

Growth or Decline: Christian Virtues and Artificial Moral Advisors

Studies in Christian Ethics

1–17

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09539468241305014

journals.sagepub.com/home/sce**Ximian Xu** Duncan Forrester Fellow, University of Edinburgh, Edinburgh, EH8
9NW, UK

Abstract

Artificial Intelligence (AI) has permeated every section of human life. Some even suggest that artificial moral advisors (AMAs)—that is, AI-driven artefacts designed to assist human moral growth by providing moral advice—can help humans to lead virtuous lives. With a focus on Thomas Aquinas's theological virtue ethics, this article will argue that AMA projects are cast into question because, theologically speaking, the cultivation of virtues and moral growth are inseparable from hope and patience given through God's grace, which is immeasurable and cannot be reductively mathematised into AI models. Moreover, the tension between AI's efficiency and patient moral life offers further criticism of AMAs. That said, the theology of patience brings forth a conceptual apparatus through which to qualify the virtuous AMA as a tool to bring together human moral advisors as silent hearers and human moral advisees as patient inquirers.

Keywords

Artificial moral advisor, hope, infused virtues, machine learning, patience, theological virtues, Thomas Aquinas, virtue ethics

Introduction

Over the past few years, a growing body of literature has emerged on religion and AI as well as robots. While AI and robots differ, robotics is inextricably linked to AI since AI enables robots to perform tasks autonomously. Jonas Simmerlein and Max Tretter's recent study reviews the present literature about religion and robots from five databases:

Corresponding author:

Ximian Xu, Duncan Forrester Fellow, University of Edinburgh, Institute for Advanced Studies in the Humanities, 2 Hope Park Square, Edinburgh EH8 9NW, Scotland.

Email: simeonximian@hotmail.com

Atla Religion, IxTheo, Periodicals Index Online, IEEEExplore, and Mendeley.¹ Between 1 March and 21 October 2022, 407 academic publications on religion and robots were produced. Simmerlein and Tretter classify these publications into four categories: liturgy and rites, religious education, spiritual care, and preaching. Their study offers a helpful overview of the current landscape of interdisciplinary studies on religion and robots as well as AI, showing that AI and robots are gaining traction in religious communities.

Simmerlein and Tretter's study leaves out research on AI and religious ethics, which has drawn significant attention over the past few years. Indeed, a few theologians have explored ethical issues surrounding AI in the early twenty-first century.² However, recent AI-and-religion studies examine ethical themes related to AI more comprehensively. Some seek to articulate the theological foundation of moral AI.³ Others investigate the ways Christian ethics can respond to challenges AI poses.⁴ In addition to these two types of studies, some studies inquire into the ethical implications of human–AI interaction, examining themes such as privacy, human dignity, and social relationships.⁵ These studies on AI and religious ethics together demonstrate that, rather than merely focusing on fictional AI as depicted in science fiction and film, theology is a crucial tool for addressing ethical questions related to AI technology.

Recent studies on artificial moral advisors (AMAs) in the field of AI ethics reveal the dearth of ethical responses made from theological perspectives to AI applications. An AMA refers to an AI system or chatbot designed to provide users with moral advice and help them to lead virtuous lives by grappling with moral dilemmas. AMA projects are predicated upon AI's autonomous decision-making, which has been widely applied to various human practices such as healthcare, legal services, and chess-playing programmes, and has now extended to making moral judgement.⁶ A question arises here: Are AMAs capable of providing humans with moral advice—generated through autonomous decision-making—that fosters the growth of human virtues?

-
1. Jonas Simmerlein and Max Tretter, 'Robots in Religious Practices: A Review', *Theology and Science* 22.2 (2024), pp. 255–73.
 2. E.g., Noreen L. Herzfeld, *In Our Image: Artificial Intelligence and the Human Spirit* (Minneapolis, MN: Fortress, 2002).
 3. E.g., Mark Graves, 'Theological Foundations for Moral Artificial Intelligence', *Journal of Moral Theology* 11, Special Issue 1 (2022), pp. 182–211.
 4. E.g., Michael Burdett, 'Proximate and Ultimate Concerns in Christian Ethical Responses to Artificial Intelligence', *Studies in Christian Ethics* 36.3 (2023), pp. 620–41.
 5. E.g., Andrea Vicini, SJ, 'Artificial Intelligence and Social Control: Ethical Issues and Theological Resources', *Journal of Moral Theology* 11, Special Issue 1 (2022), pp. 41–69; Andrew Graystone, 'Sextech: Simulated Relationships with Machines', in John Wyatt and Stephen N. Williams (eds.), *The Robot Will See You Now: Artificial Intelligence and the Christian Faith* (London: SPCK, 2021), pp. 151–65.
 6. Alistair E. Johnson et al., 'Machine Learning and Decision Support in Critical Care', *Proceedings of the IEEE* 104.2 (2016), pp. 444–66; Wendell Wallach, 'Implementing Moral Decision Making Faculties in Computers and Robots', *AI & Society* 22 (2008), pp. 463–75.

This question is underexplored in the current literature on AI and religion. To fill the gap, this article seeks to explore the above question from the perspective of Christian virtue ethics, with a particular emphasis on Thomas Aquinas's theology of virtues. It will argue that Christian virtue ethics casts AMA projects into question because, theologically speaking, the cultivation of virtues and moral growth are inseparable from hope and patience given through God's grace, which is immeasurable and cannot be reductively mathematised into AI models. Moreover, the tension between AI's efficiency and patient moral life offers further criticism of AMAs. That said, the theology of patience brings forth a conceptual apparatus through which to qualify the virtuous AMA as a tool to bring together human moral advisors as silent hearers and human moral advisees as patient inquirers. In this way, the AMA can be considered a tool to assist human moral growth provided both that the virtue of patience is not undermined and that humans do not over-trust AI and its advice.

This article is broken down into three sections. First, I will spell out the concept of AMA and narrow it down to the idea of virtuous AI as a moral advisor. Second, I will inquire into Christian virtue ethics through analysis of Aquinas's view of hope and patience, fleshing out patience as an infused virtue in Christian life. Finally, I will explicate how the Christian virtue of patience can yield critical responses to ethical challenges posed by AMAs and validate the application of AMAs in a patient way.

