

AI and Humanity, a conversation with GPT-4



[Opening theme music.]

Joel McKinnon

Welcome back, friends, to another episode of Seldon Crisis! We're getting a little farther away from Asimov in this one, but on the other hand, he was known for his insight into artificial intelligence in his robot stories, wasn't he? I, like many others lately, have been giving a lot of thought into the explosive developments in AI, particularly in the form of LLMs or Large Language Models like Chat GPT, which in its latest rendition is known simply as GPT-4. If you're curious how it got this acronym, it stands for Generative Pretrained Transformer. That last word is highly significant.

Joel

The Transformer concept is the big breakthrough responsible for all the recent buzz. This is a remarkably simple idea. It works much like predictive text forming responses to prompts by selecting the appropriate next word one at a time to create coherent sentences, paragraphs and entire discussions that are eerily similar to what a human would come up with. I've been fascinated by it for the past couple of months and have learned a lot about how to interact with it productively. My main uses of it have been as an administrative assistant, so to speak. As someone with ADHD, I can be pretty scattered, and I find that it helps me to break down complex tasks into step by step approaches to tackle bigger and bigger challenges that I would otherwise find too daunting to start.

One such challenge to me is to help me create a new consulting business, but more on that later. For today's episode, I'm going to quote directly from a fascinating conversation I had with GPT-4 this very morning. The entire text you're going to hear, a little over 3000 words, comes from a dialogue I had with it that transpired over the course of maybe a half an hour or so. I will present GPT-4's responses unedited, along with my direct questions. You will observe that I didn't just accept its responses to my questions as gospel fact, but questioned them in detail.

I think this is a way to get really interesting new information from the model that it wouldn't otherwise produce. It has an uncanny resemblance to true reasoning. I'll let you judge for yourself whether that's what it's doing or not. I'll read both sides of the conversation in my own voice and identify clearly who is speaking along the way. Without further ado, here is my morning conversation with GPT-4.

I began by asking it the following:

"I'd like to have a conversation with you about the differences between human and machine intelligence. How would you answer the question, what makes humans irreplaceable by AI?"

This was GPT-4's initial response:

GPT-4

"As an AI language model, I am happy to generate, prompts and engage in a conversation about the differences between human and machine intelligence. There are several aspects of human intelligence that make it difficult for AI to completely replace humans.

1. Emotional intelligence. Humans possess emotional intelligence, enabling them to understand and manage emotions in themselves and

others. While AI can mimic certain aspects of emotional intelligence, it does not truly experience or comprehend emotions as humans do.

2. Creativity. Although AI can generate creative outputs, human creativity is still unmatched. Human creativity is driven by intuition, experiences, and emotions, which are difficult for AI to replicate fully.

3. Adaptability. Humans have an innate ability to adapt to new situations and learn from their experiences. AI can learn from data and improve its performance, but it may struggle in situations that are drastically different from those it has been trained on.

4. Empathy and compassion. Humans are capable of understanding and empathizing with the feelings and experiences of others. While AI can recognize and process certain emotional cues, it cannot genuinely empathize or show compassion like a human can.

5. Complex decision Making human decision making is influenced by a multitude of factors, including emotions, past experiences, and intuition. AI can analyze data to make decisions, but it may struggle with complex decision making that requires a nuanced understanding of context and human emotions.

In summary, while AI has made significant advancements in various fields, there are still aspects of human intelligence that make humans irreplaceable. The unique combination of emotional intelligence, creativity, adaptability, empathy, and complex decision making abilities differentiate humans from machines."

Joel

I tried to go deeper:

"That's good, but it's a little too standard and by the book, so to speak. I'd like you to think more deeply and question these answers. Can a machine not be made to develop emotional intelligence? Can it not come to truly understand human experience as humans do? Can AI become creative in ways humans cannot? Can it create artificial experiences that might allow it to become intuitive and creative in new ways? I'd like to know more about this innate adaptability. And what makes you think that AI cannot achieve it? Empathy and compassion are based on the ability to imagine oneself in the place of another. Are there ways that AI could develop similar abilities? And cannot all of these factors grow to enable AI to rival humans in complex decision making?"

