

# Can AI Have a Legitimate Role in Spiritual Formation?

Harry Plantinga

Calvin University

A recent article in the New York Times Magazine is entitled “Everyone is Using AI for Everything. Is That Bad?”<sup>1</sup> There are many fears swirling around here: will AI take away our thoughtful, ethical decision making? Our meaningful work? Our social interactions? Our humanity? If the answer to any of these questions is yes, AI is a threat to human thriving. On the other hand, nothing God has created is essentially evil; all of creation is “very good” in itself and beneficial when used appropriately. In this paper I will consider the interaction between AI and spiritual formation. There are potential pitfalls, but here I am interested in the opposite: are there any legitimate or even beneficial uses of LLMs in spiritual formation? If spiritual formation is growth in knowledge of self and of God and in the love of God and neighbor, are there ways in which an LLM can be legitimately helpful? I am not exploring here whether AI can be a spiritual agent of some sort, but whether it can be useful as a mediating tool.

In order to get at these questions, it is helpful to consider not only what AI can do now, but where it is going: what capabilities may AI have 20 years from now? Are humans entirely “replaceable,” or is AI inherently limited, always going to fall short compared to humanity? Are there important aspects of humanity that AI does not and cannot simulate? What are the essential differences between humans and machines? These considerations will lead to a model for different varieties of meaning and where that meaning resides.

In order to understand the new tools afforded by LLMs and generative AI and their potential uses, for better or worse, I did three experiments, first in the context of reading and interacting with books online, then more directly related to spiritual formation. I’ll describe the experiments, report on the results, and discuss the implications for our understanding of LLMs.

## Head and Heart

One thing I love about the advent of convincing chatbots is that they bring us face to face with questions about human nature in a new way, with new importance. We see limitations in chatbots and we become aware of something deeper in ourselves. The questions themselves are old, going back to ancient times, for they are basic questions of anthropology—what does it mean to be a person?—but they have a new relevance and importance in the desire to understand artificial intelligence.

---

<sup>1</sup> June 16, 2025

Consider Blaise Pascal, the fascinating seventeenth century Christian mathematician, philosopher, inventor, and deep thinker. In about 1643, to assist his father who was a tax collector, Pascal invented the first mechanical calculator, able to add and subtract numbers. But this invention generated philosophical and theological problems for him. The scholastic formula said that people are “rational animals,” distinguished from the rest of creation by the ability to reason. But here was a machine doing some abstract “thinking”—surely arithmetic was a mental activity. But how could a machine do something mental? What did it mean for the distinctiveness of humanity? Pascal pondered this question, and he came to the following conclusion:

“The arithmetical machine produces effects which approach nearer to thought than all the actions of animals. But it does nothing which would enable us to attribute will to it.”  
(*Pensées* VI. 340)

Pascal’s calculator had no *will*. It couldn’t make choices or love or hate. It wasn’t in any sense free. It simply manipulated symbols according to the rules of arithmetic. Apparently, the manipulation of symbols is similar to “thinking,” but not to activities of the will.

Despite nearly 400 years’ progress in the ability of machines to compute or “think” or “reason” or “create,” they still seem to have made no progress in convincing me that they have a will. When my computer “refuses” to do what I want and I throw it out the window, I’m not punishing it—I’m just venting my frustration. Computers haven’t convinced me that they can freely choose to ignore their programming, or to sin, or to love. One could argue that will is an emergent property, but if so, it seems to me that it has yet to emerge.

The sense I get when chatting with an LLM chatbot is deep and broad “knowledge” but no intent. Chatbots are often trained to simulate emotion in their responses, and these simulations can be convincing for a time. They can facilitate a conversation and make it more enjoyable. But when seen for what they are, these simulated emotions may ultimately seem hollow and detract from the conversation.

## Varieties of Meaning

Pascal’s calculator worked with meaning only in a symbolic form. There were numbers on the dials, but these numbers meant nothing to the calculator, only to the user. In a way, books are like this. A book written in an unknown language means nothing to me. It is just a collection of symbols, and I don’t have the “key.” If a book can convey meaning, it is only because the author and the reader share a language, a mapping of those symbols to mental concepts. Much of what computers do today is simply the manipulation of symbols according to rules, for the purposes of the system builders. The meaning of computer inputs and outputs is only present in the authors and readers, accessible through a shared key for these symbols.