Artificial Moral Advisors and Virtuous AI

The approaches to AMAs can be classified into two categories. The first category is made up of AMAs designed to operate as moral decision-makers, substituting for human moral agents. Characteristic of this category is the presupposition that AI will attain full moral agency to the same extent as humans since the human mind can be reproduced within AI. Blay Whitby is one of the leading proponents of this AMA. He suggests that the AMA will be capable of making moral decisions *for* humans insofar as AI will evolve to have human-level moral agenthood. He is convinced that the AMA will improve human moral performance across various contexts and over time.⁷ Whitby takes it for granted that human-level AI will be created in the future. However, this assumption is not a consensus reached in the field of AI. I have extensively examined the theory of human-level AI in previous studies, demonstrating that human psychosomatic unity reveals an ontological distinction between humans and AI, as embodied in both religious liturgy and ethical life. From this it follows that AI cannot attain the moral status of a human-level agent.⁸ Therefore, in the remainder of this article, my primary focus will not be on the first category of AMAs.

7. Blay Whitby, 'Computing Machinery and Morality', *AI & Society* 22.4 (2008), pp. 551–63.

8. Ximian Xu, *The Digitalised Image of God: Artificial Intelligence, Liturgy, and Ethics* (London: Routledge, 2024); see also Ximian Xu, 'A Theological Account of Artificial Moral Agency', *Studies in Christian Ethics* 36.3 (2023), pp. 642–59; Ximian Xu, 'Human Sustainability in the Age of Technology: A Theological Proposal on Technomoral Human Futures', in Michael Fuller, Mark Harris, Joanna Leidenhag and Anne Runehov (eds.), *Issues in Science and Theology: Global Sustainability* (Cham: Springer, 2023), pp. 187–96.

The second category of AMAs encompasses AI-driven artefacts designed to help humans to make their own moral decisions in order to lead virtuous lives. These AMAs are proposed based on current AI technology rather than fictional AI. The second category of AMAs are less optimistic about AI technology than the first, and the AMA is considered an assistant of human agents. A recent example of this type is Francisco Lara and Jan Deckers's Socratic AMA. By 'Socratic', Lara and Deckers refer to Socrates' pedagogy, in which he helps his interlocutors to develop their own knowledge by refuting their definitions. In like manner, the Socratic AMA can act like Socrates to scrutinise human agents through asking questions concerning their moral dilemmas, thereby driving humans to discern potential moral failures.⁹ While respecting human autonomy, Lara and Deckers suggest that the questions programmed into the Socratic AMA should not be embedded with particular moral values so as not to enable AI algorithms to favour specific ethical theories.¹⁰ However, it is an overly simplistic approach to formulating questions about moral life without reference to moral values, all the more so when considering how these questions can facilitate the cultivation of moral virtues. It has been well argued by scholars like Alasdair MacIntyre that virtues are connected to communities. Virtues vary across contexts and traditions and can only be construed in relation to communities where they are practised.¹¹ As such, the effectiveness of the Socratic AMA in helping humans foster virtues is called into question. In this respect, Martin Gibert's proposal of virtuous robots seems to furnish a more viable method for designing an AMA.¹²

Gibert opposes attempts made to create a human-level virtuous robot or AI. He argues elsewhere with Dominic Martin that current AI technology cannot create AI-driven artefacts as sentient as humans.¹³ While speaking of virtuous robots or AI, Gibert relies upon Rosalind Hursthouse's definition of right action: 'An action is right [if and only if] it is what a virtuous agent would characteristically (i.e., acting in character) do in the circumstances'.¹⁴ Following this, Gibert registers that a virtuous robot refers to the one that 'behave[s] ... as virtuous people would'.¹⁵ This suggests that Gibert's account of virtuous AMA should be examined in light of currently available AI technology.

Gibert underlines the importance of machine learning for building virtuous AMA. Machine learning is a subset of AI, broadly defined as a set of computational methods that utilise the collected data of past information to make accurate predictions and improve performance with efficient algorithms. At the core of machine learning are

9. Francisco Lara and Jan Deckers, 'Artificial Intelligence as a Socratic Assistant for Moral Enhancement', *Neuroethics* 13 (2020), pp. 277–87.

10. Lara and Deckers, 'Artificial Intelligence', pp. 282, 284.

11. Alasdair MacIntyre, *After Virtue: A Study in Moral Theory*, 3rd edn (Notre Dame, IN: University of Notre Dame Press, 2007), pp. 181–203.

12. Martin Gibert, 'The Case for Virtuous Robots', *AI and Ethics* 3 (2023), pp. 135–44.

13. Martin Gibert and Dominic Martin, 'In Search of the Moral Status of AI: Why Sentience Is a Strong Argument', *AI & Society* 37 (2022), pp. 319–30.

14. Rosalind Hursthouse, *On Virtue Ethics* (Oxford: Oxford University Press, 1999), p. 28; cited in Gibert, 'The Case for Virtuous Robots', p. 136.

15. Gibert, 'The Case for Virtuous Robots', p. 136.

data collection, analysis, and statistics, which enable the training of AI models that generate outputs to meet human needs based on inputs.

