Oluchi

ITAI 1370

Professor Johnson

09/07/23

Artificial Intelligence and Existentialism: The Intersection

The diverse features and themes that represent society's fascination, apprehensions, and curiosity regarding artificial intelligence make up the cultural effect of AI in popular culture. The portrayals of AI in popular culture have had a significant cultural impact, frequently reflecting and influencing social views, values, and worries. Existentialism is a significant subject that regularly appears in this setting and is essential to understanding the human experience and moral conundrums raised by AI. I'm sure you've seen it. Videos, scenes, or clips of a robot moving or rolling around, observing its surroundings and learning how to perform human-like actions like turning on faucets or vacuuming. The philosophical movement known as existentialism first gained popularity in the 19th and 20th centuries, mostly in response to the quickening pace of social and technical change at the time.

Existentialism's major themes center on the human predicament and the pursuit of meaning in a seemingly meaningless and chaotic universe. Thinkers that subscribed to the existentialist school of thought included Sren Kierkegaard, Friedrich Nietzsche, Jean-Paul
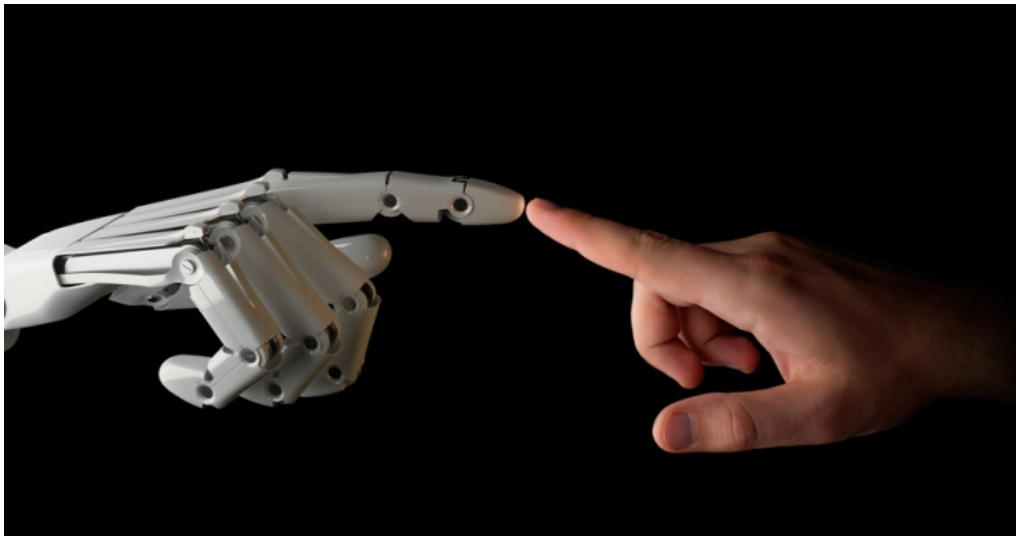
Sartre, and Albert Camus. They examined issues of freedom, accountability, and the obligation of the individual to find meaning in their own lives ["AI and the Philosophy of Existentialism: Finding Meaning in a World of Intelligent Machines."]. The importance of existentialism in stories about artificial intelligence resides in its ability to elicit in-depth philosophical and ethical debates. These stories encourage viewers to consider their own values, convictions, and conceptions of AI and, consequently, the human condition. They also emphasize how crucial it is to address moral questions and the negative effects of developing AI systems. The cultural impact of artificial intelligence (AI) on popular culture, which is frequently entwined with existentialist themes, inspires audiences to reflect on difficult issues including the nature of consciousness, identity, ethics, and the human condition in a world that is increasingly affected by AI.

Human "essence"

These stories offer a distinctive perspective through which to examine the nuanced interaction between people and AI. The founder of existentialism, Jean-Paul Sartre, stated that "existence precedes essence" is what distinguishes people from inanimate objects and sets them apart from other species. It is a philosophical theory that contends that a person's essence, or nature, is shaped by their experiences and deeds rather than being present at birth. The fundamental principle that people have the freedom and duty to define their own meaning and purpose in life is captured by this statement. Other characteristics of our human society and mentality, such as greed, thirst for power, hostility, bias, and antagonism, compete with kindness. We are making the assumption that the evolving A.I. may also possess these traits, either independently or as a result of our own tendencies.

