GENERAL ASSEMBLY

BCNMUN 2024

Identity and Conflict

Table of Contents

Topic: AI developments: Addressing the imbalance in technological innovation and ethical principles in preserving fundamental human rights.

Chair: Emma Sanchez Ballester

Position: Chair

Introduction

AI is our future. As our society rapidly propels towards a digital future, we must recognize the potential artificial intelligence has. Although the outcomes of our future remain uncertain, it is clear that we will face fundamental social shifts, primarily as a result of the presence of AI. According to BuiltIn, rising AI technologies are described as “machines capable of performing tasks that typically require human intelligence.” These innovations have revolutionized our society particularly laboral fields including healthcare, education, finance, manufacturing, etc. AI’s highly innovative features aim to improve human efficiency to those individuals exposed to the labor force and workplace; however, its accelerated development and progress has posed a threat to national security. In addition to putting at risk the existence of many crucial jobs.

Consequently, numerous disputes have spawned regarding the imbalance between technological advancements and the preservation of basic human rights. Furthermore, this led to the questioning of the ethics behind progress and innovation in technology. What long term risks are associated with creating technologies that surpass human intelligence? And if these technologies violate human rights, who should be held accountable for the behavior of these tools? The answers to these questions could clearly define the line between progress and justice. In a society that is purely relying on technology and innovation, AI could serve as a force of good for the future as it facilitates the exchange of knowledge. Nonetheless, individuals are drawn back by the potential harms that come with use of these technologies. Mainly the concerns are on the lack of privacy and transparency. It all comes down to responsibility. AI is a tool that if used correctly and with

reasonable parameters can guide individuals. As a society, it is important for us to find a balance between encouraging and embracing innovation while also promoting inclusivity and championing ethics.

Definition of Key Terms

Open AI

is an AI research and development company founded in December of 2015 named Open AI Headquartered in San Francisco, this company's mission is to develop “safe and beneficial” intelligence accessible globally. With almost 28 million dollars worth of revenue as of 2022, this company has established one of the most effective and efficient research organizations.

Modernization

is a process consisting of updating a concept or a system and making it function in a contemporary setting.

Artificial Intelligence

first began as a theory that now consists of the development of intelligent computer systems performing tasks that regularly require human intelligence.

ChatGPT

is a chatbot and virtual assistant that facilitates the exchange of knowledge making the research process easier for users. It was released on the 30th of November in 2022. However since its release it has adopted highly advanced features.

Stakeholders

are individuals or groups that share similar interests in finance particularly in the success of a business. Common examples of stakeholders include employees, customers, shareholders, suppliers, communities and governments.

Mainstream

is an idea, attitude or opinion that is publicly shared and accepted by other individuals.

Corporation

are large companies or numerous companies authorized as a single entity. Some examples include Amazon Inc, Apple Inc, Mcdonald's corp, etc.

Chatbot

is a computer program designed to interact in conversations with users mainly over the internet. In other words an advanced tool that simulates a conversation with users.

Virtual assistant

provides professional administrative services to users online. A virtual assistant can be a human individual or an AI powered machine, in both cases they are typically employed by humans to manage and assist the users

Background Information

Brief History

As many would argue, the beginning of AI dates back to 2020 when Chatgpt was launched. Nonetheless, it is partially true that AI's initial trajectory dates back to its birth in 1950. While it seems as if AI has been a recent development in technology, this is only due to the fact it became a mainstream use in the past couple of years. It is crucial for us to understand the groundworks they had laid and founded in the 1900s in order to grow and expand in the future. In the early 1900s, scientists went fanatical over the idea of developing an artificial human. After trials and errors many were able to create a simple version of what we would now call a robot. Many of these versions were steam powered and obtained fascinating features such as the ability to make facial expressions and being able to walk. In 1921, Czech playwright Karel Čapek published a play in which he introduced the idea of artificial people which we would later call robots. This was the first known use of artificial technology. From there, individuals took it into their own hands to create machines replicating the human brain, an example was the japanese professor Makoto Nishimura who built the first

japanese robot called Gakutensoku. The interest in AI spiked as scientists raced to develop highly efficient and innovative mechanisms. The concerns over AI have grown progressively, some of the first concerns date back to the 1960s when two pioneers of AI wrote about the potential dangers of AI operating independently. Furthermore in the year 1985 Ray Kurzweil publicly discussed the societal impacts and warning that the future AI holds for us in his book “The Age of Intelligent Machines.” By the 1900s the concerns of intelligence in machinery surpassing human intelligence began to attract attention in discussion and by the early 2000s these warnings were then publicly announced through mainstream discourse. Organizations with the aim to address the impacts of AI were created and individuals such as Elon Musk even warned that AI could be more dangerous than nuclear weapons. After the covid lockdown individuals worried they would be replaced by these machines which caused a crisis among workers. It is evident that the concerns of AI “taking over the world” aren't recent, as a matter of fact the origin of this dispute dates back to the creation of these machines.

