorithm:** A set of rules or instructions used in order for a computer system to perform a desired task.

**Big Data:** Large datasets which are inefficient to sort through using traditional data processing softwares, these are usually used in order to provide AI models with learning data.

**Computer Vision:** A field of computer science which enables computers to identify objects from different videos or images.

**Neural Networks:** Computer systems which are created from inspiration of how the human brain works.

# BACKGROUND INFORMATION

The rapid advancements of AI has sparked a lot of excitement around new opportunities, but also a lot of concern. This is because there are a lot of potential threats of AI. Things to consider includes:

1. **The unprecedented speed of development**
   AI algorithms are evolving at an astonishing pace. Researchers are constantly updating the software, and this rapid growth is raising concerns. It will be soon that AI will be able to improve itself without human intervention.

2. **The Chatbot Revolution**
   A critical factor is the development of advanced chatbots or large language models. These chatbots exhibit sparks of advanced general intelligence. For instance, GPT-4 outperformed 90% of human test takers on the Uniform Bar Exam, a standardised test for certifying lawyers, by reasoning and learning from vast amounts of data.

3. **Potential Risks**

- Loss of control: If AGI emerges and we get it wrong, the consequences will be disastrous. We might not have a second chance to correct any mistakes
- Security Threats: AI could assist hackers in cyberattacks or help terrorists plan chemical attacks
- Bias and Harm: AI systems can unintentionally perpetuate biases or create harmful outcomes. For example deepface videos or biassed algorithms
- Existential Risk: Some experts fear a 'nuclear level catastrophe' if AGI development goes unchecked.

# MAJOR COUNTRIES AND ORGANISATIONS INVOLVED

## United States of America

The United States of America is one of the leading countries in the development of Artificial Intelligence, most notably in 2023 when they were integral in creating the majority of new AI models. Private investment in the US is also exponentially high as there was over 62.5 billion euros invested in 2023, leading by over 55 billion euros to China. The US AI market has also been projected to reach $594 billion by 2032 which equates to a compound annual growth rate of 19% from 2023.

## United Kingdom

The UK has been another leading country in AI developments, with one of the founding fathers of computer science being Alan Turing, a British mathematician who helped the UK crack the enigma machine. The UK implemented many strategies aiming to position itself as a global centre for AI innovation. The UK has adopted a framework for regulating AI based on five core principles: safety; security and robustness; appropriate transparency and explainability; fairness; accountability and governance; and contestability and redress. In 2023, the government announced £100 million was being put towards new AI innovation and regulation. The UK government was the first to formally publish an assessment of the possible risks of advanced AI.

## Japan

Japan is a country which focuses on addressing the ethical problems of the use of AI. Due to the ageing population crisis in Japan, the labour force is very weak and small in numbers causing the country to want to implement artificial intelligence systems to fill the void left by the lack of workforce. There have been concerns that this goal of maximising AI's positive impact on society may lead to neglect in considering the possible risks. In December 2023, the Japanese government published draft guidelines for AI related businesses, in line with the G-7 international guiding principles. In February 2024, the Japan AI Safety Institute was established to advance research on methods for evaluating AI safety.

## India

India has one of the largest populations of any country in the world, and a recent IPSOS MORI survey says there are significantly higher levels of concern about AI in the country. However as a developing country AI could be a huge boost to their economy and its current focus has been on innovating the AI industry as they recognise the importance. For example India's National AI Strategy, #AIForAll is distinct in its focus on economic growth and social inclusion. India has used AI to revolutionise sectors such as healthcare, agriculture, finance and education.

## China

China is also a leading investor in the development of AI, hoping to improve its efficiency making it a worthy investment for productivity. China is expected to reach 'world-leading level' by 2025, and become 'the major innovation centre in the world' by 2030. There have been mounting concerns that China does not share the same concerns about 'AI Safety', however there is evidence to support both sides of this argument. The Chinese Communist Party released a major policy document in July 2024 which included a call to create "oversight systems to ensure the safety of artificial intelligence. China is now also exporting huge amounts of AI technology to other countries.

