Anáhuac México

Inaugural Lecture

The Ethics of AI in the Care for Life

Mons. Renzo Pegoraro

Chancellor, Pontifical Academy for Life

January 25th, 2023

www.academyforlife.va

Facultad de Bioética

It is a pleasure to spend this time with you. I offer you my lecture, my presentation in these topics, that is really a challenge, so we try to present something that is a kind of starting of a reflection of work in the area of ethics of artificial intelligence, particularly in the care for life.

It is really an epochal change and it comes with new challenges in the words of the Holy Father Pope Francis. The digital galaxy and specifically artificial intelligence is at the very heart of the epochal change. It is not a time of change, but is, it is really a change of this time and we need to understand this in order to accept these new challenges. And in particular, there are the issues concerning the application of artificial intelligence.

I will not enter into the explanation of what is artificial intelligence, I think that many of you know, what is this possibility of creating this high sophisticated interaction of many data, many knowledge and in which way with the production of algorithms we are able to combine, correlate, many information, in a very fast and short time.

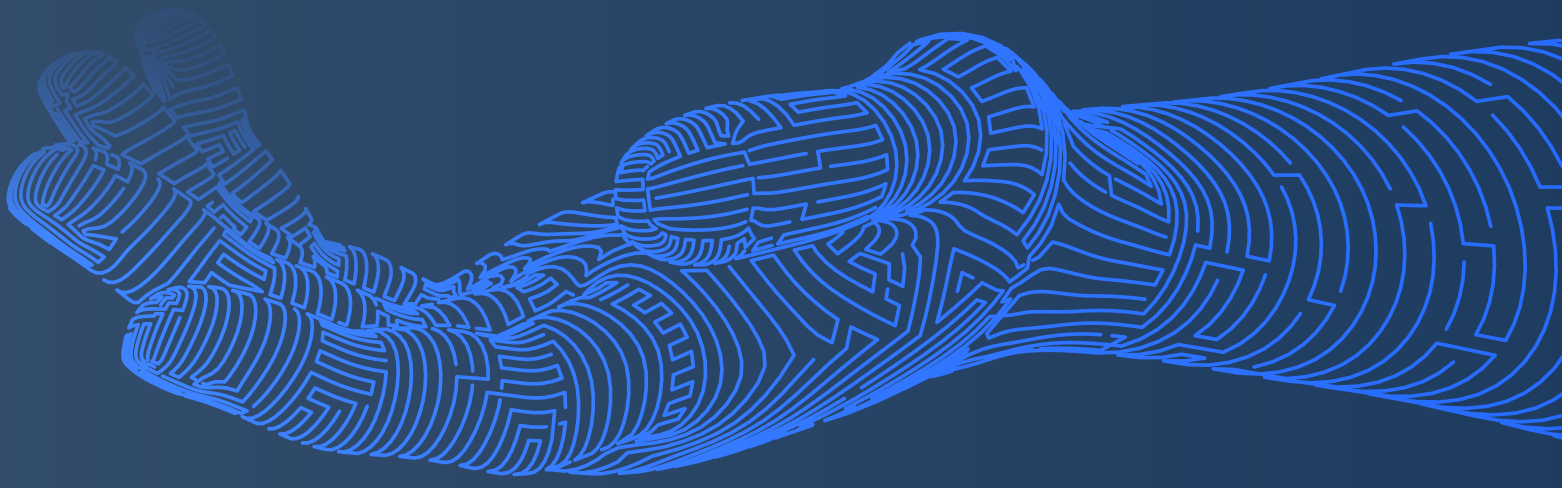
So the combination of many, many data and information, in a very quick time to elaborate a kind of combination of the information to offer some solutions; in this application of artificial intelligence, there are three areas among others that are of great importance and deserve special attention as they are particularly closer to the dimension of the person and make what is happening in the world of technology more tangible: Big data; Elderly care; and the issue of ahead and early detection of different sicknesses.

Mons. Renzo Pegoraro

So there are these three areas where we can see the positive application of artificial intelligence, but also some problems and risks, and in which way we need to elaborate and to propose ethical principles and guidelines to manage these challenges.

