



# ARTIFICIAL INTELLIGENCE

Definition

The problem

The change we want to see

How APC works on this issue

Some spaces and institutions to engage with

Read more

## DEFINITION

There is not one, single definition of “artificial intelligence” (AI) that is widely accepted. It refers to the theory and design of computer systems that can perform tasks requiring some degree of human “reasoning”: perception, association, prediction, planning, motor control, as well as systems that can learn from applying algorithms to large amounts of data. “Artificial intelligence” is actually a blanket term that could refer to varying levels and kinds of big data and algorithmic innovations. Under the rubric of AI we could include, for example, machine learning (ML), deep learning (DL) and neural networks (NNs).

While AI is a hot topic nowadays, it is not a new area. AI is a field of theory and practice that has been around for six decades. The activities involved range from machines applying logic and statistical analysis to historical data, to machines applying algorithms to large datasets and learning from them.

These technologies, however, are still in their infancy. For that reason, at the same time, there is still enormous potential to channel AI to address global challenges, while its impacts are still broadly unknown. Today there is much anxiety and concern regarding the effects of AI on social justice and on the enjoyment of human rights.

Despite such concerns, AI is everywhere – from using a virtual personal assistant to organise our working day, to travelling in a self-driving vehicle, to our phones suggesting songs or restaurants that we might like. AI is also increasingly present in the provision of public services and the design and implementation of public policies.

# THE PROBLEM

As mentioned above, there remains much promise for AI to address global challenges and promote innovation and growth. At the same time, however, there are questions about the trustworthiness of AI systems, including the dangers of codifying and reinforcing existing biases, such as those related to gender and race, or of infringing on human rights and values, such as privacy. Concerns are growing about AI systems exacerbating inequality, climate change, market concentration and the digital divide. The COVID-19 pandemic made evident the scale of AI influence in different spheres of our lives and how broad-ranging its impacts can be. Below, we discuss but a few.

- **Privacy:** The most obvious impact of AI systems and applications may be on privacy. Informational privacy may be challenged given the expansion in the extraction of personal data, as well as other information about people's lives and decisions. AI systems may facilitate and deepen privacy intrusions.
- **Reproduction and exacerbation of existing patterns of discrimination:** Contrary to popular belief that AI is neutral, infallible and efficient, it is a socio-technical system with significant limitations. As noted by Vidushi Marda, the data used to train AI systems "emerges from a world that is discriminatory and unfair, and so what the algorithm learns as ground truth is problematic to begin with." For instance, increasingly, law enforcement agencies use AI for predictive policing. Predictions made by AI systems trained with skewed data are often seen as "neutral" or "objective", which further ingrains discriminatory and

abusive practices. Biometric-based and other data-intensive systems are also being used worldwide in ways that reinforce and exacerbate structural racism and inequality.

- **Issues from a labour perspective:** Artificial intelligence, like the challenges of automation in general before it, involves the potential to both eliminate jobs, as they are taken away by machines or AI, and to create jobs, as a workforce is needed to keep systems of AI in place. Especially in the global South, says Noopur Raval, we are witnessing not total and complete automation of/in work but rather a “heteromation”, a reorganisation in the division of labour between humans and machines.
- **AI introduces censorship problems:** As AI automated decision-making systems are increasingly used by platforms to police unlawful or infringing content, there is collateral damage in the form of a chilling effect on freedom of expression. While AI has been presented as a solution to the serious harms that content moderation produces for workers entrusted to carry out this task, the increased obstacles for transparency and accountability pose a serious risk for freedom of expression online, says Luis Fernando García Muñoz.
- **The kinetic impact of AI can kill, harm or maim humans:** Artificial intelligence and automated systems have for decades now been involved in warfare and weapons systems. As the military use of AI increases, AI-based weapons and systems will increasingly have to make decisions that are ethical as they have the ability to kill human beings and start conflicts.
- **Liability:** When AI-based or automated systems make decisions that harm or negatively affect people, issues

of liability are introduced. Who is responsible when an automated system makes a decision that causes harm? The system, the coders of the system, the operators of the system, the managers? Solving problems of liability will also require explainability of systems.

