Respecting cultural diversity in ethics applied to AI: A new approach for a multicultural governance

Respetar la diversidad cultural en la ética aplicada a la Inteligencia Artificial: un nuevo enfoque para un gobierno multicultural

Autores: Emmanuel R. Goffi, Aco Momcilovic
DOI: https://doi.org/10.25058/1794600X.2135
Respecting cultural diversity in ethics applied to AI: A new approach for a multicultural governance

Artículo de reflexión

a. Dr Emmanuel R. Goffi is the co-founder and co-Director of the Global AI Ethics Institute. He also is a research fellow with the Centre for Defence and Security Studies, University of Manitoba, Winnipeg, Canada and with Z-Inspection®, Frankfurt, Germany. He is a member of the Equipe Ethique, Langue, Communication et Numérique (ELCN), Université Mohammed Premier, Oujda, Morocco and he represents France in the International Group of Artificial Intelligence, Bahrain, United Arab Emirates.

b. Mr Aco Momcilovic is the co-founder and co-Director of the Global AI Ethics Institute. He is one of the co-founders of the Alliance for Responsible AI and represents Croatia in the International Group of Artificial Intelligence, Bahrain, United Arab Emirates.

ABSTRACT

Artificial intelligence seems to be part of our everyday lives. For some it represents the promise of a better world and many improvements that would be beneficial for humanity. For others, AI is seen as threat, if not an existential threat that needs to be controlled strictly. Whatever the stance, the need to regulate AI is now widely recognized. Short of legal instruments offering a specific framework for the development and use of AI, ethics has been summoned to set standards and establish guardrails. Yet, the number of documents pertaining to ethical standards for AI has increased exponentially to reach a point where it is difficult to know how to use them efficiently. These documents have mostly been issued to promote vested interests, and the setting of a universal code of AI ethics has been seen as a solution for AI global governance. If a global governance system is required to avoid negative outcomes of AI, it appears that the idea of a universal code of ethics denies the diversity of ethical standpoints based on the diversity of philosophical cultures the world is made of. Instead of offering a legitimate and efficient tool, such a solution could lead to cultural
tensions between leading actors in the field of AI as it is the case between China and the United States. To avoid conflicting situations stemming from the denial of cultural diversity, it is more than ever necessary to put aside the idea of a universal code of AI ethics and to start thinking about a new governance system that would be respectful of the variety of ethical perspective.

KEYWORDS
Artificial intelligence, culture, governance, ethics, diversity.

RESUMEN
La Inteligencia Artificial parece ser una parte de nuestra vida diaria. Para algunos representa la promesa de un mundo mejor y muchas mejoras que serían benéficas para la humanidad. Para otros, la IA es vista amenaza, si no como un peligro existencial que necesita ser estrictamente controlado. Cualquiera sea la postura, la necesidad de regular la IA actualmente está ampliamente reconocida. Con pocos instrumentos legales que ofrezcan un marco específico para el desarrollo y uso de la IA, se ha acudido a la ética para establecer estándares y establecer límites. No obstante, el número de documentos que hacen referencia a los estándares éticos para la IA ha aumentado exponencialmente hasta alcanzar un punto donde es difícil saber cómo utilizarlos de manera eficiente. Estos documentos en su mayoría se han publicado para promover intereses particulares y el establecimiento de un código universal de ética para IA ha sido contemplado como una solución para el gobierno mundial de IA. Si se necesita un sistema de gobernanza mundial para evitar un resultado negativo de la IA, parece que la idea de un código de ética universal nega la diversidad de puntos de vista éticos que se basan en las diferentes culturas filosóficas de que el mundo es hecho. En vez de ofrecer una herramienta legítima y eficiente, tal solución podría levar a tensiones culturales entre los principios atores en el campo de la IA como es el caso de China y Estados Unidos. Para evitar situaciones conflictivas decorrentes de la negación de la diversidad cultural, es más que nunca necesario dejar de lado la idea de un código ético universal de IA y comenzar a pensar acerca de un nuevo sistema de gobernanza que respete la variedad de perspectivas éticas.

