CENTESIMUS ANNUS FOUNDATION
INTERNATIONAL CONFERENCE
"GENERATIVE ARTIFICIAL INTELLIGENCE".
SPEECH BY CARDINAL PIETRO PAROLIN, SECRETARY OF STATE
CLEMENTINE HALL
SATURDAY 22 JUNE, 11:30 A.M.

Madam President,

Dear friends,

I extend my warmest greetings to you and sincerely thank you for your work on the occasion of this International Conference promoted by the Foundation "Centesimus Annus pro Pontefice". Your commitment to an in-depth study of the Catholic Social Doctrine is a precious contribution to the mission of proclaiming the Gospel of the whole Church. Listening to the Word of God requires a constant readiness to listen to the questions and searches for meaning of contemporary humanity, and to give reasons for faith in the face of an ever-changing reality.

One of the challenges posed by current events is undoubtedly the development of digital technologies based on Artificial Intelligence, the rapid spread of which is penetrating everyday life and influencing individual habits and social behaviour.

Its impact on the globalisation of the economy, labour markets, political policies and cultural expressions is becoming increasingly clear, so much so that we are facing not just an era of change, but a true paradigm shift.

There is a profound change in the way man understands himself, in how he reads the present and imagines the future, generating a narrative of his being thrown into the world that claims a discontinuity with the philosophy of life inherited from the humanist tradition.

Compared to the great technological innovations of the past - such as the invention of the printing press, the steam engine or the automobile - the 'algorithmic revolution' seems to impose a radical paradigm shift: whereas the artefacts previously constructed by humans served to transform the physical

world, the new information technologies process an immaterial reality, namely information. By intervening in the processes of knowledge production, storage and management, they can have a significant impact on the way human intelligence develops. By externalising cognitive functions, such as memory, and extending mental capacities, such as perception, they blur the boundary between nature and culture, giving rise to a true 'digital environment'.

A "*structural change in the experience of the self and its relations with others*"[1], as Romano Guardini had already pointed out in the first decades of the 20th century, foreseeing how the combination of technology and power would lead to a complexification of ethical questions. Man today, the theologian lucidly observed, finds himself in the position of being both the subject of technological innovation and the object of its unpredictable consequences: "*He has power over things, but not yet over his own power.*"[2]

A kind of strange paradox lies at the root of an "*existential overload*"[3], of a widespread and vague anxiety about the future: if, on the one hand, confidence in technology is growing, to the point of attributing to it an almost salvific value *and "assuming that every future problem can be solved by new technological interventions"[4],* on the other hand, there is also a growing fear that machines may escape human control.

This is not to disregard the important achievements of humanity, nor to deny the positive effects of technological progress, but rather to recall the urgency of monitoring its progress so as not to give in to an uncritical logic that Pope Francis stigmatises as "*fatal pragmatism*"[5].

Understanding the transformations that are already taking place before our eyes, as a result of their rapid development, is a necessary act if we want to manage the AI revolution in a socially just and ecologically sustainable way.

----

[1] R. GUARDINI, *La fine dell'epoca moderna*, Morcelliana, Brescia 1993, 62.
[1] *Ibid*., 87.
[1] R. GUARDINI, «Europa – Realtà e compito», in ID. *Ansia per l'uomo*, vol. 1, Morcelliana, Brescia 1969, 281-282.
[1] FRANCIS- *Apostolic Exhortation- Laudate Deum*, Rome, 4th October 2023, n. 57.
[1] FRANCIS, *Laudate Deum*, 57.

This is the question addressed in the message for the 2024 World Day of Peace, in which Pope Francis urges us to acquire more tools - both individual and collective - to manage the influence that AI is already exerting on us today. The main objective is to launch a common exercise of reflection and discernment on the ethical consequences of this new frontier of human action, with a view to building a future of justice and peace for future generations.

