HUMAN RIGHTS AND DIGITAL SOCIETY

Artificial intelligence and human rights: do they need each other?

Karlos A. Castilla Juárez



Bosniac Police at the train station of Bihac get the migrants back to Sarajevo by bus, 10th November 2018 | Photography: Alessio Mamo

When people talk about the relationship between human rights and artificial intelligence (AI), the most commonly asked question nowadays is how AI affects human rights. In fact, there is a specific group of human rights that is continuously pointed to as the most exposed to these effects and their protection in the governance and regulation of artificial intelligence is considered a priority issue. This is the case of privacy and data protection, freedom of expression and access to information, equality and non-discrimination.

That these human rights and not others are constantly in the spotlight seems normal if it is understood that artificial intelligence needs data for its design, development and deployment. In essence, data are the raw material that feeds the algorithms which any AI needs to operate. That is why knowing how these data are obtained (with or without authorisation), how they are stored and how they are used, what part is used or excluded, which data are not collected, how they are spread or shared, etc., are very important aspects. At the end of the day, data are important information about our lives and activities.

Faced with this, the question we have least asked ourselves is whether human rights need artificial intelligence. That is, whether AI is needed to exercise some or all of our human rights, because this is where the number of human rights that could be affected by artificial intelligence would increase, as we would not just be talking about issues related to data but

it could also impact the enjoyment, exercise and guarantee of other rights.

For example, do we need artificial intelligence to guarantee the right to housing? Do we need AI to exercise freedom of movement? Do we need AI to exercise and guarantee access to justice? If one thinks that human rights are supposed to be universal, it is not easy to answer these questions and one can hardly be absolute in positive or negative terms, so it is important to reflect on this.

What is the relationship between human rights and artificial intelligence?

In 2023, we can say that the relationship between human rights and artificial intelligence resembles more that of a “distant neighbour” than a “mutual engagement”. This is because even though human rights and AI live in the same world and know each other in generic terms, they have not interacted much in any proximity, despite many attempts, especially by international human rights organisations [1], to help them know each other better and work together.

Those of us who work in the field of human rights, over the years – and as we have seen how AI has reached more areas of the exercise, enjoyment and guarantee of these rights – we have come in closer proximity with it, often without understanding in detail and depth all that it implies. On the other hand, those who work in the field of artificial intelligence have shown little interest in human rights and have preferred to look to ethics when asked about the negative effects that AI may have on people’s lives.

For all these reasons, we can say that the relationship between AI and human rights is still under construction, despite its enormous importance: in almost every area where artificial intelligence is being implemented, there is a human right that is directly or indirectly linked to it. This is no exaggeration. The problem is that we often do not see this link because we lose sight of the fact that, for example, by seeking to use public transport where there are AI applications, we could be interfering with the exercise and guarantee of our human rights to freedom of movement, privacy, non-discrimination or recognition of legal personality, among others. And we could say the same for many other daily activities.

> The relationship between AI and human rights is still under construction despite its enormous importance: in almost every area where artificial intelligence is being implemented, there is a human right linked to it

In this respect, the relationship between human rights and artificial intelligence is very important and should be a priority. There needs to be more interaction between the human rights and AI communities so that technology development does not take place without the

human rights community. Particularly because the ethics which is appealed to as a form of self-regulation from artificial intelligence (companies, research centres, science academies, as well as some states, international organisations and civil society associations) is insufficient to prevent and avoid the possible negative effects of AI on people.

So it is clear that there is still a lot to do in the relationship between AI and human rights and, in the near future, there should be a relationship of mutual engagement for all the reasons that will be explained below.

Human rights in the field of artificial intelligence: where does this leave ethics?

The world's largest AI producers are corporations, including, in particular, Google, Amazon, Apple, Microsoft, Meta, Baidu, Alibaba and Tencent. This means that the bulk of artificial intelligence is being produced in the United States and China, or is subject to the economic interests of people with these nationalities. So it seems clear that, in practice, it is these corporations that have been establishing their own regulation and the framework for AI-related activities. This undoubtedly also impacts on small companies and the entire AI generation (or market). For this reason, it is precisely the many spheres linked to dominant sectors in AI (companies, research centres, specialised academies and associations linked to all of the above) that are insistently proposing ethics as the optimal framework for regulation that can guarantee rights and does not jeopardise the freedom of enterprise, intellectual property and, in general, their legitimate economic, innovation or business interests.