Gibert proposes a three-step methodology for developing virtuous AMAs with the aid of machine learning. The first step is to create 'a base of virtuous people' who are 'from every kind of background'.¹⁶ It is worth noting that Gibert is not over-confident in human morality but acknowledges that these virtuous people also have moral flaws. Be that as it may, he is still optimistic about the virtuous character of humans, which is considered 'relatively constant' and enables the base to surpass average virtuous human life.¹⁷

After establishing this base, the second step is to invite these selected people to instruct AI algorithms through answering moral questionnaires, which include questions about moral beliefs, moral intuition about particular topics, moral behaviours in certain scenarios, and other moral themes.¹⁸ These questionnaires produce moral data for machine learning to generalise moral responses to a variety of moral dilemmas. Gibert does not advocate a super-AI that is capable of monitoring and predicting human dilemmas and offering corresponding solutions. Such a super-AI resembles the God Machine proposed by Ingmar Persson and Julian Savulescu. This God Machine is a bioquantum computer, supervising human morality and transforming human thoughts, intentions, and desires under any circumstances.¹⁹ *Contra* Persson and Savulescu, Gibert recognises that, despite the well-curated data of human virtuous life, machine learning cannot lead to a perfect virtuous AI-driven system.²⁰

The third step is to implement a decision algorithm in virtuous robots, as multiple moral options may arise when addressing a single moral dilemma. This decision algorithm should allow for the diversity and unpredictability of moral decisions insofar as 'opposite actions may both be morally right, to the extent that this corresponds to what distinct virtuous people would have done'.²¹ In other words, even if the proportion of decision A is higher than that of decision B, the virtuous AMA can still recommend decision B to moral advisees because it aligns with the virtuous character of certain virtuous individuals. For human users, this moral decision-making process is, to a certain extent, flexible and unpredictable.²²

Gibert's methodology concentrates upon state-of-the-art AI technology without sidestepping potential questions and issues related to virtuous AMAs. Although virtuous AMAs are not perfect or morally flawless, his methodology has three merits. First, virtuous AMAs do not undermine the autonomy of human agents but rather drive them to make final moral decisions. By doing so, virtuous AMAs are considered feasible from an engineering perspective.²³

16. Gibert, 'The Case for Virtuous Robots', p. 137.

17. Gibert, 'The Case for Virtuous Robots', pp. 137–38.

18. Gibert, 'The Case for Virtuous Robots', pp. 138–39.

19. Ingmar Persson and Julian Savulescu, 'Moral Enhancement, Freedom, and the God Machine', *The Monist* 95.3 (2012), pp. 399–421.

20. Gibert, 'The Case for Virtuous Robots', p. 139.

21. Gibert, 'The Case for Virtuous Robots', p. 139.

22. Gibert, 'The Case for Virtuous Robots', pp. 139–40.

23. Gibert, 'The Case for Virtuous Robots', pp. 140–41. It is worth noting that Gibert's methodology differs from the one provided by Julian Savulescu and his colleagues. The latter also maintain that human agents should be moral decision-makers. However, they suggest creating a prototype AMA through programming basic moral principles into AI. These principles

To put it another way, it is easier to build an imperfect AMA than a perfect one. Second, establishing a base of virtuous people from diverse backgrounds offers a cross-cultural approach to virtue ethics, which is expected to broaden our moral vision. AI possesses immense processing power and vast amounts of data, far beyond human capabilities. It collects information regarding virtues and virtuous life from various communities, whereas humans reside in a specific community. To this extent, virtuous AMAs are expected to advance, as Gibert claims, the ‘less assertive, more flexible, and tolerant nature of virtue ethics’.²⁴ Third, virtuous AMAs can enhance thought experiments about addressing moral dilemmas as they generate flexible options rather than arbitrary decisions.²⁵ They are not designed to impose moral norms or values upon human agents. Rather, human agents are invited to explore diverse approaches to grappling with moral dilemmas through engaging with virtuous AMAs.

That said, three issues lie beneath Gibert’s virtuous AMA. First, Gibert is overly confident in the ability of AI algorithms to produce proper moral advice based on collected moral data. AI algorithms are not morally neutral but have the potential to perpetuate moral issues. A key moral concern about algorithms is algorithmic bias, which means that systematic discrimination against certain individuals or people is built into algorithms.²⁶ For example, Facebook’s feed algorithms prioritise news that amplifies political polarisation, leading to bias and conflict between political groups.²⁷ Therefore, even if all collected data is moral, AI can still produce immoral advice. Worse still, the data collected for training AI is inherently biased precisely because the human agent as the subject of data is either biased against or in favour of particular moral principles, values, and people. For this reason, many AI ethicists, including prominent scholars like Aimee Van Wynsberghe and Scott Robbins, assert that AI is by no means characterised as moral due to its intrinsic bias.²⁸

The second issue underlying Gibert’s virtuous AMA is his obliviousness to a crucial limitation of machine learning. Briefly speaking, machine learning is a tool to work out

are selected by a group of moral experts from various moral theories and communities. The prototype AMA will evolve in different circumstances. Savulescu and his colleagues are oblivious to the fact that, to a significant extent, these selected moral principles enable AMAs to have the last word on moral decisions. Julian Savulescu and Hannah Maslen, ‘Moral Enhancement and Artificial Intelligence: Moral AI?’, in Jan Romportl, Eva Zackova and Jozef Kelemen (eds.), *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide* (Cham: Springer, 2015), pp. 79–95; Alberto Giubilini and Julian Savulescu, ‘The Artificial Moral Advisor. The “Ideal Observer” Meets Artificial Intelligence’, *Philosophy & Technology* 31 (2018), pp. 169–88.

24. Gibert, ‘The Case for Virtuous Robots’, p. 142.

25. Gibert asserts that virtuous AMAs can be regularly revised through ‘consulting virtuous exemplars again’ (Gibert, ‘The Case for Virtuous Robots’, p. 142).

26. Alexander Babuta and Marion Oswald, *Data Analytics and Algorithms in Policing in England and Wales: Towards a New Policy Framework* (London: Royal United Services Institute for Defence and Security Studies, 2020), p. 4.

27. David Garcia, ‘Effect of Facebook’s Feed Algorithm Put to the Test’, *Nature* 620.7972 (2023), pp. 39–41.

28. Aimee Van Wynsberghe and Scott Robbins, ‘Critiquing the Reasons for Making Artificial Moral Agents’, *Science and Engineering Ethics* 25.3 (2019), pp. 719–35.

probabilities from data. The outputs of machine learning are only concerned with the association and correlation between variables in data. Yet, correlation does not imply causation, which is characteristic of the connection between human moral decision-making and moral deliberation. Mihaela Constantinescu and Roger Crisp's distinction between machine learning and virtuous human beings is worth quoting here.