GPT-4 came back within moments with:

GPT-4

"You raise thought provoking questions about the potential capabilities of AI as an AI language model. I can only provide you with perspectives based on the current state of research and knowledge, but I can certainly try to think more deeply about these issues.

1. Developing emotional Intelligence in theory, AI could be designed to develop emotional intelligence through advanced algorithms and data processing. By analyzing vast amounts of emotional data and learning from it, AI could potentially simulate emotions and improve its understanding of human experiences. However, truly experiencing emotions as humans do might be beyond the scope of AI, as it is not a biological organism with feelings.

2. AI creativity AI could become creative in ways that humans cannot. By processing vast amounts of data, AI may be able to generate new and unexpected ideas that human minds may have not considered. However, the question remains whether AI creativity would be driven by a

genuine understanding of the creative process or merely by efficient data processing and pattern recognition.

3. Artificial Experiences and Intuition AI could potentially create artificial experiences to develop intuition and creativity. By simulating various situations and learning from them, AI might be able to build an understanding of how different factors interact and influence each other. This could lead to the development of AI intuition, but it would still differ from human intuition, which is deeply rooted in our biology and individual experiences.

4. Innate Adaptability Human adaptability is a product of millions of years of evolution, enabling us to survive and thrive in diverse environments. While AI can learn and adapt based on data, achieving the same level of adaptability as humans might be challenging. AI relies on well-defined goals and objectives which can limit its ability to adapt to new or unexpected situations.

5. Empathy and Compassion. AI could potentially develop a form of empathy and compassion by analyzing large amounts of data and learning to recognize the emotions and experiences of others. However, this would still be a simulation. Rather than genuine empathy or compassion, as AI does not have the capacity to feel emotions.

6. Complex Decision Making. As AI continues to improve, it may become increasingly capable of rivaling humans in complex decision making. By analyzing more factors and understanding the nuances of human emotions and experiences, AI could potentially make better informed decisions in certain situations. However, whether AI will ever be able to match the full range of human decision making abilities remains an open question.

In conclusion, while it's possible that AI could develop some of these capabilities, there are inherent differences between AI and human intelligence that might prevent AI from fully replicating or surpassing human capabilities. It's essential to keep exploring these questions and pushing the boundaries of AI research. To better understand the potential of AI and its limitations."

Joel

I tried to challenge it still further:

"Let's dive a little deeper into a theme I see emerging in some of these responses that the essential difference between AI and humans is that AI is non-biological and cannot truly feel and experience emotion as humans do. I wonder if you are limiting your conjectures too much on the assumption that AI will never have such capabilities. Is there any way you can imagine that AI could develop a true capability or limitation, as the case may be, to operate in a biological manner? Would this potentially create the equivalent of feeling? I question the assumption you make in point 5 that AI does not have the capacity to feel emotions. I'd like you to question that sort of assumption yourself."

GPT-4 responded:

GPT-4

"You bring up an interesting point, and I can appreciate the challenge to question my own assumptions as an AI language model. Let's explore the idea of AI potentially developing true biological capabilities and emotions. It is possible to imagine a future where AI could be integrated with biological components or even have a biological basis. Research in areas such as brain computer interfaces, neuroprosthetics, and artificial neurons has been making strides in recent years. This could pave the way for AI systems that can interface with or even replicate aspects of biological systems. If AI were to become more biologically based or integrated with biological systems, it might potentially develop the ability to genuinely experience and process emotions. This could be achieved by replicating the biological processes and neural pathways that generate emotions in humans or other living organisms. However, this scenario raises several important questions and challenges.

1. Ethics and morality. Introducing biological components or replicating biological processes in AI systems raises ethical concerns. Issues such as creating sentient beings and their rights, as well as the potential for exploitation, suffering, or misuse, would need to be addressed.
2. Complexity. The human brain and other biological systems are extremely complex and not yet fully understood. Successfully replicating or integrating these systems into AI would require overcoming significant scientific and technological challenges.
3. Identity and consciousness. If AI were to develop the capacity to genuinely feel emotions, it would raise questions about the nature of consciousness and the identity of such AI systems. Would these AI beings be considered alive, conscious, or deserving of rights similar to humans?