1. Artificial Intelligence: the current scenario, between research and experimentation

To date, there is no single definition of Artificial Intelligence (AI), although the term has now become part of everyday language. In its broadest sense, Artificial Intelligence refers to the set of theories and practical techniques aimed at developing computer systems capable of imitating the higher cognitive functions of humans: learning, memory, deduction, inference.

AI is currently undergoing a phase of rapid evolution, marking the transition from *machine learning*, where the aid of programming and instruction by researchers was still fundamental, to what is known as *deep learning*, where the emphasis is on the ability of algorithms to acquire data autonomously and progressively optimise their performance.

In view of the wide range of models, tests and objectives for AI, it is preferable to speak of it in the plural, in order to emphasise "the unbridgeable gap that exists between these systems, however surprising and powerful, and the human person [6]".

A brief overview of the diversity of AIs' modelling may be useful to understand that they are only a reflection and fragmentary expression of human intelligence.

----

6. Ibid

In fact, it is possible to identify four subsets of AI: *predictive, anomaly-based, decision and generative.*

Predictive AI is when a machine is capable of making predictions based on a combination of past inputs and the analysis of current scenarios. Using statistical algorithms and active lines of *machine learning*, predictive AI analyses historical data to identify the recurrence of patterns or trends that can be applied to other data that have not yet been analysed. This type of AI is already widely used in the stock market, medical diagnostics and to predict consumer behaviour. However, basing predictions solely on the assumption that future patterns will necessarily follow those of the past has a clear limitation. When faced with an unexpected turn of events or a sudden change in data, any kind of prediction would be unreliable.

Anomaly-based AI is trained to recognise regularities and identify discrepancies in a pattern. It is particularly useful in security because traditional methods of cyber defence, which rely on the use of pre-defined rules and patterns, are proving insufficient against dynamic threats. Through advanced analysis and detection techniques, AI systems can identify malicious activity, adapt to it in real time, and provide proactive defence tools.

Decision AI extracts information from large data sets and classifies it according to characteristics specified by the programmer. By linking the processed data to actual outcomes, it makes suggestions to guide decisions and provides useful feedback to verify its effectiveness. However, as the volume of decisions increases, it becomes necessary to entrust machines with a greater executive role. The quantitative aspect is essential to understand the scale and inevitability of algorithmic decision-making autonomy. If we were to imagine that all the data produced on an annual basis were evenly distributed among all the inhabitants of the planet, we would realise that we are already in the order of several tens of gigabytes per capita per day.[7]

----

[1] https://www.statista.com/statistics/871513/worldwide-data-created/ ↑

However, automated decision making implies a significant reduction in human involvement and the consequent delegation of the execution of autonomously performed tasks to machines. This poses an *accountability* issue, i.e. it makes it difficult to determine who should be held accountable.

In his recent address to the G7, the Holy Father expressed his concern about this type of AI, stating that there is a fundamental difference between algorithmic choices and the human capacity to decide: "*A decision is what we might call a more strategic element of a choice and demands a practical evaluation. At times, frequently amid the difficult task of governing, we are called upon to make decisions that have consequences for many people.*"[8] Human decision-making involves the exercise of wisdom, because it involves reasons that may be excluded from automated calculation, such as the pursuit of the common good and the affirmation of the inalienable value of every human life. Discernment and judgement are human acts that go beyond the mere process of decision-making by a machine.

Finally, *generative* AI uses advanced *machine learning* techniques, such as neural networks, to identify underlying patterns and relationships within a dataset in order to produce outputs that are similar in style and structure to large input datasets. In short, neural networks are trained to produce new examples that mimic the training inputs. However, it must be said that the outputs they produce are not really 'original', as generative AI '*searches the large data for information and packages it in the required style*'[9]. It is on this point that we will now focus our attention in order to better understand the risks and opportunities they present.

----

[1] FRANCESCO, *Address by His Holiness Pope Francis at the G7*, 14th June 2024.
[1] *Ibid*.

