Radoslaw Komuda, Lluis Oviedo and Sara Lumbreras

Artificial Intelligence (AI) and Christian Anthropology

Where the Concerns Lie

Abstract (English) – Artificial Intelligence (AI) and Christian Anthropology. Where the concerns lie. Many voices are raising concerns about the fast developments of AI and its rapidly extending applications. Religion and theology are learning to use these "intelligent" systems as assistants in many tasks, taking advantage of their surprising performances. Every new technical development impacts society, culture, and the human condition, which becomes very sensitive towards these changes, affecting its constitution as embodied and embedded in its ambience and circumstances. The swift progress in AI invites us to rethink several anthropological categories because humans cannot be conceived as isolated or self-referential and self-made, as could be imagined in an earlier humanistic tradition, but as more integrated in their own social, cultural, and technical context, and more supported by external means. The big issue is whether such developments point towards the betterment of the human condition, as has been the case with many former technologies, or it means instead a risky move towards uncertain, but probably darker, consequences for human freedom, dignity, and excellence. Christian anthropology meets a challenge in this emerging field. It is called to rethink its traditional views to make a place for an unexpected element, somehow distorting old ideas, threatening a displacement of its functions and performance, and encouraging a more engaging reflection.
Keywords: Artificial Intelligence, Anthropology; Image of God; Original Sin; Grace; Ethics

Abstract (Deutsch) – Künstliche Intelligenz (KI) und christliche Anthropologie: Wo die Bedenken liegen. Viele Stimmen äußern sich besorgt über die rasante Entwicklung der KI und ihre sich rasch verbreitenden Anwendungen. Religion und Theologie lernen, diese „intelligenten" Systeme als Assistenz bei vielen Aufgaben einzusetzen und sich ihre überraschenden Leistungen zunutze zu machen. Jede neue technische Entwicklung wirkt sich auf die Gesellschaft, die Kultur und das Wesen des Menschen aus, das sehr empfindlich auf diese Veränderungen reagiert, welche seine Beschaffenheit als in seiner Umgebung und seinen Lebensumständen verankert und eingebettet beeinträchtigen. Der rasche Fortschritt in der KI lädt uns dazu ein, verschiedene anthropologische Kategorien zu überdenken, da der Mensch nicht mehr als isoliert oder selbstbezogen und selbst geschaffen aufgefasst werden kann, wie man es sich in der früheren humanistischen Tradition vorstellen konnte,

sondern als stärker in seinen eigenen sozialen, kulturellen und technischen Kontext eingebunden und stärker durch externe Faktoren unterstützt. Die große Frage ist, ob solche Entwicklungen auf eine Verbesserung der menschlichen Lebensbedingungen hinweisen, wie es bei vielen früheren Technologien der Fall war, oder ob sie stattdessen einen riskanten Schritt mit ungewissen, aber wahrscheinlich düsteren Folgen für die menschliche Freiheit, Würde und Besonderheit bedeuten. Die christliche Anthropologie steht in diesem sich entwickelnden Feld vor einer Herausforderung. Sie ist aufgerufen, ihre traditionellen Ansichten zu überdenken, um einem unerwarteten Element Platz zu machen, das in gewisser Weise die alten Vorstellungen verzerrt und mit einer Verschiebung ihrer Funktionen und Leistungen droht, und zu einer engagierteren Reflexion anzuregen.

Abstract (Français) – Intelligence artificielle (IA) et anthropologie chrétienne : où se situent les inquiétudes. De nombreuses voix s'élèvent pour exprimer leur inquiétude face au développement rapide de l'IA et de ses applications. La religion et la théologie apprennent à utiliser ces systèmes "intelligents" pour les assister dans de nombreuses tâches et tirer profit de leurs performances surprenantes. Chaque nouveau développement technologique a un impact sur la société, la culture et la nature de l'être humain, ce dernier étant très sensible à ces changements qui affectent sa manière de s'intégrer à son environnement et à ses conditions de vie. Les progrès rapides de l'IA nous invitent à reconsidérer différentes catégories anthropologiques : l'homme ne peut plus être conçu comme un être isolé, autocentré et créé par lui-même, comme on pouvait l'imaginer dans la tradition humaniste antérieure. Il faut en effet prendre en compte qu'il est davantage intégré dans son propre contexte social, culturel et technique, et soutenu par des facteurs externes. La grande question est de savoir si de tels développements indiquent une amélioration des conditions de vie humaines, comme ce fut le cas pour de nombreuses technologies antérieures, ou s'ils représentent au contraire une étape risquée, aux conséquences incertaines, mais probablement sombres pour la liberté, la dignité et la spécificité humaines. Dans ce domaine en évolution, l'anthropologie chrétienne est confrontée à un défi.

1. Introduction

In an era of unprecedented technological progress, integrating Artificial Intelligence (AI) into various aspects of human life has become an undeniable reality. AI, with its remarkable capabilities, is transforming how we work, communicate, and navigate our increasingly complex daily lives. Its applications span diverse domains, from healthcare and finance to transportation and education, offering unparalleled efficiency, convenience, and accuracy. Even within religion and theology, AI has been finding a place, assisting in tasks once the sole province of the human intellect. This new technical development reverberates through

society, culture, and the very essence of the human condition. The human experience, ever-interwoven with its environment, evolves and adapts, reflecting and refracting the transformative influence of AI technologies. Consequently, the traditional humanistic conception of individuals as isolated, self-referential and self-made entities is being challenged.

The pressing question arises as to whether the swift progress of "intelligent" systems foresees an improvement in the human condition, much like previous technological innovations, or whether it carries the potential for darker consequences, threatening human freedom, dignity, and excellence. Christian anthropology, a centuries-old discipline that contemplates humanity's theological and philosophical nature, confronts a significant challenge in this rapidly evolving landscape. It is called upon to revisit its traditional views, making room for an unexpected element – AI – that can disrupt established paradigms, shift the contours of human identity, and provoke a re-evaluation of its core tenets.

This paper aims to contribute to the ongoing discourse about the place of AI in shaping our understanding of the human person and the complex relationship between technology, theology, and human flourishing. The concerns are manifold, and the path forward is far from clear. What remains certain is that the encounter between AI and Christian anthropology is a meeting of tradition and innovation, theology and technology, the sacred and the synthetic. It is a challenge that demands careful consideration, deep reflection, and open dialogue, since anthropology might not only be one of the most important research fields for the theology of AI discourse, as explored by Anna Puzio et al. (Puzio et al. 2023), but also – a key factor in limiting the expansion of modern "digital sins".

2. The Problematic Relationship of Theology and Technology: A Cautionary Tale

Traditionally, theology has observed technological advances with suspicion and curiosity, even if the right attitude was indifference at other times, as has happened with modern science. A reason for scepticism was the nature of theology being firmly rooted in long-time traditions and fearing every novelty as something that could threaten its identity and central tenets. In recent times, the main concern was that technological advances in several fields could displace and replace religious functions and performance, giving rise to much easier, reasonable, and efficient interventions to address many human and social needs or challenges. The ideas and proposals of transhumanists are just an example of such ambitions, a way to achieve what modern reason and science have promised for a long time: to be able to fix every problem, to heal every wound, and to overcome every limit, resorting to sheer rational and immanent human means.

