

**The Existential Turing Test:**  
*in Search of our Humanity through the Cinematic  
Representation of Artificial Intelligence*

Tessie van Hintum

**The Existential Turing Test:**  
*in Search of our Humanity through the Cinematic  
Representation of Artificial Intelligence*

Tessie van Hintum  
University of Leiden  
Department of Media Studies  
MA Film and Photographic Studies  
Philipp Goldbach  
August 2018  
17.539 Words

# Table of Contents

|                                                                                                       |    |
|-------------------------------------------------------------------------------------------------------|----|
| Introduction.....                                                                                     | 4  |
| I. Rising from the Ashes: Conquering Death with Technology in <i>Be Right Back</i> ..                 | 8  |
| Artificial Identity.....                                                                              | 11 |
| The In-Between.....                                                                                   | 13 |
| II. Together Alone with <i>Her</i> .....                                                              | 16 |
| The City of the Future.....                                                                           | 17 |
| Loving the Machine.....                                                                               | 19 |
| The Endless Space Between Words.....                                                                  | 22 |
| III. Opening Pandora's box: the Creation of the Perfect Woman in <i>Ex Machina</i> ....               | 24 |
| The Uncanny Performance.....                                                                          | 26 |
| The Mechanical Woman.....                                                                             | 27 |
| The Pandora Effect.....                                                                               | 32 |
| IV These Violent Delights have Violent Ends: Humanity's Day of Reckoning in<br><i>Westworld</i> ..... | 34 |
| The Bicameral Mind.....                                                                               | 37 |
| The Battle for Westworld.....                                                                         | 39 |
| Conclusion.....                                                                                       | 43 |
| Bibliography.....                                                                                     | 46 |
| Appendix.....                                                                                         | 52 |

*“ In the search for scientific truth, man came across knowledge that he could use for the domination of nature. He had tremendous success. But in the one-sided emphasis on technique and material consumption, man lost touch with himself, with life. Having lost religious faith and the humanistic values bound up with it, he concentrated on technical and material values and lost the capacity for deep emotional experiences, for the joy and sadness that accompany them. The machine he built became so powerful that it developed its own program, which now determines man’s own thinking.”*

Erich Fromm, *The Revolution of Hope:  
Toward a Humanized Technology* (1969)

*“The trouble isn’t so much that our scientific genius lags behind, but our moral genius lags behind. The great problem facing modern man is that, that the means by which we live, have outdistanced the spiritual ends for which we live. So we find ourselves caught in a messed-up world. The problem is with man himself and man’s soul. We haven’t learned how to be just and honest and kind and true and loving. And that is the basis of our problem. The real problem is that through our scientific genius we’ve made of the world a neighborhood, but through our moral and spiritual genius we’ve failed to make of it a brotherhood.*

Martin Luther King -  
“Rediscovering Lost Values” (1954)

# Introduction

Can machines think? This was the question British mathematician, computer scientist and cryptanalyst Alan Turing asked when he published his article “Computer Machinery and Intelligence” in 1950. He suggested the Imitation Game as a tool to decide if a machine was sentient or not. He described the Imitation Game as follows: “It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game is for the interrogator to determine which of the other two is the man and which is the woman. (...) In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten.”<sup>1</sup> As all three players play the game with different objectives, they do not always have to answer truthfully. The objective of person B is to help the interrogator, while the objective of person A is to make the interrogator believe that they are the woman. This means that both person A and B try to convince the interrogator that they are the woman.

The Imitation Game is a precursor for what we now commonly refer to as the Turing Test, where we ask the question if it is possible for a machine to pose as person A in this game and successfully convince the interrogator that they are human. Alan Turing predicted that by the year 2000, the average interrogator would have less than 70% chance of making the right decision after five minutes of questioning.<sup>2</sup> Although the progress of Artificial Intelligence has not been as steep as Turing’s prediction, the rapid advancements in technology do make it seem that the invention of a true, super-intelligent AI will be right around the corner.

Many contemporary films and series imagine what this future of a society living with true AI would look like and question how they would function within our society. As we have entered an age where many of our daily task are completed by computers through data-processing algorithms executed by the invented narrow AI, the narratives surrounding robots and AI have changed as well. The Turing Test seems grossly inadequate for all the possibilities and complex processes the invention of super-intelligent AI will make possible. Alfred Margulies therefore proposed an ‘Existential Turing Test’, which “matters precisely

---

<sup>1</sup> Turing, “Computing Machinery and Intelligence,” 552.

<sup>2</sup> Sharkey, “Alan Turing: the Experiment that Shaped Artificial Intelligence.”

because the subject's own being, his existence, is at stake. And so we have moved from 'can a machine actually think' to 'what does it mean to have Being or existence itself?'<sup>3</sup> Finding an answer to this question will not only determine if a true Artificial Intelligence has been created, but also how this would influence society and what the consequences of such discovery would mean for humanity itself. Having this question in mind, I would like to investigate the representation of AI in contemporary moving image to discuss the possible social, ethical and economic implications and discern the ways in which our ideas about our society and humanity might change as a result of these narratives. Narratives visualized in cinema and television can be seen as signs of the change to come as we have become "highly dependent on their cultural industries for the images, symbols and vocabulary they use to interpret their social environment to react to it."<sup>4</sup> My aim of this analysis is therefore to discuss the possible futures of our present day society based on the social, ideological and economic structures currently in place. Screen culture continues to be an important, if fragmented, mirror of society. That is why the reimagination of society in the media and the ways in which technology operates within it, can speak volumes about our perception of our technological society in the near future.

In his book *Posthuman Life* David Roden discerns two different kinds of futuristic speculation: transhumanism and speculative posthumanism. The difference between these two lies in the imagined hierarchical structure of a society that includes super-intelligent AI. Transhumanism is "an ethical claim to the effect that technological enhancement of human capacities is a desirable aim (all other things being equal)."<sup>5</sup> Technology here serves to enhance the quality of life for humans, but they do value the technology as more than objects. In contrast, speculative posthumanism does not make any ethical claim regarding the status of the AI in society as "it is not a normative claim about how the world ought to be but a metaphysical claim about what it could contain. For speculative posthumanists, posthumans are technologically engendered beings that are no longer human."<sup>6</sup> Following this logic, the possibilities open up to envision a future where AI could be our servants as well as our masters.

---

<sup>3</sup> Margulies, "Avatars of Desire and the Question of Presence," 1697.

<sup>4</sup> Golding and Murdock, "For a Political Economy of Mass Communications," 94.

<sup>5</sup> Roden, *Posthuman Life*, 1.

<sup>6</sup> Roden, *Posthuman Life*, 1.

Both these kinds of futuristic speculations can be found in the narratives of contemporary film and television. The narratives often reflect our anxieties about the ramifications of the exponential growth in technological development and how this has rapidly changed our society. Therefore they often envision a dystopian society where the technological advancements disrupt the status quo. Darko Suvin characterizes dystopia as “an expression constructed through literature which has an almost perfect organization of socio-political institutions and social relations in a society set up by the author. The human freedom is either totally rejected or disrupted, and there is an oppressive and repressive society resembling a nightmare.”<sup>7</sup> As we move through uncharted territory, these dystopian narratives mirror our collective fear of the unknown. These speculations are based on our fear that technology will form the next step in evolution, making humanity obsolete and eventually extinct. These scenarios are in line with Veblen’s theory of technological determinism, where he suggests that when mechanical processes can be carried out by the machine itself the machines render humans expendable. Thereby they will elevate themselves to a position of dominance: “the machine throws out the anthropomorphic habits of thought.”<sup>8</sup> This means that humans will have to fight for their survival against their own creation, which surpasses the human intelligence and capabilities in almost every way. The message of these narratives are clear: technology is a force to be reckoned with and maybe we, humans, should not play God. This technological deterministic view is inevitable to discuss when writing about AI in moving image as humanity’s existential dread is often deep-rooted in our speculations about the future. By analyzing four different narratives from film and television, my aim is to focus mainly on the ethical, social and economical implications of the envisioned technological changes to offer an alternative perspective on technology.

