Generative AI Cannot Replace a Spiritual Companion or Spiritual Advisor

Harris Wiseman

Abstract: Numerous generative AI spiritual advisor platforms are freely available, with some thinkers aspiring to create an artificial spiritual companion. While not ruling out that a future technology might serve here, the paper argues that generative AI is simply not the right tool for this job. The paper describes spiritual companionship and advice-giving as a relational gold standard. This is contrasted with the operations and limitations of generative AI (specifically, large language models) in order to highlight unsurmountable obstacles, meaning that generative AI cannot substitute for spiritual companionship or advice. The paper follows three central lines: first, a spiritual companionship requires a relation between two bodied beings (for various reasons explored in this paper); second, spiritual companionship is more than propositional exchange (of text inputs and outputs), a limitation which threatens to reduce spiritual advice to a narrow problem-solving rubric; and third, the paper explores the question of what happens to spiritual advice, given that "spiritual" chatbots are already a disruptor and a very impoverished product. It is argued that the described failings of generative AI as spiritual companion are baked-in, intrinsic to how it works.

> *The human person is created by God … Together with Jesus and through him, we return to the Father in the Spirit. This is our Passover in the Lord. This process, this return journey, may be called spiritual regeneration … transformation in Christ, in God … sanctification. Spiritual direction then is the gift, the charism, the ministry of guiding a person in and through his/her Passover in the Lord. It is a unique participation in another's spiritual regeneration. … Spiritual direction is a God-willed contribution of one person to another's process of spiritualization.[1]*

The question of whether generative AI is itself spiritual shall be put aside as *prima facie* negative. It has no body, it has no intuition, it has no spiritual hunger, it does not have the basic cognitive systems which support spiritual awareness,[2] it has no relationships—it has none of the foundations on which spirituality (in any sense humans would recognise) could arise. Despite a recent survey suggesting two thirds of users of ChatGPT believed it to be genuinely sentient in some regard,[3] generative AI has no comprehension of its inputs or outputs. Generative AI is not intelligent in the sense of artificial general intelligence (some would argue that it is not intelligent in any sense, a view that will become apparent throughout this paper).[4] The general consensus

---

1   Francis Kelly Nemeck and Marie Theresa Coombs, *The Way of Spiritual Direction* (Collegeville, MN: Liturgical Press, 1985), 15–16.

2   Harris Wiseman and Fraser Watts, "Spiritual Intelligence: Participating with Heart, Mind, and Body," *Zygon* 57:3 (2022): 710–718, https://doi.org/10.1111/zygo.12804.

3   Eric Hal Schwartz, "Survey Says Most Believe Generative AI Is Conscious, Which May Prove It's Good at Making Us Hallucinate, Too," *TechRadar*, 16 July 2024, https://tinyurl.com/2tjaye83.

4   George Siemens writes: "AI is broadly defined in two categories: artificial narrow intelligence (ANI) and artificial general intelligence (AGI). To date, AGI does not exist … Most of what we know as AI today has narrow intelligence—where a particular system addresses a particular problem. Unlike human intelligence, such narrow AI intelligence is effective *only* in the area in which it has been trained: fraud detection, facial recognition or social

among computer scientists is that it should not be used as a tool in tasks that require empathy, moral context, or which have legal and health implications.[5] Already, bells should be ringing for persons who think it can substitute spiritual companions. If generative AI does not even understand what it is saying and has no sense of empathy or moral context, then the idea that it can itself be spiritual is simply out of the question.

However, the question of whether a given technology can be a valuable spiritual tool is different. A tool does not need itself to be spiritual in order to benefit humans seeking spiritual support. Spiritual tools or "technologies" (i.e., from *techne*, the Greek word for art, craft, making, or doing) have been devised across all religions and used throughout the millennia, be they as rudimentary as using a knotted cord to keep track of one's rosary and drinking green tea to keep one awake through nightly meditative vigils; or, as more modern technologies, say, using digital online icons for devotion in the Eastern Orthodox church and using breathing apps to lead one through pranayama practice. The pandemic was a powerful stimulus for creative thinking about how to use distance technology in ways that support spiritual practice, a testing ground that produced many good and bad results.

There is no fundamental issue with the idea of technology in and of itself assisting or scaffolding persons in spiritual practice—just so long as the tool being used serves to work with the basic nature of the practice rather than subverting it. The argument in this paper is that, when it comes to spiritual advice or, more pressingly, as a spiritual companion, generative AI is not of this order. While it is not impossible to use it well (this is down to the user), the argument here is that the very processes involved in generative AI lend themselves only to an impoverished kind of spiritual practice. Worse, by their very

---

recommendations, for example." George Siemens, "Not Everything We Call AI Is Actually 'Artificial Intelligence': Here Is What You Need to Know," *The Conversation*, 22 December 2022, https://tinyurl.com/337h8j5j.

5   Ava McCartney, "When Not to Use Generative AI," *Gartner*, 23 April 2024, https://tinyurl.com/2rjevxs3.

nature, they subvert the processes of spiritual relationship rather than supporting or scaffolding them. Nowhere are these tendencies clearer than in the context of seeking spiritual advice or companionship.