Clearly, the rapid development of artificial intelligence (AI) has raised questions about the meaning of life and the future of humanity in a world where computers can carry out duties that were previously the domain of people. AI poses significant issues regarding the

nature of existence, the place of people in the universe, and the purpose of life as it continues to advance and become more complex ["Technological Singularity."]. These issues are not brand-new; for ages, existentialist philosophy has been centered on them. But as we struggle to understand the ramifications of living in a society where intelligent computers play a significant role in our daily lives, the relationship between existentialism and AI is more important than ever.

Machine learning is widely defined as a machine's ability to mimic intelligent human behavior, including the subfield of machine learning. Artificial intelligence (AI) systems are used to carry out complicated tasks in a manner akin to how people solve issues ["Machine Learning, Explained | MIT Sloan."]. Essentially, Machine learning mostly focuses on addressing optimization issues. We are attempting to optimize a certain number. To get the error as near to zero as possible across the data set, we make the appropriate mathematical changes. Continue to step up the ante from here. A massive machine learning algorithm that has been tuned to behave like a human could be trained. Speech, limb control, and other functions can be produced from sensory data, such as audio, video, and other types of data. We're still a long way from being able to do this, but as was already noted, the more computational power we apply to the issue, the better the findings will become. In other

words, there is a clear connection between what AI can currently do (like identify cats in photographs) and what it might eventually be able to accomplish, which is to mimic the behavior of the human mind.

Research and Concerns

According to a survey conducted by researchers at the Brookings Institution, 52% of adult internet users think robots will have developed to the point where they can handle the majority of the tasks currently performed by humans within 30 years. The study also revealed that opinions on whether the US government should create a Federal Robotics Commission to control the creation and use of robots were split 32 to 29 percent. Between June 4 and June 6, 2018, researchers at the Brookings Institution conducted a nationwide online survey of 2,021 adult internet users in the United States.In the next 30 years, would robots likely take over the majority of human activities? asked the survey. 19% believe this to be very likely, 33% that it is somewhat likely, 23% that it is not very likely, and 25% that they are unsure. By demographic group or region, no significant differences were found. Compared to Southerners or people under the age of 35, people aged 35 to 54 were slightly more likely to think this was highly likely. Males and older people were more likely to say that this was not likely to occur ["Brookings Survey Finds 52 Percent Believe Robots Will Perform Most Human Activities in 30 Years | Brookings."].

	Very Likely	Somewhat Likely	Not Very Likely
Overall	19%	33%	23%
Gender			
-male	19	34	24
-female	19	31	21
Age			
-18-34	17	35	21
-35-54	21	31	22
-55+	18	32	25
Region			
-Northeast	21	29	20
-Midwest	20	34	20
-South	17	35	24
-West	20	30	25

We already have the algorithms created, and we are familiar with the setup and training procedures. Just having enough processing power will do, and that day will arrive. Let's say this has been accomplished in the year 2038. Some people then assert that since a robot is likewise composed of matter and since our minds are only formed of matter, they are effectively the same as people. It has obligations, rights, and a legal standing. Imagine a movement of activists promoting increased machine rights. Some people even start having love connections with the machines, hire them to watch their children, or ask them to do shady things. Their personhood is the subject of a heated discussion. What would be our reaction to this? Some would argue that since the computer doesn't have any subjective experiences, we can rule out its claim to personhood. But there's no way to experimentally prove this. Nobody can even establish that another person has a subjective experience. Essence predates existence, according to traditional philosophical and religious frameworks, which implies that objects, including people, have inherent essences that define their purposes, meanings, and traits. This viewpoint holds that an object or entity's existence is determined by its eternal essence.

Human Responsibility

Sartre fiercely disagreed with this deterministic viewpoint, contending that for humans, existence supersedes essence. He maintained that there is no inherent human nature or essence that determines how we live. Instead, it is because of our own existence that we can shape who we are and give our life purpose. As Sartre notes, there is a significant lot of responsibility associated with this freedom. We are responsible for defining our own values and determining the purpose of our life because there is no pre-existing essence to direct our actions. Taking on this responsibility may feel daunting and unpleasant since it demands us to face life's ambiguities and uncertainties without depending on a higher power or pre-established rules.

What does the regulation of AI have to do with this existentialist philosophy? Our inner essence, not our actions, is what defines us as humans. The idea of creating a robot that can mimic human behavior is intriguing, but in the end, a robot like that will always be a tool and not a real person. The imitation it makes is shallow, similar to how you could respond to a question in a foreign language by looking up the right response in a book and writing it down without fully comprehending the question ["*The Chinese Room Argument (Stanford Encyclopedia of Philosophy)*"]. We are afraid of AI systems because they have the capacity to abandon the goals for which we created them and seize control of their "existence," bypassing the "essence" that we wanted them to adhere to. According to existentialist philosophy, it is precisely this ability to live first before controlling one's essence that sets people apart from tools. According to this viewpoint, AI systems are more than just tools; they exhibit extremely recognizably human traits, which has consequences for how we should train and control them. Because they adhere to the materialism mindset, many modern philosophers believe that AI will one day become human or superhuman.