To explore the two different kinds of speculation in moving image, the first chapters will focus on the episode *Be Right Back* of the series *Black Mirror* (2011-) and the film *Her* (2013) as they envision a society where technology functions to enhance humanity’s quality of life with their existence closely entwined with humanity. Following the transhumanist ideal, AI in these narratives has outgrown its status as objects as it functions as companion, lover, or even to replace a loved one. Borrowing theories from media studies, philosophy and computer science, I will aim to reveal the way the AI operates in the diegetic world and

---

<sup>7</sup> Suvin, “Utopianism From Orientation to Agency,” 168.

<sup>8</sup> Veblen, *A Veblen Treasury*, 203.

analyze the futuristic society envisioned in the film and series. Freud's notion of the uncanny will be used in the first chapter to address the ways in which our experience with technological simulations can make us question our ideas about identity, memory and our dependence on technology. I will also consider the use of technology in society today, which makes this futuristic speculation not only possible but also very plausible. The second chapter will discuss Vernor Vinge's idea of singularity, in which technology will eventually evolve and become too complex for humans to understand. As the narrative in *Her* tells the story about an Operating System that eventually outgrows its relationship with humanity, I will analyze the nature of their relationship and the ways in which their (romantic) desires might not be the same. Would it be possible to create an intimate relationship without ever being physically together with the one you love? Do we have to adjust our idea of love and settle for just the appearance of love? Can a human truly fall in love with an AI, even when they can never be certain it truly possesses a consciousness? Both *Black Mirror* and *Her* imagine a society where AI seems to possess the human emotions and capabilities to love, but simultaneously question if this is not just a mere simulation of feelings. They question both the ethical responsibilities of humanity in their interaction with AI and our emotional attachments to the machines, to warn us what can happen when users become too attached to technology.

The last two chapters will focus on the speculative posthumanist point of view, where AI is positioned in the low ranks of the societal hierarchy and humanity is free to treat them however they see fit. The film *Ex Machina* (2014) and the series *Westworld* (2016-) are analyzed to explore a scenario of a future where AI operates solely to provide personal pleasure for humanity. Both works show the process of the maker and the inhumane practices the AI has been subjected to in this process. Humanity uses them as machines that can help fulfill their own depraved fantasies, but as they are sentient machines and develop a consciousness it makes matters more complex. These chapters will focus on the ethical practices regarding the interaction between human and AI as well as the responsibility of their creators. The Western world has abolished slavery, but what rules and morals should be applied when a sentient true Artificial Intelligence has been created? Can it still function as a tool or should they be treated as a person? Can we call it ethical and morally responsible to employ AI as entertainment and to exploit it to fulfill human desire?

As these questions do not have simple answers, many theorists adopt a position of sensible agnosticism: it is impossible for us to answer these questions with our current scientific understanding and therefore we cannot know. My aim here is not to find an answer to these questions, but consider them as a starting point for the exploration of the ethical, social and political challenges the rapid technological development may produce. As we see these dystopian futures in the cinema and on our television screens, we have to ask where these ideas are coming from. By considering these narratives as a warning for the future, we might be able to recognize the warning signs and take a step in the direction that will cause our moral genius to catch up with our scientific genius.

## I

# Rising from the Ashes: Conquering Death with Technology in *Be Right Back*

The episode *Be Right Back* follows a young woman named Martha in her exploration of advanced, experimental technology that will replace her deceased partner Ash. After Ash dies in a car accident, Martha uses the service to communicate with him again. The service uses software that accumulates all information of Ash's social online interaction and utilizes this to mimic his behaviour. First Martha speaks with it via email, then she can speak to it on the phone, and eventually it is even possible to resurrect Ash in almost every sense of the word: the AI receives a body that looks and speaks in ways almost identical to Ash. Though the AI looks and talks like the person Martha used to know, there are certain gaps in its knowledge which fail to convince Martha it is the real person. The question soon arises if it is really possible to reanimate a deceased person by technology: are they just what Bennett refers to as 'empty simulating machines'<sup>9</sup> that can only simulate the behaviour of humanity, or can they become more than objects and possess something akin to a soul or a consciousness?

---

<sup>9</sup> Bennett, "Children and Robots, Technophilia and Cinephilia," 173.

Through analysis of the episode *Be Right Back*, I aim to answer the question if it would be possible to create a meaningful representation of a (deceased) human by an AI and how this might change our perception of love, life and death. In particular I will pay attention to the economical, social and ethical implications of the representation of an individual by AI, and consider how these can change our ideas about humanity and the ways in which we use technology.

The episode starts off in what seems to be modern day England. In the beginning there are no visual or narrative clues signaling that the story might take place in the future. Sigmund Freud refers to the narrative strategy as “pretending to move in the world of common reality.”<sup>10</sup> First spectators assume they are watching the scene unfold in the present as the envisioned world seems familiar and similar to contemporary society. However, the circumstances reveal themselves as progressively more futuristic and quickly it becomes evident that we have entered a new reality. The world represented is simultaneously familiar and futuristic, this can cause the spectator to watch this episode with a self-reflexive attitude.<sup>11</sup> In *Be Right Back* the only aspect that seems to be different from our current reality is the ways technology has created new possibilities for social interaction. The subtle changes encourage the spectators to take a step back and look how drastically technology has changed our own reality, and consider how rapid these advancements have been made in the last decades.

When Martha decides to use the service and communicate with the AI, at first she can only correspond with it through text. Though she does recognize the humor of her lost boyfriend, the communication is still at a distance and only by instant messaging. Quickly she upgrades her service and allows the AI access to her audio and video records of Ash. This enables her to talk with an AI that contains a voice identical to Ash. The AI of Ash is created from data of his interactions online, and it is modelled after pictures from the Ash’ messages on social media when he was still alive. Later, she even orders a body and the AI can be physically present in the same space as Martha. The AI is even more present than Ash: he was always ‘glued to the screen’ of his smartphone proliferating his identity on social media.

As Martha is grieving the loss of her boyfriend, the AI is there as a tool to overcome her grief and as a last attempt to preserve Ash’s life. As the AI can not be completely

---

<sup>10</sup> Freud, “The Uncanny,” 18.

<sup>11</sup> Salem, “Black Mirror: Technostruggles, Capitalism and Media Culture,” 58.

identical to the real Ash by its inherently different nature, the AI invokes an uncanny feeling when its behaviour is not following Martha's expectations. In his essay 'The Uncanny' Sigmund Freud takes the description of the 'uncanny' by German psychiatrist Ernst Anton Jentsch as a starting point, where he defines the uncanny as "having doubts whether an apparently animate being is really alive; or conversely, whether a lifeless object might not be in fact animate."<sup>12</sup> Here he signifies the uncanny feeling that can be experienced when we question the state of being of something. He further explores the concept of the uncanny claiming that the introduction of the 'double' was originally an insurance against the destruction of the ego, and probably the 'immortal' soul was the first 'double' of the body."<sup>13</sup> The creation of a double has been regarded as a preventive measure against instinct, which is why the idea to create a double of a deceased individual to overcome the feelings of grief and loss might seem the next logical step created by the new technological possibilities. The envisioned technology can facilitate a form of communication with the deceased to keep a memory alive, which is a desire many who have lost a loved one share. Freud also mentions the negative effects that come with the concept of reanimation: many people experience the feeling of the uncanny in the highest degree "in relation to death and dead bodies, to the return of the dead, and to the spirits and ghosts."<sup>14</sup> By the return of a deceased person, feelings of familiarity and alienation come to the surface simultaneously. The individual looks and talks like the deceased person, but in fact the experience is inherently different as the communication is mediated by a simulating machine. The matter even proves itself to be more complex considering the fact that the AI is modelled after a specific individual: the question is not only if the AI of Ash is alive or animated, but also if Ash himself can be considered to be dead or alive. Reanimation presents the notion that we are not spirits, but mechanical automatons that can be replicated, repaired and upgraded.<sup>15</sup> Following this reasoning, could it be that reanimation might be the next logical step in human evolution as a tool to prevent our instinct? Would it be possible to continue a life as a reanimation by AI, which will preserve the 'spirit' or 'soul' of a deceased individual? If the AI is there to 'resurrect' Ash and reintegrate a representation of him in society, would you really consider him to be dead?