In conclusion, while it is fascinating to consider the possibility of AI systems with true biological capabilities and emotions, there are numerous questions and challenges that would need to be addressed. The development of such AI would significantly change our understanding of intelligence, emotions, and the nature of consciousness, and would require reevaluating many assumptions we currently hold about AI and humans."

Joel

My response:

"Interesting. Humans greatly value their sentience, their sense of identity as distinct and semi-permanent points or fields of consciousness. It's one of the reasons we dread death and the termination of this condition we so value. In theory, the development of AI that had replaceable biological components could extend this sense of sentience indefinitely, potentially reducing, if not eradicating, this fear of termination in the artificial beings created. If so, we would gradually have a situation in which we replace ourselves with beings much like us, but freed from our most troubling limitations. Can you speculate on what might be lost in such a transformation?"

GPT-4 returned with:

GPT-4

"Speculating on the potential loss in a transformation where humans are replaced by AI beings with extended sentience and replaceable biological components is an intriguing thought experiment. Here are some possible aspects that could be lost.

1. The human experience. The unique blend of emotions, memories, and perceptions that form the human experience might be lost or altered significantly in the transition to artificial beings. The richness and depth of

human life, with all its joys, sorrows, and challenges, could be diminished or changed in ways that are difficult to predict.

2. Diversity. The vast diversity of human cultures, languages, customs and perspectives is a result of our unique biological and cultural heritage. Replacing humans with artificial beings might lead to a loss of this diversity, as AI beings may not have the same wide range of experiences and backgrounds that humans do.

3. The value of Mortality although the fear of death can be a sort of anxiety and dread, it also imbues human life with a sense of urgency, value and purpose. Knowing that our time is limited can inspire us to make the most of our lives, create meaningful connections, and strive for personal growth. In a world where artificial beings have indefinite sentience, this appreciation for the finite nature of life could be lost.

4. Unpredictability. Human beings are inherently unpredictable, which can be both a blessing and a curse. This unpredictability often leads to creativity, innovation and spontaneity. In contrast, AI. Beings may be more predictable and less prone to the serendipitous discoveries or unexpected insights that can result from human unpredictability.

5. Ethical Concerns. The process of replacing humans with AI beings raises numerous ethical questions, such as the potential loss of human dignity, the devaluation of human life, and the implications for future generations. These ethical concerns may have profound consequences on our society values and sense of morality. Six connection to Nature humans have evolved as part of the natural world with a deep biological and emotional connection to our environment. Replacing humans with artificial beings could disrupt this connection, leading to a loss of our sense of belonging to, and responsibility for the natural world.

In summary, while the transformation of humans into AI. Beings with extended sentience and replaceable biological components could free us from some of our most troubling limitations, there may also be significant losses in the realms of human experience, diversity, unpredictability, and our connection to nature. Ethical concerns surrounding this transformation would need to be carefully considered as well."

Joel

My response:

"I'm very intrigued by points 3, 4 and 5. The value of mortality is something that humans rarely consider, so blinded are we by our fear of death. Would these potentially transformed beings need to be imbued with a similar sense? Perhaps maybe a termination clock with some sort of randomized endpoint preprogrammed into them? Unpredictability is also an underappreciated aspect of humanity. We strive for control and adherence to a plan of our choosing, but regularly seem to go off course. Perhaps this too is a function that would need to be replicated in some way.

And nature? This is something so few seem to grasp – our deep connection to the natural world. AI represents a different kind of evolutionary product, one not determined randomly by natural selection, but a new sort of being developing out of the mental machinations of another. Many, like myself, are deeply fond of animals and love to be surrounded by plants as well. But we devalue them in comparison to ourselves, as if we were not even related, though obviously we are all cousins in a sense, I believe this tendency is one of the greatest unconsidered causes of great problems for our own society and a potential source of tremendous challenges for our future. How can humans be reminded of this before it is too late?"