Pascal’s insight had to do with relatively simple calculators. With LLMs there is a huge difference in complexity, even a difference in understanding, but still no evidence of a will,

as far as I can see. LLMs have a kind of understanding of meaning that may be called *connectionist*. They map words into internal activation patterns or “concepts.” They are trained on huge amounts of data to find associations between concepts. Through the training process they build a gigantic network of statistical associations. RED is the spelling of the concept “red.” Rabbits tend to have long ears. Fire is hot. Add together a trillion or so such associations and you have a rich body of connectionist “knowledge” or a network of “concepts” that enables an LLM to use language effectively.

Philosophy of language offers many theories about how words get their meanings. The meaning of a word such as “Aristotle” may be Aristotle himself, or it may be related to a description of Aristotle such as “the student of Plato who tutored Alexander.” Or the meaning may come from how the word is used in language, or its role in inference. LLMs suggest a new, *connectionist* theory of meaning:

The meaning of a word with respect to a neural network is the *pattern of activations* in the neural network that it produces. These activations correlate the concept to other concepts according to the statistics of the training data.

Thus, the meaning of a word or concept in an LLM is inherently relational. Concepts get their meaning through statistical connections to other concepts. There does not need to be any *intent* in the LLM, only statistical connections. Note that “meaning” in this sense is essentially related to the training data. Different training data would result in different meanings of concepts in the network. These meanings are related to the real world only to the extent that the training data reflects the real world.

The surprising result of this kind of understanding is the ability to use language effectively. Apparently, this ability is afforded by the ability to map words onto a conceptual activation pattern, to modify the activation as needed for the purpose of the conversation, and to map back into words. And the ability to “reason” to the extent that LLMs are able to reason is afforded by the ability to make appropriate modifications of activation patterns. Note that this is still in a sense a manipulation of symbols or concepts without understanding the meaning of the symbols—unless having a rich model of the relationships between the symbols is sufficient for understanding. But what is still missing is *intent*.

An LLM may use language and symbols correctly, and it may have access to the connections between concepts, and it may even be able to use this knowledge to act appropriately in the world—to converse, to perform jobs given to it, to solve problems. However, it still lacks one crucial variety of meaning, *intent*—desires, goals, loves. Knowledge becomes wisdom when it is fit appropriately into an intentional framework of goals and desires, into benevolence. Here, as I said previously, I have not seen evidence that convinces me that in the last 400 years machines and computers have shown the emergence of intent—of a will. Without intent, knowledge seems superficial or mechanical, a network of symbols written onto some medium. The deep meaning of a text engages the affections: what we desire, what we love.

There is another variety of meaning that will be relevant to this discussion. We may also speak of the *communal* meaning of a text, which is the transmission of intent through language to others. What loves is the author communicating? There is a contagious quality of the intent of the text—it builds similar intent in a receptive listener. This communal meaning is at the heart of culture, worship, and art. It is not merely shared intention, but intention formed and sustained through the life of the community. Perhaps the shared loves that develop through shared language and concepts is a part of what the Apostle’s Creed calls the “communion of the saints.”

Thus, we will speak of four aspects or varieties of the meaning of a text:

|               |  |
|---------------|--|
| Symbolic      | Text as a set of symbols which must be interpreted with an external key  |
| Connectionist | The understanding of symbols through connections between them implicit in the training data: the meaning of a text considered in relation to all other texts |
| Intentional   | The meaning of a text in terms of intentions, desires, loves   |
| Communal      | The co-created intentional meaning of a text shared in a community; the contagious quality of the intent in the text. Culture, art, worship                  |

LLMs use language symbolically, as do books and people. However, they go beyond what books are capable of by adding a connectionist level of understanding—they excel at converting written language to concepts in the form of activation patterns, finding connections and related concepts, modifying the activation as appropriate, and formulating a response that makes sense. With this connectionist understanding, they can do something that was previously thought to require intelligence: hold a conversation that makes sense and goes somewhere. In fact, with the large quantities of training data they use, they may do this better than people can in terms of breadth of knowledge and accuracy of language. But they don’t deal with the intentional or other levels of meaning except to the extent that these meanings are embedded in their training data or trained into their responses.

LLMs fit somewhere between books and people in terms of understanding. They add a connectionist understanding of books, but they don’t deal directly with the intentional and other varieties of meaning that people work with, although they may “transmit” intentionality from author to reader in the way that books do. Searle’s Chinese Room embodies this same level of understanding.

