Rome Call for AI Ethics

Introduction

"Artificial Intelligence" (AI) is bringing about profound changes in the lives of human beings, and it will continue to do so. AI offers enormous potential when it comes to improving social coexistence and personal well-being, augmenting human capabilities and enabling or facilitating many tasks that can be carried out more efficiently and effectively. However, these results are by no means guaranteed. The transformations currently underway are not just quantitative. Above all, they are qualitative, because they affect the way these tasks are carried out and the way in which we perceive reality and human nature itself, so much so that they can influence our mental and interpersonal habits. New technology must be researched and produced in accordance with criteria that ensure it truly serves the entire "human family" (Preamble, Univ. Dec. Human Rights), respecting the inherent dignity of each of its members and all natural environments, and taking into account the needs of those who are most vulnerable. The aim is not only to ensure that no one is excluded, but also to expand those areas of freedom that could be threatened by algorithmic conditioning.

Given the innovative and complex nature of the questions posed by digital transformation, it is essential for all the stakeholders involved to work together and for all the needs affected by AI to be represented. This Call is a step forward with a view to growing with a common understanding and searching for a language and solutions we can share. Based on this, we can acknowledge and accept responsibilities that take into account the entire process of technological innovation, from design through to distribution and use, encouraging real commitment in a range of practical scenarios. In the long term, the values and principles that we are able to instil in AI will help to establish a framework that regulates and acts as a point of reference for digital ethics, guiding our actions and promoting the use of technology to benefit humanity and the environment.

Now more than ever, we must guarantee an outlook in which AI is developed with a focus not on technology, but rather for the good of humanity and of the environment, of our common and shared home and of its human inhabitants, who are inextricably connected. In other words, a vision in which human beings and nature are at the heart of how digital innovation is developed, supported rather than gradually replaced by technologies that behave like rational actors but are in no way human. It is time to begin preparing for more technological future in which machines will have a more important role in the lives of human beings, but also a future in which it is clear that technological progress affirms the brilliance of the human race and remains dependent on its ethical integrity.

Ethics

All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of fellowship (cf. Art. 1, Univ. Dec. Human Rights). This fundamental condition of freedom and dignity must also be protected and guaranteed when producing and using AI systems. This must be done by safeguarding the rights and the freedom of individuals so that they are not discriminated against by algorithms due to their "race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status" (Art. 2, Univ. Dec. Human Rights).

AI systems must be conceived, designed and implemented to serve and protect human beings and the environment in which they live. This fundamental outlook must translate into a commitment to create living conditions (both social and personal) that allow both groups and individual members to strive to fully express themselves where possible.

In order for technological advancement to align with true progress for the human race and respect for the planet, it must meet three requirements. It 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 characterised by the way in which it cares for and protects the planet (our "common and shared home") with a highly sustainable approach, which also includes the use of artificial intelligence in ensuring sustainable food systems in the future. Furthermore, each person must be aware when he or she is interacting with a machine.

AI-based technology must never be used to exploit people in any way, especially those who are most vulnerable. Instead, it must be used to help people develop their abilities (empowerment/enablement) and to support the planet.

Education

Transforming the world through the innovation of AI means undertaking to build a future for and with younger generations. This undertaking must be reflected in a commitment to education, developing specific curricula that span different disciplines in the humanities, science and technology, and taking responsibility for educating younger generations. This commitment means working to improve the quality of education that young people receive; this must be delivered via methods that are accessible to all, that do not discriminate and that can offer equality of opportunity and treatment. Universal access to education must be achieved through principles of solidarity and fairness.

Access to lifelong learning must be guaranteed also for the elderly, who must be offered the opportunity to access offline services during the digital and technological transition. Moreover, these technologies can prove enormously useful in helping people with disabilities to learn and become more independent: inclusive education therefore also means using AI to support and integrate each and every person, offering help and opportunities for social participation (e.g. remote working for those with limited mobility, technological support for those with cognitive disabilities, etc.).

The impact of the 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. In the education sector, reforms are needed in order to establish high and objective standards that can improve individual results. These standards should not be limited to the development of digital skills but should focus instead on making sure that each person can fully express their capabilities and on working for the good of the community, even when there is no personal benefit to be gained from this.

As we design and plan for the society of tomorrow, the use of AI must follow forms of action that are socially oriented, creative, connective, productive, responsible, and capable of having a positive impact on the personal and social life of younger generations. The social and ethical impact of AI must be also at the core of educational activities of AI.

The main aim of this education must be to raise awareness of the opportunities and also the possible critical issues posed by AI from the perspective of social inclusion and individual respect.

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 and the underprivileged – and natural environments. The ethical commitment of all the stakeholders involved is a crucial starting point; to make this future a reality, values, principles, and in some cases, legal regulations, are absolutely indispensable in order to support, structure and guide this process.

To develop and implement AI systems that benefit humanity and the planet while acting as tools to build and maintain international peace, the development of AI must go hand in hand with robust digital security measures.

In order for AI to act as a tool for the good of humanity and the planet, we must put the topic of protecting human rights in the digital era at the heart of public debate. The time has come to question whether new forms of automation and algorithmic activity necessitate the development of stronger responsibilities. In particular, it will be essential to consider some form of "duty of explanation": we must think about making not only the decision-making criteria of AI-based algorithmic agents understandable, but also their purpose and objectives. These devices must be able to offer individuals information on the logic behind the algorithms used to make decisions. This will increase transparency, traceability and responsibility, making the computer-aided decision-making process more valid.

New forms of regulation must be encouraged to promote transparency and compliance with ethical principles, especially for advanced technologies that have a higher risk of impacting human rights, such as facial recognition.

To achieve these objectives, we must set out from the very beginning of each algorithm's development with an "algor-ethical" vision, i.e. an approach of ethics by design. Designing and planning AI systems that we can trust involves seeking a consensus among political decision-makers, UN system agencies and other intergovernmental organisations, researchers, the world of academia and representatives of non-governmental organizations regarding the ethical principles that should be built into these technologies. For this reason, the sponsors of the call express their desire to work together, in this context and at a national and international level, to promote "algor-ethics", namely the ethical use of AI as defined by the following principles:

- ***Transparency:*** *in principle, AI systems must be explainable;*
- ***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;*
- ***Responsibility:*** *those who design and deploy the use of AI must proceed with responsibility and transparency;*
- ***Impartiality:*** *do not create or act according to bias, thus safeguarding fairness and human dignity;*
- ***Reliability:*** *AI systems must be able to work reliably;*
- ***Security and privacy:*** *AI systems must work securely and respect the privacy of users.*

These principles are fundamental elements of good innovation.

Rome, September 30th, 2020