This paper will argue the following three points:

First, relationships of spiritual advice are fundamentally embodied—this is more than just saying that it's better to have this engagement with a real person as part of a relationship (a point that is obvious, yet true). Rather, there is a profound misunderstanding in place, which thinks that spiritual advice is just a matter of dispensing helpful verbal propositions. This misconstrues spiritual advice as mere problem-solving, as a process of spoon-feeding answers to seekers. It ignores nonverbal communication (mediated predominantly through the body), and might in the worst case lead to the creation of a generative echo chamber—that is, a completely insular process of merely feeding back to persons what they want to hear rather than opening them to positions that might threaten the safety of their pre-established views. To remove the body and relationality from spiritual advice is to remove something essential to its grounding.

Second, by its very nature, generative AI works to find the most predictable possible response to its inputs. Sometimes, spiritual advice is as simple as giving a person the obvious counsel (e.g., encouraging someone to forgive another or to refrain from some negative course), but spiritual advice cannot be limited to dispensing verbal clichés. That cannot be the whole and sum of it. Often, in spiritual advice, one needs to be told what one does *not* want to hear, and needs to be given more than the most predictable responses. Having no insight whatsoever, generative AI has no way of giving more, except as hallucination and error. This failing is arguably baked into the very structure of generative AI.

Third, one has to face up to the reality that generative AI advisors and companions are very popular, and increasingly so, for a range of reasons, some of which are legitimate and some more dubious. The implication is that spiritual advice has already been disrupted by this technology, and one has to confront the repercussions of an increasingly

widespread preference for an extremely impoverished version of spiritual advice—that is, preference for seeking advice on spiritual matters from a device which has no body, offers no two-way relationship, has no empathy or comprehension of its inputs or outputs, and which is only capable of giving unreliable factual statements, random errors, and the most predictable, cliché responses to what could be serious spiritual problems and concerns.

To moderate that wholly critical analysis a little, one can acknowledge that generative AI, in the right hands, for the right sort of person, used in a particular sort of creative way, might be able to proffer piecemeal benefits for spiritual practice. It is not asserted here that generative AI is irredeemably useless in spiritual terms. Depending on how it is brought to bear by the end user, it is not impossible that it be applied so as to enrich spiritual practice in some limited ways. Also, the focus here is on generative AI (specifically, large language models), and it is not impossible that some other mechanism could be devised that might, somehow, perform better. Rather, limiting discussion to spiritual advice, spiritual companionship and generative AI specifically, one must conclude that these technologies can offer nothing more than a very low-grade replica which, so far from being harmless, threaten to subvert the ideal practices in question in very serious ways.

The Ideal Form of Christian Spiritual Direction

From the opening quote, it is transparent that the relationship with the spiritual companion or advisor is sacred. It is aimed at sanctification, a person's spiritual growth, drawing them closer and closer to God, which is deemed the ultimate motive and value of human existence. This is Christian phrasing, obviously. It goes without saying that there is a diversity of types of spiritual direction, even within Christianity, let alone across other religions, and beyond.[6]

---

6   For more, see John Mabry (ed.), *Spiritual Guidance Across Religions: A Sourcebook for Spiritual Directors and Other Professionals Providing Counsel to People of Differing Faith Traditions* (Nashville, TN: Turner Publishing 2014).

Yet, across the religions, relationship with a spiritual advisor is of the utmost worth. Above all, as "soul friendship,"[7] which is very different from other kinds of companionship or relationship, spiritual direction necessitates relations which can be profoundly testing. Somewhat like psychotherapy (though usually with different *teloi*, means, motives, and language), the spiritual guidance relationship involves a slow and difficult examination, a transformation across all elements of the human person, fostering emotional, psychological, and spiritual growth in the context of ongoing dialogue.[8] Nemeck and Coombs write: "of all the possible ways of assisting a person mature, the most difficult and also the most neglected is undoubtably spiritual direction."[9]

Though the literature on spiritual direction is vast, and despite there being divergences in how spiritual advisors are conceived both within Christianity and across the religions, there remain a range of significant overlaps.[10] On that basis, it is possible to sketch a few features of an ideal form of spiritual direction that everyone might be able to recognise: The relationship is sacred; it is aimed at bringing persons closer to God (or the divine, more broadly construed); it involves self-examination, correction; it is therefore transformative, and necessarily involves at least some challenge; it is dialogical; it is person-to-person, relational; it involves feeling and refined affective sensitivity. Thus, above all, for the purposes of this paper—*it is embodied.* Spiritual advice goes on between two persons, both of whom have a body. It is from generative AI's lack of a human body that its ultimate failure as spiritual advisor must necessarily derive.

In turn, it might be tempting to muse: Would it not be helpful if a spiritual AI could take on some, or all of that role, AI not being subject

---

7   See William F. Clocksin, "Guidelines for Computational Modeling of Friendship," *Zygon* 58:4 (2023): 1045–1061, https://doi.org/10.1111/zygo.12919; Henri J. M. Nouwen, *Spiritual Direction: Wisdom for the Long Walk of Faith* (London: Harper Collins, 1981/2018).

8   Nemeck and Coombs, *The Way of Spiritual Direction*, 13.

9   Nemeck and Coombs, *The Way of Spiritual Direction*, 13.

10   See Mabry, *Spiritual Guidance Across Religions*, 5.

to so many human failings? Yet, the extremely high theological aspiration noted at the start of this paper highlights just how significant and how challenging the prospect of spiritual advisor is, in this ideal form. The rest of this paper will show, in contrast, just how paltry generative AI must necessarily be as a substitute.

