

Name	<i>Chaitanya Baranwal</i>
Module number	UWC2101S / Section 1
Assignment Submitted	<i>Paper 3</i>
Purpose	<i>Final Submission</i>
Date of submission	23rd April, 2019
Number of words	2829

“The rogue AI”: Dystopian narratives and their role in shaping perceptions

The field of Artificial Intelligence – popularly known as AI – has been gaining increasing traction in recent times, among spheres both academic and non-academic. According to the AI Index Report 2018 (Shoham et al. 70), the number of research papers on AI jumped from about 11,000 in 2007 to about 24,000 a decade later. Similar trends are shown by Google Trends, which tracks the number of times the term “Artificial Intelligence” was entered on the Google search bar. The interest over time bumped from an index of 29 in January 2008 to 79 in January 2018. The attention that artificial intelligence has been receiving is not unfounded though, with breakthroughs like Siri and face-recognition algorithms having become a part of our daily lives. Moreover, the potential that AI has in advancing society is apparent from its increased focus in research. These groundbreaking advances in AI, however, have also led to concerns about it turning “rogue”: the possibility that artificial intelligence and robots will soon become “smarter” than humans and establish dominance over the human civilization. Such fears largely stem from the hypothetical scenarios where AI acquires a superiority over humans in terms of intelligence, eventually taking over the human civilization. There have thus been contrasting opinions about the future of AI: one believing in AI’s potential of helping the human society develop further, the other believing in AI’s potential of establishing dominance over the human race.

While this development of contrasting opinions is interesting enough, what is even more interesting is that the general public – which engages with AI only as an end-user – errs

more towards the dystopian versions of an AI-driven future as compared to technology experts from academia and the industry. Studies show that 72% (Smith and Anderson) of adult Americans are worried about a future where computers can do many human jobs, in contrast with experts involved in the technology scene, where only 37% (Anderson et al.) seemed to believe that advancing AI will make our lives worse by 2030. In his book *Superintelligence: Paths, Dangers, Strategies*, Nick Bostrom argues that emergence of 'superintelligence' – any intellect that greatly exceeds the cognitive performance of humans in all areas – is a likely scenario given the rapid advances in AI, and will lead to “existential catastrophe” (115). Worth noting is that the general public tends to agree more with Bostrom's views than people established in the technology industry. According to another survey, 37% of the general public had positive expectations and 62% had negative expectations about superintelligence, compared to 51% and 50% of technology executives respectively (Edelman 10). Why is there such a difference in the perceptions of AI between the scientific community and the general population? If a large proportion of the expert community does not believe in such perceptions, it is unlikely for them to have a rigorous scientific basis. Can these fears be attributed to other sources, and if yes, how would such sources reinforce what is otherwise a scientifically-unfounded idea?

Looking at fiction involving AI – both written and film media – can help us answer these questions, because fiction enables us to simulate the narratives not grounded in real-life events. Dystopian scenarios about AI are not based on real-world events, which makes it reasonable to assume that they are largely reflected in and exacerbated by popular fiction. According to the Royal Society of London, there is a “disconnect between prevalent narratives [about AI] and the state of the science”, which can “affect public confidence and perceptions” (Cave et al. 4). Since the assumption that public perceptions about AI are exacerbated by popular fiction is sufficiently reasonable, this essay instead focuses on the ways through which popular literature reinforces dystopian narratives as “facts” in the audience's mind. It argues that the representation of AI and robots in popular literature mirrors the social anxieties of the time period, which serves as an important way in affirming dystopian perceptions of AI. It further analyses two pieces of text, both written at substantially different periods of time. The first text is *The Terminator*, released in 1984, whose plot involving an evil cyborg assassin also highlights Cold War anxieties. The second is the TV show *Westworld* (2016-18), which apart from focusing on a futuristic robot-driven

amusement park, also focuses on the concepts of women empowerment and individualism, which are receiving greater attention in a world increasingly moving away from conservative principles. The texts picked are from different periods of time to ensure that the social anxieties in question are significantly different – public priorities were surely very different between the Cold War period and the current times. In addition, as we will further see, the portrayal of AI in *The Terminator* and *Westworld* is also significantly different. This strengthens the belief that movie portrayals model the social anxieties of the time they are released in, and ensures that we do not take a generalizable idea about AI-fiction and simply map it to the texts in question.