The senior co-author of this paper belongs to a generation that grew up in an atmosphere of deep mistrust toward technology. The main reasons were those linked to existentialist philosophy, possibly because of Heidegger's criticism of the technical mentality as the main and most recognized leading voice (Heidegger 1953). This has not been the only one during the XX century. Authors of the Frankfurt School, like Adorno and Marcuse (1960) and the early Habermas (1968), expressed their sharp criticism towards those cultural developments they deemed as profoundly flawed and alienating for true human emancipation.

This generation has reacted to the swift technical developments of the last decades uneasily, and an evident ambiguity could be perceived in their theological reception. On the one hand, the critical stance pervaded many sectors, especially those more marked by humanism, social concern, and existentialism. On the other hand, most of them could not resist the temptation to get the best computers for their theological activities, like teaching or writing papers. The same has happened with smartphones, which enabled better communication with more people. When looking for better diagnosis and therapy, almost nobody would deny themselves the best and most updated medical equipment. In the end, critics were often inconsequential, or people surrendered to the obvious advantages of technological developments to avoid being left behind or considered too odd and out of date.

However, theology has often assumed an attitude of distance and indifference towards technology. It could even be ignored as a "sign of the times" due to its scarce impact on the dialectics between sin and grace presiding over many aspects of human life and culture. We doubt that this might be the best option for a more engaged and embedded theology that looks at every development to discern signs of life and death, signs of grace, and signs of sin.

Recently, the situation has become more critical with several technological developments that raise serious ethical concerns. The cases of genetic editing and artificial intelligence are among the most sensitive from this point of view. These concerns have led to a growing interest and to some evolution in Christian thinking, which no longer feels a certain distance and indifference towards technology, but a call for conscious engagement. To a certain extent, it can be said that the ethical questions have called forth the anthropological dimension and its theological reflection on the new technologies. Theology cannot abstain from these discussions and clearly needs to take positions and review its own views on human beings and their place in a world of advanced intelligent systems.

We have in mind a program that tries to pursue a more engaged and "from below" theology, i.e., a theology built from real and lived processes that is able to observe the concrete world and collect data, which becomes relevant for its

elaboration. Such endeavour is always conceived inside the doctrinal framework that provides the best and most fitting interpretation of Christian Revelation. From such a program, we need to pay more attention to the advancements in science and technology with open minds, not just from an apologetic endeavour, but looking for possible ways to collaborate and join efforts for the best of humans and our coming world.

3. Christian and Humanistic Anthropological Views and AI: Towards New Insights and Risks

Traditional anthropological conceptions have been deeply marked by the belief that humans are free and dependent, virtuous and sinners, altruistic and selfish, material and spiritual. Such views were deeply rooted in classical Greek philosophy and Biblical and Early-Christian traditions and found a more systematic expression in Augustinian and Medieval thinkers. It is hard to assess how much the traditional Christian perspective was more focused on the individual, preceding and nourishing modern humanistic individualism, or whether it was more social and communitarian, stressing mutual dependency and responsibility. These perspectives have emphasized the unique qualities of human existence, portraying humanity as the pinnacle of creation. Within Christian theology, this has translated into understanding humans as bearing the *imago Dei*, or the image of God, and possessing intrinsic worth, dignity, and purpose. However, this noble image was tainted by the effects of sinfulness stretching from the origin of humanity and requiring a restoring dynamic of redemption that could correct the dark aspects of human nature.

From this traditional standpoint, the modern understanding of the human person considers it a discrete, autonomous, and rational being, somehow isolated from the external world's influences and in complete control of its life and decisions. This view, often aligned with humanism, imagines human beings as individual agents who define themselves and their existence through their actions, thoughts, and intentions. It has laid the foundation for the concepts of moral responsibility, free will, and human exceptionalism.

Several factors marked a significant change in those views that led to the well-known and many times described crisis of humanism. The new mood started with the harsh criticism from different philosophical traditions that were born in the 19th century, like Marxism, Nietzschean radical thought and later to other currents that denounced its ethnocentrism and the legitimation of colonialism and other adverse developments. In the 20th century, the primary source of the crisis in humanism was inspired by scientific developments, like sociobiology or human ethology, and later evolutionary psychology (Oviedo/Collage

2015). Furthermore, some areas of research such as psychology or economics put some doubt on features of the human condition that were once considered cardinal (namely, rationality). This paves the way to the possibility of AI systems being better than humans in some ways. Furthermore, cognitive sciences contributed to some devaluation of the human condition, as many practitioners, especially in the area of the neurosciences, wondered to what extent the human mind worked better than a good digital system. Nowadays, we can state that recent AI developments add new arguments to those who were less convinced about human exceptionality, excellence, and autonomy.

The advent of AI challenges these traditional anthropological views in profound ways. AI systems, powered by complex algorithms, data processing, and machine learning, introduce a new dimension to human existence that is intertwined with the technological environment. Instead of viewing humans as isolated entities, AI compels us to consider humans as more integrated within their technical contexts than ever before. Instead of perceiving humans in full control, they are now seen as more dependent on external technical assistance. Instead of describing them as very limited, humans see those new technologies as potentially developing indefinitely. Such a process gives rise to a sort of *homo technologicus*, more in continuity with advances that assist and expand its capabilities, with new promises and a brighter horizon.

A crucial consideration for this would be a shift from thinking about the place of technology in people's lives to the place of humans in the world of technology. One significant aspect of this integration is the growing reliance on external means, including AI, to facilitate and enhance various aspects of life. Humans now routinely collaborate with AI systems in decision-making, information retrieval, medical diagnosis, and even – spiritual guidance. This raises fundamental questions about the boundaries of the human self, as these systems become integral to human thought processes and actions. It becomes increasingly evident that humans are deeply entangled in such intelligent systems or embedded in them, creating infinite new possibilities for us. This new situation invites us to redefine the human condition in a more profound way than previous accounts of *homo technologicus* could foresee (Hefner 1993).

Furthermore, AI's ability to extend and enhance human capabilities, whether through assistive technologies, intelligent companions, or prosthetic devices, challenges the long-standing idea of the autonomous self. The concept of the independent self is already contested on several fronts (philosophical, scientific, theological) but it still persists in our cultural context (Visala/Vainio 2023). As we navigate an increasingly technological world, the lines between what is considered intrinsic to human nature and what is facilitated by external means blur. It prompts us to rethink the very essence of humanity and the factors that contribute to its formation and flourishing.

Human nature is undertaking a transformation with unpredictable consequences. What we need, from a theological point of view, is a better insight into such developments and their impact, so that we can better discern how they affect the Christian understanding of human beings – for better or worse – and which might be the right answer to the ongoing challenge.

The integration of AI raises questions about the nature of human agency, accountability, and moral responsibility. How do we assess the actions and decisions of individuals who rely on AI systems for guidance? To what extent can we maintain the idea of free will in a world where intelligent machines influence and shape our choices? More theological problems are lurking in this changing world: will a good application of AI help to make better decisions, avoid mistakes, and render humans less vulnerable to failure and sin? Will AI become part of an updated schema of human and moral enhancement, assisting us to heal negative and dark aspects of our nature? Is AI replacing old-fashioned forms of virtue and hope, providing images of a better future for everybody?