The first one is Big Data. In this perspective for technology, the value of a human being lies in its production of a constant flow of data of different nature and sources. We are able now to collect a lot of data for different sources: coming from healthcare system, coming from a social network, coming from many other resources and it is clear that this brings out sensitive topics, that are worth mentioning.

Confidentiality and privacy for instance. Wa ask how to maintain privacy when we manage so many data, which kind of inform consent should we be using? Who is the owner of the data? Who is in charge as the possibility to use all this data is so feasible? Another issue is the possible the bias in the use of data. There is the risk to use some data for different reasons or different goals. And so it's possible to make mistakes specially in the sensitive context of health and healthcare. New technology based on artificial intelligence, collect information about patients that is strictly confidential. And it is even more important to understand how to properly store these pieces of information; therefore we should need to ensure the absolute respect for the person from which we are obtaining the data.



The second big area is senior care, there is a special interest coming from artificial intelligence for this kind group of of people. For artificial intelligence, the elderly person is a subject to be studied in order to develop sustainable solutions. This leads to sensitive topics, such as the effort to increase people's longevity and to diminish their physical and mental disabilities.

So there are good goals to support the elderly people and to reduce physical, psychological and mental disabilities, but in developing artificial intelligence, we also need the human touch as suggested by the Japanese doctor Takai. So the danger of the personalization of medicine by the use of technology is a tangible and may cause a repercussion on the quality of care. In particular patient doctor relationship may suffer from the implementation of AI. The main issue is in which way we maintain this human touch and have the machine, the robot, the artificial intelligence, supporting doctor patient relationship and not substituting it.

The aforementioned problem is specially important in elders' care, in which patients are fragile subjects, not only because of their diseases, but also because they are a vulnerable group of population. We need a special, particular protection of elderly people, and so we must be more prudent and careful when we use artificial intelligence in this context. It is important to reaffirm the centrality of the human dignity, also, when the decision making process is automatic, and delegated to artificial intelligence.

The third area I mentioned is in early detection of medical conditions and diseases. There is a positive application of A.I. in this area. Increasing results in identifying diseases such as cancer in the early stage with relevant differences in diagnosis lead to early treatment.

On the other side, there is the risk of fragmentation, of both the human being and relationship between doctors and patients into

mini-problems, this is limited fragmented problems and the derived consequence of the difficulty to have a more holistic global approach to the patient, including psychological and spiritual dimensions of the human person. We tend to focus here on a lot of specific limited aspects of the person but we are not able to have a complete holistic approach to the person.

So in this talk we approach the challenges concerning Artificial Intelligence and we wonder which kind of ethics of the user of Artificial Intelligence should we be thinking on, especially in health care.

Three years ago in February 28th, 2020 in Rome, the Pontifical Academy for Life signed a document called Rome call for A.I. Ethics. And it is a very interesting document because it was signed by the president of the academy Archbishop Vincenzo Paglia, by the president of Microsoft, the IBM's executive vice president, and the director of the FAO Dongyu Qu as well as the Italian minister of government for technological innovation and digitalization.

It implies the agreement to see which way we can offer some criteria and ethical principles in the use of Artificial Intelligence, and it is interesting that the anual AI index report, published in March 2021, by Stanford University, identified the Rome Call for AI Ethics, as one of the five most significant new stories in 2020 in the field of ethical use of AI.

Just this year, in January 10th, three representatives of the three Abrahamic religions: Christianism, Islam and Judaism, signed the Rome Call for AI Ethics along with Mons. Vicenzo Paglia, President of the Pontifical Academy for Life, Chief Rabbi Eliezer Simha Weisz, member of the council of the chief rabbinate of Israel, and Sheikh Abdallah bin Bayyah, head of the Abu Dhabi forum for peace and chairman of the United Arabic Emirates Fatwa Council.

This is interesting because of how and in which way, the main religions can contribute to define ethical criteria for AI, and they can collaborate with two very important and relevant companies as Microsoft and IBM to create a culture in the use of AI for Ethics, the respect of human dignity and the promotion of justice and solidarity.