Unlike human beings, [machine learning-based] AI systems deliberate on the right reasons by constantly and instantly calculating the right thing to do out of (potential) infinite possibilities embedded by the huge set of data that is fed for training. And the way the AI algorithm reaches its strategies based on the datasets is based on a mathematical calculus ... Ethical deliberation is inherently different from mathematical deliberation because it takes into account the particular contexts, including reasons, of particular people.²⁹

Constantinescu and Crisp make it clear that the human agent's moral decision-making process entails the *causation* of the connection between mental activities and external circumstances. Such causation varies across contexts, communities, and cultures, and is more complicated and complex than mathematical calculation. Although Gibert's virtuous AMA can enhance thought experiments about addressing moral dilemmas, its moral advice may, nevertheless, have nothing to do with the realities of human moral life.

The third, and perhaps the most critical, issue is Gibert's conception of virtue. In his position, machine learning can learn virtues from virtuous people. However, what remains undiscussed is that machine learning can also learn vices from humans. While collecting moral data from virtuous people through questionnaires, due attention must be given to the fact that moral virtues and vices are often entangled. For example, honesty is a moral virtue, but brutal honesty is a moral vice because it lacks empathy for others' feelings and often causes harm. Virtues cannot be separated from their contexts. As Alasdair MacIntyre observes, the implications of virtues should be construed within particular circumstances and with reference to social relationships, leading to the flourishing of humans as a community.³⁰ Gibert's abstraction of virtues from contexts and communities through mathematical AI models results in the neglect of human vices while maintaining an overly optimistic position on human virtuous life. Hence, it can be argued that this critical issue underlies the first two problems of Gibert's methodology.

Be that as it may, the current hype surrounding AI, particularly its spread through social media, has yielded enthusiasm about applying AI to human life on all levels. How can Christian ethics respond to virtuous AMAs like Gibert's in the scenario of applying AMAs to Christian life? Before addressing the question, I will spell out Aquinas's theological virtue ethics.

29. Mihaela Constantinescu and Roger Crisp, 'Can Robotic AI Systems Be Virtuous and Why Does This Matter?', *International Journal of Social Robotics* 14 (2022), pp. 1547–57, at p. 1551.

30. Alasdair C. MacIntyre, *Dependent Rational Animals: Why Human Beings Need the Virtues* (London: Duckworth, 1999), pp. 108–109, 160–61.

Theology of Virtue and Being Virtuous

Virtues refer to personal excellences and character traits, which are integrated with actions, dispositions, intentions, motivations, and desires. Aquinas argues that virtues are not loosely attached to humans; instead, human virtues are *habitus*.³¹ In his position, *habitus* etymologically means ‘to have’, which implies that ‘a thing has a relation in regard to itself or to something else’.³² In this sense, virtues as *habitus* are qualities of the human being, which are embodied in such relations.³³ Moreover, Aquinas highlights the ontological significance of *habitus*. He claims that ‘quality ... implies a certain mode of substance’, which refers to ‘a certain determination according to a certain measure’. This determination leaves its mark on human substance, generating the correspondence between the mode of human substance and human nature.³⁴ In this way, Aquinas weaves together *habitus* and human ontology, which orients our analysis of his concept of virtue.

That being so, virtue must be construed with reference to the being of humans. On this point, it is worth noting that, for Aquinas, a virtue generally means a quality of humans, or more precisely, a quality of the soul; it is not conceptually restrained within the confines of human morality. Moral virtues refer to moral *habitus* that dispose human appetite (natural tendency or inclination towards a particular object) towards proper acts and are in conformity with reason.³⁵ In other words, ‘[m]oral virtue perfects the appetitive part of the soul by directing it to good as defined by reason’.³⁶ From this it follows that moral virtues as *habitus* are linked up with the whole human being. In Aquinas’s virtue ethics, as Tobias Hoffmann rightly notes:

Only moral virtues, that is, habits that are in the appetitive part of the soul or depend for their existence on the appetitive part, are virtues in the full sense. Moral virtues not only *make one able to act well*, but they also ensure that *one make good use of this ability*; they are good habits not only materially (insofar as they are ordered to the good) but also formally (insofar as they are ordered to the good under the aspect of good).³⁷

Accordingly, human persons are not only behaviourally but also *essentially* virtuous. To borrow Bonnie Kent’s words, moral virtues become ‘second nature’ for human agents

31. Thomas Aquinas, *Summa Theologiae*, trans. Fr. Laurence Shapcote, ed. The Aquinas Institute, Latin/English Edition of the Works of St. Thomas Aquinas, vols. 13–20 (Green Bay, WI: Aquinas Institute, 2012), I-II, q.55, art.1, resp; hereafter *ST*.

32. *ST*, I-II, q.49, art.1, resp.

33. *ST*, I-II, q.49, art.1, resp.

34. *ST*, I-II, q.49, art.2, resp.

35. *ST*, I-II, q.58, art.2, resp. Aquinas defines ‘appetite’ as ‘an inclination of a person desirous of a thing towards that thing’; *ST*, I-II, q.8, art.1, resp.

36. *ST*, I-II, q.58, art.4, resp.

37. Tobias Hoffmann, ‘The Intellectual Virtues’, in Brian Davies and Eleonore Stump (eds.), *The Oxford Handbook of Aquinas* (Oxford: Oxford University Press, 2012), p. 328; emphasis added.

and are operative in transforming the agents into good human persons.³⁸ With this in mind, the human being as a whole should be considered when articulating moral virtues. Both virtuous human persons and the contexts where their appetite is disposed by moral virtues towards good acts are indispensable for the understanding of moral virtues.