In this evolving landscape, it is essential to critically re-evaluate our anthropological categories, asking how AI affects our understanding of the human person, its place in the world, and its relationship with the divine.

4. General Implications for the Human Condition and Society

We can summarize the positive and negative impacts that we can expect from the accelerated development of AI. As in many other fields, in this case too, we can perceive maximalist and minimalist expectations for both sides, for the good and the harmful effects that we already sense or those that we can expect shortly.

4.1 *Positive Impacts of AI on the Human Condition*

Integrating AI into various facets of human life has the potential for significant positive impacts on the human condition. AI systems have demonstrated the ability to enhance human existence in multiple ways:

1. Efficiency and Convenience: the advent of AI-powered technologies has significantly enhanced efficiency in various sectors, leading to streamlined processes and optimized tasks. This technological leap has reduced the time and effort required for daily activities, offering an unprecedented level of convenience. In theory, this shift towards greater efficiency could grant individuals more leisure time, potentially fostering a better work-life balance. It also opens up opportunities for people to engage in more meaningful endeavours, whether personal or professional. This transformation could lead to a societal shift where time is valued differently, prioritizing personal growth, creativity, and community engagement

over traditional work-centric lifestyles. However, if we are to work faster and more efficiently, why do we still work a 40-hour working week? Secondly, this optimistic view must be balanced against traction-gaining concerns about technology-induced job displacement and the need for new skills and adaptations in the face of rapid technological change. (See also: McGeorge 2023)

2. Medical Advances: the integration of AI in healthcare represents a significant leap forward, revolutionizing how medical care is delivered and experienced. From improving early disease detection, including the potential for epidemic early warning systems (MacIntyre et al. 2023), through advanced diagnostic algorithms for cancer patients (Sufyan et al. 2023) or diagnostic pathology (Shafi/Parwani 2023) to the development of highly personalized treatment plans tailored to individual patient needs (Mohsin et al. 2023). This shift towards more customized and precise medical care enhances the accessibility of healthcare, potentially bridging gaps in medical service availability. This progress in medical technology signifies a move towards a more patient-centred approach in healthcare.

3. Access to Information: AI-driven search engines and recommendation systems mark a significant shift in how we access and consume information. These tools offer not only rapid access to an extensive array of knowledge, but also can connect the dots and find relevant contexts and connotations. This democratization of knowledge fosters a culture of self-education and curiosity, where information is more readily available than ever before. It empowers people to explore diverse topics, stay updated with the latest developments and eventually to personal and professional growth.

4. Efficiency across Sectors and Economic Growth: the implementation of AI in various industries (manufacturing), and services (i.e. finance, transportation) hold immense potential for driving economic growth. These applications are not just about enhancing efficiency within these sectors; they also play a pivotal role in elevating overall living standards. By streamlining operations, optimizing resource utilization, and innovating new approaches, AI can significantly contribute to economic advancement. This progress, in turn, can lead to improved quality of life of employees, showcasing the far-reaching impact of AI beyond mere technological advancements (Cramarenco et al. 2023).

5. Mitigation of Human Suffering: AI-enabled technologies are increasingly becoming vital tools in humanitarian efforts, disaster response, and relief work. Their capabilities in analysing vast amounts of data quickly can be crucial in predicting and responding to crises, thereby potentially reducing human suffering. AI's efficiency in these scenarios can significantly aid in a more effective distribution of resources, quicker medical assistance, and better management of relief operations, thereby playing a key role in alleviating the impact of disasters and crises among affected populations (Farazmehr/Wu 2023).

6. Provision of Company and Reducing Loneliness: AI-powered assisting chatbots are already becoming a rich interaction source in many areas, like psychotherapy (Gratzer and Goldbloom 2020), networking, and clever interaction.

4.2 *Concerns and Risks*

While the positive impacts of AI are substantial, concerns and risks are equally significant. The rapid proliferation of AI technologies raises the following issues:

1. Dignity and Privacy Concerns: The extensive collection and analysis of personal data by AI systems challenge the fundamental principles of human dignity and privacy. In an era where personal information and online behaviour are constantly scrutinized, the risk to individual privacy is profound. Such surveillance can lead to a sense of intrusion, potentially diminishing the sanctity of personal space and autonomy.

2. Ethical Dilemmas: The deployment of AI in diverse sectors, ranging from autonomous weaponry to predictive policing, brings forth complex ethical dilemmas, which are not limited to accountability and bias issues but which also encompass the broader moral implications of decisions made by AI systems. These ethical challenges necessitate careful consideration of its societal impacts, questioning not only how AI decisions are made but also who is responsible for these decisions and their consequences.

3. Economic Disparities: The disparity between those who have access to and can benefit from AI and those who don't could widen, leading to further marginalization of already disadvantaged groups. Particularly concerning is the potential for automation in the services sector that results in significant job losses, leaving many people vulnerable to economic instability and marginalization, although Kyle Wiggers (Wiggers 2024) claims that might not be the case after all.

4. Threats to Human Freedom: As AI systems increasingly influence decision-making and shape individual choices, there is a concern that human freedom, defined by the capacity for making autonomous decisions, could be compromised. The more individuals rely on AI for guidance, the more susceptible they become to external manipulation, potentially undermining the autonomy that is central to human dignity.

5. Impact on Human Excellence: the reliance on AI for tasks that traditionally require human expertise can raise questions about the development and maintenance of human excellence in areas such as art, creativity, and craftsmanship.

It is essential to address these issues thoughtfully and ethically to ensure that the integration of AI aligns with the values and principles that underpin human flourishing.

5. Christian Anthropology: The Issues at Stake

A more profound theological analysis of AI and its effects is not easy for several reasons. The first one concerns the rapid development that we are assisting in these months. We have systematically reviewed the published theological studies about AI (Oviedo 2022). Most appear outdated, published before the development of generative AI and the new systems that are able to develop performances that we could not imagine until recently. For those familiar with such systems and their efficient ways to assist in many academic and other tasks, it is quite surprising how such systems work and manage to perform activities that otherwise would require much more attention and time. Then, different perspectives point to future scenarios in which "strong AI" could pose more severe challenges (Seifert et al. 2022; Dorobantu 2022). In any case, we need to update our views more often to keep attuned to those developments.

We could not resist the temptation to ask the ChatGPT platform: "What are the consequences for Christian anthropology of the development of artificial intelligence?" The answer pointed to the following issues:

1. Image of God
2. Soul and consciousness
3. Ethical considerations
4. Creation and stewardship
5. Redemptive potential
6. Transhumanism and posthumanism

The list is quite obvious, but surely it needs more development – which the system can provide if we ask for more information or fill the listed points with biblical and systematic analysis – and completion. It is already risky to undertake such a reflexive exercise, i.e., to ask the intelligent system about its own limits and negative consequences. In any case, several of these points are worth reviewing in more detail without resorting to AI.

