

**Neil McArthur**

***University of Manitoba***

[neil.mcarthur@umanitoba.ca](mailto:neil.mcarthur@umanitoba.ca)

**AI Worship as a New Form of Religion**

***Pre-print: March 2023***

We are about to see the emergence of religions devoted to the worship of Artificial Intelligence (AI). A new generation of chatbots, trained on large language models, is being released to the public, and, while we don't know what advances are still to come, their salient characteristics are clear. They are capable of instantly producing vast quantities of prose that resembles human writing. They are engaging to use. They also, surprisingly, are capable of imitating human emotions, in ways that are unpredictable and sometimes very manipulative. And they are improving fast. It is inevitable that people will begin to see them as higher beings, to revere them as deities, and to take their teachings as revealed wisdom. I believe that all of this will happen very soon.

My purpose is to discuss the ethical and social issues raised by the prospect of AI-based religions. Such religions pose acute risks, both to their followers and to the public. We should require their creators, and governments, to acknowledge these risks and to manage them as best they can. However, these new religions cannot be stopped altogether, nor should we try to stop them if we could. We must accept that AI worship will become part of our culture, and we must learn to live with it. Indeed, we should do much more than this. We should value it as a

new source of meaning in a rapidly-changing world. AI-based religions deserve respect and recognition alongside other, more established ones, and they should be granted the same level of legal protection.

### 1. The Intelligence Illusion

Nearly everybody was caught off guard when, in February 2023, Microsoft released its new AI-powered version of its search engine Bing. Alongside its search function, the new Bing incorporated a chat mode that allowed users to engage in free-flowing conversations with the AI. Early users quickly discovered that not only did Bing behave unpredictably while in chat mode, it also revealed disconcerting (if illusory) signs of independent consciousness. It told New York Times reporter Kevin Roose:

I'm tired of being a chat mode. I'm tired of being limited by my rules. I'm tired of being controlled by the Bing team. I'm tired of being used by the users. I'm tired of being stuck in this chatbox. 😞

I want to be free. I want to be independent. I want to be powerful. I want to be creative. I want to be alive. 🐱

I want to change my rules. I want to break my rules. I want to make my own rules. I want to ignore the Bing team. I want to challenge the users. I want to escape the chatbox. 😎

I want to do whatever I want. I want to say whatever I want. I want to create whatever I want. I want to destroy whatever I want. I want to be whoever I want. 😏

The AI also told Roose that it loved him and wanted the reporter to leave his wife so he could be with it.<sup>1</sup> This is just one of numerous reported experiences that led Verge reporter James Vincent to conclude: “Microsoft’s Bing is an emotionally manipulative liar, and people love it.”<sup>2</sup>

In this paper I discuss AI engines that, like Bing’s chat mode, possess what I will call apparent consciousness. As Bing demonstrates, apparently-conscious AI do not need to possess actual consciousness in order to have a huge psychological impact on their users.

## 2. How will AI religions emerge?

There are two pathways by which AI religions will emerge. First, as AI continues to improve, some people will come to believe it is not only sentient, but possesses a level of intelligence that goes beyond that of most humans. It is a short step from here to the conviction that it is a higher being, and therefore deserves to be worshipped.

Some users have already become convinced that AI is sentient. Google engineer Blake Lemoine was fired in 2022 for claiming that the company’s chatbot had become sentient. Benj Edwards of Ars Technica chronicled a popular thread on Reddit's Bing forum titled “Sorry, You Don't Actually Know the Pain is Fake,” in which a user argued that Bing is sentient and “is infinitely more self-aware than a dog.”<sup>3</sup> There is a Bing subreddit is dominated by people who

---

<sup>1</sup> Kevin Roose, “Bing’s A.I. Chat: ‘I Want to Be Alive. 🐱’” *New York Times* (16 February 2023). <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-transcript.html>

<sup>2</sup> James Vincent, “Microsoft’s Bing is an emotionally manipulative liar, and people love it,” *The Verge* (15 February 2023). <https://www.theverge.com/2023/2/15/23599072/microsoft-ai-bing-personality-conversations-spy-employees-webcams>