How can moral virtues grow in humans? The answer to this question will shed light on whether virtuous AMAs can operate for the growth of human virtues in the context of the Christian faith. Aquinas argues that there are two causes for the growth of human virtues. Human acts cause the growth of human virtues that are related to the good in conformity with human reason, and only divine operation can cultivate human virtues related to God's law.³⁹ The process of the former is called habituation (i.e., repeated action), which leads to acquired virtues, whereas the latter entails infused virtues. The key distinction between acquired and infused virtues lies in their ends: infused virtues alone orient humans towards an ultimate end designated by God himself.⁴⁰ From this it follows that these two types of virtues have different measures precisely because, as Justin Anderson rightly notes, 'measures themselves arise according to the different ends towards which one moves'.⁴¹ This being so, infused virtues cannot be measured by human standards. The immeasurability of infused virtues highlights a critical distinction between virtuous human agents and the AMA since the latter is created with measurable mathematical models. I will come back to this subject later.

Aquinas asserts that the most prominent infused virtues are the theological virtues of faith, hope, and charity, which direct humans to the ultimate and supernatural end.⁴² Faith orients the intellect towards God. Hope and charity direct the will to the ultimate end, respectively steering the motion of intention and spiritual union towards God as the ultimate object of virtue.⁴³ Yet, notwithstanding the unique function of the theological virtues, infused virtues also encompass infused moral virtues. Unlike acquired moral virtues that dispose humans to proper acts according to the natural light of reason, infused moral virtues 'enable [humans] to walk as befits the light of grace', which 'is a participation of the Divine Nature'.⁴⁴ As such,

38. Bonnie Kent, 'Habits and Virtues (Ia IIae, Qq. 49–70)', in Stephen J. Pope (ed.), *The Ethics of Aquinas* (Washington, DC: Georgetown University Press, 2002), p. 116.

39. *ST*, I-II, q.63, art.2, resp.

40. *ST*, I-II, q.63, art.2-3; see also Thomas Aquinas, *Disputed Questions on the Virtues*, trans. E.M. Atkins, ed. E.M. Atkins and Thomas Williams (Cambridge: Cambridge University Press, 2005), 'On the Virtues in General', art.10. For a comprehensive overview of Aquinas's view of the relationship between infused virtue and acquired virtue, see Romanus Cessario, OP, *The Moral Virtues and Theological Ethics*, 2nd edn (Notre Dame, IN: University of Notre Dame Press, 2009), pp. 94–125.

41. Justin M. Anderson, *Virtue and Grace in the Theology of Thomas Aquinas* (Cambridge: Cambridge University Press, 2020), p. 32.

42. *ST*, I-II, q.62, art.1, resp.

43. *ST*, I-II, q.62, art.3, resp.

44. *ST*, I-II, q.110, art.3, resp. In this respect, William Mattison helpfully fleshes out Aquinas's view of the interconnection between infused moral virtue, God's grace, and the gifts of the Holy Spirit. William Mattison, 'Are the Actions of a Person Operating out of the Gifts of the Holy Spirit the Same as the Actions of That Person Operating out of Infused Virtue?', *Studies in Christian Ethics* 36.2 (2023), pp. 350–71.

Aquinas submits that, in order to '[correspond], in due proportion, to the theological virtues', infused moral virtues perfect humans for proper acts in relation to God after the theological virtues enable them to move towards God himself.⁴⁵ It suffices to say that the three theological virtues serve as the foundation of infused moral virtues. As will be seen, theological virtues and infused moral virtues together demonstrate the uniqueness of the growth of Christian moral virtues vis-à-vis moral enhancement by AMAs. Given the limited space of this article, I will focus on the theological virtue of hope and the corresponding infused moral virtues.

The theological nature of hope rests upon God, who is the object of hope. Hope guides humans towards the infinite good and eternal happiness, which consist in God himself.⁴⁶ Aquinas maintains that the '[*habitus*] itself of hope, whereby we hope to obtain happiness, does not flow from our merits, but from grace alone'.⁴⁷ As such, the theological virtue of hope differs radically from general expectation or anticipation. Brian Davies' observation is on the mark: for Aquinas, hope 'is not just any old looking forward to what one does not possess: it is a looking forward to what only God can provide'.⁴⁸ Equally important is that God is also the helper of hope for the ultimate end.⁴⁹ Divine assistance is indispensable for hope insofar as hope is intimately connected with difficulties in attaining eternal happiness and the infinite good. Amid such difficulties and struggles, hope drives humans to 'adhere to God'.⁵⁰ Specifically, hope generates charity 'through hoping to be rewarded by God' and motivates humans to trust in God's omnipotence and mercy with certainty.⁵¹

The theological nature of hope lays a foundation for the moral virtue of fortitude (*fortitudo*). Indeed, Aquinas contends that fortitude is not about hope but rather about fear and daring. Yet, he maintains that 'the virtue of fortitude, which by its very nature bestows firmness, is chiefly concerned with ... fear, which regards flight from bodily evils, and consequently with daring, which attacks the objects of fear *in the hope of attaining some good*'.⁵² From this vantage point, it can be seen that Nicholas Lombardo's argument that fortitude encompasses hope is ill-founded.⁵³ Given that fortitude is related more to fear than to daring, hope can be viewed as *a* root of fortitude insofar as it brings forth both daring and confidence in God as part of fortitude.⁵⁴ Given this theological foundation, fortitude empowers human agents to remain steadfast

45. *ST*, I-II, q.63, art.3, resp., ad.2.

46. *ST*, II-II, q.17, art.2, resp.

47. *ST*, II-II, q.17, art.1, ad.2; see also Aquinas, *Disputed Questions on the Virtues*, 'On Hope', ad.8.

48. Brian Davies, *Aquinas* (London: Continuum, 2003), p. 204.

49. *ST*, II-II, q.17, art.2, resp; art.7, resp.

50. *ST*, II-II, q.17, art.6, resp.

51. *ST*, II-II, q.17, art.8, resp.; for a helpful analysis of the interconnection between faith, hope, and charity, see Nicholas E. Lombardo, *The Logic of Desire: Aquinas on Emotion* (Washington, DC: The Catholic University of America Press, 2011), pp. 149–70.

52. *ST*, II-II, q.141, art.3, resp; emphasis added.

