Contributions from the Catholic Church to ethical reflections in the digital era

Edoardo Sinibaldi^{1,6}, Chris Gastmans^{2,6}, Miguel Yáñez^{3,6}, Richard M. Lerner^{4,6}, László Kovács^{5,6}, Carlo Casalone⁷, Renzo Pegoraro⁷ and Vincenzo Paglia⁷

1. Italian Institute of Technology, Genova, Italy

2. Centre for Biomedical Ethics and Law, Catholic University of Leuven, Leuven, Belgium

3. Faculty of Theology, Pontifical Gregorian University, Roma, Italy

4. Institute for Applied Research in Youth Development, Tufts University, Medford, MA, USA

5. Faculty of Liberal Arts and Sciences, University of Applied Sciences, Augsburg, Germany

6. Working Group on Roboethics of the Pontifical Academy for Life, Vatican City

7. Pontifical Academy for Life (Pontificia Accademia per la Vita, PAV), Vatican City (*)

Standfirst: The Catholic Church is challenged by scientific and technological innovation but can help to integrate multiple voices in the ongoing dialogue regarding AI- and machine ethics. In this context, a multi-disciplinary working group brought together by the Church reflected on roboethics, explored the themes of embodiment, agency and intelligence.

The fast pace of digital innovation and technological progress calls for profound ethical reflection. The ethical frameworks for artificial intelligence (AI) recently proposed by private and public actors, including, e.g., OECD and UNESCO, denote increasing awareness and commitment at the international level [ref 1]. Extended transdisciplinary discussion in scientific journals [ref 2] supports the enactment of such a process, where it is crucial to preserve solidarity among people, as highlighted by the UN [ref 3]. Moreover, the need to dynamically assess the impact of digital innovation and technological progress on society [ref 4] and the environment [ref 5] is receiving growing attention, as an integral action that is needed for sustainable development. Further, it is necessary to disclose the contextual definition of shared principles, for their common understanding to be effective and to move from guidelines to actions [ref 6].

The Roman Catholic Church is proactively joining the ongoing dialogue. In his address to the 25th Anniversary General Assembly of the Pontifical Academy for Life (Pontificia Accademia per la Vita, PAV) in 2019, Pope Francis reiterated that ‘artificial intelligence, robotics and other technological innovations must be so employed that they contribute to the service of humanity and to the protection of our common home, rather than to the contrary, as some assessments unfortunately foresee’ [ref 7]. Moreover, while underlining the qualitative distinctions between machine functional automatisms and the human prerogatives of knowledge and action, Pope Francis remarked that ‘the inherent dignity of every human being must be firmly placed at the centre of our reflection and action’.

36 **Rome Call for AI Ethics**

37 The underlying difference between human beings and machines is a key point for the ongoing
38 debate, which should be nurtured by integrating the reflections offered by religious denominations
39 at large, as fully participating in pluralist societies. In this light, the Roman Catholic Church recently
40 organized a dedicated event in the Vatican entitled ‘RenAIssance. A Human-centric Artificial
41 Intelligence’ [ref 8]. This event was accompanied by the signing of the ‘Rome Call for AI Ethics’, a
42 statement regarding the need for human-centric AI ethics principles that was drafted by an
43 interdisciplinary working group coordinated by PAV. The Call was signed first by PAV’s president
44 (Abp. Vincenzo Paglia), Microsoft’s president (Brad Smith), IBM’s executive vice president (John Kelly
45 III), FAO’s director-general (Dongyu Qu), and Minister of Italian Government for Technological
46 Innovation and Digitalization (Paola Pisano).

