

Chase Frank

Sample

Humes

3 April 2023

The Matrix's Warning Against AI and Uploading

Over the course of the year thus far the humanities class has grappled with several different subjects. The majority of these subjects revolve around the question of what it means to be human. To help answer this question the class has addressed critical subjects that make up who you feel you are, among these topics are debates regarding identity such as gender, race, and even religion. Later on the notion of Artificial Intelligence and what this means for humanity as our future approaches was introduced. Upon first exposure to the topic it seemed rather profound. This is due to the way in which the debate surrounding artificial intelligence addresses something that is inevitable and largely associated with not just humanity but all forms of life which is death. Initially it feels awkward to claim that artificial intelligence will alter death because it seems when thinking about technology and about artificial intelligence the first instinct is to think about things like self-driving cars, or more advanced versions of virtual assistants. But when discussing artificial intelligence it is rare to view it as a tool that could truly alter how we as a species interpret life. However, throughout the conversation about artificial intelligence and what this potentially means for the future, the notion of having the ability to upload one's consciousness to a computer emerged. This notion would be groundbreaking as it essentially means that instead of one passing away with their physical body they would instead be able to upload their minds to a computer and live in a virtual reality. Though this set of ideas brings up numerous questions, what would a possibility like this do to society? Would people

find comfort in the notion that they could live forever? Would this technology propel humanity forward or ultimately ruin it?

Several questions regarding what this kind of technology would do to humanity are addressed quite directly in the movie *The Matrix*. The Matrix is a science fiction film directed by Lana and Lilly Wachowski from 1999 that grapples with the concept of artificial intelligence in a very intriguing way. In the movie humans have developed artificial intelligence so far that artificial intelligence takes control of our species. In relation to the argument surrounding artificial intelligence and prolonging one's life this movie parallels these notions in a fascinating way, before artificial intelligence was used commercially, Wachowski was pushing the boundaries of what artificial intelligence can do and how we as a species would respond to it. Though the *Matrix* isn't alone with its ideas about AI (movies like *A Space Odyssey* for example) this was the first movie that imagined AI being in control of our reality through feeding us sensations through a simulated reality. This paper will further discuss The Matrix later, but first it is critical to look at the conception of death and how it has evolved throughout history.

When looking for the simple medical definition of death it was surprisingly difficult to locate the sole meaning of the term. In his philosophical book *Form and Object* Tristan Garcia discusses the relationship between objects and the forms in which they take. Through this lense Garcia emphasizes the impact of technology on our understanding of death. For example Garcia discusses the criteria for death during the Renaissance which was “the last breath.”¹ He then goes on to explain that through the discovery of the cardiac function of the body, the criteria for death was altered to the heart stopping instead of breath stopping as a way to identify the dead. This revised definition of death was yet again altered in the 1960's when development in medical

¹ Garcia, “Death.” *Form and Object: A Treatise on Things*.

technology enabled respiratory assistance and resuscitation.² Once again this resulted in a further developed definition of death which now required the brain to be completely beyond the possibility of functioning. Garcia's brief summary of historical revision shows that through the use and development of technology we have continually redefined what it means to be dead. These continuous changes to a seemingly inevitable and incredibly natural aspect of life encouraged me to ask the question in what further ways can death be redefined through technology? Furthermore, what do these changes entail for humanity? (Move through this quicker and in a way that is more tied to what my thesis is) (Maybe after intro dive right into analysis of matrix and interweave these concepts, or first talk about matrix and shift gears saying where are we now)

As time goes on, it seems that the human race's dependence on technology has increased so much that it now has altered how we as humans die. Garcia addresses this notion in *Form and Object* stating that through the use of technology we have blurred the lines between the living and the dead. The example Garcia uses is the state of vegetation. If we were to endure serious injuries that fatally harmed our organs 100 years ago we would most likely be pronounced dead.³ However, with the advancement of technology, it is now possible to prolong life by artificially maintaining these organs. As Garcia points out this process of artificially prolonging one's life completely removes the natural aspect of something so natural as death.⁴ But there are plenty of things we artificially alter for our convenience, things such as child birth, the way we bathe ourselves, or even the way we eat and gather food. Just because something isn't natural, does that mean humanity should avoid it?

² Garcia, *Form and Object*

³ Garcia, *Form and Object*

⁴ Garcia, *Form and Object*

Larry Temkin deeply explores this idea of artificially prolonging life and the ethics behind doing so in his paper “Is Living Longer Better?” In this paper Temkin first discusses the aspect of fear and its correlation with death. He states “As my son Daniel powerfully put it when he was three, ‘I don't want to lie, underground, in a box, forever.’”⁵ Temkin makes it clear that this is a very common and natural feeling regarding death, though he states that it isn't necessarily that people are scared of dying but rather they simply don't want to stop living as it is something they enjoy. Though this point is understandable and to most of us relatable, it brings the question to what end? In other words, how far is too far for humanity when it comes to artificially prolonging one's life? Temkin addresses this question further in his paper when he introduces a notion called the “subjective motivational set.”⁶ This concept suggests that if we lead a life which we know has no end we will begin to lose all ambitions in life. Temkin contends that our present projects are our conditions of life, in other words we are motivated to move forward because of our awareness of an end.⁷ Though regardless of technological advancements there must always be an end right? How could we allow one to live forever beyond temporarily prolonging their life? The answer to this question lies in the world of artificial intelligence and uploading.