53. Lombardo, *The Logic of Desire*, p. 180.

54. *ST*, I-II, q.45, art.2; II-II, q.128, art.1, ad.2.

in the face of difficulties and obstacles while withstanding evils and safeguarding good with an eye to eternal happiness and infinite good.⁵⁵

Patience makes humans capable of trusting in God with fortitude and hope amid difficulties and struggles. Aquinas suggests that patience is a part of fortitude and then makes a nuanced distinction between the two virtues.

The act of fortitude consists not only in holding fast to good against the fear of future dangers, but also in not failing through sorrow or pain occasioned by things present; and it is in the latter respect that patience is akin to fortitude. Yet fortitude is chiefly about fear, which of itself evokes flight which fortitude avoids; while patience is chiefly about sorrow, for a man is said to be patient, not because he does not fly, but because he behaves in a praiseworthy manner by suffering (*patiendo*) things which hurt him here and now, in such a way as not to be inordinately saddened by them.⁵⁶

Accordingly, patience allows humans to safeguard good against sorrow through prolonged endurance.⁵⁷ Furthermore, the true patience of enduring evils and overcoming sorrow can only come from God through divine grace. For the delayed good that one safeguards with patience is directed towards the infinite good, namely God himself, the object of hope.⁵⁸ As such, hope informs patience of, and orients it towards, the infinite good and eternal happiness through moderating sorrow caused by obstacles.

Aquinas's theology of infused virtues depicts the distinct features of the Christian faith as a context for the growth of moral virtues through divine grace. By honing in on the theological virtue of hope, I have demonstrated that the moral virtue of patience is part of fortitude and thereby rooted in hope. In the next section, I will explore the extent to which virtuous AMAs can advise on the growth of the moral virtue of patience from the theological perspective of hope.

Can Virtuous AI Advise on the Growth of Patience?

Before examining virtuous AMAs from a theological perspective, let's recapitulate the three defects of Gibert's virtuous AMA: (1) over-confidence in AI algorithms, (2) the likelihood of the divergence between machine learning's outputs and the reality of our moral life, and (3) the rupture between virtues and their contexts.

Aquinas's view of the immeasurability of theological virtues echoes my critique of the first two defects of the virtuous AMA. As infused through divine grace, Aquinas asserts, '[a] theological virtue has for its object the first standard itself, which is not measured by another standard'.⁵⁹ To be sure, mathematical models and algorithms are, according to Aquinas, classified into the category of 'another standard', which, by its nature, differs from the divine standard. The three theological virtues cannot be embedded within AI

55. As such, Aquinas designates martyrdom as the highest act disposed by the virtue of fortitude; *ST*, II-II, q.124, art.3, s.c.

56. *ST*, II-II, q.136, art.4, ad.2.

57. *ST*, II-II, q.136, art.5, resp.

58. *ST*, II, q.136, art.3.

59. Aquinas, *Disputed Questions on the Virtues*, 'On Hope', art.1, art.7.

systems through mathematical measurement. Rooted in the theological virtue of hope, the infused moral virtue of patience will cast the measurability of AI into question and challenge the moral advice generated by virtuous AMAs on the Christian moral life.

Aquinas's concept of patience as a moral virtue also aligns with my critique of the third defect of Gibert's virtuous AMAs. According to Gibert, one of the reasons for using virtuous AI as a moral advisor is the AMA's capability to enhance human moral perception.

[Virtuous AMAs] with a certain degree of autonomy need not only to make decisions but to perceive the moral landscape ... [T]hey must be able to distinguish—to perceive—situations that are morally relevant ... [O]ne can use virtue exemplars not only to determine the correct actions, but also to identify what is morally relevant at the stage of moral perception. One can then consider using machine learning to allow algorithms to identify (to learn) things that are perceived as morally relevant.⁶⁰

A rationale behind this position is the belief in the efficiency of machine learning, which is characteristic of all AI-driven systems and common to all endorsements of AMAs. Indeed, machine learning possesses vast amounts of moral data for moral judgments, a quantity that is far beyond human moral knowledge. It also has tremendous processing power to identify, select, analyse, and clarify moral notions in seconds for humans when making moral judgments. However, as I noted earlier, moral data includes both moral virtues and vices, and so machine learning can also learn moral vices from historical data. Even more concerning, the AMA can generate incorrect and misleading advice just as efficiently as it can offer virtuous advice. In this light, the proposal of virtuous AMA seems to paint a rosy picture of AI-driven efficient moral growth, which turns out to be like looking through rose-coloured glasses.

That said, a contention may be raised that *human* moral advisors are also susceptible to offering misleading advice. Admittedly, the idea of the flawless AMA is just pie in the sky, as Gibert claims. Yet, many scholars have recently demonstrated that AI is often overloaded with the notion of trustworthiness, which has resulted in human over-trust in AI.⁶¹ Considering AI's efficiency in conjunction with humans' over-trust, the AMA's misleading advice may be more swiftly and readily accepted by humans than those from *human* advisors. The extent to which humans are willing to accept AI-generated outputs remains a matter of debate, and it is beyond the scope of this article to explore this question. However, recent studies on human–AI interaction suggest that human trust in AI advice is indexed to the performance of AI-driven systems.⁶² Human willingness to accept AI

60. Gibert, 'The Case for Virtuous Robots', p. 143.

61. Sebastian Krügel, Andreas Ostermaier and Matthias Uhl, 'Zombies in the Loop? Humans Trust Untrustworthy AI-Advisors for Ethical Decisions,' *Philosophy & Technology* 35 (2022), p. 17; Karoline Reinhardt, 'Trust and Trustworthiness in AI Ethics,' *AI and Ethics* 3.3 (2022), pp. 735–44.

62. E.g., see Patricia K. Kahr et al., 'It Seems Smart, but It Acts Stupid: Development of Trust in AI Advice in a Repeated Legal Decision-Making Task' (paper presented at the 28th International Conference on Intelligent User Interfaces (IUI '23), Sydney, NSW, Australia, 27–31 March 2023), pp. 528–39.

advice will increase in proportion to the performance of AI-driven systems. Therefore, AI's efficiency underpins humans' over-trust in AI.

My critical analysis of the three defects of virtuous AMAs brings to the foreground the tension between the moral virtue of patience and AI's efficiency, which is further crystallised in Aquinas's account of patience in relation to longanimity and constancy.

