

UNESCO

Artificial Intelligence Ethics

Delegate Background Guide



```
var evts = 'contextmenu dblclick  
var logHuman = function() {  
if (window.wfLogHuman) { return; }  
window.wfLogHuman = true;  
wfscr.type = 'text/javascript';  
wfscr.async = true;  
wfscr.src = url + '&r=' + Math.random();  
(document.getElementsByTagName('script'))  
(var i = 0; i < evts.length; i++) {  
  addEvent(evts[i], logHuman);  
}  
for (var i = 0; i < evts.length; i++) {  
  addEvent(evts[i], logHuman);  
}  
afe.com/?wordfence_lh=1&hid=A957C9DCB2B5F093A87A2721366313E8'");</script> <noscript> <style> .wsc-cs {<br>[r].1-1*new Date();<br>cc-E.W. parentNode: </script>
```

A Letter from Your Chairs...

Dear Delegates,

We're so excited that you chose to be a delegate in the Artificial Intelligence Ethics committee! Our committee is going to have one chair and one co-chair. Here's a bit about us:

My name is Raphael, and I am a junior at Commonwealth. I have been participating in COMMUN for a few years now, so I am so excited to chair this committee. I have always been interested in AI and its ethical consequences, and I can't wait to discuss it with you.

My name is Thomas, and I am a Freshman at Commonwealth. This will be my first year participating in COMMUN. I started participating in Model UN in seventh grade, and I went to this conference twice, so being able to host a committee is really exciting!

As you know, at this committee, you will be debating issues surrounding the ethical concerns of Artificial Intelligence. In recent years, this has been an increasingly prevalent issue as interest and innovation in AI continues to grow, and improved technology worldwide makes it easier to access AI and create more advanced systems. Currently, AI is everywhere. Whether you are unlocking your phone with face ID, scrolling through your social media feed, using spam filters, and even doing a simple google search, you are using AI. However, the ethics of its capabilities are debatable, and we will be discussing whether or not Artificial Intelligence should be used in our everyday lives.

We know this topic has a lot to offer, and can't wait to hear what everyone has to say about it. Since COMMUN is a learning conference, please feel free to reach out with any questions, comments or concerns. We look forward to seeing you all in the spring, and good luck!

Raphael Yamamoto

Chair

ryamamoto@commschool.org

Thomas Du

Co-Chair

tdu@commschool.org

Table of Contents

Problem Statement	4
History	5
Timeline	6
Discussion	8
Bloc Positions	13
Vocabulary	18
Questions to Consider	18
Works Cited	19

Problem Statement

“We are faced with a crucial question: what kind of society do we want for tomorrow? The AI revolution opens up exciting new prospects, but the anthropological and social upheaval it brings in its wake warrants careful consideration.”

— Audrey Azoulay, Director-General of UNESCO

Artificial Intelligence is an extremely powerful technology that is advancing rapidly. It is already being applied in countless fields, from driverless cars to breast cancer recognition in hospitals. With such an impactful and rapidly growing industry, it is crucial that the UN decide on a proper policy to prevent misuse and harm.

This committee’s goal will be to discuss the problems and opportunities modern AI presents, as well as come up with solutions going forward. The global conversation surrounding AI has been ripe with AI’s triumphs and failures as it is integrated into global technology.

So far, the majority of documents written to introduce ethical standards for AI have been produced by private companies and local governments. While they are a step towards considering ethics, these documents aren’t optimal: they can contain arguments written to benefit only said companies, or might not take into consideration the rest of the world. As such, it is the UN’s responsibility to standardize the ethics of AI on a global scale, to ensure that the potential misuse of AI, whether by malevolent forces or unintentionally, are mitigated and that the productive uses of AI are allowed to flourish.

AI is already in our lives, directing our choices, often in ways that can be harmful. There are some legislative vacuums around the industry that need to be filled quickly. The first step is to agree on exactly which values need to be enshrined, and which rules need to be enforced. Many frameworks and guidelines exist but are implemented unevenly, and none are truly global. AI is global, which is why we need a global instrument to regulate it.

UNESCO will be a full and active participant in this global conversation. Our organization has many years of experience in the ethics of science and technology. Our advisory bodies have already produced numerous reports and declarations on similar topics, such as the Report of the World Commission on the Ethics of Scientific

Knowledge and Technology on Robotics Ethics in 2017. The advisory bodies also have experience in developing normative instruments, including the Universal Declaration on the Human Genome and Human Rights in 1997 and the Universal Declaration on Bioethics and Human Rights in 2005.