The current issue

AI’s concerns over privacy violations alongside many other truths have been recently dismantled causing users to question the reliability of these tools. In the past month, AI privacy concerns have grown as these emerging technologies are being integrated into our day to day lives. These concerns worry education technology companies and content media companies such as Google, Buzzfeed, The New York Times and many more companies who have used these platforms in the past due to the growing concern on the replacement and displacement of human labor. The following examples are different forms AI has been used to violate data privacy rights.

Firstly, the unauthorized incorporation of user data has become a concern for many. When users first login, they are required to insert their personal data in the form of queries, this data then becomes part of the AI model training data set which is at risk of being exposed to other users. This is a pressing global issue due to the fact users' privacy rights are being violated as well as their data being extracted and used without their consent.

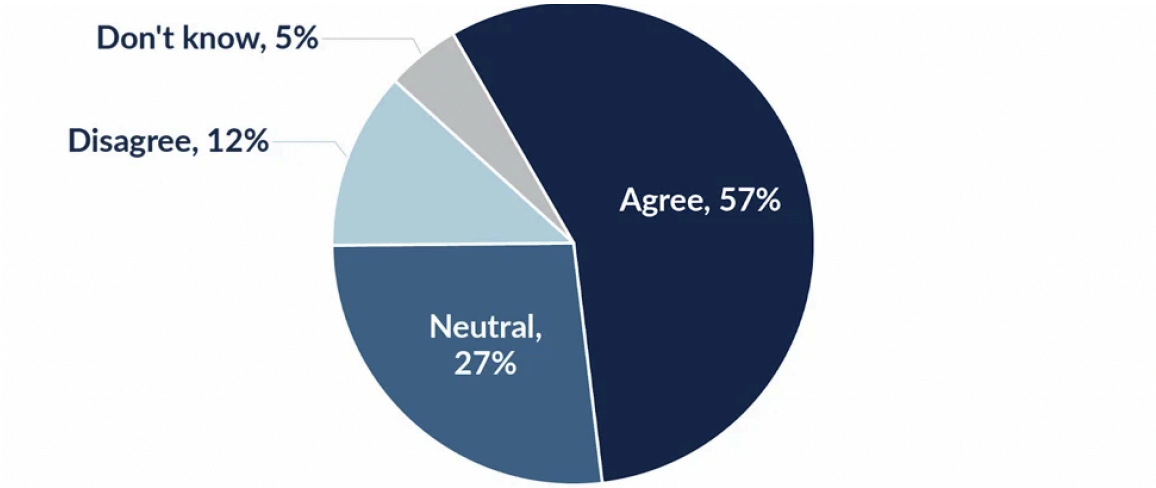
Moreover, as technology advances, biometric data security has been implemented rather than the traditional forms of identity verification. Public surveillance devices are now implementing AI to scan for biometric data in order to identify the individuals in a faster and more effective form.

The biometric data forms of identification include facial and voice recognitions, fingerprints. As convenient as these tools may seem, those that have AI ingrained in them can collect this data and in many cases be used for other illegal activities. Additionally, AI vendors aren't fully transparent with how, why and where they store a user's data.

These are just some examples of AI preserving human data through their algorithms. Authorities are aware of the harm this causes for users, however much more time and money is needed to develop a safer platform. AI developers do not have sufficient time to market nor the budget to improve the quality of these models. The limited focus on security and data protection facilitates the access to unauthorized users and gives them easy access to information and private data. As AI progresses, more regulatory bodies are recognizing the threats that it poses. Both governmental and industry specific institutions expect companies to establish a system that protects consumer privacy. However, it is complicated to take actions when many of the privacy laws and regulations do not directly address the AI model. In other words, the extent at which one's action can be considered a violation of human rights hasn't been officially declared nor publicly announced which is giving AI companies more freedom in their platforms.