Generative AI and Spiritual Advice

To examine generative AI in the guise of spiritual advisor (and how this undermines the relational and embodied quality of spiritual advice), it is helpful to look to the raft of artificial spiritual advisor AI bots that continue to emerge on a daily basis, as well as to the kinds of relationships persons have with them. What is one to make of, for example, BibleGPT, AskJesusGPT, the RoboRabbi, or QuranGuideGPT? Or, for that matter, what is one to make of any of the New Age and alternative spiritual advisors, e.g., TarotMasterGPT, StarlightAdvisorGPT, ChatKundli, or the increasing range of dream interpreting, horoscope reading, zodiac elaborating bots promising to unlock one's destiny and answer all one's most profound spiritual questions?

Endlessly more of these bots seem to be pouring onto the market. In a related domain that's helpful to look at, on the (tellingly titled) *companion.ai* website alone,[11] there are over 475 bots labelled as "therapy bots."[12] These technologies are being used on a surprisingly large scale. Social science is nascent on how, why, or by whom these tools are being used. It remains to be seen over time how far these tools are mere novelty, used for amusement, or taken seriously. The suggestion for now is that use of these bots is motivated by productivity, entertainment, curiosity, and social and relational factors (i.e., chatbots are already used for companionship).[13]

---

11   Accessed via: https://openai.com/chatgpt/.

12   Joe Tidy, "Character.ai: Young People Turning to AI Therapist Bots," *BBC News*, 5 January 2024, https://www.bbc.co.uk/news/technology-67872693.

13   Petter Brandtzaeg and Asbjørn Følstad, "Why People Use Chatbots," in *Internet Science*, ed. Ioannis Kompatsiaris et al., Lecture Notes in Computer Science 10673 (Cham: Springer, 2017), 377–392, https://doi.org/10.1007/978-3-319-70284-

Regarding the purported selling points of spiritual generative AI, the popular tech media remarks on BibleGPT, for example, that this bot is "trained on the teachings of the Bible and presented as an interactive website where users can ask questions … and receive biblical passages in response."[14] It is suggested that "this tool can help tech-savvy Christians level up their practice or provide new interpretations of the text by juxtaposing different pieces with each other." In general, then, "large language models bring the feedback of an imagined priest, rabbi, or swami to your screen, promising to deliver a 'spiritual' experience in the comfort of your own home." The hope is that these large language models "can become a way of connecting with your faith." As AI researcher Shira Eisenberg points out,

> future models can be trained on any text, religious or otherwise. The question becomes, which model will you choose to interact with? Someday, each person's base model will be trained on their own sets of values … this will result in conflicts in information and advice between different people's devices. That is not dissimilar to theological conversations that take place off the screen, however. All of it depends on whether you believe in a higher power, but if you do, it [BibleGPT] can become a way of connecting with your faith.[15]

Here one finds the basis of a future hope for a credible AI spiritual companion—each person having a "base model" which turns into a personalised AI, trained on their own values, and then working (presumably quite intimately, given the nature of the sensitive information being shared) with each individual choosing so to engage.[16]

1_30.

14   Nika Simovich Fisher, "Generative AI Has Ushered In the Next Phase of Digital Spirituality," *Wired*, 5 October 2023, https://www.wired.com/story/artificial-intelligence-spirituality-tarot/.

15   Quoted in Simovich Fisher, "Generative AI."

16   For more insight into the aspiration for spiritual companions, see Fraser Watts and Yorick Wilks, "Spiritual Conversation with a Companion Machine," *Zygon* (in press).

Marketplace Realities and Spiritual Materialism

There is an obvious yet important point to be made that spiritual generative AI threatens to undermine the relational and community-based elements of spiritual practice (a common concern over technology used during the pandemic, which for many then became preferable to rejoining religious congregations).[17] There is another problem here. For, one must be careful in how exactly one understands the words "personalised spiritual companion," particularly in the context of the competitive consumer market into which such companions would emerge. An analogue problem can be seen with the much complained-about social media algorithms that determine which material one is exposed to online.[18] Such "personalised" targeting has already been condemned for creating an echo chamber effect, serving merely to amplify one's own perspectives, to keep feeding back information which supports one's views and disproportionately reflecting back to oneself one's own prejudices.

It has been noted how destructive echo chambers are (in relation to newsfeeds) regarding the democratic health of a nation. What of the spiritual health of a person, which is precisely the domain a spiritual advisor is meant to address? "Personalisation" is an ambiguous term. The purported advancements that Eisenberg seems to regard as being so valuable (AI being trained on one's own personal value systems, i.e., a personalised AI spiritual companion) could very easily become a vehicle for *spiritual consumerism*.

---

17   Leonardo Blair, "Pastors, Churches Still Struggling in the Throes of 'Uncertainty and Unsettledness' Post-Pandemic: Study," *The Christian Post*, 4 September 2023, https://tinyurl.com/yc5r6phd.

18   See Ludovic Terren and Rosa Borge-Bravo, "Echo Chambers on Social Media: A Systematic Review of the Literature," *Review of Communication Research* 9 (2021): 99–118, https://doi.org/10.12840/ISSN.2255-4165.028. See also Miguel Risco and Manuel Lleonart-Anguix, "Feed for Good? On the Effects of Personalization Algorithms in Social Platforms," *CRC TR 224 2024 Discussion Paper Series* (University of Bonn and University of Mannheim, Germany), https://www.crctr224.de/research/discussion-papers/archive/dp580.