History

While the history of Artificial Intelligence is extensive, the first major attempt to address AI Ethics wasn't until 2017, when the Future of Life Institute held the Asilomar Conference on Beneficial AI to discuss AI Ethics and how to bring about beneficial AI while avoiding existential risk from artificial general intelligence. During that conference, twenty-three principles were created to help guide AI's continued development to offer opportunities to help and empower people in the decades and centuries ahead.

Among these principles were: Safety, Failure Transparency, Judicial Transparency, Responsibility, Value Alignment, Human Values, Personal Privacy, Liberty and Privacy, Shared Benefit, Shared Prosperity, Human Control, and AI Arms Race. However, because the conference does not have much authority, individual organizations decide the degree to which principles such as these are followed. This does not stop the people concerned, as the work of the Asilomar Conference has been continued by various other organizations.

In 2018, the first AAAI/ACM Conference on AI, Ethics, and Society was held in Louisiana. The goal of this event was to address the ethical concerns and challenges regarding issues such as privacy, safety and security, surveillance, inequality, data handling and bias, personal agency, power relations, effective modes of regulation, accountability, sanctions, and workforce displacement. The organizers believe that only a multidisciplinary effort can find the best ways to address these concerns and that we must collaborate with experts in many fields to build a community of shared concepts and concerns. This conference has been held yearly ever since, and the submissions and attendance have only increased.

In 2020, UNESCO Director-General Audrey Azoulay appointed twenty-four of the world's leading experts on the social, economic, and cultural challenges of artificial intelligence to draft internationally applicable recommendations on ethical issues raised by the development and use of AI. Addressing over six hundred submissions and fifty thousand suggestions, the experts will examine the ethical concerns. This

recommendation will be presented for adoption in 2021.

The outbreak of Covid-19 has further exposed AI vulnerabilities, the digital divide, and the disparity of opportunities. Developing countries are likely to experience more modest gains due to much lower rates of adoption of AI technologies, lack of ownership of such technologies, and often not being an active part in the design and development processes but rather recipients of technology. Thus, the UNESCO Recommendation is of utter importance to address such concerns. Ultimately, the goal is to ensure that AI is developed and used in line with the principles of inclusiveness, diversity, and transparency.

As AI grows more pervasive in our lives, and its impact on society grows, there is an increasing number of innovators. With the innovations comes the concerns about ethics. The number of organizations and experts concerned about ethics has only increased, but conferences like Asilomar do not have the authority to control the current and future innovators. However, the United Nations has the influence to change this.

Timeline

1950: Alan Turing proposes the Turing Test as a measure of machine intelligence

1951: First working AI programs were written by Christopher Strachey and Dietrich Prinz

1956: Dartmouth College summer AI conference is organized and John McCarthy coins the term “artificial intelligence” for the conference.

1956: Marvin Minsky’s views on the top-down approach, pre-programming a computer with the rules that govern human behavior, dominated

1973: The AI winter: Millions of dollars had been spent but there was little to show for it. There was much criticism, especially from leading mathematician Professor Sir James Lighthill.

1980-ish: John Hopfield and David Rumelhart popularized “deep learning” techniques which allowed computers to learn using experience.

1981: AI’s commercial value to big businesses started to be realized. Instead of trying to create a general intelligence, these systems were coded to focus on much narrower

tasks, such as helping configure new orders for new computer systems.

1982-1990: Japan invests \$400 million dollars into implementing logic programming and improving artificial intelligence as a part of FGCP. British funding is revitalized as a response to this.

1990: After years of trial and failure, AI scientist Rodney Brooks helped drive a revival of the bottom-up approach to AI

1995: “No Hands Across America”: A semi-autonomous car drove coast-to-coast across the United States with computer-controlled steering for 2,797 miles (4,501 km) of the 2,849 miles (4,585 km). The throttle and brakes were controlled by a human driver.

1997: The Deep Blue chess machine made by IBM defeats world chess champion, Gary Kasparov.

2000: Cynthia Breazeal at MIT publishes her dissertation on Sociable Machines, describing Kismet, a robot with a face that expresses emotions.

2002: iRobot created the first commercially successful robot for the home: an autonomous vacuum cleaner named Roomba

2005: The United States began to invest in autonomous robots such as BigDog from Boston Dynamics for military purposes.