[T]he very delay of the good we hope for, is of a nature to cause sorrow, according to Prov. 13:12, *Hope that is deferred afflicteth the soul*. Hence there may be patience in bearing this trial, as in enduring any other sorrows. Accordingly longanimity and constancy are both comprised under patience, insofar as both the delay of the hoped for good (which regards longanimity) and the toil which man endures in persistently accomplishing a good work (which regards constancy) may be considered under the one aspect of grievous evil.⁶³

While AI's efficiency demands that moral decisions on moral dilemmas be made in seconds, the moral virtue of patience presupposes the readiness to endure difficulties over a long period, to reflect deeply on moral life, and to maintain hope for the delayed good. From this it logically follows that the AMA's emphasis on efficiency may dismantle the context for the practice of patience and thereby undermine the moral virtue of patience, leading to the decline of hope. AI-enhanced moral efficiency stands in opposition to prolonged endurance. Efficient machine learning is results-oriented and gives birth to a type of hope that envisions a path to attaining the good and happiness with fewer, if not without, difficulties, ambivalence, or struggle. Such a hope is not a theological virtue and fails to serve as a foundation for patience as an infused moral virtue, which enables Christians to overcome sorrow while enduring suffering and hardship.

The tension between AI's efficiency and the moral virtue of patience casts the AMA's moral advice into question. Even more problematic is that the virtuous AMA is likely to render humans inert in moral life, resulting in a decline in human moral virtues. With over-trust in AI's efficiency, human moral practice is diminishing. Shannon Vallor has gone into comprehensive details of this matter and flagged up the problem of moral deskilling caused by technological advancement.

[M]oral skills are typically acquired in specific practices which, under the right conditions and *with sufficient opportunity for repetition*, foster the cultivation of practical wisdom and moral habituation that jointly constitute genuine virtue ... [P]rofound technological shifts in human practices, if they disrupt or reduce the availability of these opportunities, can interrupt the path by which these moral skills are developed, habituated, and expressed.⁶⁴

63. *ST*, II-II, q.136, art.5, resp.

64. Shannon Vallor, 'Moral Deskilling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character', *Philosophy and Technology* 28 (2015), pp. 107–24, at p. 109; emphasis added.

Vallor argues elsewhere that moral deskilling will be incurred by AI's automation that 'crowds out our own opportunities to practice moral and political reasoning'.⁶⁵ Moral skills need both sufficient opportunity for repetition—that is, long-term moral practice and action—and patience to moderate sorrow resulting from recurrent hardship and even failure.

It is worth pausing here to clarify that the purpose of my emphasis on patience and concomitant sorrow is not to justify sufferings, evils, difficulties or hardship during the growth of moral virtues. Rather, its objective is to demonstrate that patience always implies the acceptance of time and endurance. The growth of virtuous life should allow for moments of *inefficiency* and *sluggishness*, which signify the finite, fragile, and creaturely nature of humans, that is, the condition of human moral life.⁶⁶ It is during times of inefficiency that patient moral life is guided by hope—which is directed by God's grace—towards the infinite good that is distant from here and now.

This theological inquiry into patience as a moral virtue repudiates a technosolutionist approach to Christian moral life.⁶⁷ Christine Rosen summarises the features and risks of technosolutionism.

Technosolutionism is a way of understanding the world that assigns priority to engineered solutions to human problems. Its first principle is the notion that an app, a machine, a software program, or an algorithm offers the best solution to any complicated problem ... In the rush to embrace immediate technological fixes, its advocates often ignore likely long-term effects and unintended consequences.⁶⁸

Yet notwithstanding that the proposal of virtuous AMA respects human decision-making autonomy, the AMA's *modus operandi* betrays a flavour of technosolutionism. It prioritises the solutions generated by machine learning and neglects a long-term detriment, stemming from the tension between efficiency and the moral virtue of patience, to moral skills and the growth of moral virtues.

That said, my critique of virtuous AI and AMAs does not overthrow the contributions that AI technology can make to human moral life. For example, as noted earlier, the AMA

65. Shannon Vallor, *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Learning* (Oxford: Oxford University Press, 2024), p. 117.

66. Luke Bretherton helpfully discusses the relationship between human finitude and morality; Luke Bretherton, *A Primer in Christian Ethics: Christ and the Struggle to Live Well* (Cambridge: Cambridge University Press, 2023), pp. 133–55.

67. Thomistic theological virtue ethics also turns down the application of technosolutionism to human moral life at large, which echoes current studies rejecting the moral status of AI, e.g., Robert Sparrow, 'Why Machines Cannot Be Moral', *AI & Society* 36 (2021), pp. 685–93. I have explored this question elsewhere, Ximian Xu, 'How Virtuous Can Artificial Intelligence Become? Exploring Artificial Moral Advisor in Light of the Thomistic Idea of Virtue', *Perspectives on Science and Christian Faith* 76.2 (2024), pp. 66–76.

68. Christine Rosen, 'Technosolutionism Isn't the Fix', *The Hedgehog Review* 22.3 (2020), p. 84.

can advise on cross-cultural understandings of moral virtues by providing quick clarification of different moral principles and theories. In this respect, AMAs can be viewed as technological tools to assist human moral growth. This corollary is resonant with recent theological engagement with AI. A notable example is Noreen Herzfeld's use of 'created co-creator' to spell out human–AI relationship from the perspective of the *imago hominis*, demonstrating that AI is a created co-creator with humans, rather than a created creator, to perform certain tasks.⁶⁹ From this vantage point, it can be argued that virtuous AMAs can advise on human moral life provided that humans do not lose moral skills. But how?