47 In his message to PAV’s General Assembly concurrent with the Rome Call for Ethics, Pope Francis
48 recalled how digital innovation ‘affects our way of understanding the world and ourselves’, being
49 ‘increasingly present in human activity and even in human decisions’ and ‘thus altering the way we
50 think and act’ [ref 9]. Pope Francis also noted the quest for ‘a broader educational effort’, by
51 indicating that ‘it is not enough simply to trust in the moral sense of researchers and developers of
52 devices and algorithms. There is a need to create intermediate social bodies that can incorporate
53 and express the ethical sensibilities of users and educators’. Pope Francis’s address urges to promote
54 an education that aims not only to acquire technical skills, but also to understand the responsibilities
55 involved in the development, deployment and use of technologies, in order to make demanding
56 choices, even against one’s (immediate) interest, based on solid forward-looking motivations and
57 discernment. For such an educational effort to effectively permeate society and steer digital
58 innovation, it is necessary to also foster organizations (including, e.g., NGOs, professional
59 associations, cultural and social movements, and foundations) able to effectively represent and
60 convey the multi-faceted expression of users and citizens at large to developers and decision-
61 makers. While highlighting the dangers also associated, e.g., with information asymmetry, Pope
62 Francis remarked that ‘these dangers must not detract from the immense potential that new
63 technologies offer. We find ourselves before a gift from God, a resource that can bear good fruits’
64 [ref 9].

65

66 **Embodiment, Agency and Intelligence**

67 The Pontifical Academy for Life strives to interweave faith and science/technology, so as to identify
68 paths for multiple voices to respectfully walk alongside one another. Indeed, PAV’s objective is the
69 defence and promotion of the value of human life and of the dignity of the person, and the study of
70 ethical, cultural and social aspects of scientific and technological progress is integral to its mission.
71 To support its mission in relation to topics of great interest such as robotics and AI, PAV established
72 a Working Group on Roboethics. Such a Working Group gathers experts from different disciplines
73 (including natural, social and applied sciences, philosophy, and theology) and countries
74 (representing Africa, Asia, Europe, America and Australia), who are called to enter the complex
75 ongoing debate and help identify moral values and criteria of discernment, for the good of all human
76 beings. Based on meetings held during two recent PAV workshops [refs 10, 11] and additional
77 interactions, the Working Group reflected on embodiment, agency and intelligence, in particular in

78 relation to the differences between human beings and machines, in order to bring a complementary
79 contribution to the ongoing discussion on artificial systems.

80 Current research in robotics leverage the concept of embodied intelligence (sometimes called
81 material intelligence) based on the fact that robot bodies, in terms of structure and material
82 properties, can absorb or negotiate some of the challenges of interacting with the environment, with
83 benefits on robot control as well [ref 12]. According to Christian anthropology, human beings were
84 bodily created by God in His image and likeness: human beings not only have a body, but also are
85 body, in a space-time dimension. As an incarnate spirit, the human body encompasses the place of a
86 living experience: as recognized by Pope Francis in the Encyclical Letter *Laudato Si'* (155), 'our body
87 itself establishes us in a direct relationship with the environment and with other living beings' [ref
88 13]. The indivisibility of soul and human body just recalled from Christian anthropology also finds
89 correspondence in the idea of embodiment at the base of contemporary theory within
90 developmental science [ref 14], for which the human body is a three-dimensional entity: biological
91 (body as a form), psychological (body as lived experience), and sociocultural (body actively engaged
92 with the world). When pushing the research frontier by investigating, e.g., AI methods "for
93 embodying" a specific behaviour in biological entities [ref 15], or the use of invasive devices that do
94 not just readout but also stimulate body activity [ref 16], the non-reducibility of a human being to
95 their constituent parts should be kept at the centre of our reflection and action.

96 Christian anthropology also fosters the debate on machine agency. In Christian ethics, moral agency
97 is grounded on free actions and deliberate decisions taken in conscience. Human beings were
98 created with consciousness and free will, able to also form interpersonal relationships, being aware
99 of the presence of other persons (by recognizing their consciousness and free will). Although
100 machines can relate to other entities, human beings have the distinctive capacity to question the
101 criteria and principles upon which they make decisions, being capable of critical (self-)reflection and
102 ethical decision-making [ref 17]. Intrinsic intentionality, hence moral agency, can thus be strictly
103 ascribed to human beings, although also machine behaviour, resulting from the underlying
104 programming, can be positively assessed when subjected to moral judgement. Similarly, and despite
105 its widespread use in relation to machines, autonomy in the ethically relevant sense can only be
106 attributed to human beings, due to the underlying cognitive processes (entailing self-awareness,
107 self-consciousness and self-authorship according to reasons and values) closely and pre-eminently
108 identified with the dignity of human persons and human agency [ref 18]. Developmental science
109 points to human agency by also highlighting how individuals are part of a self-constructing dynamic
110 system of human life and development, for which agency and self-construction also entail
111 responsibility [ref 14].