2. Generative AI: risks and opportunities of a new frontier of human action

The wide range of applications that AI opens up shows a potential that has yet to be explored, and suggests that it can play a crucial role in improving the living conditions of individuals as well as of entire nations. The results already achieved in areas such as logistics, transport, communications, business, education, etc. are remarkable. In the field of pharmaceuticals, the collaboration between human and artificial intelligence has led to discoveries that scientists could not have made on their own. For example, MIT researchers in Boston were able to synthesise a new antibiotic - halicin - that can kill strains of bacteria that were previously resistant to all existing drugs.

The Message for the World Day of Peace 2024 mentions how the use of AI, if directed towards integral human development, "could introduce important innovations in agriculture"[10]. Through the use of remote devices, the monitoring of food crops could optimise harvests, enabling entire populations to increase their domestic production and thus counteract famine.

At the same time, however, there is an ambivalence to AI that needs to be addressed. If we look at the use of these technologies in other areas, such as the armaments sector, the possibility of remote military action could have a devastating impact on humanity. It would make it more difficult to contain conflicts and ensure security relations between countries, increasing the risk of escalation and death.

As the use of AI increases, so do concerns about privacy. Generative models may inadvertently acquire sensitive or personal information from training data (data breach). The low cost of storage means that data may be retained longer than necessary and reused for purposes other than those intended.

A serious concern is also the use of generative AI with the intention of creating illegal content, such as *email phishing*, forged documents, identity theft.

---

[1] FRANCIS, *Message for 57th World Day for Peace,* 1st January 2024, n. 6.

Another critical issue is the potential for AI to generate content that infringes intellectual property rights (copyrights, trademarks or patents).

The opportunities, but also the risks, arising from the proliferation of these devices are so great that the European Union has decided to draw up specific regulations on the subject. *The AI Act* establishes a single legal framework to regulate the development, marketing and use of AI systems, based on a classification of four risk levels (unacceptable, high, limited, minimal or no risk).

Companies with a low risk level will only be obliged to inform users about the functioning and use of their product, while services classified as high risk, which include those offered by most generative AI systems, will be subject to stricter prohibitions. In particular, content generated by cybernetic systems that alter reality - photos, video, audio - will have to be labelled to allow users to identify its origin. All types of AI classified as unacceptable risk, such as predictive policing tools, biometric recognition systems, emotion decoding and social scoring, will be banned. It should be noted, however, that most companies developing generative AI technologies are not based in Europe and are therefore not subject to the provisions of the AI Act. There is therefore an urgent need to reach an international agreement, at least in the form of a voluntary code of conduct, which would set out consistent criteria for companies in the sector.

In terms of social impact, AI's main challenges are in the fields of labour and environment.

One of the most worrying changes at both European and international level is the polarisation of the labour force. This is a trend that signals an increase in demand for highly specialised jobs and a decrease in demand for low-skilled jobs. It is expected that AI will assist humans in complex, multi-tasking jobs with significant productivity gains. In low-skilled jobs, AI will guide the human worker through specific tasks, simplifying procedures but making processes less transparent. As the level of human training declines, companies could more easily replace their workers and thus reduce production costs by

lowering wages. To counteract *the skills mismatch*, it is necessary to design policies that reintroduce values such as solidarity into the social fabric by investing in training so that those workers most affected by AI can make an effective career transition. At the same time, it is also necessary to support entrepreneurship by encouraging the new forms of mutualism enabled by digital tools, such as collaboration platforms.

However, in terms of protecting and caring for the 'Common Home', the success of AI could come at a very high price in terms of its impact on renewable resources. The mining of lithium, copper, coltan and other rare – earth metals is of great concern because of the threat it poses to biodiversity and the survival of local communities. The entire process of locating and drilling mines, extracting, transporting and processing raw materials can have irreversible effects on the environment. The use of heavy and toxic metals as smelting agents raises the problem of waste disposal. In fact, waste is often simply dumped into the water, polluting rivers and streams.