The Rome Call for AI Ethics includes three sections: Ethics, Education and Rights. About Ethics it reminds that all human beings are born free and equal in dignity and rights, therefore AI systems must be conceived, designed and implemented to serve and protect human beings and the environment in which they live.

It is important to think about in which way we create and design the AI systems, in which way we collect the data and we create the algorithms and in which way we use the algorithms. Technological advances and AI must include every human being. discriminating against no one, it must have the good of humankind and the good of every human being at its heart, finally, it must be mindful of the complex reality of our ecosystem and be characterized by the way in which it cares for and protects the planet.

Concerning Education it reflects upon the fact that transforming the world through the innovation of AI means assuming the responsibility to build a future for and with younger generations. Then we should always ask In which way we can help support younger generations in the future. The impact of transformations brought about by AI in society. work and education has made it essential to overhaul school curricula in order to make the educational motto "no one left behind" a reality. So we need to rethink and innovate something about our school curricula at all levels, to educate the people, to understand and to manage the risks AI imply for everybody.

The use of AI must follow forms of action that are socially oriented, creative, connective, productive, responsible and capable of having positive impact on the personal and social life of younger generations. It is also important helping also people with disabilities and elderly to learn and allow then to integrate A.I. in their daily lives.

Finally, concerning Human Rights, the development of AI in the service of humankind and the planet must be reflected in regulations and principles that protect people, particularly the weak, impoverished and the underprivileged ones and also protect natural environments.

The ethical commitment of all the stakeholders involved is a crucial starting point. Values, principles, and in some cases, legal regulations, are absolutely indispensable in order to support, structure and guide this process. It is very important to define some rights and to create a condition to support, regulate and improve the respect and the concrete practical guarantees of these interventions.

To achieve these objectives, we must set out from the very beginning of each algorithm's development, an *algorethical* vision, an approach of ethics from its design.

Designing and planning AI systems that we can trust, involves seeking a consensus among political decision makers, institutions, universities, NGO organizations, religions, churches; all people are invited to cooperate and create this consensus and supporting algorethics, from the starting point of design.

The expression of *algorethics*, comes from Pope Francis; in fact, in that meeting of February 28th of 2020. In order to steer AI's challenges towards respecting the dignity of every human being, the Rome Call proposes an *algorethics*, that is, an ethics of algorithms, not as an instrument of restraint but to provide direction and guidance. *Algorethics*, in the words of Pope Francis, is aimed at ensuring a competent and shared review of the processes by which we integrate relationships between human beings and today's technology. In our common pursuit of these goals, a critical contribution can be made by the principles of the Catholic Church's social teaching: the dignity of the person, justice, subsidiarity and solidarity. So the Church with its long tradition of social teaching can offer an important contribution to develop an *algorethics* for now and the future.

Following the Rome Call, I can present six principles that are contained in it:

1. Transparency. AI systems must be explainable, and now there is the problem that many algorithms are black boxes, so we don't know exactly in which way they created the algorithms and organized the data for this system. Everything must be clear so that others can learn and design it too.

2. Inclusion. The needs of all human beings must be taken into consideration so that everyone can benefit and all individuals can be offered the best possible conditions to express themselves and develop. So we must be careful with the risk that some people are not really present and included in this systems, and we want to offer this approach and concrete elaboration of the algorithms with this criteria.

3. Responsibility. Those who design and deploy the use of AI must proceed with responsibility and transparency. We need a good dialogue and a good education for the people who work in this area and collaborate, to design and deploy the use of AI with responsibility and transparency.

4. Impartiality. Do not create or act according to bias, thus safeguarding fairness and human dignity.

5. Reliability. The idea of trusting of the systems and that we are able to manage it with this approach. AI systems must be able to work reliably.

6. Security and privacy. AI systems must work securely and respect the privacy of all people. The people involved as source of data and the people who are involved in the use of algorithms must be fully respected and taken into consideration always.