2008: Cynthia Mason at Stanford presents her idea on Artificial Compassionate Intelligence, in her paper on “Giving Robots Compassion.”

2009: Google builds an autonomous car.

2011: IBM Watson defeats television game show Jeopardy! champions.

2011-2014: Apple’s Siri (2011), Google’s Google Now (2012), and Microsoft’s Cortana (2014) are smartphone apps that use Natural Language Processing to answer questions, make recommendations, and perform actions.

2014: First chatbot, Eugene Goostman, passed the Turing Test

2015: An open letter to ban the development and use of autonomous weapons signed by Hawking, Musk, Wozniak, and 3,000 researchers in AI and robotics.

2014: Amazon launches Alexa, an intelligent virtual assistant with a voice interface.

2017: Google's AI AlphaGo beats world champion Ke Jie in the complex board game of Go.

2018: Alibaba language processing AI outscores top humans at a Stanford University reading and comprehension test, scoring 82.44 against 82.304 on a set of 100,000 questions.

2020: UNESCO appoints an ad hoc committee to draft a recommendation on the ethics of AI to be presented at the November 2021 session.

Discussion

An introduction to modern AI techniques

Artificial Intelligence (AI) is traditionally defined as intelligent behavior exhibited by machines. More specifically, AI programs are usually designed to complete tasks either reserved for or exceeding the capabilities of humans. Standards of intelligence often reflect the natural intelligence shown by humans and animals: skills such as reasoning, problem-solving, perception, and learning.

While intelligence and consciousness, in humans or animals, are still not completely understood, multiple techniques have been developed to tackle these problems in computers. Some of the original approaches to AI sought to model the human thought process. One approach used predefined conditions to simulate intelligence. For example, an algorithm designed to answer whether or not one should go running might consider factors such as the weather and temperature, and then traverse through its options—almost like navigating a flowchart. Models such as these are called symbolic, as each of their actions is determined by considerations that have an assigned meaning.

Other techniques use a statistical approach. These models attempt to arrive at the desired behavior by analyzing data and attempting to find patterns. Common techniques include linear regression, logistic regression, and k-means clustering. This might be useful, for example, in creating an artificial intelligence model that can predict where different strains of Sars-Cov-2 originated or predicting someone's height from their age.

Another widely popular technique is the use of artificial neural networks (ANNs, or

NNs) that model the real neural networks found in human and animal brains. As the name suggests, they simulate a network of neurons, with each neuron governed by simple rules. When put together, these neurons can exhibit extremely complex behaviors. Take, for example, any picture from <https://thispersondoesnotexist.com/>, or, to some extent, Google Translate. Statistical and ANN techniques are called sub-symbolic, as they don't deal with explicit meaning when making decisions. Rather, they simply learn and reproduce patterns in a programmatic fashion that is largely indecipherable to humans. This can raise ethical concerns as well, as it removes a large amount of transparency from these models.

Central to many AI techniques is the concept of machine learning. Many sub-symbolic systems that aren't told explicitly which symbols matter to them rely on learning patterns from data they are given. These patterns could be something like a driverless car recognizing a pedestrian, or a system learning to identify when the stock market is about to crash. As the complexity of patterns grows, the amount of data required to teach them also grows. This can lead to ethical concerns of privacy, data collection, and data selling, as companies may collect large amounts of data to train their algorithms. Machine learning is probably the most used technique in AI at the moment.

Bias

A major concern surrounding the ethics of AI is bias. Often, we think of computers and algorithms as potentially the most impartial and unbiased decision-makers. But when an algorithm is trained on skewed data, biases can instead be propagated. Bias arises when an AI system exhibits behavior that reinforces larger biases found in society. Most often, this happens unintentionally: In the case of deep learning, an AI may "learn" biases from inherently biased training data. Therefore, it is useful to note that an AI can't create its own bias, any bias in an AI is a reflection of society's biases. Therefore, the responsibility of mitigating bias is placed on the AI's creator. Being such a powerful tool, biases in AI systems have the potential to greatly reinforce harmful attitudes, and it is imperative that they be resisted.