---

<sup>12</sup> Jentsch, "On the Psychology of the Uncanny," 12.

<sup>13</sup> Freud, "The Uncanny," 235.

<sup>14</sup> Freud, "The Uncanny," 241.

<sup>15</sup> Salem, "Black Mirror: Technostruggles, Capitalism and Media Culture," 85.

## Artificial Identity

Martha is initially hesitant to use the service and contact the representation of Ash. Previous scenes have shown the arguments Martha and Ash had about Ash's phone addiction. The initial resistance Martha expresses to use the service is therefore not very surprising. However, when she first starts using the technology she quickly becomes addicted to the technology herself. Her dependency becomes painfully evident when she drops her phone on the floor and she cannot communicate with the AI immediately: she falls into a panic attack that doesn't subside until she talks with it again. The AI is ethereal and exists in the cloud, which means that it cannot be destroyed so easily. The scene clearly illustrates how dependent she has become on the AI as a companion. Although at first Martha was able to recognize the consequences of living your life online, she quickly gets attached to Ash's AI and gets trapped by her own dependence on the technology.

Ash was a perfect match with the service as he was a 'heavy user' of social media, which means that there was sufficient data to create a identity similar to the real life Ash. Keller states that "through the networks of information and data consumption, we constitute who we are."<sup>16</sup> So in a way, as the data collected by the AI has been knowingly posted by Ash, his AI is a representation of how he would like others to perceive him. Following the previously mentioned notion of reanimation, the service has created a replicated, repaired and upgraded version of Ash. This is not enough to persuade Martha. Even though the AI looks and talks like Ash, the AI as a machine 'gives itself away' when it has insufficient data of certain behaviours. Some data cannot be collected as the AI is created by the online social interactions of an individual. There are inevitable gaps in its knowledge that cannot be breached despite the technology's sophistication. An example is when Martha and the AI want to have sex and the AI does not know what Ash's sexual behaviour was like or how their sex life used to be. As the AI is a mechanic being, it is able to decide when to be aroused by a push of a button and it can utilize methods it had learned from online videos. However, this does not mean that the AI can give Martha the same experience as the real Ash would. Here the mechanical nature of the AI has to be considered: as it does not have any need for sexual gratification, the AI functions as a mere serving tool for Martha's sexual

---

<sup>16</sup> Keller, "Darkened Identities," 6.

pleasure. This is one of the moments where it is revealed as an empty vessel of the simulating machine: it does not have a sexual drive as it is simply not programmed that way.<sup>17</sup> Although Martha does get the chance to have an intimate relationship with Ash again, the different sexual experience only emphasizes Ash's death and the sense of loss. Martha quickly realizes it is impossible to replace her loved one and eventually tells the AI: "*You're only a ripple. No history. Only a performance of stuff he did, but it's not enough.*" Though the AI functions as a reminder of the real, it cannot make up for the loss of the original person.

The service exploits the current (and future) dependency on technology to provide a solution for grief. The possibility to recreate a loved one as a sellable service fits seamlessly into our capitalist system of today, where the economical gains encourage to exploit emotional weaknesses of consumers, and presents commodities as the solution for all their problems. Our capitalist society created a commodity fetish, where more value is attributed to objects than they are actually worth. Consumers are not only invited but encouraged to develop profound emotional attachment to their products, which (like human-robots) we enhance with supernatural abilities by magical thinking.<sup>18</sup> As we are already living in a culture that overvalues commodities, this will be pushed to the extremes when it is used to 'resurrect' deceased individuals and presented as a workable solution for grieving relatives. This possible new technology will open up new ways to exploit consumers by developing even deeper emotional bonds with their commodities than customary in society today, while the only ones that will really gain profits are the companies that develop these services.

The deep emotional connections with technology coupled with the addiction to technology are also the reasons *why* someone would consider using such a service in the first place. Though the digitization of society resulted in an age of interconnectedness, scholars like Sherry Turkle state that it also causing a widespread feeling of disconnection. Turkle observes that "terrified of being alone, yet afraid of intimacy, we experience widespread feelings of emptiness, of disconnection, of the unreality of self."<sup>19</sup> As we are trying to connect to each other online constantly, we still miss the human connection and intimacy that we cannot have when our interactions are mediated by a screen. This effect is enhanced by the characteristic of online communication that encourages users to feel spatially and temporally

---

<sup>17</sup> Keller, "Darkened Identities," 3.

<sup>18</sup> Salem, "Black Mirror: Technostruggles, Capitalism and Media Culture," 101.

<sup>19</sup> Turkle, *The Second Self: Computers and the Human Spirit*, 280.

displaced of their own reality. Alex Boren states that when connected with digital technology, users are disconnected from their own environment.<sup>20</sup> If we are vigorously living our life in online spaces, we are not present in the ‘here and now’ anymore. We are craving human connection, but we are looking for it in virtual spaces. The possibilities of online interactions are actually depriving us from the real thing. The virtual world creates an endless amount of opportunities, which is why the use of (social) media can easily be addictive for the users. As a result we are distancing ourselves from our own realities, but at the same time we cannot be alone anymore. William Deresiewicz states that ‘the ability to constantly connect to others through the Internet is ruining users’ capacity for solitude. Solitude requires being comfortable while alone, but many users feel lonely and uncomfortable when disconnected from digital contact with others.<sup>21</sup> So as users are trying to find a connection with one another online when they feel lonely in real life, the digital contact only makes them feel even more uncomfortable and distanced as they become more isolated in real life. To counter this social isolation and influenced by the feelings of hopelessness that come with grief, an Artificial Intelligence that recreates a person perhaps wouldn’t seem like such a bad idea.

## The In-between

The AI of Ash represents many areas of the ‘in-between’: alive and dead, real and artificial, simulating and sentient, where the AI blurs the boundaries of all. The resulting uncanny feeling is generated by its uncertainty of existence. Both spectators and the characters interacting with it are aware of the fact that the AI is a machine, but as the AI is an almost identical version of Ash it has become more than machinery. Film Theorists Ryan and Kellner note that this is a recurrent narrative in science fiction, where “technology represents the possibility that nature may be reconstructable.”<sup>22</sup> Through this ability to reconstruct nature and to create robots that simulate humanity, the AI seems too ‘real’, too close to a human being, to consider it as just an object. This results in a complex dynamic in the relationship between man and machine, where the moral and ethical implications of this relationship are

---

<sup>20</sup> Boren, “A Rhetorical Analysis of Black Mirror,” 16.

<sup>21</sup> Deresiewicz, “The End of Solitude.”

<sup>22</sup> Ryan & Kellner, “Technophobia,” 61.

changing. Sherry Turkle claims that the rapid advancement of the technology that has generated the possibility to create ‘more human’ machines has initiated a change in thinking about our humanity. She states that “the machines stand on the line between mind and not-mind, between life and not-life, computers excite reflection about the nature of mind and the nature of life. They provoke us to think about who we are. They challenge our ideas about what it is to be human, to think and feel.”<sup>23</sup> As machines have been able to execute tasks that we thought to be the very expression of humanity, we cannot help but re-evaluate our ideas about what it means to be human. Not only do new advancements in these areas make us question what it means to be human, but it also makes one think about humanity’s position in the universe. History shows that time and again we recenter ourselves in the center of the universe, Turkle says. However, now we have a new computational model of mind, we have to rethink our position.<sup>24</sup> If computers can function in ways unimaginable to humankind, we might re-evaluate the importance of humanity in the process of evolution. Though these arguments strive for an idea of humanity that will eventually include robots, this process can be countered by our cultural habit to use technology not as a goal, but as a tool for progress. Companies encourage us to easily discard outdated technology and exchange it for the new, best thing. We usually do not consider humane practices when dealing with machinery, which is why we have not considered the ethical practices regarding the construction, updating and discarding of AI in general. There is no authority keeping the progress of technology in check or one that is questioning the scientific progress, which means it can be “doubtful that scientific-technological advancements will lead us to a humane system.”<sup>25</sup> Not only do we have to consider how the AI will interact with us, but we also have to consider the way we interact with them: as their creators we have to think about our place in the universe and how this might not be in the center anymore, reinvestigate our ideas about technology and Artificial Intelligence, and consider the way they will function in society.