In William Cheshire's “The Sum of All Thoughts” Cheshire grapples with the idea of permanently extending one's life through the use of technology. Cheshire first states that the human brain is fundamentally a computer. This leads him into the idea of uploading one's mind to a computer in order to attain a form of immortality. Cheshire claims that if the brain is as compatible and similar to computers as we think it is, “In principle, it might be possible to translate neural signals into electronic currents and, by joining dendrites with nanofibers, to

⁵ Temkin, “Is Living Longer Living Better?”, 2008

⁶ Temkin, 2008

⁷ Temkin, 2008

connect neural networks with silicon chips.”⁸ If this theory is true then this would result in all of our personal memories, our morals, our loves and fears to all be rearranged into a silicon chip inside of a computer. This would ultimately mean that the reality in which we live currently would virtually be indistinguishable from the reality we would experience in a computer. In fact this form of life would be so synonymous with our current one that Cheshire states "Substrate is morally irrelevant, assuming it doesn't affect functionality or consciousness. It doesn't matter, from a moral point of view, whether somebody runs on silicon or biological neurons (just as it doesn't matter whether you have dark or pale skin).”⁹ This statement truly emphasizes the way life through a silicone chip would be virtually identical to life in our world currently. So if one could live forever in a computer should they?

To first answer this question I feel it is critical to understand what this new technology would do to humanity collectively, an example of how this technology could potentially affect a society is found in *The Matrix*. *The Matrix* is a science fiction movie that follows the story of a computer programmer named Thomas Anderson. Anderson secretly leads a double life as a hacker named Neo. Later in the movie Neo gets pulled into a rebellion attempt against a race of artificially intelligent machines that have enslaved humanity by trapping them in a simulated reality called the Matrix. Neo teams up with a group of rebels and discovers that he is the prophesied "One" meaning that he is the one with the ability to defeat the machines and free humanity from their control. However, moving away from the plot of the movie to focus more on what this means in relation to AI, it is evident that despite the startling technological advances that are occurring now more than ever, the *Matrix* suggests that the route of AI and uploading could be a dangerous, future ending path to follow. Throughout the film there are multiple

⁸ Cheshire, “The Sum of All Thoughts”, 2015

⁹ Cheshire, 2015

warning signs about the potential dangers of AI and the way in which it could potentially corrupt society in the future.

One scene found in the *Matrix* that serves as a warning sign about AI is Morpheus's description of the origin of AI in this fictional society. Morpheus explains how AI was originally created in order to serve humans. It was developed to enhance our overall lives but unfortunately we developed it to a level that we could no longer control. Once AI was intelligent enough to develop an understanding of its own power a war between humans and machines broke out. The humans ultimately lost this war against the machines, and as a result the human species was enslaved and forced to live in a simulated reality in order to power the machines with the energy they harnessed. Instantly, the *Matrix* gives the audience reason to be wary of the development of AI. This scene emphasizes the way in which humans have a tendency to develop and create things that we don't fully understand. It's incredibly applicable to the current climate we reside in today as AI is now officially here and integrating itself into our lives more and more everyday.

Another scene from the *Matrix* that can serve as a potential warning sign can be found in Agent Smith's diagnosis of the human species. Agent Smith is an agent that worked for the *Matrix* meaning he is on the AI's side. Agent Smith stated that humans are equivalent to a virus that is destroying the planet. For this reason Agent Smith feels that the machines are far superior to humans. Though this could be viewed similarly to Morpheus's warning sign I feel that this warning is a little more cynical as it represents the AI's ability to process emotion. It is clear that Agent Smith has developed a deep hatred towards humans. The movie is showing that AI could potentially have the ability to feel emotion, and if AI can ever achieve this, humanity wouldn't be able to monitor these emotions. This of course could lead to AI developing negative emotions towards humans and starting a conflict. Additionally to the eeriness and endless amount of

danger found in the concept of an AI being able to process and perceive emotions is the way in which Agent Smith is justifying the war and enslavement of humans. This reason indicates more than just the capacity to feel emotions but rather the capability to have some sort of moral code.

