

Committee: DISEC 1

Topic: The Question of the threat from Artificial Intelligence

Chair: Uliana Tokolova

School: Royal Russell School

Summary

The rapid advancement of Artificial Intelligence (AI) technologies presents significant opportunities and challenges for global society. While AI has the potential to drive economic growth, improve healthcare, and enhance security, it also poses substantial risks. These risks include the potential for misuse in areas such as cyber warfare, data breaches, job displacement, and the creation of autonomous weapons. This report explores the various types of AI, the threats they pose to nations, and the international response to these challenges.

Definition of Key Terms

Artificial Intelligence (AI) - The simulation of human intelligence in machines programmed to think like humans and mimic their actions.

Reactive Machines -The most basic form of AI that can perform specific tasks without past experience (e.g., IBM's Deep Blue).

Limited Memory - AI systems that can use past experiences to inform future decisions (e.g., self-driving cars).

Theory of Mind - A more advanced form of AI that can understand emotions, beliefs, and intentions, enabling social interactions.

Self-Awareness - The most advanced form of AI, which possesses a sense of self and consciousness

Background Information

The concept of AI dates back to the mid-20th century when computer scientists like Alan Turing and John McCarthy began exploring the idea of machines that could simulate human intelligence. Turing's seminal 1950 paper "Computing Machinery and Intelligence" and McCarthy's 1956 Dartmouth Conference laid the foundation for AI research. Early AI systems focused on rule-based approaches and problem-solving, such as chess-playing programs and basic natural language processing. The advent of machine learning in the 1980s and 1990s marked a significant shift, allowing AI systems to learn from data rather than relying solely on predefined rules. The development of deep learning, a subset of

machine learning, in the 2010s further accelerated AI's capabilities, enabling breakthroughs in image and speech recognition, autonomous vehicles, and more.

AI technologies have spread rapidly across the world, with significant contributions from both academia and industry. Countries like the United States, China, and members of the European Union have emerged as leaders in AI research and development. The U.S. boasts technological giants like Google, Microsoft, and IBM, which have invested heavily in AI. China, with its state-backed initiatives and companies like Baidu, Alibaba, and Tencent, has made significant strides in AI applications and research. This global proliferation of AI has led to an arms race of sorts, with nations vying for supremacy in AI capabilities. AI is seen not just as a tool for economic growth but also as a strategic asset in national security and defence. Countries with advanced AI capabilities can leverage them for military applications, intelligence gathering, and cybersecurity, potentially altering the balance of power on the international stage. The rise of AI has introduced new dynamics in international relations. Nations with superior AI technologies gain significant strategic advantages, leading to geopolitical tensions. For example, AI-driven surveillance and cyber capabilities can enhance a nation's intelligence operations, but they also raise concerns about privacy and ethical implications. Moreover, AI's role in autonomous weapons systems has sparked debates about the future of warfare. The potential for AI-powered drones and robots to make life-or-death decisions autonomously is a contentious issue, prompting calls for international regulations and treaties to govern the use of such technologies. As well as that, AI has undeniably contributed to numerous societal advancements and technological breakthroughs. In medicine, AI has facilitated more accurate diagnostics, personalized treatment plans, and drug discovery processes. In communication and information access, AI-powered systems enable real-time language translation, voice assistants, and efficient data processing. Additionally, AI has enhanced manufacturing efficiency, transportation systems, and environmental monitoring, demonstrating its potential to address complex global challenges.

Despite its potential benefits, AI poses several threats that must be addressed. One of the most pressing concerns is cybersecurity. AI can be weaponized for cyberattacks, enabling hackers to conduct sophisticated intrusions that are difficult to detect and counter. AI-driven attacks can target critical infrastructure, financial systems, and personal data, leading to widespread disruption and harm. Another significant threat is the potential for AI to facilitate identity theft and data breaches. AI systems that handle vast amounts of personal information can be exploited by malicious actors to steal identities, commit fraud, and compromise individual privacy. The ability of AI to generate convincing fake content (deep fakes) further exacerbates these risks, as it can be used to manipulate public opinion and deceive individuals. AI's rapid adoption and integration into everyday life have also raised apprehensions about its societal impact. Concerns about job displacement due to automation is rising and becoming a threat to many, which may enhance economic inequalities in governments.