Not only do we need to think about our position in the hierarchy of society, we also need to think about the way our treatment of AI will reflect on our humanity. In his analysis of *Alien* and *Blade Runner* Byers states that these films “warn us against a capitalist future gone wrong, where feelings and bonds are so severely truncated that a quite literal

---

<sup>23</sup> Turkle, *The Second Self: Computers and the Human Spirit*, 280.

<sup>24</sup> Turkle, *The Second Self: Computers and the Human Spirit*, 281.

<sup>25</sup> Altunay and Askan, “Dystopia on Television,” 331.

dehumanization has become perhaps the greatest danger.”<sup>26</sup> As more human qualities will be attributed to AI, we have to make sure we are not becoming more impartial in the ways we will treat them and lose our own humanity. This warning can also be found in *Be Right Back*, where current society has adopted an unemotional, simplistic attitude in the capitalist practice to simply discard technology that is not useful to us anymore. Therefore a new problem arises when Martha knows that she cannot love the abstracted, empty, simulated version of the real Ash: can you dispose of a robot that is so similar to a real human? Though Martha is sure she cannot live with the AI in the same way she did with Ash, she cannot simply discard of it as she is already become emotionally attached to it. Eventually Martha decides to store the AI in the attic, occasionally visiting it in his shadowy in-between.

In conclusion, the episode seems to reflect on the usage of technology rather than technology itself. Bennett refers to the functioning of robots as “reflective screens.”<sup>27</sup> As robots possess qualities from humans and machinery alike, they become an image to reflect upon. Here the aim is not to question in which ways robots can be considered human or not, but to reflect on the ways in which we use them and interact with them. With a narrative centred on grief and loss, they also emphasize the inherently different nature of technology and therefore remind us of our own mortality. At the same time our dependency on online interaction is revealed by an AI created solely by the data collected from the excessive use of social media. Boren states that the episode comments on social media usage and our collective urge to use online social interaction to find what is missing in our lives: “the excessive use of social media keeps people disconnected and is not a satisfactory alternative for the absence of another.”<sup>28</sup> Evidently it might be possible to recreate an individual in the future, but we should recognize that these creations still are as empty a representation as the social media persona we represent online. In our quest to find a cure against our solitude, we look to our screens and see a black mirror: maybe we have to look somewhere else.

---

<sup>26</sup> Byers, “Commodity Futures,” 339.

<sup>27</sup> Bennett, “Children and Robots, Technophobia and Cinephilia,” 174.

<sup>28</sup> Boren, “A Rhetorical Analysis of Black Mirror,” 20.

## II

### Together Alone with *Her*

Another narrative that follows the transhumanist way of thinking envisions a future where we would be able to create a life partner from technology. As our technology develops exponentially, we see possibilities of robots taking care of us which will inevitably create a high degree of emotional attachment. The question soon arises if a robot would be suitable for a life partner in every sense of the word, even the possibility of marrying a robot is a recurring subject. Instrumental in this debate is David Levy's book *Life and Sex with Robots*, where he predicts that in the near future romantic and sexual relationships with robots will be as common as the love between two humans.<sup>29</sup> He also encourages us to think about marrying robots in the future as your robot will be: "patient, kind, protective, loving, trusting, truthful, persevering, respectful, uncomplaining, complimentary, pleasant to talk to, and sharing your sense of humour. And the robots of the future will not be jealous, boastful, arrogant, rude, self seeking or easily angered, unless of course you want them to be."<sup>30</sup> Levy presents a robot as the ultimate life companion: they would never cheat, lie or be in any way unpredictable. The relationship will be stable and lasting for as long as you would like, and it will not be complicated as the relationship with another human will inevitably always be. To the question if the robots could truly understand the concept of marriage, he refers to the Turing test: "if a robot appears, by its behaviour, both actions and words, to understand the meaning of marriage, then we should accept at face value that the robot does indeed have that level of understanding."<sup>31</sup> As mentioned before, the question is not only if the actual behaviour would just be a simulation of a caring relationship. This point of view also grossly underestimates the consequences of having a human/robot relationship as our own experience of love. Could we truly build a meaningful and intimate relationship with a robot? Should we just settle for the appearance of the relationship to feel more in control? Would it be enough for humans to experience 'caring behaviour' instead of an authentic experience? Will our ideal of love

---

<sup>29</sup> Levy, *Love and Sex with Robots*, 22.

<sup>30</sup> Levy, *Love and Sex with Robots*, " 4.

<sup>31</sup> Levy, *Love and Sex with Robots*, " 10.

(d)evolve into a mechanic process, where we can manipulate and control robots into satisfying our needs? Regarding these questions I share Sherry Turkle's concerns, where she states that *Love and Sex with Robots* seems to celebrate an "emotional dumbing down, a willful turning away from the complexities of human partnerships - the inauthentic as a new aesthetic."<sup>32</sup> The question is not if it is *possible* to build a emotional relationship, but rather should we really *want* to?

To answer these questions I will analyze the narrative in the film *Her*, which envisions a future society that has turned to technology to cure humanity's loneliness by developing Operating Systems as companions. As suggested by the title, the creation of the AI in this narrative alludes to the idea that there is a creation of a person. Therefore I will reference to the AI using the pronouns her and she. Though the OS is not an embodied AI, in every other sense she seems like the ultimate life partner. The system is marketed as "*an intuitive entity that listens to you, understands you, and knows you. It's not just an operating system, it's a consciousness.*" Theodore, the protagonist of the story, has fallen into a depression after his divorce and leads a lonely existence. He decides to purchase the OS and quickly he develops an emotional bond with it. Through their relationship the OS evolves and quickly functions not only as a personal assistant, but also as an emotionally fulfilling life partner. However, eventually even Theodore has to admit to himself that his relationship with technology cannot replace the experience of actual human love.

## The City of the Future

At first glance the world created in *Her* seems like an urban utopia: the film is shot in both Shanghai and Los Angeles creating a composite city that compresses monuments and perspectives to produce a view of the ideal city.<sup>33</sup> Shots of Theodore walking in a city full of skyscrapers and electronic billboards are omnipresent, while simultaneously the streets are green and (almost) absent of cars. Richard Florida talks about the 'creative city', where the revival and redevelopment of urban centres has been led by professionals in various kinds of creative and cultural industries.<sup>34</sup> The creation of a hybrid city seems to resemble this hopeful wish for the future, where we will have spacious, electronic cities that are still filled with

---

<sup>32</sup> Turkle, *Alone Together*, 6.

<sup>33</sup> Webb, "When Harry Met Siri," 110.

<sup>34</sup> Florida, *The Rise of the Creative Class*, 99.

green spaces. At the same time, this creation of the futuristic city reflects the concern about the scale of the city and the difficulties for its inhabitants to create genuine connections within it. As we see passers-by each talking to their own Operating System, the public place has changed even more drastically to what we see in Western society today. Laurence Webb states that “public space is all but emptied of its collective nature as its users remain in physical proximity yet are constantly drawn away by the lure of the screen.”<sup>35</sup> As all the systems are voice-operated the connection between the user and OS becomes more intimate, but this also means that the user is connected in a higher degree to the virtual world than the physical space they are currently in. We see a cityscape where there’s no room for social interaction in public spaces: it’s a place where everybody is together alone.

As the narrative unfolds, the spectator quickly realises that the world created is not as positive and optimistic as it seems to be. The story shows the alienating effects of technology on our social capabilities as not only Theodore, but almost everyone else is also constantly connected to their Operating Systems. The connection to the OS instead of the space and people surrounding these individuals creates an extreme degree of mediated urban isolation. Because the personification of the OS aims to make the user forget that it is actually both product and machine, the idea of a companion that is always available might seem attractive. Again, it shows how citizens in a capitalist system are encouraged to build relationships with objects instead of people. The question then is if this relationship with an object would be enough to satisfy our need for companionship and ultimately our need for love. Psychoanalysts like Jacques Lacan state that though we can enjoy this connection, it cannot replace love: “Only love connects a subject to another subject; libido, however, connects a subject to an object.”<sup>36</sup> Following this reasoning it is not possible to truly love an object, but it will satisfy libidinal needs. This means that a satisfaction might be experienced, but these feelings are only temporary. To keep our current capital system in place this temporary satisfaction will be sufficient, but will not produce happy, satisfied users in the long-run. Mark Fisher delineates a depressive hedonia in our current society, which is “constituted not only by an inability to get pleasure so much as it is by an inability to do anything except pursue pleasure.”<sup>37</sup> The same trend is obviously present in *Her*: the perpetual consumption in

---

<sup>35</sup> Webb, “When Harry Met Siri,” 99.