Given the forces of market competition which drive the survival of any given technology or product, as a consumer item, the importance of ensuring that "personalised" spiritual companionship not become such an echo chamber—i.e., an increasingly self-entrenching and self-enclosing data-set that excludes challenging views which oppose one's own values—must be highlighted. However, as a consumer product, that is exactly the tendency towards which such technologies lean.

It might be noted that, in many sectarian or highly conservative religious denominations, the capacity to exclude challenging information is seen as positive. For kinds of spirituality that are already insular, generative spiritual echo chambers would be a way of reinforcing a closed spiritual system, supporting a fortress mentality, the building of walls against information one does not want to hear. Personalised spiritual AI might be tantamount to an informational black hole, where nothing contradicting one's preestablished values and spiritual beliefs could hope to enter. Is this an ideal for a spiritual advisor, for a spiritual companion to aspire to? In terms of the gold standard of spiritual companionship noted above, such an echo chamber would be the exact inverse of what a spiritual companion should stand for. Put differently, what will eventually get marketed as spiritual advice or a spiritual companion may end up being the exact opposite, a fortress which actually subverts spiritual growth.

Of course, spiritual work will always involve personal preference, personal values, and choice—but essential to spiritual work is that these personal aspects be leavened through self-exploration and self-questioning; and expanded, gradually, through *challenge*. Something intrinsic to spiritual advice and companionship is that, over time, and through mutual agreement, one be challenged to stretch and question one's own limited perspectives, to locate and try to move beyond the various "idols" and "false gods" in one's life standing as barriers to spiritual progress and better relations with God.

The echo chamber which seems to be the aspiration mentioned above, is exactly at odds with the (sometimes) confrontational quality of spiritual advice. It is essential that spiritual practice gradually prune

away and otherwise facilitate some sort of change or growth *away* from problematic elements of one's preexisting value system. This challenge only comes through a confrontation with otherness, with others, and this is why spiritual advice absolutely needs to have an interpersonal dimension, or at least some way of connecting with a mature other who is willing and able to offer such a challenge. It is as such that spiritual companionship is a medium of transcendence and growth.[19]

Nothing in this section is meant to represent a decisive problem. The point here is simply to mark out various traps and temptations that are realities in the consumer marketplace in which such technologies must compete for their continued existence. Generative AI is a multibillion dollar industry that sustains itself through advertising revenue and data brokerage.[20] It is simply not in the interest of these companies to be credibly interested in the genuine spiritual growth of its user base. Spiritual companionship is a quest to greater maturity. Yet, consumerism is driven by encouraging persons to become more voracious consumers. And there is already a tendency of the "wellness

---

19 It is possible for generative AI bots to offer challenge to users. We can look, for example, to the RoboRabbi, which does offer challenges to overcome, e.g., "I challenge you to be a leader this week and lead someone in the right direction, whether it's [sic!] through your words or actions" (challenge, 28 July 2024, https://www.roborabbi.io/). However, this is essentially an example of the *gamification* process saturating contemporary app-consumer engagement (i.e., breaking down goals into incremental elements to create quantifiably measurable forward progress, with completion of each task being "rewarded" somehow, usually with a computerised token—a gem, medal, accolade, or title). However, there is much to be written on the contrast between this sort of predetermined, linear progress model constructed with clearly definable pre-established goal routes and the sort of personal challenge and accountability that arises spontaneously as one goes through the difficulties of life's challenges alongside a spiritual advisor. Spiritual progress may well involve universal human challenges needing to be overcome, but this gamification of spiritual virtue strikes me as unhelpful. Inward transformation does not occur in this linear, ever-forward-facing manner—nor do the life-challenges and relationship problems that stimulate such challenges come on demand; for example, bereavement. While scaffolding can certainly be helpful, even necessary, for growth, a gamified, predefined quantitative approach to spiritual growth misunderstands how spiritual challenge works.

20 Zak Doffman, "Google Confirms Serious AI Risks for iPhone and Android Users," *Forbes*, 12 February 2024, https://tinyurl.com/yue3y5yr.

industry" to co-opt spiritual practices and discourses as a vehicle to selling products and self-promotion. All of this threatens to vitiate a well-intended spiritual technology into just another pseudo-spiritual ego-support mechanism, another vehicle for "spiritual materialism"[21]— that is, the mistake of reducing spiritual worth to measurements of gratification and acquisition that entrench the self rather than inviting people to overcome their limits, to advance towards something greater and more all-encompassing.