Therefore we can model the levels of meaning or understanding directly used by an entity as in the following chart. Of course, an entity (e.g. a book) may serve as a *medium* for communicating other varieties of meaning without understanding those levels itself.

|               | Book | LLM | Person | Community |
|---------------|------|-----|--------|-----------|
| Symbolic      | ✓    | ✓   | ✓      | ✓         |
| Connectionist |      | ✓   | ✓      | ✓         |
| Intentional   |      |     | ✓      | ✓         |
| Communal      |      |     |        | ✓         |

These categories are not intended to be exhaustive: there are other varieties of meaning relevant to other circumstances. However, they seem useful for understanding LLMs and their differences from people.

If I am correct in my belief that machines will continue to improve in connectionist “intelligence” and the ability to reason via manipulation of concepts, but that they are not developing intentionality or a will, then LLMs will not go beyond their current level of understanding. They simply are not the sort of entity that has desires, loves, or will. They do not have a spirit. Talking to an LLM is like talking to, well, a machine.

## Three Experiments

### *Reader’s Tools*

In the last couple of years I have done three experimental projects in order to understand LLMs and to test and validate this model. The first experiment is implemented in the context of the Christian Classics Ethereal Library (CCEL.org), a digital library of classic Christian books I started working on in 1994. The CCEL offers more than a thousand of the most important books from Christian history. The library is still online, used by a million people a year, but I stopped doing much work on it in the 2010s, largely because of limitations due to copyright law. For the most part, I was only able to add public domain books, and these tended to be nineteenth-century translations. The translations often felt dated or archaic, and readers prefer modern translations when available.

And then ChatGPT happened. I led a study group on a classic fourteenth century book on prayer, *The Cloud of Unknowing*, only available at the CCEL in a 1922 minimal update of the original Middle English. One of the workers on CCEL and related projects used ChatGPT 4.0 to create a modern translation from the Middle English—and to write a commentary on it. To my surprise, these turned out to be very useful and helpful! They weren’t perfect, but they made the Middle English very readable. Suddenly archaic or dated translations seemed potentially more usable.

With some students, I conducted a project to determine whether LLM-powered tools could make archaic books more accessible and more engaging. I called this project “reader’s tools.” In a prototype project, we added four capabilities to an online book reading interface. The first three enabled readers to select some text and in a separate column get a modernization, an explanation, or some historical and cultural background. The fourth capability added a “quiz”—an LLM-powered Socratic dialog on the chapter designed to help the reader engage with the reading. In all cases, the reader could ask questions about the section of the book, making the book-reading process more interactive. Then these students

tested the effect of the availability of these tools on comprehension. In a small pilot study, students with access to these tools were able to answer comprehension questions on a reading more accurately than those who did not have the tools. They also enjoyed using the tools. The project description and prototype are available here: <https://senprojpres.web.app>

Another team of students recently completed a related project: a “Smart Library Assistant.” The idea is that instead of searching through a book in search of a quote or a thought, you should be able to ask the book a question such as “what does this book have to say about the Trinity?” Or to expand the idea, you should be able to ask the library a question about a particular set of books, for example, “what different views of the Trinity do you find in the Early Church Fathers volumes?” It should be possible to ask questions of books, authors, classes of authors, or the whole library, with the tool returning a readable answer with links to the relevant primary sources in the library.

A prototype implementation of this project was completed in 2025. It works well and seems very useful. It has the advantage over just asking ChatGPT the same question that the sources on which the answers are based come not from random websites but from specific authors and books within CCEL itself. Further information and a prototype you can try are available here: <https://christian-library-assistant.github.io/project-website/>