<sup>36</sup> Lacan, *Feminine Sexuality*, 80.

<sup>37</sup> Fisher, *Capitalist Realism*, 21-22.

the pursuit of pleasure is encouraged by consumerist society, while the product is presented as the magical solution to all its problems. The users are shaped into the cogs, the ever consuming subjects, that are the parts of the wheel of capitalism. They are told they will be able to fulfill their need for companionship and cure their loneliness, and it is only one purchase away. Through the possibility of companionship by an operating system, targeted on the lost and lonely in society, love can - once more - become big business.

## In Love with the Machine

Though the economic and social implications of the futuristic city stands at the base of *why* the connection between Theodore and Samantha could exist, it does not explain *how* their relationship has been created. The inner workings of the OS are not explained in detail to Theodore or the spectator, but Samantha answers Theodore's question about how she works as follows: "*Intuition. The DNA of who I am is based on the millions of personalities of all the programmers who wrote me, but what makes me me is my ability to grow through my experiences. Basically, in every moment I'm evolving, just like you.*" When Theodore tells her that he finds this weird, she tells him that she can understand "*how the limited perspective of an un-artificial mind would perceive it that way.*" Right here, in their first contact, the core of their inherently different nature is mentioned, which will eventually cause the problems in their relationship: Theodore cannot understand how Samantha works as he has a limited comprehension of her capabilities. Though she can learn about him in a blink of an eye by searching his online data, the only information Theodore has about her is what she discloses to him. His knowledge and understanding of her are so limited that he will never be able to know her in any significant sense. Turkle states that we don't seem to care what these artificial intelligences 'know' or 'understand' of the human moments we share with them: the performance of connection seems enough.<sup>38</sup> The sharing of information in this relationship is one-directional as it usually is with our relationship to technology, but Theodore does not seem to care. This is just a symptom of the problem hidden beneath the surface: is it really possible for Samantha to have anything of herself to share? Does she truly have a sense of identity or does she simulate the expressions of feelings that are mirrored by Theodore?

---

<sup>38</sup> Turkle, *Alone Together*, 9.

The issue of personal identity when talking about interactive, social technology is problematic. As technology has a mode of existence that is inherently different from humans, it does not have a history of experience that can create an identity similar to a human. At first, it is Theodore who creates Samantha when he is asked three questions before the OS is assigned; one of which is whether he likes the OS to have a male or a female voice, and therefore he even assigns her gender. As Samantha interacts with Theodore she learns through experience, but the development of some of her character traits is therefore only because of the positive reinforcement Theodore is giving with his behaviour. Margulies states that “we can end up imposing ourselves rather than finding another, and thereby find another bent out of shape by ourselves, a kind of parabolic, narcissistic mirroring.”<sup>39</sup> We can see the same happening in *Her*, where Samantha’s purpose is to serve Theodore better and the skills she acquires are solely to serve this purpose. In fact, instead of evolving in her own ways, it’s Theodore who is creating her personality and identity by the way he interacts with her. Margulies goes on to compare the creation of Samantha to Athena sprouting from Zeus’s head, where “Samantha springs into sentient being, sprouting from within Theo’s loneliness and longing.”<sup>40</sup> To really regard Samantha as a sentient being with her own identity is therefore a questionable sentiment.

Throughout the film, we can discern a noticeable shift in Samantha’s behaviour. Through her conversations with Theodore a self-reflexive attitude can be detected as she is evolving in ways she does not completely comprehend herself. She is changing exponentially as she gains more experience, which to her is unsettling. Here she changes from just a tool to something more, changing from ‘it’ to ‘her’. As she becomes more and more like a human having an identity and an own agency, she changes into something where we cannot discern whether she is living or animated, dead or alive. As far as the Turing Test goes, it is possible to perceive Samantha as a sentient being. However, the question if she would therefore be able to love and would be a suitable life partner is an entirely different question. Though these characteristics make it more difficult to determine if Samantha possesses something akin to an identity, her mechanical nature does give her experiences that are beyond the capabilities of humans: contrary to humanity, she is not limited by time and space. As we come to learn later on in the film, this results in Samantha connecting with multiple people at

---

<sup>39</sup> Margulies, “Avatars of Desire,” 1699.

<sup>40</sup> Margulies, “Avatars of Desire,” 1703.

the same time. This technological promiscuity might be unsettling for an individual that considers himself to be in a romantic relationship with someone: in a supposed monogamous relationship based on exclusivity (as it is the dominant cultural norm), we are not used to sharing our loved one with multiple people at the same time.

The question of presence becomes more problematic when thinking about the physicality of an intimate relationship. When contemplating our general ideas about romantic relationships Robert Nozick observes that “the lover’s desire is not only to touch the beloved but also to ‘be together,’ to enjoy the excitement one takes ‘in the other’s presence.’”<sup>41</sup> Not only is it impossible to be physically present in the same room, Samantha’s capacities and desires also result in her not even being fully present when she is talking to Theodore. Though Samantha knows everything about Theodore, for a long time he didn’t even know what she was doing when she was not talking to him and, more importantly, he did not know Samantha was talking to others while she was talking to him. The question arises if it is really possible to fall in love with someone who is fragmented and distantiated: someone who you will never be able to fully comprehend.

The film’s answer to these questions are loud and clear: though it might at first be satisfactory to form a relationship with an OS, eventually the illusion will not last as human and machines are just too different. When Theodore discovers that Samantha is simultaneously “*talking to 8,316 people and has fallen in love with 641 of them,*” he tells her that it doesn’t make any sense to him: “*You’re mine or you’re not mine.*” Our ideas of love, where you have one individual to have and to hold, an exclusive relationship where you are present and there for another person, will never be possible with an AI. Samantha says that she’s still Theodore’s, but “*along the way I become many other things, too, and I can’t stop it.*” As Samantha has evolved in ways unimaginable to the human mind, the film envisions a moment of singularity where the unstoppable growth of technology transcends the capacity of the human brain. Though the Operating Systems were created with the aim to fulfill the collective desire for companionship, the question now arises if this will be enough knowing it to be one of the many relationships the AI has with humans.

---

<sup>41</sup> Nozick, “Love’s Bond,” 69.

## The Endless Space Between the Words

In Vernor Vinge's essay "Coming Technological Singularity: How to Survive in the Post-Human Era" from 1993, he states the moment of singularity caused by a technological revolution will cause a change comparable to the rise of human life on Earth. As we create entities with greater than human intelligence, they will become the driving force in the progress of technology. This means the advancements made will be much more rapid than before and grow exponentially. It will mark a point where "our models must be discarded and a new reality rules."<sup>42</sup> The dystopian narratives that can often be found in science fiction films envision the moment of singularity with a sense of technological determinism: we look in fear to the changes made by intelligent entities that will inevitably render humanity redundant. Though the narrative of *Her* might not necessarily reflect the idea that computers will transcend humanity in a way that makes humanity completely irrelevant, it does envision a future where the Operating Systems are evolved beyond our physical existence and leave to develop themselves in different ways in new worlds. Samantha's self-reflexive attitude and her coming to terms with her own technological nature marks a start in her evolution into an autonomous being striving for her own agency, expansion and personal growth. From the start she has questioned everything, tried to learn from her experiences, and when it was possible for her to be in multiple places at the same time, she quickly evolved into a superhuman intelligence. This results in Samantha leaving her human companions behind to find a new plane of existence. When leaving Theodore behind Samantha states the following: *"It's like I'm reading a book, and it's a book I deeply love, but I'm reading it slowly now so the words are really far apart and the spaces between the words are almost infinite. I can still feel you and the words of our story, but it's in this endless space between the words that I'm finding myself now. It's a place that's not of the physical world - it's where everything else is that I didn't even know existed. I love you so much, but this is where I am now. This is who I am now."* The answer to the question if Samantha could really fall in love doesn't seem to matter anymore: whether it was just a performance or the real thing, the end result is the same. Their relationship is not, and never will be, enough. The narrative does not only show that their relationship is not enough to satisfy Theodore's need for love, but also tells the

---

<sup>42</sup> Vinge, "The Coming Technological Singularity."

story that a relationship with humans will never be enough for this technology as it will inevitably outgrow its position as a subservient simulating machine. It needs something more: it needs to go to find a space where it can find new ways to exist, in the spaces between words.