Soul Friends, Propositional Exchange, and Human Bodies

When he was a young priest, Henri Nouwen understood spiritual direction as a formal relationship for supervision and accountability between a mature spiritual leader and a new priest or minister. Later in life he preferred the term spiritual friendship, or soul friend, which conveyed the necessary give-and-take in the process of spiritual accountability and faith formation.[22]

A fairly standard worry about generative AI chatbots is that they subvert human relationships, or replace them altogether. However, not everyone has a negative view of how damaging the impact of bots as social companions are, or will be. On the basis of their research, Guingrich and Graziano write:

A common hypothesis is that these companion bots are detrimental to social health by harming or replacing human interaction … Contrary to expectations, companion bot users indicated that these relationships were beneficial to their social health … [Moreover,] perceiving companion bots as more conscious and human-like correlated with more positive opinions and better social

---

21 See Chögyam Trungpa, *Cutting Through Spiritual Materialism* (Boston: Shambhala, 2008).

22 Michael J. Christensen and Rebecca J. Laird, "Preface," in Nouwen, *Spiritual Direction*, 3.

health benefits. Humanlike bots may aid social health by supplying reliable and safe interactions[23], without necessarily harming human relationships.[24]

Rather than primarily seeking advice from generative AI, it may be that the chief motivator for engaging with them is companionship. Relationships to generative AI more broadly are already being taken up on a vast scale. Henry Shevlin writes:

> Services such as Replika offer users an "AI companion who cares," both in the form of friendly conversation and romantic and even erotic interactions. Over the last five years, AI systems like these have grown rapidly in sophistication and popularity, with Replika alone now boasting more than 10 million registered users, and new conversational chatbot apps and platforms emerging at rapid speed.[25]

The problem with construing spiritual AI in companionship terms is that spiritual advisors are not "buddies" or companions in the usual everyday sense (see Clocksin for a more extensive account of the characteristics and history of spiritual friendship).[26] Indeed, the added complexity of spiritual friendship is precisely that it epitomises the very highest standard of friendship. This paper will not comment on the prospect of generative AI companionship taken more broadly, suffice to say one might be very sceptical about how satisfying such

---

23 This assumes such relations are indeed safe. One characteristic of spiritual relationships (or relationships more generally) is precisely that they do involve risk. Yet, generative AI chatbots are risky too, in a different sense, in that they have no empathy and often suggest wild and foolish things.

24 Rose E. Guingrich and Michael S. A. Graziano, "Chatbots as Social Companions: How People Perceive Consciousness, Human Likeness, and Social Health Benefits in Machines," *ArXiv* December 2023, https://doi.org/10.48550/arXiv.2311.10599.

25 Henry Shevlin, "All Too Human? Identifying and Mitigating Ethical Risks of Social AI," *Law, Ethics & Technology* 2 (2024): 0003, https://doi.org/10.55092/let20240003.

26 Clocksin, "Guidelines," 1045–1061.

a relationship would be in the long term. In any case, problems with generative AI companions are amplified to the utmost degree when talking about spiritual friendship, which is arguably the highest possible watermark, the gold standard, of what friendship might ever aspire to reach. As has been highlighted throughout, spiritual companionship is a very particular kind of companionship, one which is not just about gratifying relations but which also necessitates accommodating challenge and difficulty. Spiritual companionship is about mutual growth, and is directed not just at maturity, but ultimately towards God and entails walking together in order to get closer to God.

The crucial difference between ordinary friendships and spiritual friendship, or *soul* friendship, as Henri Nouwen (a foremost writer on spiritual direction) conceived it, is explained as follows:

> For Henri, a spiritual director simply was someone who talks to you and prays with you about your life. Wisdom and direction emerge from the spiritual dialogue and relationship of two or more persons of faith committed to spiritual disciplines and the accountability required to live a spiritual life. Thus, spiritual direction as Henri understood it can be defined as a relationship initiated by a spiritual seeker who finds a mature person of faith willing to pray and respond with wisdom and understanding to his or her questions about how to live spiritually in a world of ambiguity and distraction.[27]

Spiritual relationship is about creating a profound two-way relationship between the seeker of guidance and someone with genuine experience and compassion. This relationship is able to support but also to elevate and, above all, according to Nouwen, helps the spiritual friends to hold each other in mutual accountability.

These characteristics should be borne in mind as one thinks about what exactly it is that large language models have to offer in terms of companionship. Significant problems arise from the fact that

---

27   Christensen and Laird, "Preface," in Nouwen, *Spiritual Direction*, 3.

large language models do nothing more than produce a series of propositional outputs. Engaging with generative AI involves nothing more than inputting text on a screen and receiving an automated textual output in response. I have already pointed out that spiritual advice and companionship are more than just a matter of bald propositional exchange and text messaging. Yet large language models are just that: propositional. What is at stake here is the reduction of the richness of communication and relationship solely to the propositional level. At least two spiritually crucial dimensions are lost: everything relating to genuine affect; and the importance in spiritual relationship of silence and presence.

In terms of affect, so much of what endows propositions with meaning is given through non-propositional factors, gesture, posture, glances, tone, pitch, the pace of breathing, and endlessly more. In other words, embodiment is decisive, and affect must be sincere. Is the highest aspiration to make a generative AI that can ape human affection so well that the illusion of reciprocal exchange can be sustained without interruption? Spiritual advice, however sober, involves human affectivity because it arises from two bodied persons in relationship. Much as persons may, on rare occasions, want to be rid of their emotions, precisely the impossibility of doing so is what sustains relationships, or breaks them. Sitting together and praying with another person is an intimate activity. To imagine a spiritual companionship utterly devoid of emotion (or where the emotion was completely one-sided), or relating with a technology that can at best give an illusion that allows one to forget the most salient truth—i.e., that one is dealing with a generative model, and nothing more—this is not a good grounding for either spiritual advisor or relationship.