The scene in *The Matrix* that best challenges this question of uploading oneself to a machine can be found when the character Cypher decides to betray his friends. In the movie the form of human life we are all familiar with is simply a simulation designed by artificial intelligence in order to harvest energy from humans. For every human in the matrix there is a real physical human, however this physical human is unconscious in some sort of preservation pod that makes them feel as if they are not where they really are and are in fact artificially experiencing “reality.” The movie follows a group of characters that have escaped these preservation pods and discovered that the reality of human life is now a post-apocalyptic wasteland ruled by artificial intelligence. Upon discovering this the group decides that they must intervene and put a stop to the destruction of humanity. Artificial intelligence is of course threatened by this group, so when the group goes in and out of the matrix the artificial intelligence has programs that follow them and appear as agents. This leads us to the scene that captures the way in which a realistic life possible through technology could potentially compromise the majority of humanity. In the movie these agents that have full control over the matrix (the simulated world) meet with Cypher at a fancy dinner restaurant. The agents are asking Cypher to sacrifice his friend, Morpheus, so that the agents can obtain access codes to the Zion mainframe which would result in the termination of this resistance group. Cypher is willing to do this if he can be put back into the matrix but in addition also have his memory wiped of him ever discovering the truth about life in the matrix. Though Cypher dies in his attempt to capture Morpheus this scene nevertheless represents some of the morals that humans would be

willing to put aside if it meant that they could achieve their dream life. This is why this scenario represented in the matrix if anything only puts the ethics behind the notion of merging life with artificial intelligence further into question.

After watching The Matrix I discovered that while it displayed many unpleasant aspects of merging life with technology, it also brings about a fine line that separates technology ruling humanity and instead humanity becoming enhanced by technology. Though I explained how Cyphers actions displayed the way in which humanity could be corrupted by this form of technology, it is necessary to explore this fine line that not only The Matrix brings about, but that the concept of artificially prolonging one's life brings about as well. A great question that The Matrix prompted, is if one could upload their consciousness to an alternate reality would they still be able to communicate with real people? Or would every interaction be simulations of people close to you? If we adopt the logic shown in The Matrix the answer to the first question would be yes. Though the reality you would experience would be artificial, the people in the simulation would be as real as you, but what would really be the difference between the two?

If I decide to upload myself to my dream reality, but 20 years prior my mother uploaded herself to her dream reality, odds are our two realities wouldn't look the same, even though both our ideal realities would include each other. A solution to this potential issue would be to have a copy of my mothers consciousness uploaded into my reality so that although this version of my mother would be a simulated version of her it would still be her in my eyes. This notion relates back to Cypher and his request to have his memory altered. Though my mother would be the exact same, just knowing that she was a mere copy of herself would taint the experience of having her in my reality. This logic leads me to believe that in order for my mother to exist in my reality in the way she exists in my life now, my memory would have to be altered to believe that

my real mother is in my reality. This method creates more holes in this concept that have to be patched. In order for me to exist in my dream reality would I want to know that I had died? Similar to Cypher, knowledge of my past life could potentially make my existence meaningless in a reality into which I had uploaded my consciousness. It seems as though this would make memory altering a necessity in order to upload one's consciousness after death. By these terms if I were to have no recollection of dying then who's to say I really did? Similar to the example I used with my mother, if I don't know she's a copy then why does it matter if she is?

All the ideas I have examined lead me to the conclusion that the notion of uploading one's consciousness into a computer is a dangerous one. The concept of uploading yourself to something where your reality is malleable creates room for many of these kinds of questions that challenge the importance of truth. I would argue that truth is something that is inherently important to us as humans. This is why when watching *The Matrix*, we see Cypher as clearly in the wrong; because of the fact that he is okay with living his life in ignorance. This is why the idea of living with a copy of my mother while she remains in a different reality, or living in artificial reality without knowing I had died feels inherently wrong. While the idea of extending your life indefinitely at face value might seem appetizing, it is clear that this technology would serve to be unproductive to humanity.

Bibliography

- Bhattacharya, Aveek, and Robert Mark Simpson. "Life in Overabundance: Agar on Life-Extension and the Fear of Death." *Ethical Theory and Moral Practice* Vol. 17, No. 2 (April 2014), 223-236. <https://www.jstor.org/stable/24478573>.
- Breitbart, William M.D. "On the inevitability of death." *Palliative and Supportive Care*, Vol. 15, No. 3 (June, 2017), 276-278.
- Cheshire, William P.J.R., M.D. "THE SUM OF ALL THOUGHTS: PROSPECTS OF UPLOADING THE MIND TO A COMPUTER." *Ethics & Medicine* 31, no. 3 (Fall, 2015): 135-141,131.
<https://login.proxy048.nclive.org/login?url=https://www.proquest.com/scholarly-journals/sum-all-thoughts-prospects-uploading-mind/docview/1719448659/se-2>.
- Garcia, Tristan. "Death." *Form and Object: A Treatise on Things*. Edinburgh University Press.
<http://www.jstor.com/stable/10.3366/j.ctt1g0b74h.26>.
- Godwin-Jones, Robert. "In a World of SMART Technology, Why Learn Another Language?" *Journal of Educational Technology & Society*, Vol. 22, No. 2 (April, 2019), 4-13.
<https://www.jstor.org/stable/10.2307/26819613>.
- Schneider, Susan. "The Age of AI." *Artificial You: AI and the Future of Your Mind*. Princeton University Press. <https://www.jstor.org/stable/j.ctvfjd00r.4>.
- Temkin, Larry S. "Is Living Longer Living Better?" *Journal of Applied Philosophy: Special Issue: The Ethics of Enhancement*, Vol. 25, No. 3 (August, 2008), 193-210.
<https://www.jstor.org/stable/24355033>.