

Open Letter to the Cardinals of the Catholic Church

Subject: The Urgent Need for Regulation of AI Technologies

Your Eminences,

On the eve of the election of the future Pope, we take the liberty of writing to you as committed Catholic faithful to draw your attention to a matter of vital importance to our modern society: the regulation of artificial intelligence (AI).

Nearly three years ago, humanity entered a new era, that of mainstream generative AI. These new technologies are not merely another advancement; they mark the beginning of a new paradigm, that of autonomous machines. These machines are now capable of learning by themselves to make complex decisions which, until recently, only humans could make through their intelligence and *conscience*.

Antiqua et Nova, published on 28 January, is a highly commendable document which reaffirms the distinction between human beings and machines in the face of the boundaries being pushed by AI. It rightly reminds us that human beings, unlike AI, must not be reduced to merely functional entities. It underlines the unique intimate character proper to humanity, which enables us to act truthfully, to form genuine relationships, and to live freely and fully in all dimensions of our humanity. In this regard, the document provides deep guidance for all who seek to understand the limits of AI.

However, the note raises a fundamental question that it does not answer: *"As AI models become increasingly capable of autonomous learning, the possibility of exercising control over them to ensure they serve human objectives may diminish. This raises the crucial issue of how to ensure that AI systems are directed towards the good of people, and not against them."* (§45)

The progressive loss of control over these tools is a reality acknowledged by experts in AI safety. This observation stands in stark contrast to the appeal in the following paragraph: *"the algorithms that govern [AI] it should be trustworthy, secure, robust enough to handle inconsistencies, and transparent in their operation to mitigate biases and unintended side effects."* (§46) Although AI systems are trained using relatively simple algorithms, the final product of this training does not itself constitute a comprehensible or controllable algorithm. At present, it is impossible to claim that such systems are "safe", let alone "transparent in their operation".

To this day, very few resources in the field of AI are devoted to enhancing our understanding of these systems. On the contrary, enormous financial and human resources are being deployed to increase their autonomy. AI is improving day by day, autonomous agents are becoming more widespread, and autonomous military weapons are also advancing rapidly without stringent oversight. Finally, the race for artificial general intelligence (AGI—an AI vastly superior to human intellectual capabilities) is now well under way. All these developments tend to diminish the role of humans in this world actively fostering our total dependence on this technology, which will likely lead to a loss of control over it.

While *Antiqua et Nova* rightly reasserts that AI “intelligence” is not equivalent to human intelligence, this does not diminish the danger posed. The purely reflexive and logical nature of AI does not negate its capacity to self-develop, replicate, escape our control, and potentially act against humanity.

AI safety experts are unanimous: the loss of control over AI, to humanity's detriment, in the coming years is highly probable. The risks of AI escaping human control and turning against us are far from negligible. Some experts estimate the probability of this occurring within the next two to three decades at around 10%. Two of the three researchers responsible for the foundational work on large language models—Geoffrey Hinton and Yoshua Bengio, recipients of the 2018 Turing Award—are among those sounding the alarm.

To date, the urgent call to regulate AI has gone largely unheard. The media have failed to relay research findings which all point in the same direction. It seems that the pressure of economic stakes and the power associated with AI are too great for any appeal to slow down or regulate this new industry to bear fruit.

At this very moment, a global race for AI supremacy is under way between nations, in which ethical considerations are being sidelined for fear that an adversarial state might develop a superior AI. Indeed, such a tool would grant the nation that possesses it strategic power which could threaten the security of others. In this context, finding common ground between nations has become extremely difficult. Reaching a binding agreement on AI regulation appears increasingly impossible. Humanity seems left to its own devices.

It is for this reason that we cry out to you, Your Eminences, through this letter. At this time of conclave, we seek to draw your attention to the particular challenges posed by AI, where the Catholic Church appears uniquely positioned to play a decisive role.

By virtue of its faith, its tradition, its history, and the enduring moral and philosophical principles it upholds, the Church is endowed with a special place and role in the regulation of AI at the international level. Politically, its neutrality—both diplomatically and technologically, since it does not develop language models itself—makes it the ideal actor to stand before other nations. Moreover, its active defence of the human person and of the common good, written into its very essence, grants it a legitimacy and authority in this matter that no other state possesses.

For all these reasons, we believe that the Catholic Church must take active steps—as it did during the Cold War with nuclear weapons—towards an international agreement to regulate AI for the common good. It is no exaggeration to compare our present moment with the second half of the twentieth century, as the risks to humanity are of a similar magnitude.

We earnestly implore you to give special consideration to the current context in your discernment and choice of the next successor to the Chair of Saint Peter—one who, faced with the challenges of AI, will strive, in faithful continuity with the Church's tradition, to promote the common good and peace on Earth.

In united prayer, we invoke the Holy Spirit to enlighten and guide you in this task. May your discernment be marked by wisdom and foresight, so that the Church may continue to shine as a beacon of light and hope for all humanity.

Please accept, Your Eminences, the expression of our deepest respect.