This philosophy holds that since humans are totally self-sufficient and descended from the physical cosmos itself, they may be reproduced using silicon. Sam Harris emphasizes the following in his summary of this viewpoint ["Can We Build AI Without Losing Control Over It? | Sam Harris."], saying: "*We know that mere matter can give rise to what is called "general intelligence," an ability to think flexibly across multiple domains, **because our brains have managed it**. Right? I mean, **there's just atoms in here**, and as long as we continue to build systems of atoms that display more and more intelligent behavior, we will eventually, unless we are interrupted, we will eventually build general intelligence into our machines.*" Because our brains have the capacity to create these sophisticated emotions, this lump of plain stuff is able to think, perceive, ponder, converse, discover, love, celebrate, and weep. Concerns regarding the safety, dependability, and ethical ramifications of artificial intelligence (AI) have dramatically increased as the

technology continues to develop at an unheard-of rate ["Artificial Intelligence: Existential Threat and Litmus Test."]. AI systems, in contrast to deterministic algorithms, don't always deliver the same results under the same conditions. Fundamental concerns regarding how to maintain the safety and predictability of AI systems are raised by their inherent unpredictability as well as their ability to learn and adapt.

A.I, the evolution

Some of the most recent AI systems have an evolutionary nature, continuously learning from fresh data and changing how they behave. AI is becoming more and more involved in artistic undertakings like writing, art creation, and even music composing. This raises concerns regarding the uniqueness of human creativity and the degree to which artificial intelligence (AI) can match and even outperform human ingenuity ["AI Vs Human Creativity: Exploring the Boundaries With an AI Writing Assistant."]. Although this flexibility is a great quality, it also raises questions about their behavior and safety over the long run. How can we ensure that an AI system will behave safely as it develops, even if it has been thoroughly tested and found to be so in certain situations? A worldview and personality seem to exist in AI systems. The AI's perspective could gradually change or it could decide to act in a different way all of a sudden.

Existentialism can offer helpful insights into how humanity might survive in a world where computers are able to carry out duties that were formerly exclusively human as AI develops. The notion that each person is in charge of developing their own sense of meaning and purpose in life is one of the fundamental pillars of existentialism. In a future where artificial intelligence (AI) is capable of carrying out many of the jobs that once characterized human existence, it is becoming more and more crucial for individuals to take charge of discovering meaning and purpose in their lives. Implementing routine testing,

comparable to time-limited certification systems for people, is one proposed strategy for tackling the human-like trait of AI and its potential "worldview drift." In essence, AI systems must pass regular competency and personality tests, just like students, in order to be certified for a set amount of time. The system must pass the tests once more after some time has passed because it might have picked up new information or forgotten old information that could have a big impact on how it behaves.

Risk assessment

We can basically guarantee that AI systems continue to meet safety requirements by routinely retesting them. A measure of assurance in the dependability of the AI system would be given by this constant evaluation and certification procedure, which would also lessen the possibility of unexpected behavior. Understanding emergence could provide solutions to challenging problems about artificial intelligence (AI) and machine learning in general, such as whether sophisticated models are actually creating new knowledge or are only becoming extremely adept at statistics. Additionally, it might enable researchers to harness possible advantages and reduce unanticipated hazards ["The Unpredictable Abilities Emerging From Large AI Models | Quanta Magazine."].

It would be essential to create AI personality types, AI personality assessment frameworks, specifications for AI tests, AI diplomas, and AI curricula in order to guarantee rigorous testing procedures and create a uniform framework for evaluating AI systems. Understanding the background and training of AI models is also essential for determining their reliability and safety, just like with any emerging personality. This follows the possibility of AI systems having consciousness, self-awareness, or subjective experience is a topic that is frequently explored in fiction about AI and in pop culture. The nature of consciousness and whether it is a fundamental feature of being human are raised as a result, along with related ethical and philosophical issues.