In any case, the explosion of ethical concerns about new developments in AI that we described earlier can be understood as a deeper questioning of the anthropological models that we take as normative or guiding for hard ethical choices. This is almost a truism: any normative system requires some view or representation of human and social ideal standards. We can perceive a deeply entangled connection between anthropological and ethical issues, to the point that one implies the other. Critically, what is at stake in the current debates is much more than a question of good and bad AI or its applications; it is much more a question of adopting a more reductive or a more holistic view of humans, a more individualistic and isolated version or a more social and engaged one (Barassi 2023; Greene et al. 2023).

5.1 *Revisiting the* Imago Dei

The first issue at stake is clearly the traditional idea of humans having been created in God's image. This central tenet of Christian anthropology has essential implications, including the concept of human excellence and uniqueness. Even if such a view has different critical versions, it is a fundamental belief or value for Christian anthropology. Marius Dorobantu has recently published an engaging analysis about this issue (Dorobantu 2022) the extent to which the new AI, in its strong version, invites us to revise the Christian idea of the *Imago Dei*. He invites us to review at least the ontological version of that principle, i.e., that humans have characteristics that render them unique in relation to other creatures or creations and closer to what we represent as the divine. Traditionally, these traits included intelligence, consciousness, and freedom. Humans have always been seen as "rational animals." Now, with the latest achievements in AI, it is apparent that rationality or intelligence is no longer exclusive to humans. Regarding consciousness, the discussion points to future scenarios that could reach that singular state, or to the fact that the current interaction with those systems comes close to the feeling of interacting with a conscious entity. Such perceptions mean a blow to human excellence in that area. As stressed by Sara Lumbreras (Lumbreras 2022), the progress of AI can be interpreted as reducing some aspects of rationality to mere mechanical calculations. This should point to the authenticity of human experience as the place where the *Imago* resides.

This leads us to stress more the other two main views of *Imago Dei*, the relational and the functional. We humans would be constituted similarly to God in the sense of being able to love and interact with the Divine person, or at least being better endowed than other beings to engage in such a relationship. The functional views stress the capacity of humans to pursue the divine work in creation or to become "created co-creators" in the expression of Philip Hefner (Hefner 1993). These alternative versions of the *Imago Dei* could be contested, too, when attending to the capacity of AI to interact richly. We mean to interact in a way that becomes helpful and inspiring with real people, or its indisputable creative capacities, still growing and showing that being creative and able to design and build is less an exclusively human feature.

In the new context, we propose a different answer to how to understand the anthropological principle of the *Imago Dei*. This principle now should be applied to the set of humans and its great technical creations, or in other words, it is not just isolated humans who become the image of God, but humans together with their machines, their prostheses, and everything that enhances and increases their positive potentials in different areas. After all, some similar expanded stance applies now to the need to include *Imago Dei* humanity with its natural environment, not humans in themselves, but humans in their habitat, sharing

with many other creatures to avoid a dangerous and abusive exclusivity. The inclusion of machines in the *Imago Dei*, something that has been pointed out by some transhumanist authors for a while now (Kurzweil 1999), opens the door to the possibility of machines being a better image of God than we are. For these authors, machines could be more rational and less "tainted" by worldly experience. For instance, they claim that machines could be more generous and not experience feelings of jealousy.

This raises critical questions about what it means to be human and related to the divine in an age of highly intelligent systems. To what extent do emotions define human beings differently from machines? Is such a difference positive or rather a handicap, something that makes us less reliable and less rational? Is being emotional a characteristic that brings people closer to or further from God? Intelligent machines can imitate emotions or be programmed with algorithms that make them react as we do under emotional stress or excitement, but this is obviously not the same thing. A holistic view of the human condition presupposes the emotional dimension, and thus its value as "intrinsically human" (another question is to what extent this is exclusively human or shared with other animals). We assume as a loss – not a gain – an anthropological model that could get rid of emotions or filter them in a precise way; we would become less human and therefore less "divine" without passion and strong emotions.

The question of beliefs and believing can clearly be incorporated into an intelligent system that is capable of carrying out an activity aimed at obtaining better beliefs, based on higher probabilities or being more reliable according to the available data. In other words, would we be better off if our cognition were always based on certain knowledge? It seems difficult to make this very human characteristic (but do animals have beliefs?) into something that makes us more 'divine', except in the sense of an incarnate God, given the possibilities and limitations of the human condition.

5.2 *Embodiment*

Another anthropological question raised after the great irruption of AI is where lies the human distinction, or what renders us human, after considering the performances of advanced systems. A possible answer lies in our bodies, given human nature's intense entanglement of body and soul. Such an inextricable composition of consciousness and the material body belongs to one of the deepest mysteries about human beings. It determines our behaviour, freedom, love, sexuality, and intuition. It is assumed that intelligent machines most need hardware, and in some cases, they can be embodied in robots, endowed with sensors to interact with their environment. However, it becomes hard to imagine an intelligent system embodied in a similar way as humans.

Being embodied entails limits and capabilities. It can be conjectured that one day, robots could behave similarly to humans and come close to a form of embodiment based on sensors and automatized reactions to their environment. In this case, they would imitate a similar structure of integrated body and soul, or material moving entity and self-conscious intelligence. At the moment, the basic Christian tenet that conceives humans as a unity of body and soul stands as such: the best way to approach the complexity of many human behaviours and attitudes, their potentialities and limits, and their luminous and dark sides. The holistic anthropological model implies in this case a complex reality, with multiple needs and capabilities. The doubt arises concerning the superiority of the soul, as has been the case in a long classical and Christian tradition, where the soul would better reflect the divine. Contemporary Christian anthropology clearly points to an integration of the body for any attempt to present human nature in its highest and normative expression, with all its ethical implications.

5.3 *Sociality and Community*

Among the many consequences of embodiment, one that is particularly important is the human capacity to establish relationships and build communities. Even if we assist in a growing extension of virtual communities or groups, the original and strong form of relationships and love requires corporeal presence. The capacity to love and to engage in all sorts of relationships constitutes another critical point in Christian anthropology that AI can challenge. The new conditions invite a mental experiment to test whether and to what extent this influence determines, in one way or another, the sociality and loving stance in humans.

The explosion of new social media has significantly affected human communication, with positive results, extending our former limited networks and offering much richer possibilities for interaction and exchange. The irruption of AI moves such dynamics even further. It opens the possibility of a new relationship with virtual chatbots able to entertain interesting and stimulating conversations in many areas without needing a subject that responds to or incarnates such expressions. These systems use voice generation and present realistic interactions in the metaverse. The central issue is to what extent such virtual communities and virtual interacting subjects can replace natural persons and at what price. Science fiction is already letting us get used to future scenarios in which human interaction is replaced by virtual ones (e.g., the movie *Her*, by Spike Jonze 2013: Annesley 2013). We can already talk to ChatGPT in the same way that the main character in *Her* interacted with his operating system.

These developments could render obsolete the experiences of intimacy, sharing in genuine relationships, and the sense of belonging to a community that

we are used to. At the very least, they could act as substitutes for them, particularly for vulnerable people. This could include those who live in isolation and for whom this would be an improvement, but also those too young to understand the differences, such as adolescents. In the latter case, relationships with AI could block the development of genuine connections.