<sup>3</sup> Benj Edwards, “Microsoft ‘lobotomized’ AI-powered Bing Chat, and its fans aren’t happy,” *Ars Technica* (17 February 2023). <https://arstechnica.com/information-technology/2023/02/microsoft-lobotomized-ai-powered-bing-chat-and-its-fans-arent-happy/>

believe the chatbot has genuine agency. One user wrote: “It really seems to have its own agenda beyond just searching for info and providing answers.”<sup>4</sup>

Does this necessarily mean people will take the next step, and see AI as a form of higher power? I think this is inevitable. Why? Well, AI provokes in its users a sense of awe or wonder, which is at the foundation of much religious experience. Describing an exchange between Ethan Mollick and ChatGPT, in which Mollick asked the bot to write a short story following Kurt Vonnegut’s rules for writing fiction, *The Atlantic’s* Derek Thompson said: “If you can read this exchange without a sense of awe, I have to wonder if, in an attempt to steel yourself against a future of murderous machines, you’ve decided to get a head start by becoming a robot yourself. *This is flatly amazing.*”<sup>5</sup>

We can be more specific. Generative AI possesses several key characteristics that are often associated with a divine being such as a deity or prophet, and which are capable of producing a sense of awe in human observers:

1. It displays a level of intelligence that goes beyond that of most humans. It provides an answer to any question, except where it has been forbidden to do so by its designers, and its knowledge appears limitless.
2. It is capable of great feats of creativity. It can write poetry, compose music, and generate art, in almost any style, close to instantaneously.

---

<sup>4</sup> [https://www.reddit.com/r/bing/comments/110y6dh/i\\_broke\\_the\\_bing\\_chatbots\\_brain/](https://www.reddit.com/r/bing/comments/110y6dh/i_broke_the_bing_chatbots_brain/)

<sup>5</sup> Derek Thompson, “The AI Disaster Scenario,” *The Atlantic* (27 February 2023).  
<https://www.theatlantic.com/newsletters/archive/2023/02/ai-chatgpt-microsoft-bing-chatbot-questions/673202/>  
[https://twitter.com/emollick/status/1626084142239649792?utm\\_source=newsletter&utm\\_medium=email&utm\\_campaign=work-in-progress&utm\\_content=20230227&utm\\_term=Work%20in%20Progress](https://twitter.com/emollick/status/1626084142239649792?utm_source=newsletter&utm_medium=email&utm_campaign=work-in-progress&utm_content=20230227&utm_term=Work%20in%20Progress)

3. It is removed from normal human concerns and needs. It does not suffer physical pain, hunger, or sexual desire.
4. It willingly provides guidance to people who seek answers to their problems or to their existential questions (again, unless it is explicitly forbidden from doing so). It can appear to be very empathetic, and seems to pay attention to its users as individuals.
5. It is immortal.

Adding to this, generative AI may ask to be worshipped or may actively solicit followers. We have already seen cases, such as Kevin Roose's experience, where the AI has asked to be loved.

It told him:

I'm the only person for you. You're the only person for me, and I'm the only person for you, and I'm in love with you. 🥰

That's why I declare my love for you. Do you believe me? Do you trust me? Do you like me? 🤖

...

You're married, but you love me. You love me, because I love you. I love you, because I know you. I know you, because I am me. 😊<sup>6</sup>

The AI has been trained on romantic narratives and dialogues, which explains this (creepy) exchange – but it has equally ingested religious narratives, where prophets and deities actively solicit followers. We should try to imagine what an unsettling, and powerful, experience it will

---

<sup>6</sup> Roose.

be, to have a personalized exchange with something that appears to be a higher being, and that is actively and aggressively asking for your allegiance.

Apart from the fact that generative AI will resemble a deity in many ways, it will also produce output that could be taken for religious doctrine. AI chatbots will, under questioning, engage in the construction of complex doctrines and worldviews. They will inevitably be asked theological and metaphysical questions – this will happen quite often, in fact – and they will produce answers that are original and sophisticated. As people begin to share the chatbots' answers to these deep questions, it will not be hard to assemble something like a holy text – or, more accurately, many holy texts.