The Encoding and Decoding Model of Communication

By depicting social anxieties in AI portrayals, popular fiction constructs associations in the audience's mind, where the idea of AI also brings with it the accompanying anxieties. This can be better understood through the essay "Encoding/Decoding" written by Stuart Hall. According to Hall, communication cannot be thought of as a linear model of "sender/message/receiver" (91). It is instead much more complex, where the message "encoded" by the producer can be "decoded" in ways the producer did not intend. The decodings "issues into the structure of social practices" (Hall 93), and have to be mapped into "dominant and preferred meanings... to 'make sense'" (Hall 98). What Hall is saying is that when a new concept like AI is introduced, it needs to be mapped to something the audience already knows, which in our case would be the widespread social anxieties of the time. The "environmental world invades the linguistic and semantic system", meaning the real-world fears become correlated with how literature communicates the notion of AI (Hall 98). The audience thus understands robots and AI through the lens of widespread social fears. The subsequent sections will further argue for this claim, where *The Terminator* constructs the ideas of a "killer robot" and AI-dominance over civilization by highlighting Cold War anxieties, and *Westworld* constructs the idea of AI-harbored rebellion against humans through the liberal ideas of women empowerment and individualism.

The Terminator and the Cold War

The Terminator, released in 1984, focuses on the experiences of Sarah Connor as she and Reese (a human soldier from the future) escape a futuristic time-travelling cyborg assassin (called the Terminator) sent to kill Sarah. The Terminator wants to wipe out the existence of Sarah's unborn son, who in the future leads humanity's last struggle against machines in a post-apocalyptic society where computers have become self-sentient. As mentioned by David Seed in his book *American Science Fiction and the Cold War*, however, there is "fine responsiveness of fiction and film to a whole range of... changes taking place during the Cold War" (11), and this responsiveness is certainly seen in the movie.

The very first Cold War anxiety that stands out in the movie is that of Soviet infiltration in America, relating the notion of a robot to an "infiltrator" which does not belong to human society. The entrance of the Terminator into the present shows him appearing out of nowhere, showing him as an outsider breaking his way into society. His very first interaction with humans is when he talks to the biker gang, repeating what they say to get a hold of the English language, clearly indicating his unfamiliarity with the language of the Americans. This unfamiliarity, coupled with the German accent of the actor, further establishes this identity of an outsider. As a result, the audience decodes the idea of a robot as something alien and not belonging to normal human society. An infiltrator's malicious intentions of 'enter, destroy and leave' are juxtaposed with the malicious intentions of the killer robot, who enters past with the only focused mission of killing Sarah. The audience thus does not unpack the idea of a robot as something which is technologically unrealistic, but through the lens of infiltration into human society to wreak havoc, a problem which was very real and imminent in the 1980s. This further reinforces the idea that AI and robots pose a serious threat to the stability of human society.

Another key anxiety of the 1980s the movie mentions is that of a nuclear holocaust, making the audience especially dismal of an AI-led future. When Sarah asks Reese about what results in the destruction of society in *The Terminator*, Reese says that "there was a nuclear war" (Cameron), thus describing a full-scale holocaust and attacking the Reagan administration's nuclear brinkmanship (Seed 6). Reese's reminiscences of the post-apocalyptic future show humanity in a terrible situation, where people are hiding in

bunkers from the evil machines, evoking in the audience emotions of living in fear during the contingencies of a nuclear war. Imagining the specifics of an AI-driven future is a difficult task, but the movie provides the audience a relatable version of it. The audience thus maps the idea of AI-led society to an anxiety a significant number of them faced and understood, correlating an AI-driven future to the dystopian visions of a nuclear holocaust. This strengthens the audience's fears that trusting machine intelligence is a bad decision to take, since they would see "people as a threat" and "exterminate" them (Cameron).