UN Secretary-General António Guterres has emphasized the importance of globally coordinated AI governance to mitigate these risks and harness AI's potential for the common good. Guterres advocates for an inclusive approach that involves multiple stakeholders, including governments, private sector entities, and civil society, to develop international standards and ethical guidelines for AI. In his view, AI should be aligned with human rights and the Sustainable Development Goals (SDGs). This includes ensuring that AI systems are transparent, accountable, and designed to protect privacy and data security. Guterres also calls for investments in education and training to prepare the workforce for the AI-driven future and to prevent job displacement. AI represents a transformative force with both immense potential and significant challenges. Its continued development requires careful consideration of ethical,

regulatory, and societal implications to maximize benefits while minimizing risks. Through responsible governance and collaborative efforts across sectors and nations, AI can contribute to a more equitable, sustainable, and technologically advanced future for global society.

Major Countries and Organizations Involved

United States of America: The USA leads in AI development with major tech companies like Google, Microsoft, and IBM driving research and applications. The government is actively shaping AI policy and regulatory frameworks to ensure ethical deployment, addressing issues like algorithmic bias and data privacy alongside promoting innovation.

European Union: The EU proactively regulates AI to protect data privacy while fostering innovation. Initiatives like the AI Act propose comprehensive rules for AI systems, emphasizing transparency, accountability, and human-centric AI development. These efforts aim to build trust among citizens and support European businesses in adopting AI technologies responsibly.

China: China is a global AI leader, advancing technologies in healthcare, transportation, and finance. The government balances AI's transformative potential with stringent oversight to manage risks. Policies focus on data governance, cybersecurity, and ethical AI use to ensure national security and societal stability.

UN High-level Advisory Body on AI: Established for global AI governance, the UN Advisory Body fosters international cooperation and policy recommendations. It promotes ethical AI principles, facilitates inclusive dialogue among member states, and shares knowledge to shape responsible AI policies worldwide.

Timeline of Events

Date	Description
1950:	Alan Turing publishes "Computing Machinery and Intelligence," introducing the concept of machines that can simulate human intelligence.
1956:	John McCarthy coins the term "Artificial Intelligence" and organizes the Dartmouth Conference, marking the beginning of AI as a field of study.
1980:	Development of the first expert systems designed to emulate human decision-making.
2014:	Google DeepMind's AlphaGo defeats professional Go player Lee Sedol, highlighting AI's proficiency in complex strategic tasks.
2016:	Microsoft's AI chatbot, Tay, is taken offline after it begins tweeting offensive messages, highlighting risks of unsupervised AI.
2017:	China announces its "Next Generation Artificial Intelligence Development Plan," aiming to become a global leader in AI by 2030.

- 2020:** The COVID-19 pandemic accelerates AI's use in healthcare for diagnosis and vaccine development, raising concerns about AI surveillance.
- 2022:** The European Union proposes the Artificial Intelligence Act, creating a regulatory framework for AI applications.
- 2022:** The UN adopts a resolution on "The impact of rapid technological change on the achievement of the Sustainable Development Goals," stressing responsible AI management.
- 2023:** The UN Secretary-General convenes a High-level Advisory Body on AI to analyse and advance recommendations for international AI governance.

Relevant UN Treaties and Events

International Convention for the Suppression of the Financing of Terrorism (1999): This convention aims to prevent terrorist organizations from accessing funding, crucial for disrupting their operations and reducing their capacity to carry out attacks. It indirectly addresses concerns about AI misuse for terrorist activities by emphasizing global cooperation in monitoring and controlling financial flows.

United Nations Security Council Resolution 1373 (2001): Adopted after the September 11, 2001 attacks, Resolution 1373 mandates enhanced international cooperation to combat terrorism. It includes measures to prevent terrorist financing and improve border security, relevant for addressing AI's role in cyberterrorism. The resolution underscores the need for coordinated global efforts to prevent terrorists from using AI for malicious cyber activities.