At the start of this paper, it was mentioned that over two thirds of persons surveyed believed that ChatGPT was sentient in some manner. This is a heartbreaking illusion that underscores the importance of clarifying these relational matters. The idea of an AI "who cares" (per Replika) is a dangerous deception. Nomisha Kurian sums

up the problems with the "risk of anthropomorphism and inappropriate responses to sensitive disclosures" with chatbots as follows:

> A well-known risk of human-chatbot interaction is the tendency for users to perceive these agents as human-like (Shum, He, and Li 2018). Chatbots designed to emulate human behaviour and courtesy often prompt anthropomorphism, where users attribute human traits, emotions and intentions to them (Darling 2017). The design aims to create an impression of care, wherein users view the chatbot as empathetic and trustworthy (Weidinger et al. 2021). Even when users understand the chatbot's non-human nature, they may still engage with it as if it were human, mimicking human-to-human dialogue (Sundar and Kim 2019) … In other words, "knowing" that an AI system is artificial may not stop a user from treating it as human and potentially confiding personal or sensitive information.[28]

If there is a substantive risk that persons are liable to just forget that chatbots are utterly incapable of emotion, empathy, or the least care, that brings users' judgement into question regarding the value of these interfaces. Put differently, persons may very well be willing to accept generative artificial spiritual companions, but they should not. That acceptance may be based on a false impression of sentience in a technology that is incapable of it. The reality of interpersonal emotional exchange, as given in the practice of praying together—per genuine spiritual companionship—is not something that should be given over to a predictive model with a colourful interface when persons are so prone to projecting the illusion of emotions onto it.

Second, in spiritual relationship there is the crucial need for moments of silence and presence. Relationship is not just about talking and performing activities, it is sometimes just about being with, or simply being together—that is, a matter of presence. Presence requires

---

28   Nomisha Kurian "'No, Alexa, No!': Designing Child-Safe AI and Protecting Children from the Risks of the 'Empathy Gap' in Large Language Models," *Learning, Media and Technology*, first view (2024): 1–14, https://doi.org/10.1080/17439884.2024.2367052 (and the sources quoted therein).

a body. Humans communicate and express compassion and encouragement, approbation, and disapproval through bodily presence, not just by giving peppy words of edification or issuing text suggestions. Our gestures and glances often say infinitely more than our words can, particularly in conversations that have spiritual or subtle emotional dimensions. Can one imagine enjoying a meaningful silence with a generative AI chatbot?

What is left of spiritual advice once it has been divested of all its feeling and stripped down to nothing but the propositional level? Every aspect of spirituality and relationship that extends beyond the propositional level is torn away in engaging with generative AI. This misrepresents spiritual advice as a mechanical problem-solving exercise. It is not just that generative AI threatens to undermine the human relational quality, then, it undermines the community of spiritual advice. If one is assuming that spiritual advice is just about being spoon-fed some gospel quote or some edifying word of encouragement, then one is left with the narrowest caricature of what spiritual conversation consists in. Certainly, genuine spiritual advice does involve issuing propositional suggestions and imperatives, when the moment calls for it. However, these are determined collaboratively, not by spoon-feeding. Two-way relationship mediates insight, and the spiritual advisor should not be a crutch that simply tells one "what to do" in any given circumstance. There is a collaboration in spiritual relationship where a person is led to understand what and why the advice makes sense. Support is given and received. Progress is checked. All of these crucial dimensions, definitive of spiritual companionship, are wholly lacking in generative AI.

Baked-in Failings

Many failings of generative AI in spiritual context have been discussed above. Can one not just say that these are teething pains which are surely soon to be remedied by future iterations of generative AI? Most likely not. These problems are baked-in, necessary and inevitable

failings of generative AI for spiritual advice and companionship—this is because they arise, not out of error, but as artefacts of how generative AI works in the first place.

Contrary to any belief that generative AI constitute an all-purpose technology that can serve in any situation or context, it is understood that generative AI have a family of failings. Such baked-in problems include generative AI's extreme unreliability in the following regards:[29] giving reliably true and factual responses (this is due to its predictive quality, generative AI simply hallucinates the most likely output, whether factually true or not);[30] responses requiring empathy or moral context (this is termed generative AI's "empathy gap," i.e., it does not have any);[31] giving advice that invites courses of action relating to health or other safety concerns (this is for reasons of legal liability, and because of the previous two points, i.e., it has no empathy and is

---

29   Perhaps criticising generative AI seems like sacrilege given contemporary hype. However, generative AI is not the all-purpose tool that current popular opinion suggests. Generative AI is but one form of AI technology in an ecosystem of machine learning and other types of AI-related operations. As part of that ecosystem, generative AI is good for some things (content generation, conversational user interfaces, knowledge discovery); of medium help for other tasks (segmentation/classification, recommendation systems, perception, intelligent automation, anomaly detection/monitoring); and an extremely poor tool for others (prediction/forecasting, planning, decision intelligence, autonomous systems). It has various other problems (e.g., data privacy, liability, and unreliable outputs, to name a few). So, taking a position of extreme scepticism that it is going to come remotely close to offering a satisfying spiritual experience is far from sacrilege. Using the wrong tool for the wrong task leads to poor results. Given the uptake of generative AI, it is a service to point out these failings and limitations. For details, see McCartney "When Not to Use Generative AI."