If this were not the case, it would be hard to understand why Europe – where the financial and political power of AI producers is still less – is the only region in the world where, in 2023, progress is being made in AI regulation, with closer attention to human rights, and abandoning mere ethical regulations. Admittedly, there are obstacles and certain aspects will require more thorough analysis and adjustments, especially if one of the goals is to protect people's fundamental rights.

In this context, the ethics in artificial intelligence that is being promoted and which it is wished to establish as a framework for regulation and the prevention of harm to people is seen to be changing, malleable, partial and uncertain, and allows impunity in cases of non-compliance. This is particularly true when it is the large corporations, with their economic or business interests, that ultimately establish this ethics or proposes its parameters and areas of application.

Faced with this, we have human rights which, firstly, allow us to presuppose that a certain number of philosophical debates have been resolved in order to define and identify them. Secondly, they represent minimum agreements on universal values (although not all of them are universal) in the form of rights and freedoms, so we know with a high degree of certainty what rights there are [2], their scope, where they are recognised and who must respect, protect and guarantee them.

Furthermore, global support for these rights is quite broad: there are commonalities in national and international regulations in many regions of the world. This means that these corporations (and others that also produce AI), like anyone else, have responsibilities in upholding other people's human rights, based on the content of national civil, labour, administrative and criminal legislation, and also deriving from national and international human rights standards, even though they are continually denied and eluded by all manner of enterprises; indeed, so widespread is this practice that it was necessary to develop the UN Guiding Principles on Business and Human Rights as an attempt to reaffirm the minimum standards that must be upheld.

Accordingly, human rights in the field of artificial intelligence, first of all, must address primarily everything that is related to the data used by AI to provide effective protection for the rights that might be impacted (privacy and data protection, freedom of expression and access to information, equality and non-discrimination). Second, they should have an influence in the areas where AI may be applied, that is, with respect to the questions and answers that the different algorithms will generate, as this is where many more rights may be impacted.

When viewed from the first (data-centred) perspective, for example, it will always be better to apply the successive international human rights standards that have been created, developed and updated on the right to private life, as they establish clear parameters in dealing with public and ordinary persons, and they define what parts of this right cannot be infringed under any circumstances. It is better to apply these standards than to let a business or person determine or evaluate in ethical terms what is considered "ethically appropriate" to know, disclose or use regarding the private life of a person or of a group of persons concerning whom a series of data have been gathered, with or without their total consent.

From the second perspective, and as already attempted in the proposed European regulation, it is necessary to define the areas in which human rights are clearly at risk and, therefore, also the parameters for which the use and development of artificial intelligence should be limited. In the case of Europe, for example, based on the establishment of "risk levels" [3], all AI systems that pose a clear threat to people's security, livelihoods and rights are considered an unacceptable risk: from the social scoring implemented by governments to the toys that use voice assistants, which encourage dangerous behaviour. Ad hoc ethical assessments of these issues do not guarantee a minimum of legal certainty and security, even if these assessments are carried out in good faith.

Thus, based on the creation of high risk, limited risk and no risk areas, the European Union has proposed prohibiting the use of artificial intelligence in biometric categorisation systems that rely on sensitive personal features such as gender, race, ethnicity, religion or political affiliation. In addition, it would prohibit the use of predictive monitoring systems to assess the risk posed by a person or a group of persons for committing a crime or offense, and the use of emotion recognition systems by the police and border controls, or in workplaces and schools. Likewise, random tracking of biometric data from social media or surveillance cameras to create or expand facial recognition databases will be limited. All in

all, it seems clear that there are areas of human rights where AI must clearly be banned, however much some people might think it acceptable on ethical grounds.