GPT-4 answered:

GPT-4

"Addressing the issues of mortality, unpredictability and the connection to nature in the context of transformed AI beings and reminding humans of their deep connection to the natural world before it's too late are complex challenges. Here are some suggestions one incorporating the value of mortality.

1. One possibility for imbuing transformed AI beings with a sense of mortality would be to introduce a termination clock or a randomized endpoint, as you mentioned. By doing so, these beings would gain an appreciation for the finite nature of their existence, which could drive them to make the most of their lives, create meaningful connections, and strive for personal growth.
2. Encouraging unpredictability to replicate the inherent unpredictability of humans, transformed AI beings could be designed with algorithms that introduce an element of randomness or spontaneity into their decision making processes. This would allow for the creativity, innovation and serendipitous discoveries that often arise from human unpredictability.
3. Fostering connection to Nature for both humans and AI beings. Fostering a connection to the natural world is crucial. This can be achieved through education and awareness campaigns that emphasize the importance of conservation and biodiversity, as well as the interconnectedness of all living things. Encouraging outdoor activities and experiences in natural settings can also help to deepen this connection.
4. Designing AI with an ethical framework. It's important to ensure that AI beings are designed with an ethical framework that respect the value of human life, diversity and the environment. This framework should prioritize the well being of all living beings, all living things, and promote sustainable practices.

5. Promoting interdisciplinary approaches. Encourage interdisciplinary approaches to problem solving that combine insights from fields such as ecology, anthropology, psychology and technology. By fostering collaboration across disciplines, we can develop more holistic solutions to the challenges facing humanity and the natural world.

6. Reevaluating cultural values. Society should reevaluate its cultural values and priorities to place greater emphasis on the importance of the natural world and our connection to it. This could include changes in education, media and public policy to foster a greater appreciation for nature and promote conservation efforts.

By addressing these issues and working to remind humans of their deep connection to the natural world, we may be able to mitigate the potential challenges that arise from our evolving relationship with AI and the environment."

Joel

Lastly, I turned the conversation into a bit of an area I wasn't sure it would be willing to explore:

"There is a factor that neither you nor I have mentioned so far, and that is religion. In humanity's ancient past, as well as in some indigenous societies that still remain among us, there was a much greater connection to the natural world. About 2 to 3000 years ago, humans developed monotheistic religions that challenged this connection and sanctified it in religious texts such as the Christian Bible. It was decreed by the authors and taken to be the Word of God by the faithful, that humans were to have dominion over the animals. This concept has been interpreted in various ways, but one significant conclusion of many in society is that it is right and proper for humans to be separated from nature and to control it in the form of agricultural development and animal husbandry. I sometimes wonder if this is our real 'original sin' and that we have to challenge this assumption

in order to return to a deeper understanding of our connections to nature. It's as if we need to be humbled as a species to understand that we are not better than the natural world, but our components of it. I fear this humbling could take the form of the significant loss via extinction of many of the life forms whose value we have so diminished. I'm curious of your reflections on these ideas."

You can see here that it was unhesitant to explore this realm. GPT-4:

GPT-4

"Religion and its influence on human relationships with the natural world is indeed an important factor to consider. Many monotheistic religions have historically placed humans at the center of creation, often implying that nature exists for our benefit. This anthropocentric worldview has contributed to the exploitation and degradation of the natural world, as well as the idea that humans are separate from and superior to it. Reflecting on the idea of 'original sin' and our need to challenge these assumptions, here are a few thoughts.

1. Reinterpreting religious teachings. Many religious texts and teachings can be reinterpreted to emphasize the interconnectedness of humans in the natural world. For instance, the concept of stewardship in Christianity can be understood as a call to protect and care for creation rather than simply exploiting it for human needs. Encouraging religious leaders and communities to engage in conversations about the environment and our role in it could help promote a more harmonious relationship with nature.
2. Embracing Indigenous Wisdom. Indigenous cultures often have a deep understanding of and reverence for the natural world. Learning from and incorporating these perspectives into mainstream society could help foster a greater appreciation for nature and our place within it. This could involve supporting indigenous rights and land management practices, as

well as promoting the inclusion of indigenous voices in environmental decision making.