An example of bias in AI was Amazon's (unused) recruiting engine. The engine was designed to simplify and automate the hiring process; an algorithm would read through the resumes and rate each one on a five-star scale. However, it was discovered that the AI was consistently rating male candidates higher than female candidates. Reportedly, the algorithm penalized resumes containing the words "women" and mentions of candidates attending female colleges. The bias arose because the system

was trained on Amazon's hiring pattern from previous years, which itself was skewed towards men. This way, the AI learned biases from the outside world. In other situations with bias, AI engineers might modify their training data to better represent the desired outcomes.

For a more deliberate example, consider the voice assistants found in almost every modern device: Siri, Alexa, Cortana, Google Assistant, among others. By default, these assistants are "female," potentially reinforcing gender biases of servility and submissiveness.

Transparency

Another ethical concern with AI is visibility and transparency. Machine Learning (ML), as opposed to "typical" algorithms, isn't hand-coded. Instead, AI developers train an AI on data until it exhibits the desired behavior. Internally, this entails adjusting up to potentially billions of number variables in an enormous function. This means that the behavior of an ML model isn't easily explainable; the computer learns patterns without assigning specific meaning (see the statistical approaches listed above).

A defining factor that influences the behavior of ML models is the data it's trained on, as the training data represents the desired behavior. With such an integral part of the model—and often a potential source for mistakes—having access to training data is extremely useful for cultivating transparency. However, training data is almost never made available to the public.

Developed on the scale it is currently, AI requires vast amounts of training data. For companies, this creates competition. For example, if one company created data for their own product, they wouldn't want to undermine its value by letting other companies use it. As with any prevalent technology, it is necessary for consumers to be able to understand and investigate AI to be able to use it safely.

Transparency is not only a concern regarding the algorithms themselves but also regarding developers and companies. To increase the visibility of the market and allow for more collaboration between AI developers, developers should also consider their level of transparency. This could include, for example, sharing how an algorithm was made or making training data publicly available.

Literacy and Education

As AI becomes cemented further into technology and science, it is important that society be aware of its implications. Often, the concepts of "data literacy" and "AI

literacy” are cited as the necessary skills to remedy this. Literacy often includes knowledge about how intelligent systems procure, process, and utilize data, as well as where and how AI is commonly used, and its potential dangers. This knowledge is useful to understand how AI systems work, as well as to learn what to be wary of and what to trust.

AI has been threatening to incite increasingly large amounts of labor displacement. This displacement presents an issue for education, as displaced workers would need to be retrained and the way students are educated for career paths would need to be modified. Similarly, training for individuals entering the workforce might need to change to accommodate all the ways AI can affect the hiring process and the specific occupation. (Recommendation engines like Indeed, automated hiring, etc.)

The education of AI engineers is also a concern. It is important for AI developers to be aware of the ethical issues surrounding AI, such that they can make design choices that promote safety.

Responsibility

When an AI exhibits harmful behavior, who should be held responsible? For example, consider a self-driving car that hits a pedestrian. Should we hold the AI developers accountable? The company behind the car? Often, decisions can be attributed to the person that made them, and decisions can be explained. However, the decisions made by AI aren't always explainable, and machines can't be held responsible. Questions of responsibility tie closely into considerations of laws regarding AI.

Misinformation and Social Media

The internet and social media present enormous amounts of information to billions of people. As such, it is important for users not only to be able to parse this information but also to make sure that it is truthful and not harmful. Misinformation can be especially dangerous in places where large groups of people have significant influence, such as popular elections and the stock market.

AI systems show promise in being able to detect misinformation in social media—for example, they might be able to fact-check information, detect if two posts contain the same (false) information, or if a user is a bot. This shows promise for large companies, who must vet extremely large amounts of information constantly: with an automated system, less misinformation might be propagated to users, and companies wouldn't have to rely on human editors as much. Mark Zuckerberg, for example, frequently referenced AI in his 2016 testimony as the solution to misinformation issues. With the

complexity needed for such a system, bias must also be a concern, as well as moral considerations such as freedom of speech.

Potential Benefits

Artificial intelligence shows promising potential to improve many areas of technology. As Audrey Azoulay, director-general of the UNESCO, argues, “Artificial intelligence (AI) is undergoing exponential growth and finding new applications in an ever-increasing number of sectors, including security, the environment, research and education, health, culture, and trade.” It is a potential solution and a potential danger to many topics that are core to UNESCO’s considerations. For example, it has the potential to change education for both students and teachers, as well as the potential to aid in culture. It has proven useful during the COVID-19 pandemic for contact tracing and data crunching for potential solutions. It also has the potential to help combat climate change, perhaps through data analysis and ecosystem management. These promising advancements further show how indispensable AI will be as a technology, and how important it is to ensure it is developed safely.