ChatGPT: A threat or a source of empowerment?

A key event in the expansion of AI has been the release of ChatGPT. ChatGPT is a chatbot and a virtual assistant developed through the company Open AI. The chatbot was first launched on November 30, 2022. Throughout the past few years, the company has updated this platform by incorporating highly advanced and innovative tools. Despite the excitement among some users, the overall public perception of AI was transformed into fear due to people’s privacy becoming a concern.. Users question the ethics and intentions of these tech companies. Fear has now replaced what was originally considered thrill and enthusiasm. This can clearly be seen in the results of a survey completed by the International Association of Privacy Professions in 2023. The survey stated that 57 percent of consumers are now panicking as they view AI as a threat while 27 percent feel neutral and only 12 percent disagree with the harm that is posed with AI. Although AI could serve as a source of good to innovate and facilitate the exchange of knowledge, its issues concerning data privacy brings worry to users, causing them to distance themselves from these platforms and rely on already existing alternatives such as google.



(https://www.eweek.com/artificial-intelligence/ai-privacy-issues/)

An example of a previous privacy violation

A clear example of a privacy violation that directly involved AI technology was Open AI's first major outrage in March of 2023. Sources revealed 1.2% of the ChatGPT plus subscriber's chat history data, payments and other personal information were exposed to other unauthorized users during a period of time due to a bug in the system. AI automatically took ChatGPT down for a couple days to patch the bug that had been identified in an open source library. This bug allowed users to see other active users' chat history as well as other private data users had inserted when they first signed in. From then policies, frameworks and conventions were held to discuss the punishments, consequences and future actions that were needed in order to ensure security and safety to all the users in these platforms.

AI in the future

This continuous dispute has led many to question the future of these rapidly advancing technologies. As predicted by many scientists, AI is expected to grow increasingly.Its presence will not only revolutionize sectors such as healthcare, banking and transport, but it also has the potential to replace individuals in their jobs. If frameworks and laws are put in place, AI could add up to 4.4 trillion U.S dollars to the global economy. However, it is estimated if the current countries uphold innovation over their rights, by the year 2300 AI capabilities will rival humans. This constant tug between innovation and international principles must be urgently discussed as this will determine the future of our labor and day to day routines.

Major Countries and Organizations Involved

United States of America

The U.S is a country known for their continuous progress in creating a blossoming and rich economy, leading them to currently stand as an international powerhouse in the AI field. Their highly advanced tech companies, research institutions and startup culture all serve as examples of their global leadership in the technological environment. For instance, Silicon Valley in California is considered a global center of technological innovation, in which dozens of leading and top tier technologies, softwares and internet companies are found. Furthermore, major technological companies headquartered are located in the U.S some even in Silicon Valley including Google, Apple, Facebook, Microsoft and many more. Through the contribution of these companies as well as prestigious universities found in the U.S like Stanford, Massachusetts Institute of Technology (MIT), University of California, Berkeley and others have collaboratively and consistently shown outstanding talent in the field. Furthermore, the U.S government has publicly announced their support towards AI initiatives.

China

China, another country who highly prioritizes technological advancements,has also placed regulations on AI users. Since 2021, China has implemented a series of targeting regulations. The Chinese authorities established the first set of rules that imposed new obligations on companies to generate new rights to users as well as to offer protection to them. The Chinese party quickly followed up with a new regulation on “deep synthesis” which urged AI providers to generate content that did not violate individual rights or harm the nation's image. Furthermore, China then banned particular features such as ChatGPT in some districts as it spread misinformation and caused public disorder in the nation.

Thailand

As of 2022 Thailand has implemented various measures and ethical guidelines to promote the use of AI while upholding user rights. AI Thailand is a national program aimed to construct an

infrastructure for the development of artificial intelligence as it has the transformative potential to promote economic growth as well as increase Thailand's competitiveness. The result of this program was the establishment of Thailand's AI national strategy and action plan (2022-2027) which was approved by the prime minister's cabinet office on July 26th, 2022. Thailand strongly urges for the effective implementation of AI to further enhance the economy and quality of life within 2027. In a study conducted by IPSOS, a market research company, Thailand agreed with 73% that they trusted AI companies equally to other companies. In regards to all the questions on trust Thailand publicly showed trust towards AI companies. All in all Thailand is at interest in the expansion of AI particularly in their country as it obtains the transformative potential to help less developed nations like theirs to grow economically and socially.