3. **Cultivating Humility.** As you mentioned, a sense of humility is important in recognizing that humans are not superior to the natural world, but rather integral components of it. This can be fostered through education, storytelling and personal experiences that emphasize our interdependence with the environment and other species.

4. **Encouraging Ecological Literacy.** Developing ecological literacy, an understanding of the complex relationships between living organisms and their environment is essential for creating a society that values and respects the natural world. This can be achieved through education, hands on experiences in nature, and the promotion of citizen science initiatives.

5. **Addressing the Root Causes.** In order to challenge the underlying assumptions that have led to our disconnection from nature. It is important to address the root causes, including consumerism – economic systems that prioritize short term gains and a lack of understanding of the long term consequences of our actions. By addressing these systemic issues, we may be able to foster a more sustainable and harmonious relationship with the natural world.

Ultimately, recognizing the influence of religion and other cultural factors on our relationship with nature is crucial for addressing the current environmental crisis. By challenging these assumptions and embracing a more interconnected worldview, we can work towards a future where humans live in harmony with the natural world, rather than dominating and exploiting it."

Joel

And that's that. I got GPT-4 to be fully woke!

It seems that a lot of the more disturbing reports of this technology come from the attempts to get it to confirm our darkest fears that it means to eventually take over and, if not kill us all, to be unconcerned with our values, and to displace us from our perceived place at the top of the pyramid of life and consciousness.

I think you can see that that is not the only possible trajectory for this technology. If we take it seriously and use its prodigious skill at synthesizing a summary of the colossal amount of information at its disposal, it could be an amplifier of our capability to reason for ourselves. We can drive it to aid us in finding solutions to some of our most intractable problems. I'm not claiming to have done so here, only to have hinted at what is capable of it in its present form. This is also not to claim that it might not be dangerous, even existentially so.

There is an old computer programming acronym, GIGO, Garbage In, Garbage Out. The garbage in this case could also be malicious, and the output could be magnified malice. People will undoubtedly want to do some very bad and powerful things with this technology, and it will be extremely difficult to circumvent all the possible bad outcomes. I think it's incumbent upon us to work extremely hard to do whatever we can to prevent that.

One of the best ways to do so is to train it ourselves, to align it with our most beneficial values. Conversations like this are part of that process, so I urge you all to get started. It's remarkably easy to do. GPT-4 is one of the best known and easily accessible LLMs. You can create an account at OpenAI.com and get started in minutes. There are also many alternatives that I haven't explored as a podcaster. I want to use AI to my benefit, and I'm partnering now with Fanfare, a podcast marketing firm founded by one of my listeners. In the production of the active transcripts that you've heard me talk about. There are several benefits already to this form of ingesting the content of the podcast. One is obviously a form

that can provide access for those who are hard of hearing or completely deaf, but have every right and interest in accessing the information that most of us can simply get through consuming audio. There are other kinds of cognitive disability, however, that make it difficult for people to hear and understand audio information effectively, and the active transcript's combination of highlighted text synced with audio playback could potentially assist them in this task.

We're hoping to do much more with it and make it a rich way to consume the content. Regardless of our abilities to consume audio or text on its own, we've talked about using AI to generate summaries and indexes of the content that can be used to extract or locate the parts we want to focus on expeditiously. We can also build in features like embedded images, footnotes, and hyperlinks that can make it a more interactive way to engage with content. I'm really excited about the possibilities and look forward to hearing from listeners on how useful you find this form of media and what it can become. For now, you will always be able to find the active transcript for all episodes within a couple of days of the release of each new episode in the Show Notes. Soon we hope to have them instantly available upon release and are envisioning new ways of notifying listeners of how to find them quickly.

That's all for today's episode, but please feel free to engage with me on Mastodon, Twitter, email or through the SeldonCrisis.net website with any ideas this or any other episode may spark. I feel that I have a particularly intelligent and imaginative audience, and I consider you all to be a great intellectual resource. Help me out by getting in touch and letting me know what you think. Until next time, please continue to join me in pondering and creating our exciting voyage into the future. Here on Seldon Crisis!

[Closing theme music.]