112 Christian anthropology articulates a vision of human beings called to cultivate, develop and increase
113 creation, establishing a future-oriented ethics open to and responsible for development. Such ethics
114 promotes an attitude toward science and technology that is fundamentally confident and welcoming
115 of innovation [ref 17].

116

117

118

119 **Artificial intelligence, machine intelligence, or machine behaviour**

120 Reflecting on machine intelligence, Pope Francis said that ‘we need to understand better what
121 intelligence, conscience, emotionality, affective intentionality and autonomy of moral action mean’,
122 by remarking that ‘designation of “artificial intelligence”, although certainly effective, may risk being
123 misleading’ [ref 7]. Actually, machine intelligence is often described in terms of machine behaviour
124 [ref 19], yet this presumes a double passage, namely a technical step from artificial intelligence to
125 machine intelligence, followed by a more subtle transition from machine intelligence to machine
126 behaviour. Indeed, machine intelligence integrates AI, which is a rigorously defined branch of
127 computer science, with the embodied intelligence approach above introduced for artificial systems.
128 The expression ‘machine intelligence’, while reflecting a human desire to understand and mimic
129 human intelligence through artificial systems, is thus grounded on a nominal derivation. Yet talking
130 of machine intelligence does not move the fact we are able to access it, and experience it, only
131 through the limits imposed by our sensory channels and language. This implies a shift towards a
132 phenomenological perspective, fed by the observation and the investigation of machine actions in
133 context, which, in turn, activates two complementary processes. On the one hand, there is a
134 reduction of intelligence to only behaviour. On the other hand, interactions with and amongst
135 “behaving machines” allows them to be simultaneously shaped by and contributors to sociocultural
136 context, thus enforcing the connection with human beings. This is consistent with the fact that, by
137 considering the ability to provide comprehensive explanations of chosen actions as emblematic of
138 intelligence, some research is looking for suitable forms of explanation of machine behaviour to
139 enhance human trust in artificial systems [ref 20].

140 Leveraging Christian-anthropological reflection, however, human experience, including intelligence,
141 cannot be reduced into categories fully accessible by machines. Furthermore, the use, in narratives
142 about machines, of language directly taken from human experience, should be carefully
143 reconsidered, as also noted by researchers from cognitive science [ref 21]. Moreover, the search for
144 AI explanation (or interpretation) methods should systematically consider the difference between
145 correlation and causation. Accordingly, the dialogue already started between technologists and, e.g.,
146 researchers in social sciences should be extended to also encompass philosophical and theological
147 reflection, with expected benefits for all.

148 **Conclusion**

149 Digital innovation and technological progress must contribute to human development and benefit,
150 with respect for society, especially those potentially left behind, and the environment. Potential
151 benefits, like those being investigated, e.g., in healthcare in terms of assistive technologies [ref 22],
152 therapeutics [ref 23] and restored functions [ref 16], must be implemented through equitable
153 applications. Global cooperation and inclusive dialogue also encompassing religious denominations
154 at large must be fostered, also through scientific events and journals.

155 The Pontifical Academy for Life is proactively participating, with courage and discernment, in this
156 social and historical context that is undergoing a profound transformation. The Church is challenged
157 by scientific and technological innovation; yet it can help foster a new alliance to respectfully listen
158 and integrate multiple voices, keeping in mind that voices are necessarily plural. Yet, the human
159 family, our common good, our common home, is one.