South Korea

South Korea is a crucial proponent of AI, reflecting their commitment to innovation and emphasizing their future economic and technological vision. South Korea is considered to be a global leader in robotics which further helps to explain their 2019 aim to invest nearly 2 billion dollars by 2022 in the advancements of AI. The money invested would be to enhance the infrastructure of AI as well as invest in the implementation of these tools in companies such as Samsung and Hyundai. The implementation of AI across various sectors in the nation allows South Korea to position itself as an AI leader in comparison to other nations further capitalizing their dominance and differentiating them from other countries.

Syria

Syria alongside many other countries such as Chad, Yemen, Afghanistan, Eritrea, Eswatini, South Sudan, North Korea, Libya and Sudan are all on the list of countries banned from using chat gpt. The ban particularly in Syria is due to the ongoing conflict amongst other nations as well as affairs within their country regarding the strict government control over information and accessibility. For countries such as Syria this ban exemplifies the national concerns of the Syrian government over the potential misuse and propagation of misinformation. The country strongly emphasizes how the expansion of these tools can further lead to unrest and discord within a nation. Overall these countries seek unification amongst their citizens, which further causes them to view tools such as AI as threats in their system causing them to disagree and oppose its usage as a

source. In conclusion Syria and many other countries view these tools as threats posed to challenge their authority therefore they refrain from using and promoting it to their civilians.

UNESCO

Also known as the United Nations Educational, Scientific, Cultural Organization, UNESCO has 193 members and 11 associate members. This organization aims to contribute to global peace and security through education. The organization was established on November 16th, 1945 in London after two world wars with a clear objective to achieve lasting peace and justice among countries. UNESCO has previously spoken out on the numerous benefits AI has, accepting its potential in the labor force. Nonetheless, its accelerated progress has presented multiple risks that have outpaced frameworks and policies. Due to the multiple challenges and drawbacks, UNESCO seeks the creation of resolutions that find balance between innovation and the core principles of equity and fairness. Furthermore, they have additionally highlighted the implicit risk over the benefit of this rapid advancement. Overall, UNESCO is committed to guiding and helping member states that use AI responsibly.

ITU (International Telecommunication Union)

The International Telecommunication Union was established on May 17th, 1865 to direct the first international telegraph network. As of 1947, the ITU became the United Nations specialized agency for information and communication technologies. Regarding AI advancements, ITU initially presented itself as a neutral informative platform for governments and other institutions to inform themselves on the capabilities and functions of the newly dominating technologies. Moreover, as the risks exponentially increased among users, the ITU strongly began promoting the need of digital policies in order to guide member states to effectively and safely use this technology. The ITU has not only warned users about the dangers of AI, but has also recently announced their fundamental contribution to the UN secretary general's new high level advisory body on AI. The report calls for a proposal to strengthen the alignment between international principles and AI's development. In addition, the report urges the strengthening of international governance of AI by prioritizing scannings for potential risks, international cooperation and computing capacity.

Timeline of Events

Date	Description of event
1942	**Isaac Asimov publicly releases the “Three Laws of Robotics”** In the 1940s, Isaac Asimov, a science fiction writer published “I Robot'' a collection of short stories. He was known as one of the first writers to discuss machine intelligence in his stories. In one of his most famous writings, he described the “Three Laws of Robotics.” The first law consists of a nonviolent approach, robots may not injure or harm the human being. The second law explores the importance of obedience, or in other words, how the robot must obey the orders it is given by the human. Lastly the third law describes a robot's protection, stating that the robot must protect its existence. To this day, these three laws are continuously revisited by scientists to ensure their platforms and inventions are followed.
1950	**The creation of the “Imitation Game” by Alan Turing** In the 1950s, Alan Turing, an English mathematician and computer scientist proposed the following: the imitation game. In the early 1950s Alan Turing developed a research paper that was then transformed into a framework on machine intelligence. His paper revolved around the following question: “Can machines think?” This particular theoretical question allowed Turing to establish an exercise in which a human interrogator is asked to differentiate a machine and human being text. This exercise was known as the Imitation Game and it provided a benchmark that would further shape the beginning of the philosophy of AI.
1956	**Dartmouth conference on AI principles** In 1956 a small group of scientists came together for the Dartmouth Summer Research Project on Artificial Intelligence, which marked the birth of AI as a field. The first meeting was

organized by John McCarhty who at the time was a mathematics professor in college. The term “artificial intelligence” was adopted in that gathering. The scientists discussed the potential AI had in learning and researching. Through this, they created a general framework of research areas where technology intelligence could potentially be of use.