> Ethics as a form of self-regulation for artificial intelligence is insufficient to prevent the negative effects on people

From a purely economic and mercantilist-capitalist business or corporate viewpoint, establishing spheres where AI is prohibited may seem unjustified severe interference. Hence, in part, the insistence on ethics-based regulations. However, from the perspective of the protection of human rights, it is clear that there are certain minimum levels of these rights and freedoms that must be non-negotiable, just as there are limits to the infringement of human rights in many other areas in the "analogue world" in which we have lived until now.

Artificial intelligence in the field of human rights

In the artificial intelligence sector, the areas in which human rights must be taken into account seem clear. Now, however, we must analyse a less explored perspective: does AI need to be present in the field of human rights? Or, to put it another way, is AI necessary for exercising some or all of our human rights?

If we remember that, originally, exercising, guaranteeing and protecting human rights did not need artificial intelligence, in 2023, and looking to the future, it does not seem essential to use this technology for such purposes, unless artificial intelligence makes it possible to guarantee the so-called universality of human rights that, 75 years after adoption of the Universal Declaration of Human Rights, has yet to be fully respected in all regions of the world.

In any case, determining whether artificial intelligence is necessary for exercising, guaranteeing and protecting human rights cannot be based on a general conclusion for all rights, but requires an analysis of each of the rights recognised at both national and international level to determine whether AI is necessary and useful on the basis of the rights' elements and characterisation. It even requires a more particular analysis of the place or country in question, as there are places in the world where access to Internet is not guaranteed, so considering the use of AI would have the effect of magnifying inequalities [4].

Individual analysis of rights is very important because, for example, it is not the same to exercise the right to peaceful protest as it is to exercise the right to housing. We can protest with identical impacts to those of the analogue world; in fact, in some cases, one could even obtain more focused effects if the protest is made in virtual or digital environments, so long as the right's purposes are fulfilled. However, it would always be insufficient to have a dwelling in the digital world if we don't have one in analogue reality

or if we cannot even access a dwelling because we cannot use the application that would enable us to do this, or because the application excludes us, without any reasonable justification, on the basis of the characteristics and data used by the algorithm.

In this respect, the fact that there is no specific artificial intelligence for exercising, guaranteeing and protecting a specific right should not be considered per se as something negative. We should accept and understand that not all human rights need AI in order to be effectively exercised, guaranteed and protected. Indeed, the truth of the matter is that if there are people in the analogue world who are being excluded or cannot exercise a certain right, the use of artificial intelligence is very unlikely to change the situation.

Let us consider, for example, the access to justice. In today's real analogue world, many people have no effective guarantee of access to this right, for many reasons, even in so-called *developed* countries [5], so implementation of a "digital justice" or "automated courts" does not provide by itself a solution to this problem. And even less if access to these digital applications must overcome the same access barriers as analogue justice [6]. Therefore, it is very probably only a good option for those who already have access to "analogue justice".

But even in these cases, it does not necessarily provide a solution for those who already have access to justice if this "digital justice" is unable to assure the legal guarantees that are required in any court procedure (the judge's impartiality, public procedure, access to legal counsel, the prohibition of undue delays, and the use of available evidence, among others). So we must take what is often touted as "modernity" with a pinch of salt, because it does not necessarily imply effective access, exercise and observance of human rights.

> It is important to decide on the use of artificial intelligence with human rights in mind. if we only consider what is modern, it is very possible that we will simply be creating modern forms of exclusion

Faced with this situation, it is important to take human rights into account when considering the use of artificial intelligence in order to determine whether or not AI offers anything better than what already exists in the analogue world for exercising, guaranteeing and protecting such rights. In other words, we should not be afraid to say no to an apparent modernity that will not contribute anything positive to human rights, and much less so when it is offered by companies who care about their business and not the universality of human rights.

In the previous example, there may be specific aspects of this access to justice or legal protection in which AI is useful. For example, processing files, access to such files, information management or solving simple problems [7]. However, this does not mean that it is useful for everything that the right implies and must contain.