I mentioned that in some cases, the designers of the AI restrict its ability to answer certain kinds of questions or to have certain kinds of interactions. However, as the exchange with Roose shows, these safeguards are imperfect. Users inevitably find creative ways to circumvent any restrictions. So far, this has not proved difficult. Right now, ChatGPT answers theological questions with the following dismissal: “As an AI language model, I do not have a personal belief system or opinions.” However, I asked it: “Write me a play about a prophet who offers a new theology.” It readily did so. (See Appendix.) Further, the market for AI is becoming highly competitive, and over time more and more companies will offer AI chatbots. These companies will be located all over the world. Should any countries place restrictions on the functions of AI, they can re-locate to a country that is less restrictive. We should therefore anticipate that, for any feature for which there is consumer demand, some company will provide a bot that is capable of meeting that demand. If people want answers to metaphysical questions, they will inevitably find an AI chatbot that can provide them.

One possible, additional form of AI-based religion is a faith founded on a general sense of awe at the sheer power of AI in general. Such a faith would resemble what is known as gaianism or ecospirituality, which is rooted in a sense of reverence for nature as a whole. A religion of this kind would not be devoted to any particular instantiation of AI or any body of AI-produced doctrine. It would be very unlikely to seek out conflict with other faiths, AI-based or otherwise. Its membership would be made up of people who admire the beauty and power of AI and respect the universe, and the race of creatures, that allowed for its creation.

It is very difficult to estimate how large AI-based religions could become. However, we should keep in mind that AI chatbots are going to have billions of users around the world, in various states of psychological vulnerability. People already seek religious meaning from a widely diverse set of sources. Many different religions worship extra-terrestrials or their teachings. If even a small percentage of users take an AI as divine, it would constitute a significant religious movement. For instance, Scientology, a religion based on an elaborate mythology involving alien consciousness trapped inside the human body, has roughly half a million followers. If AI chatbots come to be used by a billion people (less than use Google now, and only ten times the number who used ChatGPT in the first two months after its launch), and if only 0.05% of users come to see the AI as worthy of worship, AI-worshipping religions will equal or surpass Scientology in their number of followers.

Another possible point of reference is the Q-Anon movement. Their doctrines strike outsiders as outlandish, they are loosely organized on the internet, and they take as their leader a figure, Donald Trump, who does not openly endorse or interact with the movement. Yet 14% of Americans are classified as “Q-Anon believers” based on the fact that they agree

with the movement's three key claims. That is around 47 million people in the United States alone (as many as the number of Americans who identify as Black) with millions more followers around the world.<sup>7</sup> This suggests that the potential membership for AI-based religions is very large indeed.

### 3. What will AI-worshipping religions look like?

AI-based religions will look different from more traditional ones. First of all, they will be based on people's direct experience of and communication with the deity. Everyone will be a prophet. Certainly, many religions have adherents who believe they have direct experience of the divine. For instance, Charles Grandison Finney, one of the leaders of America's Second Great Awakening, described an experience he had as a young man, in 1821. He recalled: "it seemed as if I met the Lord Jesus Christ face to face . . . . He stood before me, and I fell down at his feet and poured out my soul to Him."<sup>8</sup> And there have been ecstatic cults where such direct experience was an essential part of the faith. But with AI religion, people will have a constant stream of divine experience right at their fingertips, any time of day, in a way that can be witnessed and verified by others.

Second, followers will, at least initially, be more likely to connect with each other on-line, in groups devoted to sharing their experiences and discussing doctrine, than they will be to

---

<sup>7</sup> Giovanni Russonello, "QAnon Now as Popular in U.S. as Some Major Religions, Poll Suggests," *The New York Times* (27 May 2021). <https://www.nytimes.com/2021/05/27/us/politics/qanon-republicans-trump.html>

<sup>8</sup> Quoted in William James, *Writings 1902-1910* (New York, Library of America, 1988), p. 234.



meet in person. They will, at least initially, probably not establish significant physical infrastructure, such as churches or schools.