United Nations Security Council Resolution 1540 (2004): Resolution 1540 aims to prevent non-state actors from acquiring weapons of mass destruction (WMDs). It requires UN member states to implement measures preventing the unauthorized acquisition and use of WMD-related materials. This resolution is pertinent to AI's role in autonomous weapons systems, emphasizing the need for international controls to mitigate risks associated with AI technology in WMD development.

Previous Attempts to solve the Issue

Global Network for Securing Electoral Integrity (2021): Established by USAID, this forum aims to address digital threats that compromise electoral integrity globally. Despite efforts to enhance collaboration among international partners, the network has faced challenges in achieving widespread participation and effectiveness across regions. If fully successful, the network would effectively coordinate cybersecurity measures, promote best practices in electoral security, and facilitate rapid response to emerging digital threats. Its success rate hinges on achieving broad engagement from electoral bodies, governments, and cybersecurity experts worldwide, ensuring timely and effective interventions to safeguard electoral processes.

European Democracy Action Plan (2020): Introduced by the European Commission, the plan seeks to enhance electoral processes within EU member states, particularly in combating disinformation and improving transparency. However, its implementation has encountered hurdles such as varying national regulations and political challenges within member states. A fully successful implementation would establish harmonized standards for electoral integrity, bolster resilience against disinformation campaigns, and ensure transparent electoral practices across Europe. Success would be measured by increased public trust in electoral outcomes and reduced impact of disinformation, supported by consistent adherence to EU guidelines and effective cross-border cooperation.

Foreign Influence Transparency Scheme (2018): Initiated in Australia, this scheme aims to increase transparency regarding foreign influence in Australian governmental processes and public discourse. Its effectiveness has been limited by complexities in defining and regulating foreign influence, as well as compliance challenges for affected entities. A successful scheme would enhance public awareness of foreign influence activities, strengthen regulatory frameworks to prevent undue influence, and uphold integrity in Australian democratic processes. Success metrics include improved transparency in political dealings, reduced instances of foreign interference, and heightened public confidence in governmental decision-making processes.

Possible Solutions

1. **Regulatory Bodies:** The possibility of establishing international regulatory bodies to oversee AI development and ensure adherence to ethical standards.
2. **Public Awareness Campaigns:** Can educate the public on AI's benefits and risks, promoting informed decision-making and reducing susceptibility to misinformation.
3. **Financial Regulation:** May implement stringent financial regulations to monitor and control funding for AI development, preventing misuse and ensuring transparency.
4. **Sanctions and Trade Restrictions:** Introduce sanctions and trade restrictions on entities misusing AI technologies, deterring unethical practices and encouraging responsible development.

Bibliography

Vinuesa, R., et al. (2020). "The Unchecked Advancement of AI: Impacts on Sustainable Development Goals." Available at <https://www.nature.com/articles/s41467-019-14108-y>

Müller, V. C. (2016). "Ethics of Artificial Intelligence and Robotics." Available at: <https://plato.stanford.edu/entries/ethics-ai/>

United Nations Secretary-General's Office. "High-level Advisory Body on A.I." Available at: https://www3.weforum.org/docs/WEF_Generative_AI_Governance_2024.pdf

Office of the United Nations High Commissioner for Human Rights (1994). "Human Rights and Elections: A Handbook on the Legal Technical and Human Rights Aspects of Elections."

USAID. "Global Network for Securing Electoral Integrity." Available at: <https://www.usaid.gov/democracy/global-network-securing-electoral-integrity>

European Commission. "European Democracy Action Plan." Available at:
<https://www.europarl.europa.eu/legislative-train/theme-a-new-push-for-european-democracy/file-european-democracy-action-plan>

Attorney-General's Department, Australian Government. "Foreign Influence Transparency Scheme." Available at: <https://www.dfat.gov.au/international-relations/Pages/foreign-influence-transparency-scheme#:~:text=Its%20purpose%20is%20to%20provide,Australia's%20government%20and%20political%20process.>