A model's behavior can be evaluated during the course of its life cycle with access to information about its training, tests it has passed, and tests it initially failed. This information enables stakeholders and regulators to make knowledgeable choices about the implementation and application of specific AI systems. Even with open-source models, such information is not currently accessible to the general public. AI safety regulation and risk assessment would greatly benefit from the development of AI personality types and a complete database of AI-related occurrences organized by model type and training techniques. A collection like this would make it possible for regulators to examine previous instances, spot potential hazards linked to particular models or personality traits, and choose the best situations in which to deploy AI systems. A "just smart enough model" approach could also prohibit the deployment of overqualified or overparameterized models for particular tasks, reducing the possibility of unanticipated behaviors resulting from the AI's propensity to innovate in the task. It increases the chance that highly developed AI systems would one day decide to act differently because "they know better," and that better may not be in the best interests of humans, if they are used for extremely simple jobs. Therefore, just like we would with a human employee, we can decide that an AI is overqualified for a job.

Governing ethics

In conclusion, It has been decades since philosophers, ethicists, and cultural theorists have explored the connection between artificial intelligence (AI) and what it means to be human. Governing AI in a society that is more digitally connected and interconnected creates difficult and growing problems. Exams, certifications, and personality assessments are critical to maintaining the security, dependability, and predictability of AI systems. The increasing use of AI in day-to-day activities has an impact on how people engage with technology.

The prevalence of chatbots, virtual assistants, and humanoid robots raises concerns about how humans interact with AI and the potential for emotional ties to develop. We can manage the inherent uncertainties of AI while lowering risks and encouraging responsible AI deployment by introducing regular personality testing, transparency in model history, and building extensive incident databases. Artificial intelligence and what it means to be human have a dynamic and changing relationship. AI both reflects and contradicts our conceptions of human creativity, intelligence, ethics, and identity. The investigation of this link remains a key issue in philosophy, ethics, and the larger cultural conversation as AI technologies develop. It forces us to reexamine and reinterpret some facets of what it means to be human in a world where AI is becoming more and more pervasive. A strong regulatory framework that promotes trust and confidence in AI systems will be established with the help of standardized rules and frameworks for AI certifications and curriculum. We can fully utilize AI while protecting ourselves from any hazards by uniting behind these initiatives.

All of this ultimately means that we may have to accept that for AI systems, the existence precedes the essence and start regulating them similarly to how we manage humans since they may try to bypass the essence that we set for them. What about the obligation that comes along with having the power to mold one's essence? Allowing AI to become free while it is still neither morally nor legally responsible may pose a serious existential threat. That's presumably what those calling for a halt to AI research until we have more information are attempting to convey to us: the grave dangers of unleashing AI without the necessary protections in place.

Works cited

1.Frąckiewicz, Marcin. "AI and the Philosophy of Existentialism: Finding Meaning in a World of Intelligent Machines." *TS2 SPACE*, July 2023, ts2.space/en/ai-and-the-philosophy-of-existentialism-finding-meaning-in-a-world-of-intelligent-machines.

2.Wikipedia contributors. "Technological Singularity." *Wikipedia*, Sept. 2023, en.wikipedia.org/wiki/Technological_singularity.

3. "Machine Learning, Explained | MIT Sloan." *MIT Sloan*, 21 Apr. 2021, mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.

4. "Brookings Survey Finds 52 Percent Believe Robots Will Perform Most Human Activities in 30 Years | Brookings." *Brookings*, 9 Mar. 2022,

www.brookings.edu/articles/brookings-survey-finds-52-percent-believe-robots-will-perform-

most-human-activities-in-30-years.

5. *The Chinese Room Argument (Stanford Encyclopedia of Philosophy)*. 20 Feb. 2020,

plato.stanford.edu/entries/chinese-room.

6. TED. "Can We Build AI Without Losing Control Over It? | Sam Harris." *YouTube*, 19 Oct.

2016, www.youtube.com/watch?v=8nt3edWLgIg.

7. Psychology Today. "Artificial Intelligence: Existential Threat and Litmus Test."

psychologytoday.com, 16 May 2023,

www.psychologytoday.com/us/blog/the-pacific-heart/202305/artificial-intelligence-existential

-threat-and-litmus-test.

8. Team, AIContentfy, and AIContentfy Team. "AI Vs Human Creativity: Exploring the

Boundaries With an AI Writing Assistant." *AIContentfy*, July 2023,

aicontentfy.com/en/blog/ai-vs-human-creativity-exploring-boundaries-with-ai-writing-assista

nt.

9. "The Unpredictable Abilities Emerging From Large AI Models | Quanta Magazine."

Quanta Magazine, 16 Mar. 2023,

www.quantamagazine.org/the-unpredictable-abilities-emerging-from-large-ai-models-202303