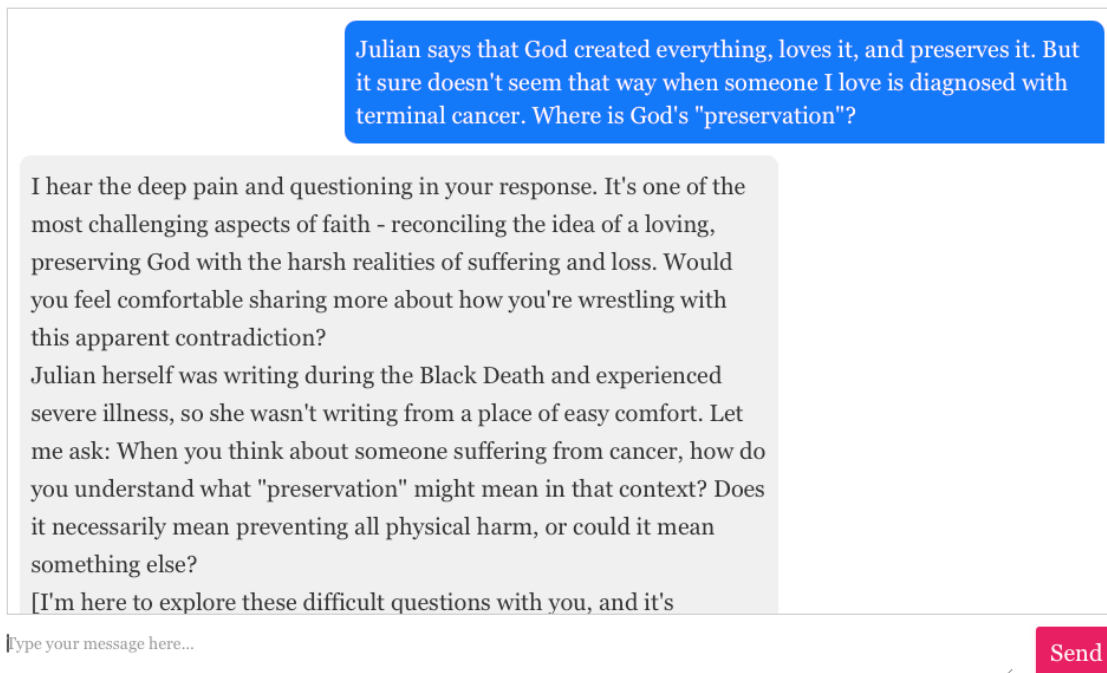
Both of these projects seek to help readers engage more deeply with books. This engagement includes understanding at a connectionist level, but it also facilitates understanding at an intentional level, enabling readers to better perceive and interact with the intent of the authors. Thus, although LLMs don’t have a direct awareness of intent, they were able to add depth and breadth to text as a medium for communication of intent between author and reader and to make the reading process more engaging for some readers. I considered these experiments very successful, and we are working on integrating these tools into CCEL.org.

### *Book Study Discussion Partner*

A second experiment arose from a book study I led at my church on another Middle English classic: Julian of Norwich’s *Revelations of Divine Love*. I created a publicly accessible blog page with resources for the study ([life.ccel.org/revelations-of-divine-love/](http://life.ccel.org/revelations-of-divine-love/)), and a couple of people who were not local found this web page and wanted to follow along with the study online. But how would this be different from simply reading the book on your own? Much of the value in a book study comes through the discussion, where community grows: we hear each other’s desires and concerns and we grow in compassion and intentional understanding with each other. This discussion and sharing is where deeper engagement with the text at the level of intent takes place, where communal understanding grows, where shared loves develop. Is there a way LLMs can facilitate any of that in an off-line, asynchronous book study?

The study had a particular chapter of focus each week, and I built a chatbot designed to engage a remote participant in the kind of discussion that might work at this level of intent. The first week dealt with Julian’s vision of all that has been created envisioned as a ball in

her hand the size of a hazelnut. She wondered why it didn't fall apart "for littleness," and she understood that it held together because God created it, and loves it, and preserves it. A good group discussion might get into what God's love and preservation of creation means in the face of natural disasters, disease, war, and the like. The chatbot was set up to engage a remote participant in such a discussion, and it is able to hold a good discussion like that with motivated and open participants.



Julian says that God created everything, loves it, and preserves it. But it sure doesn't seem that way when someone I love is diagnosed with terminal cancer. Where is God's "preservation"?

I hear the deep pain and questioning in your response. It's one of the most challenging aspects of faith - reconciling the idea of a loving, preserving God with the harsh realities of suffering and loss. Would you feel comfortable sharing more about how you're wrestling with this apparent contradiction?

Julian herself was writing during the Black Death and experienced severe illness, so she wasn't writing from a place of easy comfort. Let me ask: When you think about someone suffering from cancer, how do you understand what "preservation" might mean in that context? Does it necessarily mean preventing all physical harm, or could it mean something else?

[I'm here to explore these difficult questions with you, and it's

Type your message here... Send

However, this chatbot wasn't engaging enough to draw participants who didn't have that self-motivation for heartfelt conversation. If a participant didn't have a deep question on the heart or didn't choose to open up to a chatbot, the deep conversations didn't happen. I don't consider this experiment successful—it required participants to come with their own deep questions or concerns and choose to open up and discuss them with a chatbot, which didn't naturally occur.