So far, I tried to go a little deeper in the contributions coming form the Catholic Church, I remember the principle of dignity of the human being, solidarity, justice, but I focus a little bit more on an anthropological approach. The anthropology is very important in the relationship between human beings and machines, or intelligent learning machine. And there are two references that you can find, this article with Sinibaldi and myself as authors, *"Contributions form the Catholic Church to ethical reflections in the digital era"* it was published at Nature Machine Intelligence. And the second one form the Pontifical Academy, *"The good algorithm? Artificial Intelligence, Ethics, Law, Health"*.

For Christian anthropology, there are seven issues that I briefly present. For Christian Anthropology, the human body is a three dimensional entity, biological, psychological and sociocultural. And that is important because with AI there is the risk to to loose the idea of the unity of the person, because the machines and AI, don't have body, and there is the risk to reduce the body to one dimension, the biological and forget the complete and integral view of the human body.

Secondly, Christian anthropology also fosters the debate on machine agency. What is really the work, the activity, the kind of ability of the machine for it to be intelligent. We need to clarify that the machine is not an artificial human being, it is what is is: a machine, with a lot of possibilities, strong, relevant and high level of capacities and abilities, but is is mandatory to maintain the clear distinction between the machine and the human person. It can be useful to say that human beings have the distinctive capacity to question the criteria and principles on which they make decisions, being capable of critical self reflection and ethical decision making is inly for humans, machines can not accomplish this.

So what is specific of the human being is the capacity of critical self reflection. capacity of decision making with the values and ethical principles, we call this area moral conscience. In the other hand, autonomy in the ethically relevant sense, can only be attributed to human beings, due to the underlying cognitive processes and pre eminently identified with the dignity of human persons and human agency. So the ethical autonomy and the capacity of discernment of decisions are specifically of the human person, the machine can calculate and offer solutions but it is not the same process of moral discernment and moral decisions.

Christian anthropology articulates a vision of human beings called to cultivate, develop and increase creation, establishing a future oriented ethics open to and responsible for human development. Such ethics promote an attitude towards science and technology that is fundamentally confident and welcoming of innovation. So our ethical approach is open to the future, to innovation and therefore, we try to walk together with science technology but there is the risk that sometimes science and technology move very fast and ethics walks slower than them. The concern, is, then, in which way can move ethics along with science and technology.

Another important consideration is that from Christian anthropological reflection, human experience, including intelligence, cannot be reduced into categories fully accessible by machines. Furthermore, the use, in narratives about machines, of language directly taken from human experience, should be carefully reconsidered, so we must be careful in talking about intelligence, because it is a more logical combination and correlation of data, the machine is not explaining but describing a correlation, it has no memory but record, no decision but a logical calculation. Sometimes we use a metaphorical language, and we apply

to the machine experiences that only humans can have, the machine doesn't hear or see but is able to register the essentials, it does not have capacity of perception, feelings or emotions.

The Church is challenged by scientific and technological innovation, yet it can help foster a new alliance to respectfully listen and integrate multiple voices, keeping in mind that voices are necessarily plural. Yet, the human family, our common good, our common home, is one. The challenge is in which way can we stress that we are in the same family and we work together, using technology for our future in our common home.