Global Developments

Many countries have already begun crafting policy initiatives to address AI. These include funding, national strategies, as well as other plans. For a detailed list, the OECD has gathered many policy documents in one place: <https://oecd.ai/dashboards?selectedTab=policyInstruments>. The Future of Life Institute has also gathered a few national and international strategies: <https://futureoflife.org/national-international-ai-strategies/>.

UN Action

The UN has been similarly active, launching several events and plans regarding AI. For example, the International Telecommunication Union (ITU), a specialized agency of the United Nations focused on communication technologies, has organized the AI for Good Global Summit, held annually since 2017. These summits have been opportunities for AI Innovators to convene, and have focused on safe and accessible use of AI, as well as ways AI can help to achieve the Sustainable Development Goals (SDGs). The UNESCO, in particular, has hosted multiple conferences. These include an International Conference on Artificial Intelligence and Education, as well as an International Event where the Regional Forum of Artificial Intelligence in Latin America and the Caribbean debated AI and its potential for achieving the SDGs.

One of UNESCO’s major AI initiatives has been to create a comprehensive standard-

setting instrument regarding AI. A group of AI experts was selected and tasked with researching and creating the Recommendation on the Ethics of AI. The Recommendation will outline principles and concrete policy to ensure that the development and use of AI are focused on supporting humanity and the environment. A final draft of the Recommendation is planned to be adopted in November 2021, but a first draft can be read here: <https://unesdoc.unesco.org/ark:/48223/pf0000373434>. The UNESCO has launched online global consultations to discuss the first draft, and are currently (as of Jan 2021) revising it. The Recommendation has a large potential to promote global standards where there aren't any currently. As Audrey Azoulay, the Director-General of the UNESCO has said, "It is our responsibility to lead a universal enlightened debate in order to enter this new era with our eyes wide open, without sacrificing our values, and establish a common global foundation of ethical principles for artificial intelligence."

Bloc Positions

United States

The United States is one of the global superpowers leading the AI race. As one of the wealthiest countries, the number of people with access to cutting-edge technology is enormous. The US has always been a country with innovative technology, and that has continued with its advancements in AI. The issue of AI ethics has always been a concern in the US. There have been many conferences and papers published about the issue throughout the years, and the concern has only grown. Recently, in 2019, former President Donald Trump issued an executive order launching the American AI Initiative, which is guided by five principles: driving technological breakthroughs, driving the development of appropriate technical standards, training workers with the skills to develop and apply AI technologies, protecting American values including civil liberties and privacy and fostering public trust and confidence in AI technologies, and protecting US technological advantage in AI, while promoting an international environment that supports innovation.

Canada

Likewise, Canada, although not a global superpower, has some of the world's leading researchers in AI and has built its strategy in the race around retaining and producing new talent. According to Canada's global investment attraction and promotion agency, Invest in Canada, from 2016 to 2017 Canada invested 1.3 billion dollars towards AI research and development. Furthermore, Toronto has the highest concentration of

AI start-ups and Montréal has the highest concentration of students and researchers studying AI in the world. The talent and money backing Canadian AI, and its dedication to ethical concerns are impressive. Canada recently released an evolving Directive on Automated Decision-Making that pushes the government to commit to utilizing AI while maintaining “transparency, accountability, legality, and procedural fairness. The Canadian government believes that before adopting the technology, they must address the bias and responsible use requirements. The Pan-Canadian AI Strategy, which represents the first national AI strategy developed by any country, is a research and innovation initiative to strengthen Canada as an AI research hub. The strategy also aims to unite thought leaders to examine the societal implications of AI before implementation.

Russia

Although Russia is another global superpower in the AI race, the country has lagged behind others. President Putin has released a national AI strategy that includes a long list of goals and primary objectives for the development of AI, including some to be achieved by the year 2024, and some by the year 2030, but nothing too groundbreaking has come out of Russian innovators yet. These goals include improving the availability and quality of data, increasing the availability of hardware, and creating appropriate standards and a regulatory system that guarantees public safety and stimulates the development of AI technologies. The strategy also lists principles for the development and use of artificial intelligence technologies and states that their observance is obligatory during the implementation of the strategy. The principles include: the protection of human rights and liberties, security, transparency, technological sovereignty, innovation cycle integrity, reasonable thrift, and support for the competition. Recently, a released draft proposal said that legal and ethical rules should be created to govern the interaction of an individual with AI, determine the distribution of responsibility between owners, developers, and suppliers of data for damage caused using AI systems, and clarify the regulation of the circulation of results of intellectual activity using AI.