## Russia

Russia is a country which is interested in AI due to its military capability, unfortunately it is currently difficult to estimate whether the country's Ministry of Defense actually has begun to utilise systems and weapons that use AI, including on the Ukrainian battlefield. However it is confirmed that presently the Russian military establishment is investing in developing, researching and testing AI - with these investments shaped both by the understanding of where such emphasis is placed among likely competitors, such as the United States and NATO, and where resources should be allocated based on the ongoing complicated combat in Ukraine. There is a fear beginning to grow that this interest in AI enabled weapons could lead to an AI arms race.

# Minor Countries

## United Arab Emirates

In 2017, the UAE appointed the world's first AI minister and then in 2020 created a dedicated AI research university. Furthermore in 2023 they launched an open-source large language model known as Falcon. In March this year the UAE's Artificial Intelligence and Advanced Technology Council launched a technology investment company known as MGX, - with Mubadala (a UAE based investment company) and G42 (a large AI tech company) as founding partners. In September, this new company - alongside BlackRock and Microsoft - announced plans to support the construction of data centres and related energy infrastructure by investing $30 billion USD.

## Singapore

In 2019 Singapore became one of the first countries to adopt a national AI strategy. In December last year the Deputy Prime Minister, Lawrence Wong (who has since become prime minister!) updated these initiatives and further expanded upon them in a policy framework billed as "National AI Strategy 2.0." The government has allocated $743 million over the next five years to boost the country's AI capabilities. Singapore also has several clear policies in place to try and protect from threats of AI - the Model AI Governance Framework which was updated in 2020 provides detailed guidance to private sector organisations about addressing ethical and governance issues when deploying AI solutions.

## Israel

Israel is well known for its technology sector as well as innovation in AI, having a keen focus on cybersecurity and defence in the past. This interest has only grown recently with the outbreak of further violence within the region. High-ranking officers in the Israel Defence Forces (IDF) recently acknowledged that the use of AI-based tools as part of Israel's military arsenal is on the rise, with the IDF now deploying AI-based systems for defensive, reconnaissance and aggressive needs. They recently used an undisclosed AI-powered database that identified 37,000 potential targets based on their apparent links to Hamas, according to intelligence sources involved in the war.

Israel's use of powerful AI systems in its war on Hamas has entered uncharted territory for advanced warfare, raising a host of legal and moral questions, and transforming the relationship between military personnel and machines. One intelligence officer who used Lavender - the name of their AI system - said that they had more faith in a "statistical mechanism" than a grieving soldier. "Everyone there, including me, lost people on October 7. The machine did it coldly. And that made it easier.".

# TIMELINE OF KEY EVENTS

| Event | Date |
|---|---|
| **The Enigma machine was cracked** using "Christopher" , an AI built by Alan Turing. | 1941 |
| **Alan Turing published "Computing Machinery and Intelligence,"** introducing the Turing test and opening the doors to what would be known as AI. | 1950 |