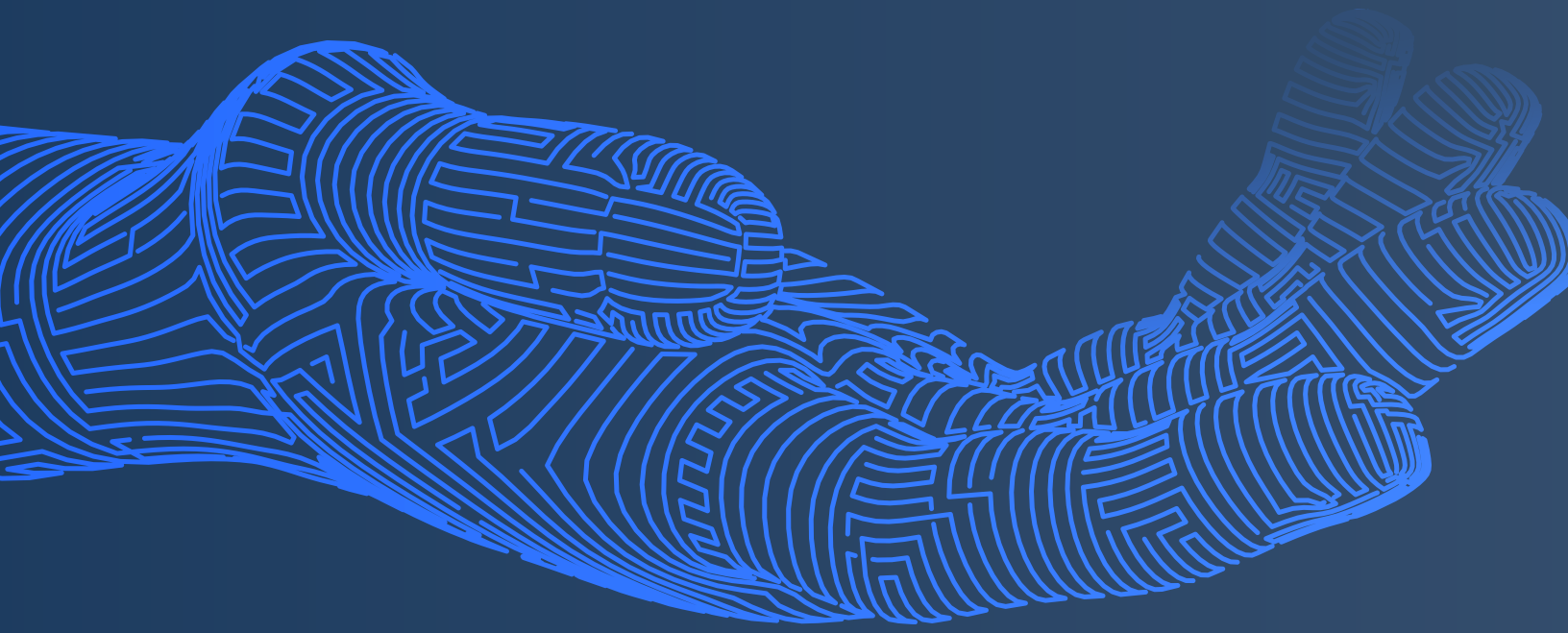
And I conclude with the reminder that there is an interesting contribution coming from the group of the G20, that had a meeting in 2021 in Italy with focus on human centric Al. In the second paragraph of the abstract it mentions that, "to achieve this, we recommend an ethics-by-design methodology for adoption by G20 Governments, which can drive the design, development, and deployment of trustworthy Al/digital ecosystems in each State.

This methodology should be harmonised and supported by a multinational approach inspired by the recognition of trustworthy Al as a common good. "It is interesting that we can find many words and concepts coming from Christian Anthropology and the teaching of the Catholic Church to have an ethics by design, to guarantee a common good, and to offer a system that can guarantee human centric Al, with the human being at the center of al technologies and development of Al". https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf

Digital innovation and technological progress must contribute to human development and benefit, with respect for society, especially those potentially left behind and the environment. Potential benefits, like those being investigated, for example, in healthcare in terms of assistive technologies, therapeutics and restored functions, must be implemented through equitable applications. Global cooperation and inclusive dialogue, also encompassing religious denominations at large must be fostered, also through scientific events and journals.

To paraphrase the Gospel, it is not man for the machine, but the machine for man.
Therefore, the ethical commitment of all those involved, scientists, philosophers and theologians, the Catholic Church, health workers, and many other people, is to promote an AI for the respect, promotion and care of human life from its beginning to all stages of growth and of course in the end of life.

Special attention must be paid to the most vulnerable: to nascent life, to the disabled and the elderly. The knowledge, education and training of AI users, students, teacher, doctors and nurses, especially in the healthcare system, is crucial. We need to inform the people and educate everybody on what is the real meaning of AI and in which way to use it: this is that machines and the AI are for the human beings not the human beings for the AI.



UNIVERSIDAD ANÁHUAC MR

VINCE IN BONO MALUM

MÉXICO

Facultad de Bioética