Archpriest Dr Georgios LEKKAS

Counsellor at the Representation Office of the Church of Greece to the E.U. (Brussels)

ARTIFICIAL INTELLIGENCE AND AN APPLIED ETHICS. BUT WHAT KIND OF ETHICS?

We warmly welcome the initiative of the European Commission and the High-Level Expert Group on Artificial Intelligence to hold a public debate around the *Draft Ethics Guidelines for Trustworthy AI* which should be observed by designers of artificial intelligence systems for a 'credible artificial intelligence *made in Europe*'.

The ever more advanced systems of artificial intelligence, which inevitably promote an implicit ethical view, will through their repeated use become a means whereby generations of European citizens are educated. For this reason this is indeed a commendable attempt by the European High-Level Expert Group on Artificial Intelligence to set out in a Statement of Principles the ethical values and principles that the designers of primarily European artificial intelligence systems should establish and promote.

The authors of this draft have rightly undertaken to derive a system of ethical principles and values from the current legal framework on fundamental human rights, as stated in the EU Treaties and the European Charter of Fundamental Rights, and then to advocate its implementation by everyone involved in the operation of artificial intelligence systems.

However, the European legal system on human rights can be interpreted either within the framework of an individualistic morality, designed to protect the individual from the society in which he lives, or through a collective ethics by which fundamental rights are recognized in every human being as a necessary prerequisite that permits him to live in a society of peace and love with everyone else (= the social principle).

The proposed European draft ethics for artificial intelligence presupposes the conception of the human being as an autonomous rational and free entity, who is obliged to engage with his or her counterparts only in order to serve his or her own complex social needs. However, this proposed draft of the ethics, based on the philosophical model of a human as the self-referential being par excellence, being autonomous and ontologically sufficient in itself, extends and applies to the field of artificial intelligence - and before long also autonomous artificial intelligence - the individualistic conception of the human being whose predominance on a global level has already produced devastating effects for every individual human being who faces the spectre of isolation from others as well as the destruction of our planet.

Is the idea of the 'autonomous' human being to which the draft under discussion often refers adequate for an 'human-centric' approach to artificial intelligence? As an Orthodox Christian my answer is 'no'. The experience of two thousand years of Orthodox tradition says that human beings are not merely autonomous rational entities who relate to others out of the need to survive, but free and intelligent loving hearts which by virtue of their own nature require others in order to be free. From this perspective, the others (our Creator-God, fellow human beings, the Cosmos) are necessary for my freedom, simply because without any of them I will not have any choice at all. In

Orthodox tradition, human beings are not considered to be units that need to coexist peacefully within society merely because this serves the separate individuality of each of them; instead they are conceived of as members of a common body where the condition of each member necessarily affects the health of the whole body and the health of the whole body has beneficial consequences for the proper functioning of each member - that is what I mean by the term "social principle".

For Orthodox tradition, human beings are free to think or not, they are free to love or not, to act or not, but their freedom cannot be formulated in algorithmic terms because it exists prior to reflective thought, since at heart it is a consequence of the *ex nihilo* creation of mankind, that is, the origin of mankind solely and exclusively in the free creative will of our Creator-God. That is why the human being is not in danger of being destroyed by autonomous systems of artificial intelligence, from which human intelligence seeks to protect itself through Guidelines such as the one we are now discussing. Since the mystery of human freedom - but also of the dynamic entity that constitutes the human being - is hidden in our deep and ontological relationship with others, and especially with our Creator-God, human beings are only in danger of being destroyed by themselves. For the possibility of such an outcome being ruled out today, it is not enough to draw up ethical guidelines for artificial intelligence; we must, among other things, immediately criminalize investigations into the merging of human capabilities and machines as crimes against humanity which must be punished by the gravest of the penalties provided for in our European legal system. We consider it equally urgent that legislative initiatives are ratified within the EU to prevent the anthropomorphic simulation of artificial intelligence systems to such an extent that it becomes difficult to discriminate between - or even establishes a societal belief in the equivalence of - human and machine.

The question is not, however, how to prepare ourselves to resist the impending autonomous systems of artificial intelligence, but how to use them in the service of our ontological interrelation with everyone and with everything (our fellow humans, the Cosmos, our Creator-God) in order to attain a bliss from which the machine is excluded, by virtue of its nature. We are human beings - *anthropos* is the Greek word from which the English words *anthropology* and *anthropological* are derived - because we have been made to live in relation with others, first with our Creator and God – *ano* as a prefix of the Greek word *anthropos* means someone or something which is higher than we are - and then with all the other human beings, brothers and sisters, regardless of colour, race or religion. Therefore, in accordance with such an understanding of the human as an essentially relational being, we have been created to love in freedom and with all our heart and with all our mind our Creator-God, as well as to cherish in freedom all other human beings, just as we should love ourselves - the two prerequisites of our bliss.

If human happiness presupposes a deep and lasting association with others, the recognition of the 'social principle' as a necessary principle for the operation of artificial intelligence systems is essential. Such a principle dictates that the operation of these systems ALWAYS serves, in the short or long term, the ontological need for a deep coexistence between all human beings within the single body of mankind, otherwise any such systems will be rejected, since each time the relational value of the human being is put in danger, the human being risks, willingly or not, being turned into something far inferior to man, into a beast or a man-machine. The technical and non-technical methods which need to be called upon for the application of the social principle for the design and operation of artificial intelligence are the task of the scientific community. Nevertheless, the question of what sort of ethics is applied in the area of Artificial Intelligence must be a decision

arrived at through the broad consensus of civil society and its organizations and the cooperation of intellectuals and scientists – for this reason we warmly welcome this debate as a necessary step in the right direction - so as to avert the risk that the individualism which Europe has inherited from the previous century is placed on a pedestal, and results in a new situation in Europe where my other half is my robot!

Brussels, 21.1. 2019