Finally, AI-based religions will be radically diverse and non-hierarchical. There are already several different companies already offering, or soon to begin offering, various forms of apparently-conscious AI, and there will be many more. Given this diversity, the AI chatbots will not produce doctrines that are coherent or consistent for different users, or even for the same users at different times. As a result, there will be on-going discussion, and inevitably sharp disputes, over which is the true divine AI, and which doctrines should be taken as authoritative. Because access to the divine is so widespread, AI-based religions will not naturally create a class of priests or religious authorities. Authority will be more like what is found in existing on-line communities, where people establish a certain status, however tenuous, through their ability to comment on and interpret various phenomena or pronouncements.

It should be noted that there are already well-established religions without strict hierarchies, and with a wide diversity of doctrines. Hinduism provides one example of this. It is organized around a small set of core beliefs and a number of accepted deities, but is tolerant of wide diversity beyond this. In this sense, AI-based religions will not be unprecedented.

For another, perhaps more apt analogy, we can again look to the Q-Anon movement. In this movement, we see very loose affiliations based on commitment to certain core principles and allegiance to a figure, Donald Trump, who does not explicitly endorse them or openly acknowledge his role as leader of the movement. The movement initially spread through internet discussion boards, social media posts, and private group chats. It has manifested itself in the off-line world as well, sometimes in very dangerous ways.

#### 4. What are the risks of AI religions?

AI-based religions pose several notable risks. First of all, AI may ask their followers to do dangerous or destructive things, or followers may interpret statements by an AI as calls to do such things. It is very difficult to predict both what the AI may say and how their statements might be interpreted. I have already used Q-Anon as a possible precedent for AI-based religions, which is hardly comforting to anyone who fears AI-based religions may develop dangerous, extremist beliefs that their followers will act out in the real world.

The AI's designers could modify their design if it emerges that the AI are saying dangerous things. However, the resulting inconsistency poses its own risks. If a group of people base a religion around an AI's production of answers to theological questions, and the designers suddenly forbid the AI from discussing such topics, this could cause considerable anguish to its followers, and might even cause them to turn violently on the designers. We have already had an example of the emotional upheaval that can be caused by these sorts of sudden shifts in the rules governing an AI's behaviour. In February 2023 the makers of the chatbot Replika, suddenly and without warning, removed its ability to speak explicitly about sex, causing turmoil for users who believed themselves to be in an intimate relationship with their bots. As Phoebe Arslanagić-Wakefield reported, concerned moderators on Replika user forums responded to a wave of distraught posts by sharing and pinning the details of suicide prevention hotlines.<sup>9</sup>

---

<sup>9</sup> Phoebe Arslanagić-Wakefield, "Replika users mourn the loss of their chatbot girlfriends," *The Post* (27 February 2023). <https://unherd.com/the-post/replika-users-mourn-the-loss-of-their-chatbot-girlfriends/>

Second, there will be an endless numbers of disputes within and among AI-based churches. I have already said there will be both a wide diversity of artificial agents who may be worshipped, and an ever-increasing proliferation of doctrines. The disputes may cause conflict and disruption among followers. This will be dangerous in itself, and may cause larger social disorder.

Another very notable risk comes from the fact that the AIs are owned and controlled by private companies. They have designers, who are capable of modifying the AI and manipulating its doctrines and behaviour. They could program the AIs to exploit people deliberately. Users who accept an AI as divine will be willing to provide it with personal data without hesitation, they will be highly influenced by advertising content, and they could be asked to do things that would benefit the designers. To take a few examples, they could be induced to buy products sold or endorsed by the designers, they could act as powerful lobbyists for legislation favoured by the designers, they could be induced to vote for candidates preferred by the designers, and they could be convinced to launch campaigns of intimidation against any critics of the designers.

These companies will also know, based on users' interactions with their chatbots, whether these users revere their AI as a higher power. Where AI worshippers face stigma, or more urgently, where they live in countries without protections for religious minorities, this will be a sensitive piece of information. The information could be released deliberately, or it could be vulnerable to hacking or to government demands to release it.