| | |
|---|---|
| **The first proposal of artificial intelligence** was created by Allen Newell, Cliff Shaw and Herbert Simon. This first concept of AI was known as the Logic Theorist, a machine which could prove different mathematical theorems from Russel's Principia Mathematica. The term "Artificial intelligence" was later birthed at the Dartmouth Summer Research Project hosted by Marvin Minsky and John McCarthy in the same year. | 1956 |
| The **General problem solver** was created by Newell and Simon, this machine could find answers to many fairly structured problems. | 1957 |
| **Frank Rosenblatt developed the perceptron**, an early ANN that could learn from data and became the foundation for modern neural networks. John McCarthy developed the programming language Lisp, which was quickly adopted by the AI industry and gained enormous popularity among developers. | 1959 |
| **James Lighthill released the report "Artificial Intelligence: A General Survey"**, which caused the British government to significantly reduce support for AI research | 1973 |
| **Danny Hillis designed parallel computers for AI** and other computational tasks, an architecture similar to modern GPUs. | 1981 |
| **IBM's Deep Blue defeated Garry Kasparov in a historic chess rematch**, the first defeat of a reigning world chess champion by a computer under tournament conditions. | 1997 |
| Rajat Raina, Anand Madhavan and Andrew Ng p**ublished "Large-Scale Deep Unsupervised Learning Using Graphics Processors"**, presenting the idea of using GPUs to train large neural networks | 2009 |
| **DeepMind introduced deep reinforcement learning**, a CNN that learned based on rewards and learned to play games through repetition, surpassing human expert levels.<br>Google researcher Tomas Mikolov and colleagues introduced Word2vec to automatically identify semantic relationships between words. | 2013 |
| **Facebook developed the deep learning facial recognition system DeepFace**, which identifies human faces in digital images with near-human accuracy. | 2014 |
| **British physicist Stephen Hawking warned**, "Unless we learn how to prepare for, and avoid, the potential risks, AI could be the worst event in the history of our civilization." | 2017 |
| Developed by IBM, Airbus and the German Aerospace Center DLR, **Cimon was the first robot sent into space** to assist astronauts.<br>OpenAI released GPT (Generative Pre-trained Transformer), paving the way for subsequent LLMs. | 2018 |
| **Google AI and Langone Medical Center's deep learning algorithm** outperformed radiologists in detecting potential lung cancers. | 2019 |
| **The University of Oxford developed an AI test called Curial** to rapidly identify COVID-19 in emergency room patients. | 2020 |

| | |
|---|---|
| **OpenAI released ChatGPT** in November to provide a chat-based interface to its GPT-3.5 LLM. | 2022 |
| Elon Musk, Steve Wozniak and thousands more signatories urged a six-month pause on training "AI systems more powerful than GPT-4." | 2023 |

# PREVIOUS ATTEMPTS

## European Commission AI act

The European Commission AI act aims at addressing risks created through the application of AI. It determines which spheres of AI application are high risk, and defines four levels of risk for AI systems:

1. Unacceptable Risk: AI use is impossible (those of high risk which do not meet their specific obligations);
2. High Risk
   a) Critical infrastructures which can put citizens life and health at risk)
   b) Safety components of products (robot assisted surgeries)
   c) Law enforcement
   d) Essential private and public services (e.g. Credit or insurance scores)
   e) Migration (e.g. visa issue based on automated AI examination).
   f) Administration of justice and democratic processes (e.g. AI solutions to court rulings);
3. Limited risk: AI systems with a specific transparency obligation, per example, ChatGPT;
4. Minimal Risk: AI systems do not pose a significant risk.

## UK AI Regulation

In March 2023, the Government released its strategy for regulating AI with a focus on fostering innovation. Instead of having one set of strict rules for all AI development, the paper suggested a more flexible approach where different regulatory bodies in specific areas are in charge of AI governance, with support from central AI regulatory roles. These regulatory bodies are starting to reveal their plans for incorporating AI in their specific industries. One instance is when the Financial Conduct Authority (FCA), a financial regulatory body in the UK, has established its Consumer Duty to guarantee that all entities under its jurisdiction provide services that are fair value for consumers. These regulations prioritise the need for efficient AI management, particularly in competitive markets that focus on customer needs, such as insurance pricing.

## US AI Regulation

Similar to the UK, the US also lacks a comprehensive AI regulation, however, there are various frameworks and guidelines in place at federal and state levels. President Joe Biden signed an executive order for the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence just days before the UK's Safety Summit in October 2023. The order states, "Harnessing AI for good and realising its myriad benefits requires mitigating its substantial risks. This endeavour demands a society-wide effort that includes government, the private sector, academia, and civil society". The

executive order aims to require companies to test and disclose information on AI systems, and these requirements will probably be incorporated into rules at both federal and state levels.

## Japan AI Regulation

In 2019, Japan's government published the Social Principles of Human-Centred AI. According to this document, Japan's goal is to realise the world's first "AI-ready society" through a set of principles:

- AI must not infringe upon fundamental human rights
- Basic education and literacy must be ensured
- Privacy must be protected, security ensured, and a fair, competitive environment guaranteed
- Fairness, accountability and transparency are necessary
- AI innovation will be promoted through the close collaboration and cooperation of domestic and international stakeholders

These principles play out across three core philosophies:

- Respect for human dignity
- A sustainable society
- A society where diverse backgrounds support individual wellbeing

## India AI Regulation

Presently India lacks dedicated laws and policies to regulate AI but it has established a series of initiatives and guidelines.