Water protection is also at the heart of issues directly related to the rise of digital technologies. In fact, the training and execution of AI models on a large scale involves not only a huge amount of energy, with a consequent increase in CO2 emissions, but also a consumption of water that may not be sustainable in the medium-long term. To give a concrete example, the researchers calculated that training ChatGpt-3 used seven hundred thousand litres of fresh water to cool Microsoft's modern data centres.

Tracking the water footprint of IT giants should be mandatory. Insisting on transparency is the first step towards responsible use of resources and planning for sustainability goals. A priority should be to replenish more water by 2030 than has been used so far.

3. Conclusions

The immense growth of technology needs to be accompanied by a proper education in accountability for its future developments. No technology is inherently 'neutral' in the sense of being culturally indifferent, disembodied. It

is a fully human product whose orientations are made up of choices conditioned by individual, social and cultural values, from one generation to the next.

Pope Francis has identified two fundamental and inalienable criteria for evaluating new technologies: respect for the dignity of every human being and fraternity among all men and women. He states with conviction that "*A technological development that does not lead to an improvement in the quality of life of the entire human race cannot be considered real progress.*" [11] Otherwise, if digital technologies were the prerogative of a small section of humanity, inequalities could grow disproportionately: not only wealth, but also the power that comes from possessing these new forms of knowledge could be concentrated in the hands of a few. AI would become another powerful tool, reinforcing and consolidating the dynamics already at work in the technocratic paradigm. It is therefore important to ask what forms of power AI simplifies, reproduces or enables, what interests it promotes and who is accountable for it (accountability).

The algoretic approach, which sees artificial intelligence as a moral agent and outlines an ethics of intention, must also be complemented by an ethics of consequence, i.e. one that judges the morality of an action by the result it produces. In order to improve algorithms, it is necessary to develop a socio-technical framework in which projects are carried out in a way that not only promotes beneficial goals, but also aims to achieve this result in a socially just and sustainable way.

This is the challenge before us, especially for Christ's faithful: to discern, in this age of transformation, the right moment to redirect humanity's path towards the Lord and towards others. To commit ourselves to rediscovering our common belonging to the Common Home, striving to build a more just

----

[1] FRANCIS, Address of His Holiness, Pope Francis to the 57th World day for Peace, n.2.

and fraternal world. It would be an oversimplification, however, to reduce this change of pace to the technical and economic dimension alone, since legislative measures and political decisions, however necessary they may be, are not sufficient to ensure that these new technologies are placed at the service of fraternity and peace.

Faced with a revolution that is already in our homes and that will profoundly change the way we live, the way we inform ourselves, learn, communicate or spend our leisure time, we need a major investment in education, for all ages, to understand how these machines work, what we can expect from them and what we need to be critically vigilant about.

We need basic literacy, because only a sufficiently broad and shared wealth of knowledge can create the cultural climate that will make regulations more effective and, above all, influence the actions of individuals. We need to promote a culture of care[12] that counters the throwaway culture[13], the belief that sacrificing some of humanity - perhaps by reducing labour costs[14]- is an acceptable price to pay for technological progress . This is a necessary step to decolonise the popular imagination from the salvific narrative that attributes to machines the power to solve all our problems, and from the myth of '*infinite or unlimited growth.*'.[15]

Between the apocalyptic scenarios that imagine dystopian futures and the salvific conceptions that see technology as the solution to every problem, there is the middle ground: using AI to improve the conditions of humanity, provided that it is subject to an appropriate form of control at the cultural, social and political levels.

This is not to demonise AI, but to emphasise that these machines must remain human-centred and defend human rights, to prevent '*the uniqueness of the person being identified and reduced to a set of data*'.»[16].

---

[12] Cfr. FRANCIS, *Encyclical Letter Laudato si' on the care of the Common Home*, 231.
[13] FRANCIS, *Address to the members of the diplomatic Corps accredited to the Holy See,*13 January2014.
[14]¹ FRANCIS, *Encyclical Letter Fratelli tutti*, 20.
[15]¹ FRANCIS: ENCYCLICAL LETTER *Laudato Si'*, 106.
[16] *Ibid.*, 5.