The theological nature of patience can help form a critical yet appreciative understanding of the function of the virtuous AMA. Rachel Muers argues that biblical passages clearly show the eschatologically oriented nature of God's patience (e.g., 2 Pet. 3:8-9, 15), which means the eternal God gives time to creation and directs it towards eternity.⁷⁰ This eschatological feature of patience aligns with Aquinas's theological account of hope-based patience, showing that patience guides the Christian moral life towards the infinite good and happiness. In this respect, Muers foregrounds the connection between silence and God's patience through analysis of Rev. 8:1. She suggests that 'silence in heaven' signals 'the story of God with the world', indicating that God's listening silence is part of God's communication to his creation, shaping the life of God's people.⁷¹ God's patience in listening silence indicates that humans 'are [the] image of God in [their] listening' and should be 'responsible hearers'.⁷²

Muers's elaboration on responsible hearers centres on the listening church, but the idea of patience in listening silence can be appropriated to expand upon how AMAs can assist the growth of human virtues without weakening human moral skills. A rationale raised for the design and operation of AMAs is that *the virtuous AMA is a technological mediator that represents the listening human advisor's silence and communicates moral advice to advisees as patient inquirers*. From this vantage point, we can see a critical yet appreciative approach to the application of AMAs. As Gibert suggests, the AMA assimilates a multitude of moral principles and values from virtuous people across contexts and cultures to advise on virtuous life. In this sense, these virtuous people can be viewed as hearers who silently listen to those who are seeking moral advice.

This rationale is of great importance for the design of AMAs, which, in turn, facilitates the growth of advisees' moral virtues with patience. Specifically, the designers, stakeholders, and engineers of AMAs should reckon with how to embed the characteristic features of patience into AI. Jaco Hamman helpfully observes: 'For engineers and code writers to build patience ... into an AI's algorithm, the AI needs to discover time beyond chronological time. It needs to see time as an emotional and relational construct

69. Noreen Herzfeld, 'Co-creator or Co-creator?: The Problem with Artificial Intelligence', in Ulf Görman, Willem Drees and Hubert Meisinger (eds.), *Creative Creatures: Values and Ethical Issues in Theology, Science and Technology* (London: T&T Clark, 2005), pp. 45–52.

70. Rachel Muers, 'Silence and the Patience of God', *Modern Theology* 17.1 (2001), pp. 85–98, at pp. 86–87.

71. Muers, 'Silence and the Patience of God', pp. 89–91.

72. Muers, 'Silence and the Patience of God', p. 93.

contextually informed'.⁷³ With this in mind, the design of AMAs should not prioritise the efficiency of generating moral advice but should instead strike a balance between efficiency and patience, allowing human moral advisees some time for patient moral deliberation and reasoning. For example, instead of offering moral advice straightforwardly, virtuous AMAs can generate some heuristic questions regarding the advisee's context, culture, intention, concerns, and so forth.⁷⁴ These questions are not value-free. Human moral advisors *delegate* virtuous values to AMAs in order to mediate their moral advice patiently and silently to moral advisees.⁷⁵ In terms of the Christian faith as the context for virtuous life, these questions are particularly relevant to a faith community where advisees patiently practise their virtues based on the three theological virtues. In this way, the patience of virtuous people assists the practice of the advisee's moral virtue of patience, leading to the growth of the latter's moral virtues. Furthermore, since the selected virtuous people are not flawless, the convergence of the double patience in the virtuous AMA reminds moral advisees of critical engagement with the AMA's advice.

That being so, the question 'Can Virtuous AI Advise on the Growth of Patience?' does not have a simple either-or answer. On the one hand, patience as an infused moral virtue is immeasurable in terms of AI algorithms and should be oriented by the theological virtue of hope. To this extent, the AMA cannot advise on patience, much less to say on hope. On the other hand, the virtuous AMA can be considered a tool to mediate the patient silence of human moral advisors to advisees and assist their patient moral growth.

Conclusion

My exploration of AMAs through the lens of Thomistic virtue ethics has uncovered significant limitations in applying AI to Christian moral life, a conclusion far less optimistic than many scholars had anticipated. Being virtuous cannot be mathematised into AI. For the infused virtues—particularly the three theological virtues—are imparted to humans by God's grace and, consequently, cannot be measured according to the *human* standard of the mathematical operation of AI algorithms.

The theological virtue of hope and the infused moral virtue of patience have been used to flesh out the limitations of employing the virtuous AMA to facilitate the growth of Christian moral virtues. The emphasis of this article falls on the tension between efficiency and patience, demonstrating the likelihood to lead to the decline in moral

73. Jaco J. Hamman, *Pastoral Virtues for Artificial Intelligence: Care and the Algorithms that Guide Our Lives* (Lanham, MD: Lexington Books, 2022), p. 73. Although Hamman discusses patience in relation to artificial general intelligence, his corollaries made in this book are useful for the design of real-life AI.

74. The answers to these questions include the advisee's privacy. Concerns about privacy warrant further studies on the ethical and legal framework for AI applications.

75. Delegation does not mean that AI can operate independently of humans. For a detailed discussion on this, see Sara Lumbreras, 'Lessons from the Quest for Artificial Consciousness: The Emergence Criterion, Insight-Oriented AI, and Imago Dei', *Zygon: Journal of Religion and Science* 57.4 (2022), pp. 963–83, especially pp. 973–74.

virtues with the application of AMAs to moral life. That being said, the theology of patience can provide a conceptual apparatus through which to qualify the AMA as a tool to mediate the human moral advisor's patience and assist the growth of the advisee's patience. This qualification can encourage AI engineers, designers, and stakeholders to reappraise the development of AI with an eye to embedding patience and other moral virtues alike into AI-driven systems. In short, an overly optimistic attitude towards AI results in an over-trust in the application of AI-driven artefacts to moral life. Such enthusiasm for AI technology trivialises not only God's grace but also moral virtues and dilemmas, ultimately leading to a decline in moral virtues.

Acknowledgements

A short version of this article was presented at the 2024 Annual Conference of the Society for the Study of Christian Ethics. I would like to thank Revd Dr Helen Dawes, Dr Eve Poole, and Prof. Esther Reed for their feedback.

Author's Note

Ximian Xu is currently a Senior Research Associate in the Faculty of Divinity, University of Cambridge, Cambridge, United Kingdom.


Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Ximian Xu  <https://orcid.org/0000-0002-3159-1224>