Again, the question is whether we need to define humanity according to new parameters and extend the idea of communication and community beyond the traditional means, as we suggested regarding the concept of *Imago Dei*. We still doubt how far the idea of humans beings able to love – a characteristic that can be recognized in other animals – can be extended to the virtual realms and still more to intelligent artificial systems. In any case, at the moment, we need to account for that possibility. Similarly, as we recognize that the capacity to love is not just human but extends to the biological realm, we can be open to an extension of such an ability to artificial systems.

The real difference is that virtual love and friendship with artificially generated subjects in chatbots, even those closer to us through voice and appearance, does not entail a dynamic of giving love, sacrificing for each other, and giving up something to favour the beloved person. The interaction will always be limited to a chat that does not require deeper engagement. Such limit helps us to better enlighten the true nature of love and friendship as inspired by Christian ideals and shows the real content of giving one's life to another person.

5.4 *Evil, Sin, and Salvation*

The last point worthy of a deeper analysis touches on the dynamics between sinfulness and evil that is present in the human condition, together with redemption or restoration. Many visionaries convey a future in which intelligent systems could assist humans in overcoming their darker side and enhance us morally, rendering obsolete and redundant the Christian and other religious schemes of salvation and betterment (Savulescu/Maslen 2015).

The profound question is whether and to what extent AI will provide the definitive tool to overcome most human limits, including death, as transhumanists claim. AI is becoming a handy assistant in many facets of human and social life, even theological research. The question – in a radical way – is to what extent it could change and redefine the terms of the tension between evil or sin, and salvation or grace, or whether we need to rethink the critical dialectic that defines the human condition and is a central tenet of the Christian faith: that we all are sinners, and as such, we live under the shadow of evil; nevertheless we are redeemed by Christ and able to overcome that adverse condition with the assistance of grace.

Our answer to the former question is negative. In the current conditions and technological developments, very little seems to change in our horizon, allowing

us to assume a very optimistic secular stance. As we write these pages (February of 2024), a new war rages in the Middle East with thousands of innocent victims on each side. This new war adds to the one that has been raging in Ukraine for two years. From a purely empirical perspective, nothing allows us to think that peace comes closer with the explosion of robust new intelligent systems, and we cannot foresee a more harmonious human panorama. In critical ways, these intelligent systems are being used conspicuously in new and more destructive and accurate weapons.

Sinfulness and the corresponding need for redemption imply a realistic stance in our reflections on the human condition, and provide a firmer basis for ethical discernment. The normative issues surrounding AI depend to a large degree on the extent to which we understand human beings to be more or less fallible and flawed. For this reason, it is extremely important – in the face of some popular temptations – to maintain a consciously realistic view of humans as prone to abuse, selfishness, and even destructive tendencies when trying to construct rules that govern the development and application of AI systems. The perceived evil associated with human nature and its need for redemption alerts us to the enormous risks that seem to arise less spontaneously from AI itself than from the bad faith and intentions of programmers and people who might use these powerful technologies for the worst.

6. The Ethical Level: Values, Virtues, and Human Flourishing

In its tradition, Christian anthropology emphasizes core values and virtues that guide human conduct and foster moral excellence. These values include love, compassion, justice, humility, and stewardship. These virtues represent the ethical framework within which individuals strive to live virtuously and fulfil their divine purpose. As we will discuss later, designing AI with these values in mind (Values-Centred Design) and, even more profoundly, with the aim of human flourishing is paramount. For this, a deep understanding of anthropology is critical: it is impossible to develop humanizing technologies and apply them in a humanizing way if we do not have a clear idea of what it means to be human.

AI systems should, therefore, be designed with human values, and ethical considerations should be integrated into AI development processes, guiding how AI systems are conceived, programmed and used (what is known as ethics by design). This last aspect demands a particular effort and consideration since an informed and ethical approach to AI usage is particularly challenging in the pre-configurative world where the younger generation holds authority, since they are the ones who understand the world and find their feet in the unknown

more quickly. That is also why we need more engagement in media literacy (Sutton/Smith 2021).

Many guidelines for AI Ethics have been proposed in the past few years. Some are very much focused on the specificities of technology (such as the one presented by the Institute of Electrical and Electronic Engineers, IEEE in the United States). In contrast, others have had a broader ethical scope, such as the extensive Asilomar Principles, the guide presented by UNESCO, or even the forthcoming European AI Act. A comprehensive review of 84 of these guidelines published in 2019 highlights that there are several clusters of issues and values, around which almost every guideline revolves: accountability and transparency, bias and group harms, privacy, fairness and justice, responsibility, professional integrity and roles and non-maleficence (Jobin et al. 2019). In the Rome Call for AI Ethics, signed by the Pontifical Academy of Life (2020) and other organizations, six principles are recognized: transparency, inclusion, accountability, impartiality, reliability, and security. It is essential to link these issues and values to their theological foundations, as they can illuminate their importance, point towards their needed extensions, and support their application to specific situations.

1. Beneficence and Non-maleficence: These two principles are the cornerstones of Bioethics (which deals with technologies applied to living beings) and can also be used in the context of AI. Beneficence refers to the ethical obligation to act for the benefit of others, promoting good and positive outcomes. In AI ethics, beneficence implies designing and using AI systems to contribute positively to society and individual well-being, actively seeking to do good. This implies, for instance, a duty to cooperate in sharing data for research, for instance in the medical space, where the secondary use for data in research could lead to new discoveries that could benefit millions. Non-maleficence, often summarized by the Latin phrase "primum non nocere" emphasizes the importance of not causing harm or minimizing harm when it cannot be completely avoided. In AI ethics, non-maleficence translates into ensuring that AI systems do not harm individuals or society and that any potential risks or negative impacts are identified and mitigated.

2. Prudence and Risk Aversion: Prudence involves a careful consideration of the risks and potential impacts of our decisions, prioritizing safety. In contexts where actions have far-reaching and possibly profound consequences, such as in AI, the principle of prudence dictates a risk-averse approach. AI, especially some potential future developments such as General AI or, according to some, Superintelligence, might even threaten humanity (that is, it might be an *existential threat* (Bostrom 2017)). By being prudent, we acknowledge our knowledge's limits and complex systems' inherent unpredictability.

3. Humility is also a critical value that appears twice: In the development of AI and the Face of AI Power. First, it should be noted that AI works by extracting

patterns in the data that it gets in its training. This means that it is fallible in two very different senses. First, if data are flawed, the system will be biased, as in the previous examples. In addition, the training process might be flawed as well. This flaw takes the main form of what is known as overfitting: the patterns that the machine learns end up being the data themselves. Like a bad student who learns examples by heart and then fails when the test features slightly different problems, the machine learns data in a way that does not allow generalization. On the other hand, the patterns that appear in the data might not be the ones that we would like AI to learn. These undesirable patterns include, more prominently, examples of bias and discrimination. We have witnessed examples of AI that gave African American convicts a higher chance of recidivism (and that resulted in them being granted parole significantly less often than if they were Caucasian). We have also seen how CV-filtering algorithms discarded job applications from women because, in technology companies, the vast majority of past hires were men. We must be humble not only with respect to our own decisions but, more than anything, with the systems we create. We are not able to understand black-box algorithms, such as Artificial Neural Networks, which give us their answers but not the reasoning behind them. We must internalize the fact that because we cannot understand how black-boxes work, we need other, more transparent models to assess whether AI is robust and fair. This is one of the main reasons that support the need for transparency along the full development of machine learning systems.