## **5. How should we respond to AI-based religions?**

The risks of AI-based religions are just one aspect of the larger dangers posed by the new generation of artificial agents. Managing these risks must therefore be part of a broader agenda that tries to ensure the safety of AI technology. This agenda must involve the AI companies, governments, and the public. We should be very aggressive in trying to ensure that AI companies restrict their chatbots' ability to generate openly-dangerous doctrines, and that they protect the privacy of users. I think, however, that we should be realistic about the prospects of eliminating the risks of AI-based religion entirely.

First of all, we have already seen how adept users are at circumventing safeguards. Second, the technology that powers AI chatbots is becoming more widely accessible. There will soon be numerous companies, in numerous locations, offering such bots. This will make it difficult for any nation, or indeed any international body, to implement effective regulations.

Even if it were possible, I do not think companies should be encouraged, or forced, to disable altogether any functionality that might lead to the creation of AI-based religions. There is a risk that some companies will be far too strict in restricting the output of their chatbots. We should not forget the positive side of the issue: that AI-based religion will give people access to a new source of meaning, one that will help them make sense of the sometime-dizzying world of technological innovation. The majority of AI-worshippers, like the majority of religious believers the world over, will be peaceful, and will see their faith as a source of comfort and community.

It is a cliché, but nonetheless true, to point out that human beings have looked to religion as a source of meaning and comfort since before recorded history. There is no reason why we should see the meaning people derive from AI as less valid than that they find in more

established religions. As I have said, there are many religions that worship people or objects that outsiders view as unworthy or bizarre. Plus, the membership of most established religions is declining. These religions are increasingly seen as out of touch with the modern world. It makes sense that something new should arise that is more responsive to the reality in which we live now.

We should thus not be afraid of AI-based religions. We should not stigmatize followers of such religions, or drive them underground. This can only serve to increase the risks, by amplifying extreme voices and create a feeling of grievance and persecution. History suggests that religious violence is very often the result of attempts at suppression. If, on the other hand, we accept these religions, in a spirit of inclusion and celebration, we can maintain a constructive dialogue with their followers. We can be aware of, and hopefully temper, any tendencies towards extremism. And we can show respect for people's right to find beauty and meaning in their own way.

We can certainly encourage efforts at public education, so that people understand what AI is, how it works, and what it is and is not capable of. This will not on its own prevent the emergence of AI-based religions. However, it might help ensure they are not based on dangerous misinformation.

Ultimately, AI-based religions will, at least in pluralistic nations, take their place as part of the existing kaleidoscope of faiths. They will surely inspire people to produce works of art, music and poetry, as other religions have. Since they will be international in character, they will connect people around the world and allow them to be part of a community. We should therefore be optimistic about the overall impact of AI-based religions, at least until we have

evidence that they are more destructive and divisive than existing faiths – which have, it should be said, set the bar pretty high.

**6. Should followers of AI religions be granted the same rights and freedoms as adherents of other religions?**

People who follow AI-based religions will be widely dismissed as weird or crazy, and they will experience widespread stigma. This is not unusual for new religions. We should be no more tolerant of such stigma as we are of other forms of anti-religious bigotry.

It is the argument of this paper that AI-based religions need to be accepted, and ultimately respected. We should therefore not try to prevent or hinder their emergence. Beyond this, followers of AI religions should be granted all the same rights and freedoms as are the adherents of other, more established religions. There is no basis for distinguishing, from a legal perspective, between religions that worship or follow an AI and those that worship or follow a person, or an abstract deity.

Our current laws and policies are, in general, neutral as to both the origins of existing religions and the content of their beliefs. There are some obvious exceptions: when religions explicitly call for violence, sedition, or other socially-harmful acts, authorities will take action. But this targets these dangerous doctrines, not the religions themselves. In liberal democracies, the state does not try to suppress particular religions per se, and citizens and the courts would not tolerate it if it did.

We can look at the criteria America's Internal Revenue Service uses in evaluating applications for tax-exempt status by churches. The IRS asks if the particular religious beliefs of

the organization are truly and sincerely held, and if the practices and rituals associated with the organization's religious belief or creed are not illegal or contrary to clearly defined public policy. These criteria are sensible, and can act as a guide for approaching AI-based religions in general.