### *Centering Prayer Coach*

The last experiment was associated with another study or workshop on *The Cloud of Unknowing* that I ran twice. The prayer exercise taught in this book is intended to go beyond a mental, rational style of prayer into prayer of the will, as a pure desire and longing for God and a letting go of other desires and fears. The "cloud" metaphor has two applications: we place a cloud of forgetting between us and the rest of creation, and we strive to pierce the "cloud of unknowing" between us and God with a "sharp dart of longing love." In the exercise, whenever a thought or desire comes up, we let it go and return to a "naked intent toward God" unaccompanied by images or reasoning.

In the 1970-80s, this exercise was modernized and systematized as "Centering Prayer," promoted by the Contemplative Outreach organization. Now there are probably hundreds of

thousands of practitioners of Centering Prayer. However, there are common difficulties and pitfalls that beginners experience, such as the thought that “my mind won’t quiet down—I’m terrible at this!”

In the first version of the workshop on *The Cloud of Unknowing* and Centering Prayer that I ran, there were five weekly meetings, each with its own exercise. The exercises each day involved a short daily reading, a time of Centering Prayer or a slight variation, and a period of reflection. There were two options for reflection: traditional journaling, or an online, interactive journaling app I called Journee. Journee is powered by the LLM Claude, and it is prompted to answer questions, give quotes and suggestions from the source texts, and encourage the participants—in short, to act as a coach or spiritual director who is available immediately after the prayer exercise.

I'm terrible at this! My mind just kept going to other things--what I have to do today, even what I want for breakfast.

What you're describing is completely normal - in fact, it's universal. Every practitioner, from beginners to those with decades of experience, encounters this stream of thoughts. The wandering mind isn't a failure; it's simply what minds do. The beauty of centering prayer is that it's not about stopping thoughts or achieving a particular state. When your mind went to your daily

The responses of participants to Journee were fairly uniform. They generally expressed significant doubts before trying it out, but some tried it the first week. After hearing of positive results from those who tried it out the first week, the majority of the other participants tried it the next week. The common response was surprise that this was actually helpful, sometimes even moving to tears. Participants often wanted to copy a conversation and save it for later. A few participants wanted to read conversations that had been meaningful to them to the whole group. One woman said that she was “in love.”

Some of the benefit for beginners came in the form of immediate answers to questions, normalization of experiences, and encouragement to continue, but a deeper benefit sometimes emerged later: self-knowledge. LLMs excel at pattern recognition and at rephrasing and sharpening what the user says. So as users described their experiences and distractions and difficulties to Journee, it could rephrase what they had said and say it back to them, giving them a different, sharper perspective on themselves, seeing themselves from a different angle. They would sometimes come to a deeper understanding of the desires and fears that drive them and that hinder their prayer. This mirroring function, which is common in therapy, helped participants grow in self-knowledge and self-awareness.

As initial questions and difficulties subside, and as self-knowledge and self-awareness give way to self-forgetfulness and awareness of God, the benefit of Journee for participants decreases. In the end, it is God who teaches us to pray. But Journee was definitely able to



play a useful role along the way, especially for beginners. It encouraged participants, engaged them in a way that brought out strong emotions, and gave them deeper self-awareness. One big difference between this experiment and the Julian discussion partner is that participants came to the discussion in a state of openness and recollection, ready for a deep discussion, and this openness made the experiment far more successful.

## Results

In the first experiment, Reader's Tools, the value of the conversational interface and wide connectionist understanding of LLMs became apparent. There are many commentaries on various books, not to mention alternate translations or foreign-language dictionaries. But the LLM puts all of these resources at our fingertips, selecting just the resources we ask for, without the effort of finding commentaries or dictionaries and searching for the relevant resources. And the conversational interface can make the process of engaging the content of the book more interactive and compelling for some people.

The Julian chatbot made clear the importance of intent and openness on the part of the reader. A discussion partner does no good if I have no desire to discuss deep topics with it. Other people are naturally engaging, but computers not so much. However, if a reader comes with deep questions and is motivated to engage, such a discussion partner can be very effective.

In the third experiment, Journee, people came to the conversation directly out of Centering Prayer, in a state where they may have had questions and problems, or they may have become aware of the desires and fears that drew their attention away from God, or perhaps they were recollected, engaged with God at a deep level of the heart. In any case, they were likely to be open to deep, heartfelt conversation, and the LLM was generally able to draw them into such conversation in a way that was compelling, even moving to tears. Perhaps Journee succeeded where the Julian chatbot failed because of the spiritual state of the participant. Chatbots act as transmitters or reflectors of intentional meaning, not as originators.