United Kingdom

At the 2018 World Economic Forum Annual Meeting, Prime Minister Theresa May announced, “As we seize these opportunities of technology, so we also have to shape this change to ensure it works for everyone...This includes establishing the rules and standards that can make the most of Artificial Intelligence in a responsible way.” As a wealthy country that ranked number one in the 2017 Government AI Readiness Index developed by Oxford Insights due to measures including digital skills training

and data quality, the UK has the potential to compete with countries like the US and China in AI. In 2017, Parliament established the Select Committee on AI to further consider the economic, ethical, and social implications of advances in artificial intelligence, and to make recommendations, and in 2018, they published a 183-page report. It suggests the country may have a competitive advantage in considering the ethics of AI, and proposes five principles: AI should be developed for the common good and benefit of humanity; AI should operate on principles of intelligibility and fairness; AI should not be used to diminish the data rights or privacy of individuals, families, or communities; all citizens have the right to be educated to enable them to flourish mentally, emotionally, and economically alongside artificial intelligence; the autonomous power to hurt, destroy, or deceive human beings should never be vested in artificial intelligence.

France

In 2017, France developed a national strategy for AI titled “AI for Humanity,” with several different entities and an AI Policy Initiative was launched in 2017. It culminated in a one hundred forty-seven page report in 2018, detailing seven prime themes: developing an aggressive data policy, targeting four strategic sectors (healthcare, environment, transport, and defense), boosting the potential of French research, and investing in talent, planning for the impact of AI on labor, making AI more environmentally friendly, opening up the black boxes of AI, and ensuring that AI supports inclusivity and diversity. Also in 2017, France’s National Commission for Information Technology and Liberties was tasked with the organization of an ethical debate on AI by the Digital Republic Bill. Its objective included ensuring “characteristics that must be at the heart of the French AI model: respect for privacy, protection of personal data, transparency, accountability of actors and contribution to collective wellbeing”. Throughout 2017, CNIL held more than forty-five debates and events involving more than 3000 people. More recently, French President Emmanuel Macron announced that he would invest one point five billion euros into AI research until the end of his term in 2022.

India

Despite having the second largest online market in the world, India has not been a massive contributor to the advancement of AI, which doesn’t give them a strong position on AI ethics. Their largest AI project was in 2018 when the Indian government defined a national policy on AI in a working paper, the National Strategy for Artificial Intelligence. The NITI Aayog paper identifies five focus areas where AI development could enable both growth and greater inclusion: healthcare, agriculture, education,

infrastructure, and transportation and mobility. The paper also discusses five barriers to be addressed: lack of research expertise, absence of enabling data ecosystems, high resource cost and low awareness for adoption, lack of regulations around privacy and security, and absence of a collaborative approach to adoption and applications. The paper includes several recommendations but is positioned as an essential pre-read intended merely to begin the conversation for what will be an evolving National Strategy for Artificial Intelligence.

China

Having the largest online market in the world and arguably leading the race in AI to make China one of the most influential countries in the world. In 2017, The State Council of China released the New Generation Artificial Intelligence Development Plan, which outlines China's strategy to build an AI industry worth one hundred fifty billion dollars in the next few years and to become the leading AI power by 2030. This marked the development of the AI sector as a national priority and it was included in President Xi Jinping's grand vision for China. In 2019, the New Generation AI Governance Expert Committee released principles of artificial intelligence governance, which include harmony and friendliness, fairness and justice, inclusiveness and sharing, respect for privacy, security and controllability, shared responsibility, open cooperation, and agile governance. In 2019, the Beijing AI Principles were released by a coalition including the Beijing Academy of Artificial Intelligence, Peking University, Tsinghua University, Institute of Automation and Institute of Computing Technology in Chinese Academy of Sciences, and an AI industrial league involving firms like Baidu, Alibaba, and Tencent. The 15 Principles call for "the construction of a human community with a shared future, and the realization of beneficial AI for humankind and nature." The principles are separated into three sections: Research and Development, Use, and Governance. They include focus on benefitting all of humanity and the environment; serving human values such as privacy, dignity, freedom, autonomy, and rights; continuous focus on AI safety and security; inclusivity; openness; supporting international cooperation and avoiding a malicious AI race; and long-term planning for more advanced AI systems.