- **Lack of transparency:** These systems are often implemented without accountability or community participation in the decisions around their implementation or in the evaluation and oversight of their impacts, further limiting the detection and remedy of undesired outcomes. The term “algorithmic transparency” has been used a lot to describe the concerns and issues with regard to transparency and AI. The focus on algorithms, rather than the transparency of the whole system, may not be helpful, as it will put the power in the hands of “experts” (e.g. mathematicians, statisticians and cryptographers) to explain algorithms, rather than the underlying systems. Many algorithms are already open and accessible. How these algorithms are rolled out into systems, products and services is just as important as the transparency of the algorithms themselves. As Comninos and Konzett argue, the concept of “system transparency” may be more useful.

As with so many uses of digital technology, most of the negative impacts of AI, in particular its use to violate rights, disproportionately affect those already in a situation of vulnerability.

Cybersecurity is an important recurrent theme in AI policy initiatives. Trustworthy systems are required to unleash the potential of AI for good.

## **THE CHANGE WE WANT TO SEE**

We want to see a world in which AI is compliant with data and privacy protection. We want the potential impacts of AI on other human rights to be recognised and addressed before the development and deployment of new technologies, infrastructure and products.

We believe that AI systems and algorithms must be transparent and offer people the ability to exercise their right to an explanation when they are affected by automated decision making. The notion of algorithmic/system transparency is in line with APC's belief in open source software: the more systems are open, the more explanation they will be able to provide, and the more opportunity we will have to create better systems.

The OECD has proposed interesting principles that should guide states in addressing AI-related ethical and human rights concerns:

- AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.
- AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards – for example, enabling human intervention where necessary – to ensure a fair and just society.
- There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.
- AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed.
- Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.

As is the case with all emerging technology, the AI context is constantly evolving. There are many areas that still need further research and analysis.

## HOW APC WORKS ON THIS ISSUE

APC does not focus on AI exclusively, but rather on the implications of AI systems on human rights, social justice and sustainable development.

APC contributes to policy discussions on AI by producing research on the positive and negative impacts of AI systems on privacy, security, freedom of expression and association, access to information, access to work, among other issues, with a specific focus on the global South.

Some of APC's proposals in relation to this agenda include:

- Any AI regulations adopted at the national level should promote the protection of the right to privacy and other human rights.
- Robust data protection regimes should be implemented by states.
- A moratorium should be established in relation to AI technologies and systems that can impose high risks on the enjoyment of human rights, including biometric recognition technologies.
- Human rights due diligence must be conducted for the design, development and deployment of AI systems.
- Both governments and businesses must implement transparency policies in relation to AI uses and allow for independent auditing.
- Explainability in relation to AI systems should be greatly expanded.
- Users, in particular those representing vulnerable groups, should be involved and actively participate in AI-related decision making.

## SOME SPACES AND INSTITUTIONS TO ENGAGE WITH

- The United Nations Human Rights Council and its Special Procedures
- The International Telecommunication Union (ITU), through, for example, the AI for Good initiative, study groups, and the preparatory process for the World Telecommunication/Information and Communication Technology Policy Forum (WTPF-21), where AI is one of the subthemes
- The Council of Europe
- The Freedom Online Coalition (FOC)
- The Organisation for Economic Co-operation and Development (OECD)
- UNESCO
- The global Internet Governance Forum (IGF) and national and regional IGF initiatives (NRIs)
- Digital rights events and conferences such as RightsCon
- National AI strategy discussions at the national level

## **READ (AND WATCH) MORE**

[Global Information Society Watch \(GISWatch\) 2019 – Artificial intelligence: Human rights, social justice and development](#)

[How I'm fighting bias in algorithms](#)

[FABRICS: Emerging AI Readiness](#)

[Artificial Intelligence, Human Rights, Democracy and the Rule of Law: A Primer](#)

[The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems](#)

[Racial discrimination and emerging digital technologies: a human rights analysis \(Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance\)](#)

[The right to privacy in the digital age \(Report of the United Nations High Commissioner for Human Rights\)](#)

[Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights \(Report of the Human Rights Council Advisory Committee\)](#)

[Towards Regulation of AI Systems: Global perspectives on the development of a legal framework on artificial intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law](#)