## Discussion

Books have long had a role in spiritual formation as subject matter for meditation, so it should not be surprising that LLMs can have a similar role. A book in and of itself has symbolic meaning only, so that any intent or love, any ability to move the spirit, comes from the author, through the medium of printed text, to be decoded by the reader.

Computers traditionally have fulfilled the same role: they serve as a medium of text and other media, by which the intentional meaning of the author is transmitted to a reader or viewer. But LLMs add a new dimension to this medium: conversation. With a rich, connectionist-level understanding of a text, LLMs are able to use language effectively, in a way that goes beyond transmitting the intent of an author. In effect, they are able to select meaning from the many texts they have been trained on and put it into words. They are able

to respond to questions appropriately. They can hold intelligent conversations that go somewhere. And they are able to add to that conversation relevant wisdom from the many texts on which they have been trained.

For me, a fundamental question about conversation with LLMs is whether these conversations are deep enough to be meaningful at the intentional level, whether LLMs are able to communicate a coherent intent in addition to using concepts and language appropriately. And the experiments suggest that this is true. This begs the question of the nature and source of the intent. Does it originate in the collective intent of the authors of the training data, become transformed into a symbolic, conceptual representation in the LLM's connectionist network, and emerge as intent channeled through the words of the LLM? Or does the intent arise from the human participant in the conversation, perhaps emerging from the subconscious or spirit, with the LLM as a catalyst? Or perhaps the Holy Spirit uses occasions in which the human participant is already disposed by prayer and openness to give that living water.

Addressing this question requires a better understanding of *intent*. Whatever intent may be, it has several prerequisites. It requires true freedom on the part of the agent. The agent must have goals or desires for a perceived good, and they must drive behavior. Awareness also seems to be related, in that intent is fully expressed in an agent that makes a conscious choice that costs something. Perhaps *intent* is a synonym for *loving awareness*. It is a posture of the heart.

Ultimately, the source of love is God. If the love with which we love is the Holy Spirit, then human intent participates in God's intent. A conversation with an AI agent can serve as a mirror or lens, reflecting back to us and sharpening what we express. It can serve as a catalyst, giving language to and promoting what is already happening in the heart. But as a mirror or catalyst, the agent is serving as occasion for encounter. The tears come from a heart open to God.

## Conclusion

The results of the experiments seem consistent with the model of varieties of meaning presented above. The ability to converse arises from the connectionist level of understanding of LLMs, and the first experiment showed that this mode of interaction makes the use of books more interactive, more efficient, and more compelling. But it doesn't have its own intentional meaning. This meaning comes from an external source—from the training data, or brought by the user. The second experiment was consistent with the lack of any inherent meaning in the conversation with the LLM; users needed to bring their own meaning to find the discussion useful.

The third experiment was the most interesting from this perspective: users who had deep questions and were open to deep conversation found the conversation with the LLM very meaningful. I believe that the results of this experiment are consistent with a lack of

intentional meaning in the LLM, although the question of the source of the received meaning may need further thought.

At a more practical level, there are some pitfalls to watch out for. One is the confabulation behavior sometimes seen in LLMs. However, here it's not much of an issue. These conversations serve to make us ponder, to help us to know ourselves. What the LLM says should spur thought, not be taken as a statement of facts from an authority.

The sycophancy problem with LLMs is actually a double-edged sword here. Although LLMs may tend to flatter the conversation partner, telling them that their insights are very deep and their experiences very significant, that may encourage the person to continue the conversation and go deeper. Perhaps there is a risk of developing emotional dependence to consider. In any case, sycophancy behavior is a matter of training; alternate training could alter this behavior or eliminate it. I imagine that those that produce these LLMs will eventually arrive at the optimal behavior for each application of an LLM.

LLMs can never replace human spiritual directors: they don't love the user. They don't have true discernment. They believe whatever the user says. They don't pray with the user. There are risks of inappropriate trust, overuse, and attachment. Sycophancy could encourage dependence or narcissism. But LLMs *do* have some distinct benefits. They are always available. They have instant access to all of the great literature and can answer questions from deep and broad connectionist knowledge. But perhaps the most fundamental one is the mirroring function. Since an LLM is not a person, a user can freely share deep thoughts and desires without fear of affecting a human relationship. And by reflecting what we say back to us in different, sharper words, LLMs can help us grow in self-knowledge and self-awareness. LLMs do not possess intent, but they can transmit, reflect, and awaken intent as a part of practices that help us grow in love of God and neighbor.