- 160 1. Jobin, A., Ienca, M. & Vayena, E. Nat. Mach. Intell. 1, 389-399 (2019).
- 161 2. Rich, A. S. et al. Nat. Mach. Intell. 2, 2-9 (2020).
- 162 3. Luengo-Oroz, M. Nat. Mach. Intell. 1, 494-494 (2019).
- 163 4. Calvo, R. A., Peters, D. & Cave, S. Nat. Mach. Intell. <https://doi.org/10.1038/s42256-020-0151-z>
- 164 (2020).
- 165 5. Hao, K. in MIT Technol. Rev. (2019, 6 June).
- 166 6. Theodorou, A. & Dignum, V. Nat. Mach. Intell. 2, 10-12 (2020).
- 167 7. Pope Francis' Address to Pontifical Academy for Life's General Assembly for the 25th Anniversary
- 168 (Vatican, 2019, 25 February). Holy See Press Office
- 169 <https://press.vatican.va/content/salastampa/en/bollettino/pubblico/2019/02/25/190225d.html>
- 170 8. Rome Call for AI Ethics. <https://romecall.org/>
- 171 9. Pope Francis' Address to Pontifical Academy for Life's General Assembly for the signature of the
- 172 Rome Call for AI Ethics (Vatican, 2020, 28 February). Holy See Press Office
- 173 [https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2020/02/28/0134/00291.html#](https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2020/02/28/0134/00291.html#eng)
- 174 [eng](https://press.vatican.va/content/salastampa/it/bollettino/pubblico/2020/02/28/0134/00291.html#eng)
- 175 10. 2020 PAV Workshop on Artificial Intelligence (The "Good" Algorithm? Artificial Intelligence:
- 176 Ethics, Law, Health; Vatican, 2020, 26-28 February). Pontifical Academy for Life
- 177 <http://www.academyforlife.va/content/pav/en/news/2019/intelligenza-artificiale.html>
- 178 11. 2019 PAV Workshop on Roboethics (Roboethics: Humans, Machines and Health; Vatican, 2019,
- 179 25-27 February). Pontifical Academy for Life
- 180 <http://www.academyforlife.va/content/pav/en/events/general-assembly-2019.html>
- 181 12. Mengüç, Y., Correll, N., Kramer, R. & Paik, J. Sci. Robot. 2, eaar4527 (2017).
- 182 13. Laudato si' - Encyclical Letter of the Holy Father Francis on care for our common home (Libreria
- 183 Editrice Vaticana, 2015)
- 184 [http://www.vatican.va/content/francesco/en/encyclicals/documents/papafrancesco_20150524_enc](http://www.vatican.va/content/francesco/en/encyclicals/documents/papafrancesco_20150524_enciclica-laudato-si.html)
- 185 [iclica-laudato-si.html](http://www.vatican.va/content/francesco/en/encyclicals/documents/papafrancesco_20150524_enciclica-laudato-si.html)
- 186 14. Lerner, R. M. Concepts and Theories of Human Development, 4th ed. (Routledge, 2018).
- 187 15. Kriegman, S., Blackiston, D., Levin, M. & Bongard, J. PNAS 117, 1853-1859 (2020).
- 188 16. Nat. Mach. Intell. 1, 385-385 (2019).
- 189 17. Robotization of Life. Ethics in view of new challenges (Commission of the Bishops' Conferences of
- 190 the European Union, 2019) [http://www.comece.eu/comece-publishes-reflection-on-robotisation-of-](http://www.comece.eu/comece-publishes-reflection-on-robotisation-of-life)
- 191 [life](http://www.comece.eu/comece-publishes-reflection-on-robotisation-of-life)

- 192 18. Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems (European Group on
193 Ethics in Science and New Technologies, 2018) [https://op.europa.eu/en/publication-detail/-](https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source78120382)
194 [/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source78120382](https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source78120382)
- 195 19. Rahwan, I. et al. Nature 568, 477-486 (2019).
- 196 20. Edmonds, M. et al. Sci. Robot. 4, eaay4663 (2019).
- 197 21. Shevlin, H. & Halina, M. Nat. Mach. Intell. 1, 165-167 (2019).
- 198 22. Moyle, W. Nat. Rev. Neurol. 15, 353–359 (2019).
- 199 23. Schneider, P. et al. Nat. Rev. Drug Discov. <https://doi.org/10.1038/s41573-019-0050-3> (2019).
- 200
- 201
- 202
- 203 Correspondence to: Renzo Pegoraro, Pontifical Academy for Life, Vatican City.
- 204 E-mail: cancelliere@pav.va Submitted as: Comment Competing interests: The author declares no
- 205 competing interests