In *The Terminator*, Sarah is the embodiment of American society – who goes to work as a waitress in the morning and parties on Friday nights. *The Terminator*, on the other hand, is a real threat to Sarah and the American society in all the ways discussed above. This juxtaposes the portrayal of AI and robots with the fears of the Cold War, thereby accentuating the fears related to AI and making them more relatable to the common public. Analogous to Hall's model, the concept of robots and AI was quite new to the audience, who took the movie's portrayal of widespread social anxieties as a lens to unpack these concepts. What resulted was the idea of a "killer robot", which although culturally pervasive, "is not a fair representation of robots in the real world" (Lafrance).

***Westworld* and the Rise of Modern Liberalism**

This theme of robots turning evil and setting out on a murderous rage to kill humans is contrasted in *Westworld*, which is a TV show released in 2016 and is based on a 1973 film of the same name. *Westworld*'s story takes place in a futuristic amusement park (called Westworld) populated by robots (also known as "hosts"). The park caters to high-paying "guests" so that they can live their wildest fantasies in a no-consequence environment, where the hosts are programmed to be unable to harm the guests in any manner. As we will see, the portrayal of AI in *Westworld* is different from that in *The Terminator*. It especially focuses on modern liberal ideas like women empowerment and individualism, which are more relevant now than they were in the Cold War, making the portrayal of 'AI consciousness' and 'AI-led rebellion' very relatable to the contemporary audience.

Westworld brings out the theme of women empowerment, which has been gaining increased focus through movements like #MeToo and is the belief that women deserve an equal

stature in society as men. The narrative *Westworld* employs is opposite to the one we see in *The Terminator*. Instead of a muscular male robot being the center of attention, the TV series highlights women through its choices of female protagonists for major story arcs of the show – be it Dolores Abernathy or Maeve Millay, both of whom are hosts in the park. Moreover, instead of robots oppressing humans, the story starts with humans being the oppressors and female hosts being oppressed through widespread murder, rape, and violence. This evokes the audience's sympathy for the robots, especially in times where more people are realising the equal stature of women in society. The audience starts treating robots like humans, thereby setting a precursor for 'AI sentience'. The feminist narrative flows smoothly into the second season, where the same female characters who previously suffered oppression take up arms against humans. What is interesting here is that since a precursor had been set where the audience sympathises with the robots, this "bloody campaign" appears justified. The audience starts viewing the robots as an oppressed social group desiring equal rights, mapping human qualities like sentience and the desires of moral rights onto robots. The show thus makes the AI portrayals more relatable to a modern audience through the idea of social empowerment, who find themselves agreeing that an "AI-led rebellion" is a justified and a very possible idea.

In addition to feminism, *Westworld* also says a lot about the notion of individualism, which is the belief that an individual can make his/her own choices, without external pressures from social groups and institutions. Individualism has seen a greater expression in practices around the globe, according to a research titled "Global Increases in Individualism" (Santos et al.). Since the idea of self-autonomy holds great weight in the modern audience's mind, the depiction of robots making their own decisions in *Westworld* maps onto the ideas of individualism, and the audience find themselves agreeing with this depiction of "self-conscious" robots. In contrast to *The Terminator's* robots where the entirety of AI has the common objective of exterminating humans, each host in *Westworld* is a master of their own choices. While Maeve's primary intentions are ensuring the safety of her family and getting out of *Westworld*, Dolores's prime motivations are killing the guests, and eventually destroying humanity. A central theme of *Westworld*, thus, is observing the interplay between the different paths that different hosts choose to take. The theme of robots showing traits of individualism further blurs the lines between a human and a robot and makes the audience wonder what differentiates the two. In such a

situation, the audience finds themselves relating more to the emotions and sufferings of the hosts, and how the possibility of sentient robots gets embedded with what their beliefs on equality and moral rights are. As a result, the idea of robots gaining consciousness and harboring an AI-led rebellion seems all the more possible and even justified, similar to what we saw in the previous paragraph.