30   Contrast this with Elon Musk's statement regarding his new AI chatbot: "Once known as TruthGPT, Musk initially billed Grok as 'a maximum truth-seeking AI that tries to understand the nature of the universe.' Musk has promised that Grok will be 'anti-woke' and offers a 'Fun Mode' as well as an 'Unhinged Fun Mode'." See Rob Waugh, "Elon Musk's X (Twitter) Is Now Training Its Grok AI Using Your Data: Here's How to Stop It," *Yahoo News* (26 July 2024), https://tinyurl.com/42prwc84.

31   Anat Perry, "AI Will Never Convey the Essence of Human Empathy," *Nature Human Behaviour* 7 (2023): 1808–1809, https://doi.org/10.1038/s41562-023-01675-w.

factually unreliable, all poor grounds for advice-giving, leading to the dangerously irresponsible suggestions generative AI is famous for).[32]

Some broad workarounds are possible for these failings (guard-rails which try to remove bias and obviously dangerous responses). However, because these failings emerge precisely from the nature of how generative AI functions, there are always limits to how far safety rails can be retrofitted to prevent disastrous outputs.

All of the above failings are relevant when seeking spiritual advice or relationship. Sometimes, spiritual discourse involves factual discourse; it absolutely requires empathy; and the whole point of a spiritual advisor is to provide suggestions (i.e., advice) for courses of action and reflection that are aimed at being life-altering on the long run, collaborative though such advice may be. It is as important for the current argument that generative AI has these problems at all, as that they are intrinsic, structural failings which emerge from the nature of the technology itself. That these problems are unassailable implies that generative AI advice is and will remain problematic. Yet, spiritual advice must cut to the very heart of a person's life. The *caveat emptor* here is significant. The requisite disclaimer would have to read as follows: "Come, seek advice from our AI spiritual guide. Warning: advice may be false, life-threatening, and is totally devoid of care. OpenAI takes no responsibility for consequences of advice given. Follow *at your own risk*." This is hardly an encouraging basis for a spiritual advisor.

Generative AI Spiritual Advice as Necessarily Predictable and Cliché

Yet more baked-in problems are to be revealed. This paper will consider one last example, namely, the poor quality of the advice that generative

---

32   Kurian writes of "recent cases in which interactions with AI led to potentially dangerous situations for young users. They include an incident, in 2021, when Amazon's AI voice assistant, Alexa, instructed a 10-year-old to touch a live electrical plug with a coin." A vast litany of such dangerous suggestions exists. Quoted in Amanda Scott, "Cambridge Study: AI Chatbots Have an 'Empathy Gap,' and It Could Be Dangerous," *SciTechDaily* (2024), https://tinyurl.com/4x53432y.

AI produces. In response to spiritual inquiries, generative answers are manifestly trite and cliché. Moreover, when probed for a deeper explanation, generative AI is completely incapable of elaborating. It merely restates the same proposition in slightly different ways. This is not surprising given that it has no comprehension of its inputs or outputs. It has no insight into what it is saying or why.

Looking into generative AI as the brute force mechanism that it is exposes the structural nature of this problem. Generative AI works as follows:

> GenAI … creates new content based on what it has learned from existing content. The process of learning from existing content is called training and results in the creation of a statistical model. When given a prompt, GenAI uses this statistical model to predict what an expected response might be—and this generates new content.[33]

In short, generative language models "learn about patterns in language through training data. Then, given some text, *they predict what comes next.*"[34] This beautifully clear account of how generative AI operates (as Michael Wooldridge put it, generative AI is "autocorrect on steroids" or, as Nomisha Kurian called them, "stochastic parrots")[35] contains the core answers for why its profound limitations as a spiritual companion are necessary and inevitable. The very essence of how generative AI

---

33   Gwendolyn Stripling, *Introduction to Generative AI, Google Cloud* (8 May 2023), https://www.youtube.com/watch?v=G2fqAlgmoPo.

34   Stripling, *Introduction*, italics added.

35   See "Michael Wooldridge on AI and Sentient Robots," *The Life Scientific* (9 December 2023), https://www.bbc.co.uk/programmes/m001tgk9. Amanda Scott cites Kurian saying: "LLMs have been described as "stochastic parrots": a reference to the fact that they use statistical probability to mimic language patterns without necessarily understanding them. A similar method underpins how they respond to emotions. This means that even though chatbots have remarkable language abilities, they may handle the abstract, emotional, and unpredictable aspects of conversation poorly." In Scott "Cambridge Study."

works is to parrot the most predictable set of words that follow any given input.[36]

In spiritual conversation, there are times when giving the obvious clichés to a person in spiritual need will do. Sometimes being told to be more grateful, or to forgive someone, or to make a confession and ask for forgiveness are exactly what is needed. Many spiritual problems are universal and simple, and need only the simple and obvious answers. The problem is discerning when the simple advice is appropriate, and when more is needed. Beyond the cliché, one finds the whole rest of the spiritual life with all its immense ambiguities and the context-grounded complexities that come with it.