NATO

As rival powers are leveraging new technologies to pursue the goals of greater economic competitiveness alongside greater military capabilities, the Allies, consisting of Belgium, Canada, Denmark, France, Iceland, Italy, Luxembourg, the Netherlands, Norway, Portugal, the United Kingdom and the United States, Greece, Turkey, Germany, Spain, the Czech Republic, Hungary, Poland, Bulgaria, Estonia, Latvia, Lithuania, Romania, Slovakia, Slovenia, Albania, Croatia, Montenegro, and North

Macedonia, face a range of challenges as they seek to exploit emerging and disruptive technologies, with AI at the center. For AI, the production factors are high-skill specialist talent and Information and Communication Technologies infrastructure for computing and storage, and data is the key input. Nato believes that supporting measures for the production factors, talent, and infrastructure, should include sound and flexible human resources and contracting policies to attract and nurture the best human talent, as well as the deployment of relevant and secure computational and data storage capacities. The Alliance's success with AI will also depend on new and well-designed principles and practices relating to good governance and responsible use that would be adopted.

D10

Led by the UK, the D10 is an alliance of ten democracies: Canada, France, Germany, Italy, Japan, the United Kingdom, the United States, Australia, India, and South Korea. They plan to focus on norm development, market supply, and other issues around 5G telecommunications technology and artificial intelligence. By excluding both China and Russia, they are trying to create alternate suppliers of 5G equipment and other technologies in order to avoid relying on the Chinese telecommunications giant Huawei. While not explicitly focusing on AI, the attempt to halt reliance on AI made from Chinese equipment is notable. Beyond the objective of exploring alternatives to Huawei 5G technology, the future actions of the coalition are not clear.

GPAI

Unlike the D10, the GPAI's official motives for formation (in May and June of 2020) cannot be explicitly pinned to a concrete objective. However, like the D10, both China and Russia are excluded and all of the G7 countries are members, but it also includes India, the Republic of Korea, Singapore, Slovenia, and the European Union. The GPAI seeks to foster a multi-stakeholder approach bringing together leading experts to collaborate across four themes: Responsible AI, Data Governance, The Future of Work, and Innovation & Commercialization. The GPAI is designed to serve as a forum for discussion and coordination on technical and policy research that furthers the shared democratic vision of the member state. Since it was recently founded, there is not yet an apparent mission statement on short, medium, or long-term goals that this initiative seeks to achieve; no clear plan of action or work plan; and no clear sense of how the GPAI will add to other collaborative efforts in the AI space.

Vocabulary

algorithm: A set of instructions given to a computer to solve a task. (Keep in mind, however, that AI algorithms aren't told exactly what to do. Rather, behavior is learned.)

Machine Learning: Algorithms that can learn patterns from data and make decisions based on what it has learned.

Neural Network: A powerful machine learning technique loosely based on the human brain. Neural networks consist of any number of interconnected "neurons" that operate under simple calculations.

deep learning: Machine learning using neural networks.

bias: Anomalous behavior in the output of an AI system. These can arise from biased assumptions during the algorithm's design or flawed training data.

training data: Data from which Machine Learning algorithms learn.

transparency: AI that is explainable. This includes being able to understand why an AI algorithm made a decision it did, as well as how an AI was designed and on what data it was trained.

ai/data literacy: The skills and knowledge to be able to use and think critically about data and artificial intelligence safely.

misinformation: Inaccurate information. Social media and the Internet can allow misinformation to spread quickly and freely.

Questions to Consider

How should UNESCO encourage the international community to take an ethical approach to AI? You might consider the difference between concrete rules and moral principles, and where each might be necessary.

How should countries mitigate bias and misinformation in AI? How should they promote transparency?

What is the best way to increase AI literacy and foster AI and data education?

How should responsibility be handled when an algorithm displays dangerous behavior?

How can AI be used to solve some of UNESCO's goals? Consider, for example, global warming, the eradication of poverty, culture, education, science, communication, and information.

Works Cited

<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

<https://www.forbes.com/sites/cognitiveworld/2020/12/29/ethical-concerns-of-ai/?sh=32cdcfe623a8>

<https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/#>

<https://research.aimultiple.com/ai-bias/#>