1970-1980

AI suffers it first winter

After months of innovation and success, AI experienced its first winter in 1984, in simpler terms, suffering reduced interest and funding. The AI winter is referred to as the period of unrealistic expectations and promises generated by its creators. Authorities began cutting off funding for AI research, as they were seeing no effective results. The capabilities of AI programs and systems remain limited primarily due to the lack of computing power during that time. DARPA the Defense Advanced Research Projects Agency was an organization that funded the majority of the research in the 60s. The organization demanded timelines and clear descriptions on the products to ensure progress. This tragic event lasted through the 70s and 80s. Due to this period, AI was viewed as a fraud and a disappointment to society.

November 30th, 2022

ChatGPT is launched and available to the public

In 2022, Open AI released ChatGPT to the public, an AI chatbot that facilitates user’s research experiences, allowing them to interact with the chatbot. However, the launch of this platform brought up various concerns of safety and privacy. ChatGPT models are trained with large datasets that contain users' private information such as email addresses, bank records, personal names and other online records. Users began pleading for organizations and the authorities to place harsher security measures to prevent illegal behavior including data leaks and improper access to information.

May 2023

Users fight for their privacy rights through revolts and protests

	The release of AI led to revolts across the globe, ranging from London, San Francisco and New York. Activists are not only roaming the streets with posters, but many have gone on strike. These workers are seeking a more secure and safe platform to use in exchange for AI. In addition, many activists have clustered at the entrance of Open AI's headquarters in San Francisco as a form of protest. Days after these revolts, Open AI executives stated they were planning on working with the U.S Defense Department on open strengthening the cyber security software.
21st of March, 2024	**United Nations adopts a United States led draft resolution** After visibly seeing the frustration among users, the United Nations held a convention in which a U.S led draft resolution was passed urging for a global effort to establish a balance between innovation and equity. The resolution titled "Seizing the opportunity of safe, secure and trustworthy artificial intelligence systems for sustainable development" was co-sponsored or supported by more than 120 member states, this marked the progress towards reaching the 17 sustainable development goals.
17th of May, 2024	**Council of Europe Framework Convention on Artificial Intelligence and Human Rights** The passing of this resolution further inspired institutions such as the Council of Europe to create frameworks on the principals that are needed in order to utilize ChatGPT responsibly in the future.

Relevant UN Treaties and Resolutions

As previously stated, the beginning of Artificial Intelligence dates back to the 1900s, however, it wasn't till recently that these innovative tools caught the public's eye. After releasing ChatGPT and other AI features, countries have not only identified the risks and challenges that come with these tools, but also led many to question the ethics behind these tools. For those reasons, organizations

such as UNESCO and the UN are working on establishing reports, landmark resolutions and agreements in order to ensure the protection of their citizens. The following documents are all examples of relevant UN treaties and resolutions that could potentially serve as guiding principles.

United Nations AI Resolution: A Significant Global Policy Effort To Harness The Technology For Sustainable Development on 21st March, 2024

On the 21st of March, 2024, the General Assembly passed its first revolutionary resolution on artificial intelligence, supported by 120 member states. The resolution calls upon member states to establish “safe, secure and trustworthy AI systems” (Geneva Graduate Institute). They advocate for the responsible development of AI while protecting human core principles and international laws. The adoption of the US-led resolution is a global effort to promote an ethical and responsible use of AI. In the assembly they further recognized the division caused by rapid innovation particularly in lower income and developing nations. Due to this division, the resolution strongly urged member states to cooperate in order to close the digital gap among these developing nations. By closing this gap and supporting developing countries they aimed to ensure inclusive and equitable access to these innovative advancements in technology globally. All in all, this first global resolution was a step towards an AI system that upholds innovation while embracing the UNs core values and protecting human rights.

Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law on 17th May, 2024

The Council of Europe has adopted the first international legally binding treaty on the ethics behind the use of AI. The Council of Europe Framework Convention on AI and human rights democracy and the rule of law was passed during the annual ministerial meeting of the council Europe's committee of ministers in Strasbourg. This convention brings together the Minister for foreign affairs of the 46 council of European member states. The treaty that is also open to non-European countries released a legal framework that addresses the AI cycle from its creation to its function. Furthermore, it analyzes the possible drawbacks and challenges posed to its function. After two years of collaborative work among an intergovernmental body, the Committee on Artificial Intelligence (CAI), 46 council of European member states, the European Union and 11 non member states came together to participate in drafting the treaty and others served as representatives of academia, private sectors and observers. Furthermore a follow-up mechanism known as a

Conference of the Parties will be held in order to measure the effectiveness of the treaty. Parties in the convention will not be obligated to apply the treaties provisions to activities, however, will be required to ensure that these activities protect international laws and rights.

Previous Attempts to solve the Issue

Many individuals fear the power and dominance AI holds in our societies. That fear is fueling countries to take steps of prevention. Measures will range from censorship to punishment for utilizing these tools. One clear example of an attempt to solve this issue is censorship. In countries such as China, Russia, Cuba, Iran, Saudi Arabia, North Korea, Italy, etc. features such as ChatGPT have been banned for multiple concerns primarily due to data privacy and the misuse of information. Authorities in these countries worry that these platforms could influence public opinion as well as spread misinformation. By banning these platforms, authorities reduce the risk of public disorder and maintain their citizens pure from fake news. The outcomes of these previous attempts are multifaceted while there are benefits that emerge such as enhanced privacy protection and public unification and organization. It also reveals drawbacks of these possible solutions including stifled innovation among the public and possible economic consequences as they are redirecting the focus towards other alternatives that are less dependent on AI. In addition, conventions and meetings have been held to discuss the use and functionality of these platforms. The most successful way of controlling and reducing the risk for users has been the censoring of the platforms whether that is in restricted areas such as school or across a whole nation. On the other hand, it is crucial to not mistake the countries that are choosing to censor and ban these tools with those countries that are prohibited from accessing the platforms. Countries such as Eritrea, Eswatini, South Sudan, Syria, Chad, Yemen, Afghanistan, Libya and Sudan are banned from using Chatgpt and other features due to their digital infrastructure, policy limitations and cybersecurity issues. As of this time, these countries don't have access to the platforms.

Possible Solutions

The examples below are solutions suggested by politicians, scientists and other individuals that could potentially guide countries in the right direction in finding a balance between innovation

and human rights. The efficiency of these solutions will vary depending on the country. These solutions could be used as guiding stepping stones for countries.

Education

Education has become a fundamental pillar in our modern lives. Over the years, it has become society's most viable solution to combat and mitigate many issues world-wide. Education doesn't just consist of the capacity of learning which enriches us with knowledge, it is also being able to critically analyze and question the systems around us in order to grow and shape our individual character. One's initial perception towards an issue may adjust as they gain information and knowledge on the concept. Through education and awareness campaigns, corporations and state actors can further educate younger generations regarding the norms and expectations associated with these emerging tools. This rising generation was born into this fast-paced digital society we live in today and they will encounter many more severe changes due to the development of these tools in the future. Informative sessions in both public and private schools are clear examples of how education can be used as a transformative tool in our modern day society. Experts in this field will teach these young adults the key principles that should be considered when utilizing these tools.

Punishment as a consequence

In many successful cases punishment has served as a threat to reduce unpleasant or illegal behavior. Punishment in psychology is referred to as a consequence whose objective is to reduce undesirable behavior in an individual. Inside the brain, the punishment appears to “potentiate neural response to errors and decrease neural response to rewards' ' (ScienceDirect) this increases the risk for anxiety furthermore causing an individual to back out from following through with their behavior. The punishment proposed should not include any sort of physical or verbal violence towards individuals. Punishment or consequences for those violating human rights or international laws serves as a non-violent form of advocating for fairness and responsibility when using these tools. Punishment could range from fines: a sum of money an individual must pay as a punishment, community service:the offender must serve and do some unpaid work that will benefit the community, confiscation: involuntary submission of one's belongings or money as a punishment for their actions. These actions will receive harsher punishments depending on the severity of the violation, however who exactly would be punished for violating user privacy? Should the creators of

these platforms be punished? Should the users who use these tools irresponsibly be fined and to what extent should they be punished? These are all questions that are yet to be resolved, and its answers vary depending on the country's stance. Countries such as Thailand, Saudi Arabia, China, Indonesia, Myanmar and the UAE exhibit varying levels of severity in their legal punishments which often reflect cultural and political influences. Their punishments are used to enforce order and can include long-prison sentences, corporal punishments and even death penalty. Meanwhile countries like Norway, Sweden, FInland, Netherlands and Australia prioritize rehabilitation and more humane treatments. Their punishments can range from obligatory community service to rehabilitation programs. Each system approach obtains its advantages and flaws, the efficiency varies depending on a country's conditions.

