



United Nations
Educational, Scientific and
Cultural Organization



International Research Centre
on Artificial Intelligence
under the auspices of UNESCO



Opinion Series Report

AI Recommendation Regional Consultation UNESCO

Note from the editor

Throughout 2020 the IRCAI programme committees authored an Opinion Series on a range of topics across AI and sustainability as part of a wider portfolio of IRCAI's international initiatives. This opinion series explores what AI currently means to researchers across the world and across a variety of disciplines both in AI and sustainable development, and was led by the IRCAI executive. The views and opinions expressed by authors are their own and do not reflect the position of IRCAI, but are simply an illustration of the various opinions reflective of the diverse initiatives that IRCAI is pursuing, from basic research to policy creation in and around AI.

As IRCAI is working from home for the foreseeable future, it means we are not organising in person events. Introducing this new quick report format has the same spirit – original stripped-down research pieces, an attempt to an intellectual dialogue – just a different space.

Davor Orlic, Chief Operations Officer at IRCAI

Abstract

In light of recent global events, the question of the proper interplay between ethics and emerging technology constitutes perhaps one of the most pressing concerns of our era. As the world's population slips ever deeper into virtual life and depends ever more on the many tools technology can offer, the moral status of these relations, as well as the technology that permits them, can seem at once mysterious and dire. Nevertheless, these ethical questions are hardly novel.

Authors

Katherine Evans, Sorbonne University, Paris
Gregor Strojín, President of the CAHAI
Emma Ruttkamp-Bloem, University of Pretoria, COMEST member

December 2020

In as early as 2001, the philosophical community had already entered into what has since come to be known as the uniqueness debate: a discussion which questioned whether the realities of emerging ICTs constituted a true expansion of moral categories such as good and bad, right and wrong; or whether instead these technologies simply provided novel variations on a familiar humanist tune. Retrospectively, the case studies which informed this debate appear almost archaic: the role of the internet in a new form of ‘cyberstalking’, or the potential of a computer virus to cause some new type of moral harm to a user. With examples such as these, it often seemed reasonable to conclude that these technologies played an expansive role in the moral agency of human beings, without going so far as to claim that this newly augmented agency presented anything truly exotic to traditional ethical thought.

the environment, and future generations.

Most acutely, these discussions exposed the true impact AI has already made on human agency, and the vast amount of conflicting and heterogeneous claims that stakeholders can be seen to hold over the future of its governance.

In one sense, the results of the consultation then appear to tell the story of humanity grappling with the collective governance of a new and powerful tool. In another sense, however, the consultation appears to point to a veritable revolution in the structure and content of human moral conduct. Never before have the interests at work behind certain forms of algorithmic practice posed such serious threats to the sanctity of human rights such as individual privacy and dignity, or to such important institutions as democracy; and thus in some ways, never before has the scalability, connectedness and opacity of an individual actor had such an

Nearly twenty years later, the question of the ethical uniqueness of emerging artificial intelligence has not only received global interest, but also requires an urgent and global response. In this spirit, the first draft of UNESCO’s Recommendation on the Ethics of Artificial Intelligence, and the Regional Consultation for the European Region (Group II) of which this report is a part, should be seen as an inter-disciplinary and multi-stakeholder attempt to provide such a response.

In effect, the results of the Regional Consultation shed much light on the ethical complexities which underpin this recommendation. Most broadly, the discussions made plain the necessary transition from a micro-level ethical analysis of AI (from the perspective of individuals acting in a specific society), towards a more macro-level appraisal of the impact AI has, and will continue to have on the planet as a whole: the connectedness of individuals, and the value of living in harmony with each other,

anonymous, yet resounding impact on individual lives. Further still, never before has technological progress so strongly required a reiteration and commitment to the full gamut of human values, and never before has a human invention required us to so seriously advocate for the intrinsic moral worth of these.

It would then appear that artificial intelligence requires us to identify and resolve wholly unique ethical problems, in ways which will likely

have a fundamental impact on what it means to be human in the virtual world and beyond. In this sense, the problem with which we are faced is less a question of ‘constraining’ technology so as to align with our common conception of human good. Instead, it becomes a question of discovering the proper guidelines --and moral bright lines-- of a new, technologically infused form of human agency, one whose scope and scale extends far past the purview of traditional moral paradigms. A new level of decisional transparency is necessary, not only among those who hold the authority to impact human life, but also amongst those machines which manage ever-increasing portions of it. Simultaneously, the users of these new technologies must be brought to understand how their data may come to define them across their lives, and must be assured a meaningful place at a truly diverse and inclusive bargaining table. In this vein, current approaches to multi-stakeholder and inter-disciplinary consultations not only provide insight into how we view technology and its impact, but may also foster

understanding as to the ‘how’ and the ‘why’ of human decision-making and its necessary evolution.

In short, a new form of moral marketplace has been ushered in by nascent artificial intelligence, posing both a challenge and an opportunity for human moral progress. For its full potential to be reached, this progress must not pass through a limited series of consultations, or a particular field or jurisdiction. It follows then, that the artificial intelligence of today is not easily reducible to the causal force or interests of any familiar target such as Silicon Valley, nor does it appear to be a force that can be countered by any one institution; if only because the nature and technical qualities of artificial intelligence are themselves ever-evolving. Instead, artificial intelligence presents itself as a medium through which a global community interacts, and the ethical guidelines which inform human behaviour within it ought to reflect this entirely unique reality.