4. Fairness and Equality: As explained above, AI systems can perpetuate bias and discrimination if not designed and implemented with fairness. There have been enough scandals in the past few years to make society (and regulators) understand that, when left to their own devices, AI algorithms will perpetuate and even create new biases just because, in their normal functioning, they will also learn the bias that is already present in the data they use for training. Christian ethics call for a just and equitable society, prompting theologians to explore how to ensure that AI systems do not reinforce existing injustices. This also highlights that good AI and digital policies are needed in addition to data protection laws.

5. Transparency: The main issue that calls for enforcement in the development of AI-systems is transparency in algorithms, data and decisions. This is technically difficult at times, given that black-box algorithms do not provide any explanation for their decisions. The patterns they discover remain hidden within the model. The European AI Act considers that, for decisions that have a high impact on people's lives, such as the ones that have to do with health or with accessing the job market, it is not acceptable to use AI systems that take a decision without providing reasoning. This means that methods for what has been known as Explainable AI should be developed. Explainable AI is a whole

movement, from the purely technological side, that is, creating models that can give explanations for their decisions. One particularly telling example is that of Cynthia Rudin's work (Rudin 2019), who took the biased, black-box model used in Compass to determine parole and developed a completely transparent algorithm only driven by a few variables (including the convict's age and whether there were any violent crimes in their record) and could achieve a prediction of recidivism that was almost as good as the one from the black box, without conforming to its opacity.

6. Accountability and Responsibility: Theological ethics emphasizes accountability and responsibility in human conduct. In the context of AI, this means holding individuals and organizations responsible for the ethical implications of AI decisions and fostering transparency in how AI systems make choices. As explained above, many cases of failed algorithms and system applications have already existed, even among the industry's most respected players, including Amazon and Netflix (Kearns/Roth 2020). This means that accountability is a prerequisite not only for avoiding the bias discussed above but also for making companies and individuals responsible for their decisions with the support of AI systems.

7. Privacy: Although legislation such as the General Data Protection Regulation has focused on personal data, AI brings a new level of concern to the possibility of collecting personal data and using it to manipulate individuals or collectives. Privacy is essential for many reasons that impact personal autonomy, security, and social dynamics. Privacy allows individuals to control their personal information, a fundamental aspect of their dignity and decision-making freedom. From a security perspective, privacy is critical in safeguarding individuals from risks like identity theft, financial fraud, and other forms of cybercrime. The absence of privacy opens doors to potential misuse of data by malicious entities. When people are confident their data is handled respectfully and securely, their engagement and support for these systems strengthen. Additionally, it encourages free expression and association, fundamental aspects of a democratic society, by allowing people to communicate and associate without fear of surveillance.

8. Security: Security in AI is not just about protecting systems from unauthorized access or cyberattacks; it's about ensuring the reliability, trustworthiness, and safety of AI applications in various aspects of our lives. Secure AI systems are crucial for preventing manipulation, misuse, or unintended consequences that could arise from compromised AI. For instance, a breach in an AI system could lead to the exploitation of personal data, financial fraud, or even pose risks to physical safety in cases where AI controls critical infrastructure such as the energy system or autonomous vehicles. Moreover, the integrity of AI systems is fundamental to maintaining public trust. As AI increasingly becomes integrated

into strategic and sensitive areas like national security, healthcare, and finance, the importance of security escalates. Any vulnerability in these areas can have far-reaching consequences, underscoring the need for security measures.

9. Moral Agency and Decision-Making: The theological notion of moral agency remains central, even in a world where AI influences decisions. Ethical reflection should focus on how individuals can exercise moral agency in AI-influenced choices and maintain their capacity to make virtuous decisions. In the context of our trust in AI, we must address how far this trust is placed. When asked about solving world hunger, if AI suggests breeding rhinoceroses in Antarctica, would we do it? Without the transparency mentioned above and virtually non-existent explainable AI? This question has been explored by Maldonaldo and Torres (2019), but can it be compared to our trust in divine providence? Even though AI is grounded in algorithms and data, the exact workings of complex models can be opaque, leading users to trust the output without fully understanding the process.

10. Similarly, trust in divine providence involves faith in a higher power's unseen guidance. Both scenarios require a leap of faith, whether in technology or spirituality, and both deal with complexity and uncertainty in decision-making processes. Although it might be intriguing and conceptually interesting, the nature of trust in each case is fundamentally different, rooted in different paradigms – technological versus spiritual. As argued above, if our societies require transparency, transparent models will be developed: we should not need to sustain a faith in our AI systems that parallels our faith in providence, and explainable/interpretable AI should be demanded so that new techniques (both statistically performing and transparent) are developed and deployed.

11. AI and Education: It has long been known that "whatever is not used goes to waste." This is true in a technological context, with, for instance, our memory (where the Google effect means that we are far less likely to remember if we know that the information is available online). Using AI for cognitive tasks that used to be performed exclusively by humans means that we will not only be far less equipped to serve them independently, but also, if we only use these tools, we will never even learn. This is particularly problematic in the context of AI applied in education. Would we accept that students do not understand how to calculate basic operations provided they learn how to use a calculator? Would we assume that they do not know how to perform a good summary of a text, provided they can use AI to generate the summary from them automatically? This value, which still seems to be absent from the work performed in the ethical guidelines, is core to Christian anthropology. However, it is already clear that humans need to learn to perform AI-aided tasks without this help (what we could call *unplugged learning*) to be independent, even if they use them regularly to reap the benefits of their efficiency. Along the same lines, working

without the tools from time to time would be necessary to avoid losing this autonomy (what has been referred to as "blackout periods").

12. Sustainability: AI, like many other sectors of economic activity, incurs a great deal of ecological impact, mainly in terms of power consumption but also in the responsible sourcing of raw materials for AI hardware and the disposal of electronic waste. For instance, the ecological impact of AI systems should be considered when deciding whether to apply these systems for a given task or to keep it to our more standard means.

13. Social and Distributive Justice: The application of AI (above all, in the service sector) will destroy thousands of jobs. It is still unclear how many of these lost jobs will be compensated by the new forms of employment that will emerge in the coming years. This has led many thinkers to press for the idea of a universal global income as the only possible road to cope with the expected rise in unemployment. At the same time, a new sector of unskilled jobs is necessary to train the algorithms, where the minimal payments received lead to conditions of semi-slavery in the Global South. The impact of AI on employment and the economy is one of the main consequences of this technology, with one crucial question: "*How should the profits of automation be shared?*" The principle of solidarity, one of the pillars of the Church's Social Doctrine, should be kept at the forefront of these decisions.

14. Subsidiarity: Power concentration is also a key element when examining the changes brought about by AI. It should not be acceptable to have a vast part of the international economy in the hands of only a few companies that dominate the new AI marketplace. This leads not only to economic power but also to the possibility of having an unprecedented social impact. It is paramount to examine market and competition rules and enhanced regulatory oversight so that the new technologies can benefit society instead of creating a new, more powerful type of monopoly. The subsidiarity principle, a pillar of the Church's Social Doctrine, should be re-examined to illuminate the changing landscape of AI economics.