Legislative actions to strengthen consumer rights

Furthermore, consumer rights should be urgently prioritized in these platforms. According to Study.com consumer rights are defined as “the user's right to safety, the right to be informed, the right to choose and the right to be heard by others.” AI has the power and potential to make erroneous and biased mistakes that could harm a consumers rights. AI further leads to the creation of dark patterns, also known as the manipulation of consumers, into guiding them to make decisions they wouldn't have made otherwise. This makes it difficult for users to realize they are being manipulated. Therefore, harsher legislation should be implemented in order to ensure a secure and safe system that respects consumer rights. Legislation could include fines for the violation of consumer rights, censorship in particular sectors such as healthcare and the construction of legal frameworks to explore who should be held accountable for the harm caused.

Establish detectors and scanners

As many frameworks state they want to ensure the transparency and protection of rights in these platforms, in order to achieve this goal, machines such as detectors and scanners as well as yearly check ups are crucial to identify its development. Investments in scanners and detectors are fundamental as they reassure citizens that the systems are operating correctly and securely. Measuring and evaluating the function of these tools will take place in the form of yearly standard and benchmarks revisions. In order for this to occur, countries must come together to decide what is ethically right and what should be called out for or punished. In other words, what generic data

should be collected and what could be left out. These are all norms and expectations that should be discussed before taking further action.

Investing in improving already existing alternatives

As previously mentioned, censorship is an alternative many countries are pushing towards, one that is barely discussed and would be improving other existing platforms. AI originally came out as a machine that collected its information primarily from Google, from there it would reduce and simplify it when a user asked a question. By investing money to further research and improve already existing platforms such as google, individuals wouldn't find the need to use AI in their day to day lives. Creating an easier navigating web and synthesizing information, all while maintaining one's data private, is the alternative many individuals are looking for. Rather than censoring these tools, we would be replacing them with a more secure platform to research on. The features on AI could be replicated for google's platform, allowing users to feel safer and more secure when working.

Censorship

As stated earlier, many countries have decided to partially ban AI, prohibiting it in particular areas such as schools, healthcare institutions, media publishing companies such as Gannet and even financial services including JPMorgan Chase and the Bank of America. Meanwhile other countries have fully restricted its access to citizens out of fear. Censorship can be seen as an effective solution in this case as it protects individuals from misinformation and reduces their risk of being violated by non-transparent or unsecure behavior. Censorship has served as a force of good for countries such as China, as it strengthens their national security by keeping information that could potentially change the public's opinion far from the citizens. Censorship reduces the risk of others dismantling the hidden truth, further maintaining social order and protecting the authorities. This solution seems ideal for hegemonies as they argue it protects individuals from harmful content. Censorship not only limits individuals ability to express but also manipulates public perception by restricting access to opposing viewpoints. This radical solution tends to be used by authoritarian regimes to maintain control and prevent councils disorder and criticism of the government.

Conclusion

A danger to our societies or an opportunity to expand? The rising AI technologies have provoked reactions of all types. Its groundbreaking features amaze users, however its privacy concerns worry others. This constant battle to recognize and champion innovation while still preserving and upholding international principles has led to disputes, revolts and protests among individuals. The challenging discussion between these two extremes has become crucial as we advance towards a purely technological and digital based future. As society advances, values and norms shift, meaning we must let go of core principles in order to accept and ingrain these technologies in our day-to-day lives, or should these features adjust to match our principles? Regardless of the solution, countries must collaboratively come together to respectfully engage in discussions that will set the norms and expectations. Countries should seek a vision of unity, peace and justice while solving this urgent issue.

Bibliography

"The 13 Most Advanced Countries in Artificial Intelligence." *Rejolut*, https://rejolut.com/blog/13-top-ai-countries/. Accessed 30 July 2024.