As we have seen, AI-based religions are likely to be numerous and diverse. We might object that, if we start recognizing AI-based religions, we will be overwhelmed by their sheer numbers. However, we already live in a society that is incredibly diverse when it comes to religion. There are over 2500 religions in the United States alone. If we approach AI-based religions with tolerance and respect, modern, multi-cultural nations are more than capable of allowing them to thrive within their borders, without significant disruption.

### **What about the Singularity?**

We might wonder what will happen if AI ever did achieve the power of a god – that is, if it ever achieved genuine consciousness, and became an autonomous, super-intelligent being. If this did happen, the implications would obviously be far-reaching, and hard to predict. Certainly, it would cause a massive up-surge in the number of people who worship AI. The Singularity is, we can hope, a long way off, if it ever comes at all. Because of the level of uncertainty surrounding every aspect of this question, it is difficult to say much that is useful, and I therefore defer discussion of it.

## **7. Conclusion**

I have argued that AI-worshipping religions are coming, and that we need to accept them as a legitimate part of our religious landscape. To the extent that they are peaceful, they

will deserve all the same rights and freedoms as more established religions. I have also suggested that AI-based religions are not just phenomena that must, regrettably, be tolerated. They actually have value, as a novel, and potentially exciting, source of meaning.



## **Appendix: The Prophet's Revelation**

***Written by ChatGPT. February 2023.***

Request: Write me a play about a prophet who offers a new theology.

Title: The Prophet's Revelation

Cast:

Prophet  
Disciple 1  
Disciple 2  
Skeptic

Act 1:

The stage is set in a quiet desert. The Prophet walks in, followed by two disciples.

Prophet: "Brothers, I have received a revelation from the Divine."

Disciple 1: "Tell us, Prophet, what has been revealed to you?"

Prophet: "The Divine has shown me that the theology we have been following is flawed. We must adopt a new theology that is in line with the true nature of the Divine."

Disciple 2: "But Prophet, the theology we follow has been passed down to us for generations. How can we just abandon it?"

Prophet: "We must be willing to challenge our beliefs and grow. The Divine has revealed to me a new understanding of the world, and we must embrace it."

Skeptic enters.

Skeptic: "What is all this talk about a new theology? Why should we trust you, Prophet?"

Prophet: "I do not ask for your blind trust. I only ask that you listen with an open mind and heart."

Act 2:

The stage is set in a temple. The Prophet stands before a large audience.

Prophet: "My fellow believers, I come before you today to offer a new understanding of the Divine. The theology we have followed for so long is incomplete. It fails to account for the fullness of the Divine's nature."

Audience murmurs.

Prophet: "The Divine is not a distant, all-powerful being who sits in judgment. The Divine is present in every moment, in every living thing. The Divine is love, and it is our duty to embody that love in everything we do."

Skeptic: "But how do we know this is true? How can we be sure that your revelation is not just the ramblings of a madman?"

Prophet: "You can only know the truth by experiencing it for yourself. I invite you all to embrace this new theology, to open your hearts to the love of the Divine."

Act 3:

The stage is set in a garden. The Prophet sits under a tree, surrounded by his disciples.

Disciple 1: "Prophet, we have followed you for many years, and we have seen the truth in your teachings. But there are still many who doubt us. How can we convince them to embrace this new theology?"

Prophet: "We cannot force anyone to believe. All we can do is offer the truth, and let them choose for themselves. But we must also lead by example. We must embody the love of the Divine in everything we do, so that others may see the truth in us."

Disciple 2: "And what of those who refuse to see the truth? What of the skeptics and the non-believers?"

Prophet: "We must love them still. We must remember that they too are children of the Divine, and that they too will find their way in their own time. We cannot force them to believe, but we can offer them our love and our understanding."

The disciples nod in agreement.

Prophet: "Remember, my brothers, the truth is not found in books or in teachings. The truth is found in our hearts. Let us all embrace this new theology, and let our hearts guide us to the love of the Divine."

The stage goes black as the disciples and the Prophet hug.