PALABRAS CLAVE
Inteligencia Artificial, cultura, gobernanza, ética, diversidad.

INTRODUCTION
Artificial intelligence (AI) has flooded into our everyday lives at fast pace to a point where it has become subject to all kind of phantasms. Seen as an existential threat by some, AI is also considered as a beneficial tool for humanity by others. Whatever the stance, the concerns and fears raised by AI inevitably led to the question
of its regulation and then to its governance at the international level.

Considering people’s apprehensions, public authorities and private companies started to release a large number of documents – codes, regulations, recommendations, charters and alike – to frame the development, deployment and use of AI.

The multiplication of regulatory instruments has naturally led to the need for a global governance, aiming at the harmonization of regulations [1], [2], [3].

Thus, the idea of a global/universal governance based on ethical principles applicable to AI has blossomed among international organizations such as the Organisation for Economic Co-operation and Development (OECD) [4], the United Nations Educational, Scientific and Cultural Organization (UNESCO) [5], the European Union (EU) [6], the G20 [7], or the Council of Europe [8].

Yet, the interpretation of these calls for cooperation in the field of AI governance have been interpreted in a very restrictive way leading to a tendency to favor the setting of so-called “universal principles” and Western ideas regarding values. The possibility of these principles stems itself from “the more foundational belief that these ideas are universal” [9], and that consequently some kind of universal ethics is reachable and desirable.

Therefore, some international institutions such as the UNESCO and the EU have decided to promote the implementation of universal ethical norms as the way to ensure the global governance of AI.

If the intention is laudable, the execution is not without posing problems since universal ethics has been so far reduced to a narrow Western perspective. Incidentally, the Institute of Electrical and Electronics Engineers (IEEE) [10] has clearly identified the “urgent need to broaden traditional ethics (…) beyond the scope of ‘Western’ ethical foundations”, stressing that “the reality of the situation reveals a monopoly of, and a bias toward, established Western ethical value system” and calling for the inclusion of “other traditions of ethics”.

Given the rise and appetite of AI non-Western actors, China leading the way, this Western monopoly could reinforce existing, or even give birth to new tensions, thus spoiling all attempts to establish an international system of AI governance.

Then, to avoid conflicting situations, it seems more than ever necessary to step back to the initial call for multilateral policy based on open discussions respectful of ethical particularisms.

**ETHICS APPLIED TO AI: THE WESTERN WAY**

**The myth of control**

The need of a global governance of artificial intelligence is widely recognized. The potentialities of this technology and systems fitted with AI are strategic, yet their outcomes can be either beneficial or conversely detrimental. The current polarization of the debate between technophobes and technophiles is illustrative of the incredible range of possibilities opened by AI. The difficulty lies now in our ability to overcome a Manichean perspective on AI to reach a nuanced stance offering us a way towards a development and a use of AI systems that would be as much controlled as possible, keeping in mind that full control is illusional.

So far, the narrative on artificial intelligence regulation has been built around the premise that human beings will remain in control promoting human oversight to prevent or minimize risks associated with AI systems [11]. This stance is highly questionable, and the assertion regarding human control is misleading. The belief in the capacity of human beings to keep control over AI is itself a very cultural standpoint stemming from the Aristotelian Scala Naturae (natural ladder or ladder of life) [12], [13] and from a specific cosmogony in which humans are at the top of the hierarchy of all sentient and non-sentient beings. Given this ascendancy over the world, humans thanks to techne, are entitled to control their environment and to change it for their benefit. AI is nothing else but the last avatar of this belief in progress through technology seen as a mere instrument in the hands of human beings, and, as such, supposedly under their control.

However, some voices have raised to warn against this deluding perception stressing that we are “utterly blind to the essence of
technology” [14], that technical objects have their own dynamics [15], and that according to the substantivist theory, technology “is not simply a means but has become an environment and a way of life” which “constitutes a new cultural system that restructures the entire social world as an object of control” [16].