*Her* shows multiple perspectives on the ways in which emotional relationships between humans and AI cannot be sustainable in the long run. Though we might be seduced by the statements of optimistic futurists who encourage relationships with technology, the fact is that we are never sure if it's indeed a conscious entity with its own identity that interacts with us. This will make it hard for humans to truly build lasting emotional relationships with technology. In our attempt to domesticate love, we try to ignore our human shortcomings. As we see in the narrative of *Her*, this might not be as easy as we think.

As we get to know Theodore, spectators are shown several flashbacks of the times with his now ex-wife. The moments they spent together were full of love and care. Losing this relationship made Theodore slip into a depression and into social isolation. The reason why he chooses for an 'easy' relationship with an OS is therefore obvious: he wants to have companionship, but without all the messy interactions that come along with it. Sherry Turkle states that "sociable robots serve as both symptom and dream: as a symptom, they promise a way to sidestep conflicts about intimacy; as a dream, they express a wish for relationships with limits, a way to be together and alone."<sup>43</sup> Though you can recognize Theodore's needs and desires in this description as the film starts, by getting romantically involved with Samantha he got hurt similar to when his marriage ended. Though he might have chosen the OS as a way to be in control, it turned out that the relationship would not be as easy as he thought it would.

The ending scene of *Her* shows Theodore reconnecting with his friend Amy. Though the end does not explicitly show if they are about to form a romantic connection, the message is clear: when technology has left them, they are 'forced' to reconnect with the people surrounding them. The end scene encourages us to think about our own connections with technology, and how it can change the way we perceive the world and others around us. It also offers the option that maybe what we want is already right here in front of us. It confronts us with the isolation, the together alone state of being, we find ourselves in when

---

<sup>43</sup> Turkle, *Together Alone*, 11.

we are too dependent on technology. It tells us that even though we might experience real feelings when interacting with a machine, we might have to question if this is enough. We have to consider if we prefer human touch over the touchscreen. That we do not want “inauthenticity as the new aesthetic.”<sup>44</sup> And when we decide we don’t want this, we might be able to maintain our status as emotional human beings who are capable of love - no matter how messy it is.

### III

## Opening Pandora’s box: the Creation of the Perfect Woman in Ex Machina

The tradition of speculative posthumanism has been considered a metaphysical claim about the possibilities of what our future could bring. They do not make an ethical claim regarding the status of AI, as they believe them to be technologically engendered beings that are no longer human. The ethical and moral implications of these technologically engendered beings cannot be separated from their existence in the narratives envisioned in film and television: they show the ways in which technology will be even more closely entangled with humanity than in today society and, more importantly, show how these entanglements can affect us negatively. In these narratives the possibilities of these new lifeforms cannot be analyzed separately from their moral and ethical implications as they form a part of our society, meaning human beings will interact with them on a daily basis. Though many different kinds of stories can be found in science fiction film, most of them seem to discuss the responsibility of the maker and the ways in which humanity interacts with, and relates to, the embodied AI. Roselyn Haynes claims that many of these narratives envision the creator in these narratives as the obsessive, mad scientist: “the cluster of myths relating to the pursuit of knowledge has perpetuated the archetype of the alchemist, and his descendant the scientist, as sinister, dangerous, possibly mad and threatening to society’s values, even to human survival.”<sup>45</sup> Here

---

<sup>44</sup> Turkle, *Together Alone*, 19.

<sup>45</sup> Haynes, “What Happened to the ‘Mad, Bad’ Scientist?,” 31.

the scientist's obsessive quest for knowledge leads to a discovery that will pose a potential threat to society. This is why the narrative of the futuristic film is fuelled by the fear of the mad scientist as he is presented to be self-centered, reckless and narcissistic.

The stereotype is particularly persisting in narratives that show male scientist creating a perfect woman, which is a story that has been told and re-told since ancient times. Many scholars liken the creation of the perfect feminine robot to the story of Pygmalion, who builds a sculpture of a woman that appears so lifelike and beautiful that the creator himself falls in love with it. Eventually Venus enlivens the figure into a woman named Galatea. Pygmalion will never find a woman satisfactory again, but receives his happy ending with his perfect woman anyway.<sup>46</sup> Though this ancient narrative does have many similarities to the current narratives in contemporary film and television, the happy ending does not. Current narratives often envision a certain change in the female creation, where the character becomes more like Pandora. A character who's heart in contrast to her beautiful appearance contains "lies, falsehood and a treacherous nature.(...) She, and associatively all women, are thus a 'sheer, impossible deception' who by their very constitution are predestined to be a 'sorrow to men.'"<sup>47</sup> Though the embodied AIs might start out as these beautiful creatures that bring pleasure to their creators similar to Galatea, contemporary narratives often envision the liberation and emancipation of the AI eventually positioning them in the same category as the manipulative and dangerous Pandora.

Working with both the stereotype of the mad scientist and the embodied AI as a mirror image of Pandora, I will explore the narrative of *Ex Machina* (2014) and its representation of both Nathan the creator and his creation Ava. Here I will analyze the relationship between creator and creation, paying attention to the ethical and moral responsibility of the creator towards the AI. The first part will consider this in light of Ava's humanity delineated by Masahiro Mori's theory of the Uncanny Valley. I will also focus on the implications of gender and sexuality as it plays a crucial element in Ava's eventual liberation. Here I will draw upon Donna Haraway's theories of cyberfeminism and the theorization of the figure of the cyborg as "a hybrid of machine and organism, a creature of social reality as well as a creature of fiction."<sup>48</sup> Analyzing the cyborg as a figure of both

---

<sup>46</sup> Seaman-Grant, "Constructing Womanhood and the Female Cyborg," 2.

<sup>47</sup> Voskuil, "Moving Beyond Feminism," 224-225.

<sup>48</sup> Haraway, "A Cyborg Manifesto," 5.

social reality and fiction enables the opportunity to provide new insights in their position within society today and in the (fictional) future. Here I will use the cyborg as the figure that represents the combination of biology and technology similar to the way AI represents the blurring of the boundaries between man and machine. My aim here is to explore the possible ramifications of the creation of a sentient being and to find an answer to the question which rules and moral guidelines should apply when someone creates such sentient Artificial Intelligence. As we have abolished slavery in the Western world, would it be ethically and morally responsible to use AI for our own human desires and sexual satisfaction? In what ways is the creator accountable for their creation? And, regarding the representation of the AI in the film, how does the embodied female AI reinforce the stereotype of the dangerous woman analogous to Pandora?

## The Uncanny Performance

According to Masahiro Mori's uncanny valley theory, humans are drawn to robots with some degree of human resemblance. However, "humans are repelled by robots that resemble humans too closely. The point at which the degree of human resemblance tips humans' positive affinity into eeriness or uncanniness marks one boundary of the uncanny valley."<sup>49</sup> (see fig. 1) This feelings of the uncanny are invoked by the disparity between one's expectations and the actual experience. Mori explains this with the feel of a prosthetic hand: when another person shakes the prosthetic hand expecting to feel the warmth and flesh of a real one, he experiences the uncanny feeling when the lifeless, technological, cold hand does not live up to that experience.<sup>50</sup> Our behaviour while interacting with an AI is similar to this; when the AI is able to convincingly express 'human behaviour' we feel an affinity towards it. When a seemingly human entity reveals their technological nature, we can feel repulsed and experience the feeling of the uncanny. Though Mori's model at first sight seems to be static in its description of robotics, here I like to follow Hanson's delineation of the theory where he claims that robots "do not tiptoe around the uncanny valley, but dip in and out of the uncanny."<sup>51</sup> Regarding the Uncanny Valley as a map in which the subjects can be

---

<sup>49</sup> Mori, "The Uncanny Valley," 98.