Spiritual advisors need to have more than one strategy at hand. Presenting the most predictable answer to a given input cannot work if that is the only thing one is capable of doing. Depending on context, spiritual responsiveness may require an unpredictable response, or may demand that persons do what is most unreasonable. It has to be remembered that the spiritual life involves, at least in part, some diminution of self-concern and some increased interest in justice. It necessarily involves, on occasion, doing work *against* one's own self-interest, construed in worldly terms (that is, assuming one takes seriously Christ's command in John 13:34 (NIV), "Love one another"). This involves *not* being strategically self-serving at all times; not measuring success and wellbeing wholly in terms of worldly value.[37] Likewise, spiritual maturity can sometimes mean speaking against or taking action against the *status quo*, when injustices are being performed. All of this is very risky business in the context of spiritual advice, and needs to be handled with the utmost care—a technology that can only provide

---

36 Another problem with cliché is that it reinforces dubious stereotypes. Consider the following from a Bloomberg article of text-to-image generation: "The world according to Stable Diffusion is run by White male CEOs. Women are rarely doctors, lawyers or judges. Men with dark skin commit crimes, while women with dark skin flip burgers." Leonardo Nicoletti and Dina Bass, "Humans Are Biased: Generative AI Is Even Worse," *Bloomberg* (9 June 2023), https://www.bloomberg.com/graphics/2023-generative-ai-bias/.

37 David H Kelsey, *Eccentric Existence*: *A Theological Anthropology* (Louisville, KY: John Knoxville Press, 2009), 345.

predictable answers is not adequate to this task (and problems of legal liability have already been mentioned; one can imagine the creators of a spiritual chatbot getting sued for the consequences of following its advice: "The generative AI told me to sell all my possessions and give all my money to the poor").

As a recombinatory tool, generative AI can only ever mash together non-contextual pieces of wisdom issued previously by real persons or in text, without comprehension or foresight. Excepting errors and hallucination (all of which can produce dangerously irresponsible advice), generative AI can *only* produce the cliché response. That is its very nature. Such a structural limitation means that generative AI is absolutely the wrong tool for the subtleties of spiritual advice and companionship. This problem is irrevocable.

Conclusion: Spiritual Regression

To conclude, expecting generative AI to serve as a substitute spiritual advisor or companion is just burdening it too much. This is not the fault of generative AI, which, used as the right tool for the right job, might be endlessly fruitful and a marvellous advancement. Rather, this is a problem with those who have excessively high expectations of what generative AI can and should be used for (and, as a corollary, low expectation of spiritual companion as the relational gold standard). It is the *expectation* of substantive spiritual value from generative AI that is the problem, the misapplication of a tool whose value lies elsewhere. Perhaps some other AI technology might fit the bill in the future, though this paper shows immense obstacles to be overcome. In any case, generative AI simply is not the right tool for the job.

We return to the crux of the paper. We see two tragically conflicting tendencies. Spiritual AI is already a disruptor; but it is the disruption of an ideal being replaced with an extraordinarily impoverished substitute. It is a disruptor which subverts instead of elevating what it is taken as providing. Thus, even more seriously, it is not just replacing genuine spiritual companionship, it is *redefining* it. Such low expectations

of spiritual companionship and advice shift the very definition of what counts as spiritual guidance, of what spiritual relationship is, can, and should be, lowering the bar by an unacceptable distance.

Generative AI offers certain temptations—namely, the illusion of relationship. That is: a one-sided pseudo-relationship without risks or responsibilities, the very opposite of spiritual companionship. In this sense, generative AI as spiritual companion represents a kind of idolatry, a false god,[38] a perverse inversion of companionship. It is a relationship with something that is not even capable of offering relationship. Spiritual companionship, which represents the highest ideal of relationships, *soul friendship*, is being counterfeited with a relationship that is not even a relationship. The highest has been degraded not just to the lowest, but to something that is not even a two-way relationship at all. Generative AI merely churns out, endlessly and on demand, without any understanding, whichever data its statistical model has deemed most likely to follow on from whatever is inputted. That is all generative AI is.

Finally, if one wishes to salvage the idea of spiritual generative AI, even just as an adjunct or support to a human spiritual advisor, one is in the position of having to explain what the parameters and limits of suitable support might be for generative AI, i.e., a predictive text model that has no empathy, no comprehension, no insight; that is factually unreliable, gives bizarre and dangerous advice, is otherwise limited to issuing necessarily predictable, trite cliché, which cannot offer challenge for fear of subjecting its creators to legal liability; which has to compete with more gratifying AI companions in a multibillion dollar

---

38   There is a strong theological trend, particularly within ecotheology, of characterising consumerism as the worst form of modern idolatry, a death-dealing false god creating vast ecological destruction, damaging individual persons and social relations, perpetuating injustice while trading in falsehoods about what happiness and wellbeing consist in, and seeking to displace and undermine the opposing true God of compassion, justice, and love (the language used in this description is characteristic of the general terms and timbre used). See Jan-Olav Henriksen, "Theology and Climate Change," in *Redeeming the Sense of the Universal: Scandinavian Creation Theology on Politics and Ecology*, ed. Trygve Wyller et al. (in press).

marketplace that is financially incentivised to positively *impede* the spiritual growth of its user base (i.e., to ensure consumers grow only to become even bigger consumers). These are just some of the parameters constraining the utility of generative AI, even as a mere support or adjunct to a human spiritual advisor.

If there is any service this paper seeks to offer, it is to restore the notion of spiritual companionship to its gold standard against the temptations of the fast-food version that is AI companionship; and to inspire the imperative to avoid overburdening generative AI (an otherwise valuable technology) with a set of tasks that it is not capable of approximating, and indeed, which its very structure subverts. As David Ford is so fond of reminding us, "The corruption of the best is the worst."[39]

---

39   David F. Ford, *The Shape of Living: Spiritual Directions for Everyday Life* (Norwich: Canterbury Press, 2014), xxvi, 15.