Interestingly, this thirst for control, and consequently for power, is not universally shared. Many cultures do have a quite different view on the place and role of humans in the wider natural ecosystem, giving birth to a cosmogony where humans are not at the top of the hierarchy but at the same level than any other living or non-living objects. For instance, spiritualities such as Shinto, Buddhism, Confucianism, Animism or even aboriginal wisdoms in the Pacific and American First Nations beliefs, do not position humans at the top or even at the centre of the world, but as cogs of a larger ecosystem in which harmony must be reached through respectful relations. In most cases, nature is even regarded as being above humans.

This understanding of our place in the world has shaped our perceptions regarding our relation to the world we live in. The faith in technology seen as an instrument allowing us to be in control, has led us to think about AI as a panacea with which, more than ever, “[t]he desire for power consecrates the machine as a means of supremacy” [15].

But as Juvenal asked, quis custodiet ipsos custodes, who is to guard the guards themselves?” [17] The development and use of AI to “improve”, if this ever makes sense, our lives have come with the need to control the tool through which we control: the need to guard the new artificial guardian of our well-being.

Ethical norms: one size fits all?

Norms have naturally imposed themselves as a tool of choice to frame artificial intelligence and alike systems.

Short of any consistent legal instrument, ethics has appeared to be a fallback solution and turned into cosm-ethics, namely a communication rhetoric aiming at reassuring users by building trust artificially trough smart wording and performative speech acts [18]. Even if “normative models and principles” are seen as necessary to make sure AI will be developed and use in a human-centered approach [19], some actors do not look on legal framework with a favourable eye. Quite the contrary, many actors promoting vested interests are reluctant to welcome legal norms [1], [2], [20] and actively shaping the discourse to establish a “moral background” that will limit the debate to their own concerns [21].

This strong interest for ethics applied to AI has given rise to a plethora of documents aiming at regulating the development and use of AI systems. The ideas of trustworthy AI and responsible IA have become common and widely used in many public and private circles to demonstrate their will to promote ethically acceptable AI. Nonetheless, behind words can hide agendas that are less attractive than speeches.

In this context, it is worth reminding that most codes and alike documents pertaining to ethics applied to AI (around 60% of identified documents), have been released by Western stakeholders, mainly by private companies and political bodies in Northern America and Europe [1], [22]. It is then obvious that the setting of these codes and documents is strongly culturally influenced and built based on a Western appraisal of the context, the concerns, and the solutions to be offered.

This narrow lens through which the ethical acceptability of AI is assessed is problematic in that it does not recognize the diversity of ideas, beliefs, philosophical stances, and context analysis that results from specific perceptions. While AI systems are heavily contested for their biases, it seems that no one cares about the cultural bias pervading the reflection on ethics applied to AI, and its potentially dangerous outcomes.

FROM PRAISEWORTHY THEORY ...

Respecting cultural diversity: a legal demand

Then, so far, ethics applied to AI has focused on the Western perspectives about both concerns and solutions.

The whole narrative about ethics applied to AI, misleadingly named AI ethics, has been designed with the aim of “constructing and imposing a shared ethical frame on a contentious conversation” [21]. This “moral background” of ethical AI/ML discussions”, as Greene et al. call
it, legitimates both practices and policies set unilaterally by the West, without having to enter in the complexity of the subject and consequently diving into the intricacies of cultural diversity.

Yet, many documents released at the international level are calling for respect for cultural diversity, stressing its fundamental importance for security, human development, dignity, freedoms, fruitful international cooperation, communities’ empowerment, to cite but a few fields impacted.

Back to 1948, article 22 of the Universal Declaration of Human Rights (UDHR) made a clear link between culture and dignity stating that everyone “is entitled to realization, (…) of the economic, social and cultural rights indispensable for his dignity and the free development of his personality” [23]. To this end, the United Nations Charter explicitly calls for the promotion of “international co-operation in the economic, social, cultural, educational, and health fields, and assisting in the realization of human rights and fundamental freedoms for all” (art. 13.2).