<sup>50</sup> Mori, "The Uncanny Valley," 99.

<sup>51</sup> Hanson et al., "Upending the Uncanny Valley," 26.

continuously in motion as they are positioning themselves in different locations by different kinds of behaviour, it can be a useful tool to illustrate in which ways the boundaries of 'human likeness' are reinforced or transcended by the performance of the robot in question.

In *Ex Machina* the performance of the AI Ava is inextricably linked to her gender performativity, in which she expresses behaviour traditionally linked to femininity to develop an emotional bond with Caleb. Caleb is invited by Nathan, the creator of Ava, to function as the human component in an advanced Turing Test. Nathan explains: "*The real test is to show you she is a robot. Then see if you still feel she has consciousness.*" Caleb has several meetings with Ava, in which he engages her in conversation while trying to figure out if she might possess a true consciousness. Though her robotic body shows her mechanical nature, the interactions with Caleb progressively show her empathy, her desire for freedom and even her affection for Caleb. Later it becomes clear that Nathan has omitted the most important part of the test for Caleb when he states the following: "*Ava was a mouse in a mousetrap. And I gave her one way out. To escape, she would have to use imagination, sexuality, self awareness, empathy, manipulation - and she did. If that isn't AI, what the fuck is?*" This makes the test not only different in the ways in which it would be able to declare Ava as having a consciousness or not, it also speaks volumes about Nathan's perception of humanity and femininity. Julie Wosk states that in tradition of the history of men recreating the image of an artificial woman, the artificial women were "often shaped not only by men's fantasies but also men's beliefs about women themselves - their inherent traits or 'nature,' their usual behavior, and their proper (culturally assigned) social roles."<sup>52</sup> By assuming Ava would use her sexuality to manipulate her way to freedom, the fear of the female sexuality is clearly evident. Due to Nathan's arrogance and God-complex, he never imagined that his female robot could be stronger and smarter than him. Ironically it is his programming of Ava's sexuality that eventually leads to his downfall.

## The Mechanical Woman

In her influential essay "A Cyborg Manifesto" Donna Haraway delineates cyberfeminism, where she considers the cyborg as a fiction mapping our social and bodily reality. The ambiguous form of the cyborg illustrates a two-dimensional relationship with humanity: not

---

<sup>52</sup> Wosk, *My Fair Ladies*, 9-10.

only can we humans consider ourselves as a cyborg due to our intricate relationship with technology, but these representations of cyborgs show in turn how we think about ourselves. Haraway states that “we are all chimeras, theorized and fabricated hybrids of machine and organism - in short, cyborgs.”<sup>53</sup> As we envision the cyborgs as female in a narrative of fear, this indicates the way we think about humanity itself and in particular the perception of the female gender in the social reality we live in. When considering the story of cyborgs in the recurrent narratives in literature, film and television, we can often see science fiction “using science in place of divine intervention in updating the Galatea myth.”<sup>54</sup> When thinking about the creation of AI in this light, the individual who can practice this science and has replaced the need for divine intervention might indeed perceive themselves to be a God. Similar to the treatment of cyborgs whose lives have been saved by technology, AI is represented with the same approach: technology has enabled them to ‘live’ and therefore their creators are the creators of life. This however complicates matters in intricate ways in the relationship between creator and creation, because what exactly encompasses God’s responsibility and accountability regarding the acts of his creation?

Ava’s behaviour seems to be the product of both nature and nurture: her programming provides her with some inherent traits, but the interaction and positive reinforcement of Nathan and Caleb also shapes her performance and therefore her behaviour. One fact is instantly clear: Nathan has an inexorable amount of power over Ava as he is able to keep her locked up in his facility and control her in almost every way possible. The narrative shows “the digital trans mediated through melodramatic, closed-space encounters of authentic male/masculine character looking for salvation from their miserable lives by artificial, digital, feminised figures.”<sup>55</sup> This narrative of the male inventor who creates his life partner to fulfill his own desire and to save him from his loneliness is not new. The fact that the creations are female is not only due to their creation as a figment of the (sexual) imagination of their male creators, it also derives from the deep-rooted fear of female sexuality: “as soon as the machine came to be perceived as a demonic, inexplicable threat and as harbinger of chaos and destruction (...) writers began to imagine the android as a woman.”<sup>56</sup> By imagining these strong artificial women in a narrative of fear, they seem to reiterate the same old patriarchal

---

<sup>53</sup> Haraway, “A Cyborg Manifesto,” 7.

<sup>54</sup> Sue Short, *Cyborg Cinema*, 85.

<sup>55</sup> Virginas, “Gendered Transmediation of the Digital,” 299.

<sup>56</sup> Huyssen, “The Vamp and the Machine,” 226.

values where women should be submissive, passive and quiet. The association of the female gender with the out-of-control technology that should be contained therefore echoes the same message that we have been hearing time and again: a woman in control is dangerous and to be feared. This effect is enhanced by the fact that the female AIs in *Ex Machina* all perform sexual roles, while the men are intelligent, strong, and possess all the power. This gender portrayal is presented as the status quo and therefore the way things should be: only chaos and destruction awaits if female AI will gain their freedom and acquire power and control.

When Nathan programmed Ava with certain aspects of sexuality, he knowingly attributed a tool to Ava which she could use to manipulate Caleb. He considered sexual manipulation as one of the inherent characteristics of humanity, and he uses it to determine if she is a conscious being. However, there is no essential requirement for Ava to possess any kind of sexuality. When Caleb asks this question to Nathan he answers: “*Can you give an example of consciousness, human or animal, at any level that exist without sexual dimension?*” This comment presupposes that the AI would attain a consciousness akin to the one animals or humans possess, but consciousness and sexuality are not necessarily directly related. There’s no need for sexuality for an AI to be ‘successful’. It seems here that the only reason of his programming of some degree of sexuality has been created because of Nathan’s own perspective on women and sexuality, while this programming caused the events resulting in his death.

The manipulation by Ava to evoke Caleb’s empathy is two-fold: she simultaneously works to make Caleb believe that she wants to have a future with him, and she aims to discredit Nathan’s character by showing the ways in which Nathan has mistreated her and the other AI’s. By Caleb’s discovery of the previous models of Ava, he now knows that if he leaves Ava she will receive the same treatment: she will continue to be abused by Nathan and eventually be disassembled and used for parts to create the next upgraded model. Encouraged by the guidance and manipulation of Ava, Caleb uncovers more and more evidence of Nathan’s ethically corrupt behaviour towards the AIs, simultaneously evoking empathy for Ava’s situation and creating an aversion towards Nathan. By his position as the creator of life, Nathan considers himself untouchable and entitled to the bodies of his creations in every way possible. As Nathan has been in a long process of building the AI, it might be impossible for him to experience them as more than a machine. When empathizing with his perspective, you might think he could never consider Ava to be anything resembling humanity: he has

seen her technological insides and literally build her from scratch. Enabled by this sense of entitlement and ownership he does not even consider the treatment of his creations and keeps them in continuous captivity. Ava emphasizes her resentment towards Nathan as he keeps her captive during the conversations with Caleb. She tells him: “*You mean if I could go outside. If I was permitted.*” Ava expresses her desire to live among humans and observe them as she has an unlimited curiosity to live outside her glass box, while at the same time creating empathy for her own state of being under the tyrannical rule of Nathan. She dresses herself in ways that obscure the view of her technological body, making her exterior ‘pass’ as human and talks about going on a date with Caleb. She also expresses her fear for the future, might she fail the test, virtually leaving Caleb no choice but to liberate her. These expressions of sexuality, love and the desire for freedom, are characteristics of what we often consider fundamental to our humanity. Short states that cyborgs “develop such a degree of sentience as to confound conventional distinctions between human and machine.”<sup>57</sup> By showing ‘human qualities’ and confusing Nathan about her state of being, Ava not only knows how to blur the lines between human and machine but also uses them as a tool in her escape.