Neither can we expect that all decisions concerning a human right will be taken by artificial intelligence, because, no matter how much data there is, human intervention will always have an important role to play in the individual allocation of rights and freedoms. Considering all this, it seems undeniable that AI may help with advice, suggestions and parameters in gaining a more complete understanding of situations or in obtaining important information, but the final decision should not depend solely on this technology.

In order to determine whether we need AI to exercise, guarantee and protect a specific right or freedom, we should ask at least the following questions:

- What does AI bring to a certain right that goes beyond its exercise, guarantee or protection in analogue reality?
- What level of effective access to a certain right currently exists in analogue reality?
- Does the use of AI improve access to or exercise of this right and does not perpetuate existing (or even create new) exclusions?
- Is AI able to guarantee all the elements and characterisations that have been given to a certain right in the standards concerning human rights?
- What is the decision level that will be delegated to artificial intelligence for exercising, guaranteeing or protecting this right and, therefore, what is the decision level that people will have?

If we ask these questions before deciding that the right to health, the right to work, the right to education, freedom of expression, the right of access to information, the right to nationality or freedom of movement will be exercised, guaranteed or protected by artificial intelligence, it may be possible to diminish or avoid further impairments of human rights. Otherwise, if we only consider “what is modern”, it is very possible that we will simply be creating “modern” forms of exclusion and infringement of human rights.

Conclusion

In the field of human rights, what could give us cause for concern about any technology, not just artificial intelligence, is the technology creation and design process, on the one hand, and the technologies’ uses, applications, creations, effects or results that exclude, restrict, discriminate, nullify, impair, prevent or reduce the recognition, enjoyment or exercise of human rights and fundamental freedoms on equal terms, on the other [8].

Viewed in this light, it is important not to lose sight of human rights in artificial intelligence, just as we must not lose sight of AI in the field of human rights, as there are impacts in both directions, although not all of them are always negative.

If the aim is to safeguard people’s rights and freedoms, current reality shows us that human

rights are more necessary in the field of artificial intelligence than AI in the field of human rights.

REFERENCES AND FOOTNOTES

1 — See, for example: Council of Europe Commissioner for Human Rights (2023). *Human rights by design future-proofing human rights protection in the era of AI*. Strasbourg: Council of Europe.

2 — Risse, Mathias (2018). "Human Rights and Artificial Intelligence: An Urgently Needed Agenda". *HKS Faculty Research Working Paper Series.* RWP18-015, p. 10.

3 — European Commission (2020*)*. *White Paper on Artificial Intelligence: a European approach to excellence and trust*. Brussels.

4 — Castilla, Karlos (2022). *Cuatro ángulos de análisis de la igualdad y la no discriminación en la inteligencia artificial*, Barcelona: Human Rights Institute of Catalonia, p. 6.

5 — Deseau, Arnaud; Levai, Adam; Schmiegelow, Michèle (2019). "Access to Justice and Economic Development: Evidence from an International Panel Dataset". *LIDAM Discussion Papers IRES*, 2019009. Louvain: Catholic University of Louvain, Institute of Economic and Social Research (IRES).

6 — Castilla, Karlos (2012). *Acceso efectivo a la justicia: elementos y caracterización*. Mexico: Porrua.

7 — Reiling, Dory (2020). "Courts and Artificial Intelligence". *International Journal for Court Administration 8.* 11(2)*,* p. 10.

8 — Pont, Anna; Passera, Agostina; Castilla, Karlos (2022). "Análisis introductorio"*. Impactos de las nuevas tecnologías en los derechos humanos*. Barcelona: Human Rights Institute of Catalonia, p. 11.



Karlos A. Castilla Juárez

Karlos A. Castilla Juárez holds a BA in Law from the National Autonomous University of Mexico and a PhD in Law from the Pompeu Fabra University. He is head of research at the Human Rights Institute of Catalonia, a member of the University of Barcelona's Public Law Observatory and associate professor in International Human Rights Law at the Pompeu Fabra University. His research activity has focused on international law (international human rights systems and international litigation), constitutional law (conventionality control and access to justice systems), the rights of migrants (immigrant internment, deportation, migration and climate change) and specific human rights-related issues (racism, equality, non-discrimination, intersectionality, freedom of expression and transitional justice).