In February 2021 NITI Aayog drafted the Principle for Responsible AI which examines the ethical considerations, categorised into either system and societal considerations. It outlines 7 overarching principles for the responsible governance of AI: safety and reliability; inclusivity and non-discrimination; equality; privacy and security; transparency; accountability; protection and reinforcement of positive human values.

The Ministry of Electronics and Information Technology has instituted committees on AI tasked with delivering reports on AI developments and safety and ethical concerns. Similarly, the Bureau of Indian Standards has set up a committee dedicated to AI, which is in the process of drafting Indian Standards for the industry.

## China AI regulation

China is in the midst of rolling out some of the world's earliest and most detailed regulations governing artificial intelligence (AI). These include measures governing recommendation algorithms, as well as new rules for synthetically generated images and chatbots in the mould of ChatGPT. China's emerging AI governance framework will reshape how the technology is built and deployed within China and internationally, impacting both Chinese technology exports and global AI research networks.

Tese regulations deserve careful study on how they will affect China's AI trajectory and what they can teach policymakers around the world about regulating the technology. Even if countries fundamentally disagree on the specific content of a regulation, they can still learn from each other when it comes to the underlying structures and technical feasibility of different regulatory approaches.

# POSSIBLE SOLUTIONS

There are many possible solutions to the global threats of Artificial Intelligence, and this list is just a guide on a few of the most known solutions.

1. Collaboration across borders and stakeholder groups
    - Improving collaboration is essential as AI transcends national boundaries, and shared understanding and cooperation can help us address common challenges

2. Human centric policies
    - Developing policies that prioritise human well-being is critical. We should ensure that AI technologies are directed toward augmenting humans and serving the common good
    - Policymakers can create guidelines that encourage responsible AI development, transparency, and fairness. This includes addressing biases, privacy concerns, and ethical implications

3. Empowering individuals in the "Race with the robots":
    - As AI transforms the job market, we must empower individuals to adapt. Education systems need to emphasise lifelong learning and reskilling
    - Shifting economic and political priorities toward supporting workers in this new landscape is essential. We want people to thrive alongside AI, not be left behind

4. Responsible AI development and regulation:
    - Encouraging responsible AI practices means developers are more likely to consider ethical implications, bias detection, and explainability
    - Governments can establish regulatory frameworks that balance innovation with safety. These frameworks should evolve as AI technologies do

# BIBLIOGRAPHY

The Guardian: Israel used AI to identify 37,000 Hamas targets, from
https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes

White & Case: AI Watch: Global regulatory tracker, from
https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-singapore

Embassy of the UAE: The UAE and US: Forging the Future of Artificial Intelligence and Advanced Tech, from
https://www.uae-embassy.org/uae-us-cooperation/economic/artificial-intelligence#:~:text=The%20UAE%27s%20AI%20Ambitions,large%20language%20model%20in%202023

India Briefing: India's Regulation of AI and Large Language Models, from
https://www.india-briefing.com/news/india-regulation-of-ai-and-large-language-models-31680.html/#:~:text=Presently%2C%20India%20lacks%20a%20dedicated%20regulation%20for%20AI%2C%20but%20instead,regulatory%20landscape%20for%20AI%20technology.

Barna: Hesitant & Hopeful: How Different Generations View Artificial Intelligence, from
https://www.barna.com/research/generations-ai/

European Parliament: Economic impacts of artificial intelligence (AI), from
https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf

Caltech: How Do AI and Machine Learning Differ?, from
https://scienceexchange.caltech.edu/topics/artificial-intelligence-research/artificial-intelligence-vs-machine-learning

McKinsey Digital: The economic potential of generative AI: The next productivity frontier, from
https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#introduction

Copilot: AI Threats and Concerns, from
https://copilot.microsoft.com/