Now, fundamental freedoms encompass “freedom of speech and belief and freedom from fear and want” as asserted in the UDHR. These four freedoms stem from the human security concept as initially described in the Human Development Report, which stresses the tie between freedoms from fear and from want, on the one hand, and peace, on the other hand [24]. So far, the United Nations has clearly instituted direct relations between cultural rights, human dignity and freedom, and their necessary interconnection to attain human security and personal achievement of each individual. In other words, freedom from fear or from want cannot be detached from respect for cultural diversity. This tight relation had been already highlighted in the 1966 International Covenant on Economic, Social, and Cultural Rights, which asserted that “the ideal of free human beings enjoying freedom from fear and want can only be achieved if conditions are created whereby everyone may enjoy his economic, social and cultural rights” [25].

A matter of human security

Even more, the Commission on Human Security published a report in 2003 in which the overlap between human security, human rights and human development was underlined [25]. According to the Commission, human security refers to the protection of “the vital core of all human lives in ways that enhance human freedoms and human fulfilment”, as well as of “fundamental freedoms”. Human security aims at “protecting people from critical (severe) and pervasive (widespread) threats and situations”, and request the creation of “political, social, environmental, economic, military and cultural systems that together give people the building blocks of survival, livelihood and dignity”. Through this definition, the Commission demonstrates that human development, rights, fulfillment, security, dignity and cultural systems are deeply intertwined, and that any action on one of these elements has an impact on others. Eventually, any infringement to culture is an infringement to security; any breach to security is a breach to freedom; any violation of freedom is a violation against dignity and so on so forth. Basically, cultural systems must be respected to avoid any domino effect that could heavily impact freedoms, security, or human rights.

Furthermore, in 2000, the Millennium Declaration was adopted by he UN General Assembly which affirmed the existence of fundamental values essential to international relations, among which freedom, namely the fact that people have the right to remain free “from the fear of violence, oppression or injustice”, and tolerance, that is respect for human beings “in all their diversity of belief, culture and language”, even stressing that “[d]ifferences within and between societies should be neither feared nor repressed, but cherished as a precious asset of humanity” [26]. One year later, the UNESCO released the Universal Declaration on Cultural Diversity in which cultural rights were recognized as “part of human rights” [27]. Furthermore, article 4 of the Declaration establishes that “[t]he defence of cultural diversity is an ethical imperative, inseparable from respect for human dignity”. Finally, in 2012, the UN General Assembly adopted resolution 66/290 recognizing “[t]he right of people to live in freedom and dignity” [28].

Aligned with the objectives of the UN Charter, these documents evidently emphasize that the play of international relations cannot occur in a setting deprived of respect for cultural diversity, and that any attempt to contravene this right might lead to the weakening of human security, to
increasing tensions, and, potentially, to breaches to international peace and stability.

As one can see, the will to integrate cultural diversity for the sake of international stability pervades international reflections calling for respect towards cultural diversity considered as “an essential component of human development, representing a source of identity, innovation and creativity for all” [29] and “a precious asset of humanity” [26].

... TO HAZARDOUS PRACTICES

Avoiding the ethical Western centrism

Yet, current processes towards a global governance of AI are led by Western actors, following a Western logic based on the assumption of the existence of universal values on which a universal set of ethical norms applicable to AI could be built. The quest for a universal code of AI ethics has led nations to divert from initial praiseworthy intentions to fall into the trap of philosophical Western-centric approach of ethics applied to AI.

This centrism is nothing new. Ethical cultural centrism or even eurocentrism has been identified by intercultural philosophers [30], [31], [32].

Thus, even the UNESCO [27] which promoted the defence of cultural diversity to the rank of “an ethical imperative” has entered the normative race and proposed a set of recommendation arguing that the organization “has a unique perspective to add to this debate given its strong comparative advantage thanks to its universality in membership” [5]. Yet, the universality of the UNESCO is subject to debate in the sense that there is inevitably a strong conformist tropism among diplomats and members of the Special Committee. The representativity of the members of the Committee must not be taken for granted. Especially since, “[m]any UN agencies suffer from a technocratic bias: they believe in technical fixes without much regard for cultural and social factors” [4].