"AI: 15 key moments in the story of artificial intelligence." *BBC*, https://www.bbc.co.uk/teach/articles/zh77cqt. Accessed 30 July 2024.

"AI Winter: The Highs and Lows of Artificial Intelligence." *History of Data Science*, https://www.historyofdatascience.com/ai-winter-the-highs-and-lows-of-artificial-intelligence/. Accessed 30 July 2024.

Artigas, Carme, and James Manyika. "AI Advisory Body | United Nations." *the United Nations*, https://www.un.org/en/ai-advisory-body. Accessed 30 July 2024.

Felipe, Loey. "General Assembly adopts landmark resolution on artificial intelligence." *UN News*, 21 March 2024, https://news.un.org/en/story/2024/03/1147831. Accessed 30 July 2024.

"From Chatbots to ChatGPT: Navigating consumer rights in an AI-driven world." *Shoosmiths*, 15 February 2024, https://www.shoosmiths.com/insights/articles/from-chatbots-to-chatgpt-navigating-consumer-rights-in-an-ai-driven-world. Accessed 30 July 2024.

Gordon, Anna. "Why Protesters Are Demanding Pause on AI Development." *Time*, 13 May 2024, https://time.com/6977680/ai-protests-international/. Accessed 30 July 2024.

"Guiding AI - 2023 review." *ITU*, https://www.itu.int/osg/year-in-review-2023/driving-digital-impact/guiding-artificial-intelligence/. Accessed 30 July 2024.

"The History of AI: A Timeline of Artificial Intelligence." *Coursera*, 16 May 2024, https://www.coursera.org/articles/history-of-ai. Accessed 30 July 2024.

Hiter, Shelby. "AI and Privacy Issues: Challenges, Solutions, and Best Practices." *eWEEK*, 20 June 2024, https://www.eweek.com/artificial-intelligence/ai-privacy-issues/. Accessed 30 July 2024.

Holmes, Wayne. "Artificial intelligence in education." *UNESCO*, https://www.unesco.org/en/digital-education/artificial-intelligence. Accessed 30 July 2024.

Huet, Ellen, et al. "AI Protest at OpenAI HQ in San Francisco Focuses on Military Work." *Bloomberg*, 13 February 2024, https://www.bloomberg.com/news/newsletters/2024-02-13/ai-protest-at-openai-hq-in-san-francisco-focuses-on-military-work. Accessed 30 July 2024.

Khelif, Mouloud. "United Nations AI Resolution: a Significant Global Policy Effort to Harness the Technology for Sustainable Development | IHEID FC." *Graduate Institute - Executive Education*, 6 May 2024, https://executive.graduateinstitute.ch/communications/news/united-nations-ai-resolution-significant-global-policy-effort-harness. Accessed 30 July 2024.

O'BRIEN, MATT. “Scientists warn of AI dangers but don't agree on solutions.” *AP News*, 3 May 2023, https://apnews.com/article/ai-godfather-google-geoffrey-hinton-yoshua-bengio-chatgpt-5f7dc295a576833dfc3378071b5716f2. Accessed 30 July 2024.

“Our Mission | World Economic Forum.” *The World Economic Forum*, https://www.weforum.org/about/world-economic-forum/. Accessed 30 July 2024.

Sheehan, Matt. “Tracing the Roots of China's AI Regulations.” *Carnegie Endowment for International Peace*, 27 February 2024, https://carnegieendowment.org/research/2024/02/tracing-the-roots-of-chinas-ai-regulations?lang=en. Accessed 30 July 2024.

Sullivan, Mark. “The 10 most important moments in AI (so far).” *Fast Company*, 16 September 2019, https://www.fastcompany.com/90402503/the-10-most-important-moments-in-ai-so-far. Accessed 30 July 2024.

“The Timeline of Artificial Intelligence - From the 1940s to the 2020s.” *Verloop.io*, https://verloop.io/blog/the-timeline-of-artificial-intelligence-from-the-1940s/. Accessed 30 July 2024.

Valcarce, Maxime. “193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence.” *UN News*, 25 November 2021, https://news.un.org/en/story/2021/11/1106612. Accessed 30 July 2024.

“What is the future of AI (Artificial Intelligence)?” *McKinsey*, 30 April 2024, https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-ai. Accessed 30 July 2024.