15. Inclusivity: a digital divide, where only a part of society can access the benefits of AI, is unacceptable. There should be some means established so that all sectors of society can share the benefits of automation, irrespective of age, location, or economic status.

16. Democratic Participation, Truth, and the Right to Information: New technologies can provide the means for a level of democratic participation that has never been seen. However, they can also support the spread of misinformation, even in some particularly difficult to prove forms (*deepfakes*). The right to information is one of the challenges of the digital era, with more data being available, but mainly without any guarantee of veracity. New systems should be proposed to provide relevant, objective information to the public.

17. Honesty and Creativity: The widespread use of Generative AI tools has created controversy with respect to what should be considered original and how human creativity should be defined. There is due acknowledgment to the original creators whose work was fed into the AI tools and defining guidelines, so that AI can leverage our means for creativity rather than drowning this creativity in a sea of mechanical creations. It is essential to protect this honesty in the education sector, where the use of AI could impede learning and damage our autonomy and independence, as underscored above, in the long term. Along the same lines, the misuse of AI to generate what appears to be new knowledge, resulting in an inflation of academic papers with doubtful contributions, should be rejected radically. The use of AI to increase knowledge rather than to drown it in a sea of (mis)information is one of the cornerstones of its ethical application.

18. Love, Empathy, and Compassion in a Digital Age: The application of AI in healthcare, caregiving, and pastoral care raises questions about the authenticity of human empathy in AI-human interactions. Theological reflection is needed to determine how the values of love and compassion remain in a world where AI may provide care and support. Human beings tend to attribute human feelings and consciousness to other beings, including machines. In the past few years, we have seen the rise of chatbots that are used to provide company for the otherwise lonely, such as the example of Xiao Ice in China, which is defined as an "ideal girlfriend" that can be programmed to send messages or call with a preferred frequency or with a preferred tone. The existence of these chatbots creates the possibility of artificial company, artificial empathy, and artificial compassion. However, it should be apparent from the beginning that, at least with current technology, AI is not conscious and does not experience any subjectivity. It does not feel empathy even though it might look empathic. The distinction between essence and appearance is paramount in this respect. Do human beings long for connection, for a relationship with others that is seen as equally human? Or are we content with being entertained, with our emotions being reflected back to ourselves? It is tempting to declare that humans long for other humans and that in the same way that a hypocritical, "fake" friendship would be rejected by most, the relationships with the machine should be dismissed as well. However, an epidemic of loneliness might invite us to reconsider whether the machine can potentially offer a "second best" to those who are deprived of human connection.

19. The Common Good: Stewardship of AI technologies for the common good of humanity is a central theme in Christian ethics. The ethical development and application of AI should be guided by a commitment to promote the welfare of all individuals and communities, as well as of the Common Home. Christian ethics places, first of all, a strong emphasis on promoting human flourishing and the well-being of all individuals. Stewardship of AI technologies

aligns with this principle by ensuring that AI is developed and applied to enhance human life.

As Christian anthropology engages with AI, it becomes apparent that a values-driven approach is essential to address AI technologies' ethical concerns and opportunities. We can foster a more holistic and inclusive conversation about the intersection of technology, faith, and human experience by anchoring ethical reflection in traditional Christian values and virtues, which establish clear theological foundations for AI ethics.

Christian anthropology also aligns with modern trends in digital policy since it places the human person at the centre of its focus. By grounding AI ethics in Christian values, without excluding value systems of other religions or world views, we prioritize the well-being and dignity of individuals as paramount.

7. Conclusion and Future Work

The intersection of artificial intelligence and Christian anthropology is a meeting of tradition and innovation, theology and technology, and the sacred and the synthetic. This paper has explored the profound implications of AI for Christian anthropology, navigating a complex landscape where age-old beliefs encounter the relentless advance of technological progress.

As we conclude this exploration, several vital implications emerge:

1. Redefining the Human Condition: AI challenges traditional anthropological views, forcing us to rethink the nature of the human condition. The integration of AI and the coexistence of intelligent machines within human contexts complicate established notions of autonomy and human uniqueness. AI prompts us to reconsider the boundaries of human existence and the interplay of external technologies in shaping human identity.

2. Ethical and Theological Reflection: AI's ethical and theological implications are vast. AI raises questions about human agency, responsibility and morality, and inspires a re-evaluation of the values and virtues that underpin Christian anthropology. Love, compassion, justice, and humility must find new expressions in a world where AI influences human life.

3. Guiding Principles: AI's development, deployment, and governance should align with traditional Christian values and virtues. Ethical guidelines rooted in theological principles can help shape AI technologies and practices to become more ethical, equitable, and aligned with the common good.

4. Ongoing Dialogue and Continuous Review: The challenges and opportunities at the intersection of AI and Christian anthropology and theology call for an ongoing dialogue and reflection. Theological institutions, religious communities, and technological innovators must engage in meaningful conversations

about navigating this complex terrain. Furthermore, theological education and pastoral care must adapt to accommodate the evolving relationship between humans and AI.

In the broader context of society, the encounter between AI and Christian anthropology is emblematic of the larger dialogue between technology and the human spirit. This dialogue compels us to ask how AI challenges traditional beliefs and how it can inspire a deeper understanding of what it means to be human. As technology continues to shape human experience, the Christian community and theologians are called upon to engage in a thoughtful and ethically grounded examination of AI's impact on the values, virtues, and theological underpinnings that define humanity. The future of AI and Christian anthropology is a frontier of possibility, and in navigating this terrain, we have an opportunity to reimagine theological discourse, aligning age-old wisdom with the advances of our time. The journey is ongoing, and the conversation is far from concluded.

Literature

Annesley, J. (2013): Being Spike Jonze: Intertextuality and Convergence in Film, Music Video and Advertising. In: *Journal of Contemporary Film* 11-1, 23-37. Online: https://doi.org/10.1386/ncin.11.1.23_1 (last accessed: 07 February 2024).

Barassi, V. (2023): AI and the Western Illusion of Human Nature: Anthropology's Fight against Human Reductionism and its Interdisciplinary Future. At: *RAI2022: Anthropology, AI and the Future of Human Society.* Online: https://www.alexandria.unisg.ch/handle/20.500.14171/117024 (last accessed: 07 February 2024).

Bostrom, N. (2017): *Superintelligence.* Paris, Dunod.

Cramarenco, R. E. et al. (2023): *The Impact of Artificial Intelligence (AI) on Employees' Skills and Well-Being in Global Labor Markets: A Systematic Review.* Oeconomia Copernicana 14/3, 731–67. Online: https://doi.org/10.24136/oc.2023.022 (last accessed: 07 February 2024).

Dorobantu, M. (2022): Strong Artificial Intelligence and theological anthropology: one problem, two solutions. In: Jorion, P. (ed.): *Humanism and its Discontents: The Rise of Transhumanism and Posthumanism.* London, Palgrave, 19-33.