The best and certainly most concerning illustration of Western-centrism is the work done by the European Union trying to impose itself as a normative actor and acting as a moral crusader through cosm-ethics, that is to say a narrative using words of ethics without doing ethics [18], [33]. A recent report even highlighted the fact that the European Union could not keep on focusing on the ethical narrative in a highly competitive setting. In other words, the cosm-ethical “marketing narrative” [34], will not be able to hide for long the real geopolitical dimension of AI as an "external policy tool" [35].

Denying the reality of cultural diversity by the imposition of a narrow Western perspective on ethics applied to AI enshrined in a supposedly universal set of norms, has the potential to erode human dignity, contravene the right to cultural diversity and represent a “critical and pervasive threat” to vital core of human security [25].

The risks with the Western normative tyranny

This is not pure theoretical reflection. As mentioned by the IEEE [10], the Western monopoly over ethics applied to AI is a reality, and this monopoly will soon be confronted to the will of some raising stakeholders to become leaders in the domain, China leading the way. If one does not take care, this monopoly might turn into a “tyranny of Western norms” that would deny “the right for cultures to express their peculiarities” [33].

The denial of the right to dignity and its associated right to cultural diversity can be considered as a threat by some communities or even nations. The mere imposition of ethical norms that do not fit with local cultural perspective can be experienced as a form of violence against which it would be worth standing up nay fighting.

The risk is even bigger if ethics is used as a mere communication tool aiming at ensuring a position of dominance for some actors. In that matter, the European normative proselytism is highly disputable and its blindness in front of the diversity of ethical systems [33] will inevitably lead to strong reactions from future AI leading actors. At the end of the day, the European normative tyranny could end up with a risky deregulation instead of the expected establishment of a framework for global governance.

The reality of cultural diversity cannot be ignored. Even more, it must be considered and respected. Some communities or countries already do not align with ethical principles set
by the West. Some might even not share the same interests in ethical considerations: some countries are legitimately focused on financial and strategic benefits, much more than on fundamental rights. That does not mean that they do not have ethical standards. That just means their standards and priorities are different. Ethical debate is a Western luxury that some human communities cannot afford. In some cases, ethics is not even a concern. Grassroot people from cultures based on spiritual hierarchy such as Confucianism or Islam, for instance, do not question the ethical acceptability of decision made by legitimate authorities.

Thus, the imposition of a unique set of Western norms as a ground for AI global governance, could lead to a cultural tyranny that would disrupt and destructure national communities, forcing people to chose between their own traditional culture and Western culture, and potentially destabilize some countries leading to domestic and international tensions.

BACK TO REAL MULTILATERALISM

The need for real multilateralism

The only option to avoid potential frictions and even digital conflicts, seems to be multilateralism.

Instead of trying to set universal ethical norms which could be considered as a threat to specific values and interests in the long run, it is worth moving back to theoretical stances calling for respect towards cultural differences. As mentioned by the Commission on Human Security [25], to existing problems, the new Millennium has added “a fear that existing institutions and policies are not able to cope with weakening multilateralism, falling respect for human rights”. Looking closely at the interplay between nation-states in the field of environment, terrorism, international law, defence and security, for instance, this fear is perfectly legitimate.

 Nonetheless, even if multilateralism is not a panacea and if international actors, public and private, will keep cooperating or opposing each other based on their own interests, multilateral action is the only viable way to attain a consensus on global governance for AI.

Following, the work done by the OECD, along with eth G20 and G7, global governance cannot be established without due reciprocal respect for cultural particularisms. This must be nonnegotiable for it is a sine qua non condition for compromise and peaceful and fair treatment of all standpoints from all stakeholders. According to the OECD [4], “AI has pervasive, far-reaching and global implications that are transforming societies, economic sectors and the world of work, and are likely to increasingly do so in the future”. This pervasiveness can be seen in the normative dimension of AI. This is why the OECD recommends that mechanisms and safeguards, adapted to the context, should be implemented in order to protect human rights, among which the right to dignity. There is no way to think that “universal” standards based on a very superficial approach of deontology [33] would be adjusted to contexts. We can already see that incapacity to adjust to the reality of the diversity of perspectives when it comes to address biases, privacy, video surveillance, autonomous vehicles, or even lethal autonomous weapons.