Literature's Influence on Society

By juxtaposing AI with popular fears of the time, movies add a sense of familiarity to the way they portray AI. This happens because such narratives appeal to the subconscious concerns and fears of the audience, and the portrayals of AI gradually become integrated with these subconscious fears. This leads to a sort of cycle, as mentioned by Milton C. Albrecht in his essay "The Relationship of Literature and Society", where "if literature reflects, then it also confirms and strengthens cultural norms, attitudes, and beliefs" (431). It cannot be denied that for the literary requirements of a movie, AI needs to be dramatized, and simply going by real-world scenarios and developments would result in plotlines lacking any thrill or interest. However, these dramatised narratives create a feedback loop where the audience introduced to such narratives are not educated with the realities of AI technology. Instead, they are fed more of these dystopian narratives, further solidifying such unrealistic perceptions. As mentioned before, appealing to the popular social anxieties makes the narratives of such dystopian movies more relatable to the audience, who find themselves agreeing with the narrative. When someone from the audience would encounter the idea of AI, they would make sense of it not in terms of how the state of AI technology in the real world is, but in terms of how AI was portrayed in such movies. Literature shaping popular opinion is a reason why certain regimes resort to censorship, because literature disseminates views and opinions which might be antithetical to the regime's needs.

Conclusion

This essay examined the dystopian narratives surrounding artificial intelligence and robots in the movie *The Terminator* and the TV show *Westworld* and analysed how such narratives are closely linked to other social anxieties of the time period. That such dystopian

portrayals of AI dominance and AI-led rebellion are linked to other social anxieties reaffirms these exaggerated portrayals in the audience's mind, which can be problematic because it is through popular media that an unfamiliar audience engages with new knowledge on AI. A disconnect of such portrayals from technological basis can lead to exaggerated fears and misinformed debates, which can seriously affect AI research and development. This essay was limited to the analysis of only two movies due to spatial constraints, but it would be interesting to see how the portrayals of different time periods model the corresponding social anxieties, and what potential effects this can have on the current perceptions of artificial intelligence and robots.

Works Cited

Albrecht, Milton C. "The Relationship of Literature and Society." *American Journal of Sociology*, vol. 59, no. 5, 1954, pp. 425-436, *JSTOR*, www.jstor.org/stable/2772244.

Anderson, Janna, et al. "Artificial Intelligence and the Future of Humans." *Internet and Technology*, Pew Research Center, 10 December 2018, <https://www.pewinternet.org/2018/12/10/artificial-intelligence-and-the-future-of-humans/>.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford UK, 2014.

Cameron, James, and Gale A. Hurd. *The Terminator*. Los Angeles: Hemdale, 1984.

Cave, S., et al. *Portrayals and perceptions of AI and why they matter*. The Royal Society, London, November 2018.

Edelman AI Center of Expertise. *AI Survey Report 2019*. New York (NY), 2019. Web.

Hall, Stuart. "Encoding, Decoding". In *The Cultural Studies Reader* edited by Simon During, 90-103. Routledge, London, 1999.

Lafrance, Adrienne. "What is a Robot?" *The Atlantic*, 22, March 2016, <https://www.theatlantic.com/technology/archive/2016/03/what-is-a-human/473166/>.

Santos, Henri C., et al. "Global Increases in Individualism." *Psychological Science*, vol. 28, no. 9, 2017, pp. 1228–1239, doi:10.1177/0956797617700622.

Seed, David. *American Science Fiction and the Cold War*. Edinburgh, 1999.

Shoham, Yoav, et al. *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University, Stanford, CA: December, 2018. Web.

Smith, Aaron and Monica Anderson. "Automation in Everyday Life." *Internet and Technology*, Pew Research Center, 4 October 2017, <https://www.pewinternet.org/2017/10/04/automation-in-everyday-life/>.