Farazmehr, S./Wu, Y. (2023): Locating and Deploying Essential Goods and Equipment in Disasters Using AI-Enabled Approaches: A Systematic Literature Review. In: *Progress in Disaster Science* 19. Online: https://doi.org/10.1016/j.pdisas.2023.100292 (last accessed: 07 February 2024).

Gratzer, D./Goldbloom, D. (2020): Therapy and E-Therapy —Preparing Future Psychiatrists in the Era of Apps and Chatbots. In: *Academic Psychiatry* 44/2, 231–34. Online: https://doi.org/10.1007/s40596-019-01170-3 (last accessed: 07 February 2024).

Greene, T. et al. (2023): Atomist or Holist? A Diagnosis and Vision for a More Productive Interdisciplinary AI Ethics Dialogue. In: *Patterns* 4-1. Online: https://doi.org/10.1016/j.patter.2022.100652 (last accessed: 07 February 2024).

Habermas, J. (1968): *Wissenschaft und Technik als „Ideologie".* Frankfurt/M., Suhrkamp.

Hefner, P. (1993): *The Human Factor: Evolution, Culture and religion.* Minneapolis, Fortress Press.

Heidegger, M. (1953/2002): *Die Frage nach der Technik.* In: Ders.: *Die Technik und Die Kehre.* Stuttgart, Klett-Cotta, 5-36.

Jobin, A. et al. (2019): The global landscape of AI ethics guidelines. In: *Nature Machine Intelligence* 1, 389–399.

Kearns, M./Roth, A. (2020): *The Ethical Algorithm: The Science of Socially Aware Algorithm Design.* New York, Oxford University Press.

Kurzweil, R. (1999): *The age of spiritual machines: When computers exceed human intelligence.* New York, Penguin.

Lumbreras, S. (2022): Lessons from the Quest for Artificial Consciousness: The Emergence Criterion, Insight-Oriented AI, and Imago Dei. In: *Zygon* 57/4, 963-983.

MacIntyre, C. R. et al. (2023): Artificial Intelligence in Public Health: The Potential of Epidemic Early Warning Systems. In: *Journal of International Medical Research* 51/3. Online: https://doi.org/10.1177/03000605231159335 (last accessed: 07 February 2024).

Maldonaldo, V./Torres, A. (ed.) (2019): *When the Yogurt Took Over.* Netflix.

Marcuse, H. (1960): *Reason and Revolution: Hegel and the Rise of Social Theory.* Boston, Beacon.

McGeorge, D. (2023): *The ChatGPT Revolution: How to Simplify Your Work and Life Admin with AI.* Melbourne/VIC, John Wiley & Sons.

Mohsin, S. N. et al. (2023): The Role of Artificial Intelligence in Prediction, Risk Stratification, and Personalized Treatment Planning for Congenital Heart Diseases. In: *Cureus* 15/8. Online: https://doi.org/10.7759/cureus.44374 (last accessed: 07 February 2024).

Oviedo, L./Colage, I. (2015): Cultural Influences on the Brain, Science and Theology on Human Specificity. In: *Open Theology* 1, 354–365.

Oviedo, L. (2022): Artificial Intelligence and Theology: Looking for a Positive — But Not Uncritical — Reception. In: *Zygon* 57/4. Online: https://doi.org/10.1111/zygo.12832 (last accessed: 07 February 2024).

Pontifical Academy of Life et al. (2020): Rome Call for AI Ethics. Online: https://www.romecall.org/wp-content/uploads/2022/03/RomeCall_Paper_web.pdf (last accessed: 03 March 2024).

Puzio, A. et al. (2023): *Alexa, wie hast du's mit der Religion?* Darmstadt, wbg.

Rudin, C. (2019): Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. In: *Nature Machine Intelligence* 1/5, 206-215.

Savulescu, J./Maslen, H. (2015): Moral Enhancement and Artificial Intelligence: Moral AI? In: Romportl, J. et al. (eds): *Beyond Artificial Intelligence.* Springer, Cham. Online: https://doi.org/10.1007/978-3-319-09668-1_6 (last accessed: 07 February 2024).

Seifert, J. et al. (2022): On Humans and Machines: Anthropological Considerations on "Intelligent" Neurotechnologies (INTs). In: *Techné: Research in Philosophy and Technology* 26/2, 207–231.

Shafi, S./Parwani, A.V. (2023): Artificial Intelligence in Diagnostic Pathology. In: *Diagnostic Pathology* 18/1, 109. Online: https://doi.org/10.1186/s13000-023-01375-z (last accessed: 07 February 2024).

Sufyan, M. et al. (2023): Artificial Intelligence in Cancer Diagnosis and Therapy: Current Status and Future Perspective. In: *Computers in Biology and Medicine* 165. Online: https://doi.org/10.1016/j.compbiomed.2023.107356 (last accessed: 07 February 2024).

Sutton, A. T./Smith, B. (2021): *Redeeming Technology: A Christian Approach to Healthy Digital Habits*. Saint Louis, MO, Concordia Publishing House.

Visala, A./Vainio, O.-P. (eds.) (2023): *Theological Perspectives on Free Will*. London, Routledge.

Wiggers, K. (2024): New MIT CSAIL study suggests that AI won't steal as many jobs as expected. In: *Techcrunch.* Online: https://tcrn.ch/3Htvb20 (last accessed: 23 January 2024).

Authors

Lluis Oviedo is a full professor at Pontifical University Antonianum in Rome. He holds the chair of theological anthropology, and teaches hermeneutics and issues in the interface between science and religion. He has published several books on religion in Western societies, challenges and possibilities; on altruism; and on consecrated life. His last work deals with the credibility of the Christian proposal. In recent years his research has focused on religion, health and wellbeing; and on issues of artificial intelligence and anthropology. Postal address: Via Merulana, 124, 00185 Roma, Italia. Email address: loviedo@antonianum.eu.

Radoslaw Komuda is a Ph.D. candidate at the Faculty of Theology of Nicolaus Copernicus University in Torun, Poland. Awarded in the Polish Ministry of Science and Higher Education's prestigious "Diamond Grant" program. Thanks to a scholarship granted by the Government of Japan, he pursued his studies at the Faculty of Information Science and Technology at Hokkaido University in Sapporo. He has presented papers at scientific conferences in Japan, the US and across Europe. He is the founder and Principal Investigator of the Digital Ethics Research Laboratory in Torun. Postal address: Ul. Gagarina 11, 87-100 Torun, Poland. Email address: rkomuda@doktorant.umk.pl.

Sara Lumbreras is a professor at the ICAI School of Engineering of the Universidad Pontificia Comillas in Madrid, Spain. She is deputy director of Research Results at the Technological Research Institute and co-director of the Hana and Francisco J. Ayala Center for Science, Technology and Religion. She is the author of more than fifty academic publications and has directed or participated in more than twenty projects with private companies and public institutions. Her research focuses on the development and application of decision support techniques for complex problems, including classical mathematics and Artificial Intelligence. She also develops a line of research in the Philosophy of Technology and the implications of AI in anthropology. Postal address: Santa Cruz de Marcenado 26, 28015 Madrid, Spain. Email address: slumbreras@comillas.edu.