As mentioned in the IEEE [10] report, the full benefit of AI systems “will be attained only if they are aligned with society’s defined values and ethical principles”. This point is also stressed by the UNESCO [5], when it calls for the promotion of social justice and fairness, while considering cultural systems. Normative actors are no exceptions. Their potential impact on the international stability goes along with responsibilities regarding the enforcement of basic obligations related to human rights. AI stakeholders, normative actors included, need to be aware of “cross-cultural ethical variations while also respecting widely held international legal norms” [10].

Ethical universalism: a slippery slope

To that end, AI stakeholders must free themselves from the current “ethical universalism” tendency pervading in the codes, guidelines and other documents pertaining to ethics applied to AI, and that are ultimately mere statements and “not mass mobilization documents” [21]. In fact, as stated by Greene et al., “[h]igh-profile values statements are powerful instruments for constructing and imposing a shared ethical frame on a contentious conversation”. There are not efficient tools universally accepted and abided by.

The Western normative hegemony eventually appears to be a threat to cultural diversity. As such, it will give birth to way more issues than it
will solve problems. That is why an intercultural debate on ethics is more than ever needed. Intercultural ethics offers a forum for discussion and a way to compromises that can lead to the establishment of common ethical standards. These standards might be local, giving birth to regulatory clusters built on shared perceptions regarding the benefits and drawbacks of AI, common vision on what is acceptable and what is not regarding AI, and mutual interests. Reaching such local regulation, such as the one set in the European Union, would be the first step towards a global governance monitored by a neutral mediation body, inspired by the settlement mechanism implemented by the World Trade Organization. Such a neutral third-party would allow regulatory clusters to interact with each other and meet in the middle of the bridge.

Admittedly, this option is far from being perfect, but it is worth being studied and improved.

CONCLUSION

Norms can be a powerful weapon to either regulate the development and use of AI. They can also, as any other weapons, be used in a harmful way. The current will to establish universal standards framing AI can lead, and is already leading, to the use of norms as an instrument to promote vested Western interests and to impose a cultural hegemony that goes against international norms regarding human rights.

Implicitly denying the fundamental importance of cultural diversity in the ethical appraisal of AI, the West is playing a dangerous game that could end up with tensions between actual and up-and-coming leading actors in the field of AI. These tensions could add to existing ones and turn into severe confrontations. The current opposition between China, which has clearly stated that it would take the lead on artificial intelligence, and the United States, could be for instance exacerbated by cultural questions. In March, the harsh exchange between Beijing and Washington representatives in Anchorage regarding democracy, is one illustration among many other of the coming contest of the ethical wills.

It is concerning to see that the potential risks associated with the imposition of Western ethical standards, based on Western values is so poorly studied and analyzed. Yet, norms can be harmful for cultural diversity, and consequently to human dignity. They can lead to a feeling of threat, to which some actors could answer by violent means at different degrees.

Moving back to multiculturalism and due respect for cultural diversity is the only sustainable way towards global governance. Abandoning the idea of a supposedly “universal code of AI” must be a first step, followed by multilateral debates at both local and global levels. The main reasonable aim should be the establishment of regulatory clusters in which willing countries would set their own standards aligned with their religious and philosophical beliefs, political interests, shared perceptions and so on.

Then, the final step to allow these clusters to communicate and, when necessary, work together, would be the implementation of a mediation neutral body, that would ease exchanges and provide advice and support to parties, and eventually help to settle disputes.

Obviously, multilateralism is not a fix-all solutions and suffer from some weaknesses. However, short of a better solution and of further work on the risks associated with the setting of universal standards, it remains the least bad option.

BIBLIOGRAFÍA


Philosophy and Sanskrit (2002).