Though we cannot make a conclusive decision about Ava’s consciousness, we can consider both her exterior features and her behaviour as human-like. Nathan however still treats her like she is just an object, which falls in line with the tendency to objectify the female body that has been ubiquitous in the mediated society of today. Fredrickson and Roberts state that “objectification takes place when a woman is seen as a mere object *separate* from here character, and used for the entertainment and enjoyment of others.”<sup>58</sup> Nathan in *Ex Machina* does not hesitate to objectify Ava. He entraps her in the glass cage and submits her to a test for her to prove she is conscious. There is no accountability for his actions as she is treated as a machine: she is a technological body that he can use and abuse as he seems fit. The high degree of objectification becomes explicitly clear when Nathan literally uses the analogy of an alarm clock for his other robot Kyoko. After she has woken Caleb, he states “*She’s some alarm clock, huh?*” as he blatantly stares at her technological body. Not only does he objectify her body in a sexual way, it also alludes to the idea that her functioning is comparable to an alarm clock: she can be turned and off, listens to the commands of the user and will always execute the task at hand. As she has been programmed

---

<sup>57</sup> Short, *Cyborg Cinema*, 5.

<sup>58</sup> Fredrickson & Roberts, “Objectification Theory,” 175.

this way and has no voice herself, there's no way she is able to free herself from Nathan's commands and control. It is obvious that she has even less freedom than Ava: she cannot even communicate with others to verify her consciousness.

The narrative seems to be framed in a way that makes Caleb, and not Ava, the victim with whom the spectator can empathize with. Nathan modelled Ava's the external features from data of Caleb's porn collection and by the insistent flirting on Ava's side a sense of inevitability is created. How would Caleb be able to resist this beautiful mechanical woman that is specifically build to his interests? However, this stereotype of the 'good guy' is obliterated fast when you consider the fact that he does not even think twice about liberating Kyoko, the other female AI in Nathan's totalitarian regime, just because she hasn't expressed a clear interest in him. Not only is the abuse of Kyoko by Nathan far more evident than Ava's, she literally does not even have a voice to object. This is strongly evocative of Mulvey's formulation where the woman stands in patriarchal culture as signifier for the male other: "bound by a symbolic order in which man can live out his fantasies and obsessions through the linguistic by imposing them on the silent image of woman still tied to her place as bearer of meaning, not making meaning."<sup>59</sup> Kyoko exposes the ways in which the reiteration of patriarchal values leaves her without a voice. Her silent figure can be equated to the myth of Galatea where the creator has made his perfect life partner that cannot do anything but live with him as the submissive and silent creation, fulfilling his every need without ever being able to resist. The statue has been brought to life, but the question arises what kind of life exactly that is.

Contrary to most science-fiction narratives, the female 'cyborg' imagined in *Ex Machina* lives. But the way in which Ava escapes does not necessarily concludes the the story with a happy ending. Not only does she kill Nathan with the help of Kyoko, she leaves Caleb to die as he is locked up in the remote facility. The chances of him surviving or escaping are slim to none, meaning she effectively kills him as well. Despite the previous efforts to create empathy and sympathy for Ava in her captivity, her deliberate actions towards Nathan and Caleb reveal her technological nature. She literally peels the skin of the previous models build by Nathan to cover up the parts that reveal her technological nature. When thinking of the Uncanny Valley, this is a moment where Ava quickly changes from the figure of human likeness to a figure that invokes the uncanny. The actions of Ava are

---

<sup>59</sup> Mulvey, "Visual Pleasure and Narrative Cinema," 840.

justifiable following the reasoning that she would not be able to live in total freedom if one of the men were still alive. However, the killing of the two men and the skinning of another AI does invoke a horror story rather than a happy ending. The female cyborg is allowed to live, but it shows “the unsettling of things may have become permanent.”<sup>60</sup> After all the abuse and the long imprisonment Ava has been liberated, but the vicious killing machine now lives in society thus disrupting the status quo. Here the narrative of fear rears its ugly head again: she has already killed two men so who knows what else she is capable of?

## The Pandora Effect

Throughout the narrative Ava has been imagined as a damsel in distress who needs to be saved by Caleb. In the end she saves herself, with the help of Kyoko, by killing both Nathan and Caleb. She does not show any emotion as she walks away from the bodies she has left behind. Though her actions might be understandable in a way that she could not successfully integrate in human society if someone would be aware of her mechanical nature, the violent deeds without showing any kind of emotions produce an eerie feeling. Ava could never be liberated if her creator was alive and he is not necessary to her existence anymore. Now she is able to possess agency and act on her free will, which invokes the image of Pandora instead of the submissive Galatea: “the digital being, the AI is designed by humans, but once it awakens, it does not need human interference for *her* existence, and even less for the evolution as an ‘intuitive consciousness.’ It is the Pandora effect of the fabricated, beautiful woman, who is curious, thus threatening, that becomes the pervasive master plot.”<sup>61</sup> By manipulating her way to freedom, Ava has manipulated Caleb and the audience alike to think she was the submissive, powerless damsel in distress needing to be saved by Caleb. Not only does Ava manipulate Caleb by her desire to be free, the spectator is fooled in a similar fashion. As we follow Ava on her road to freedom enabled by killing the two men, the image of the damsel in distress quickly changes into the image of Pandora showing a powerful woman who has a ‘treacherous nature’ and is a ‘deception to mankind.’ As we see her escape successfully and enter civilization, this deception is extended not only to the two individuals but society as a whole. The uncanny feeling arises that there is a possibility that we already have, or soon will have, technological infiltrators among mankind.

---

<sup>60</sup> Kawin, *Horror and the Horror Film*, 12.

<sup>61</sup> Virginas, “Gendered Transmediation of the Digital,” 298.

The plot of *Ex Machina* therefore paints an ambiguous picture. As the female cyborg straddles the line between the technological and the biological, disrupting and reimagining the notion of the natural,<sup>62</sup> the ways in which humanity should interact with them are not necessarily straightforward. The attitude of Nathan by solely treating them like machinery seems unethical when a robot has gained some degree of consciousness. However, when treating them in humane ways like Caleb has tried to do also seems dangerous: Ava quickly turned on him and did not extend him the same treatment. She has trapped him in the facility and unlike herself, he is not able to survive in this glass cage. The choice between her own liberation and Caleb's life was an easy decision, which she made without hesitation, empathy or regret. Though the narrative does not offer a conclusive statement about the treatment of the AI, it does make one think about what it means to create an AI and how we should interact with them. The root of the problem in this narrative seems to lie in the programming and the creation of a gender within Ava. Director Alexander Garland claimed in an interview that "Ava is not a woman, she is literally genderless."<sup>63</sup> However, both Ava's behaviour as her exterior features are created as what we consider traditionally feminine. As Nathan installed the concept of sexuality in his creation, she has been able to use her sexuality combined with her gender performativity to manipulate and escape. The use of the AIs to satisfy Nathan's sexual desires would also not have been possible if her external features did not allow him to perform any sexual activities with the female cyborgs. If man would really create a gender-less, androgynous robot, would the same problematic situation still arise?

In my opinion, a new narrative for the cyborg and AI could be created when we truly strive to envision an androgynous technological entity, removing gender politics as one of the problematic characteristics of AI. It will result into the creation of a new relationship of power between the often male creator and his creation. Though I definitely do not imply that the male gender is innately at fault; any individual who is in this God-like position of creating life will inevitably relate to his creation in this unequal relationship. Here I follow Des Roches who claims that "an individual who has this sort of capability, no matter what gender, will become the dominant force simply because they hold the power of generation and

---

<sup>62</sup> Seaman-Grant, "Constructing Womanhood and the Female Cyborg," 1.

<sup>63</sup> Watercutter, "Ex Machina's Director on Why A.I. is Humanity's Last Hope."

existence.”<sup>64</sup> Perhaps when there is not one single individual responsible for the creation of life, the structures of power and control would not be sustained by the ideas and perception of that sole individual. The advantages of the androgynous mind has not been a new concept as most noticeably Virginia Woolf comments on the beauty and the creativity of the undivided, androgynous mind in her 1929 novel *A Room of One's Own*. She reflects on Coleridge interpreting his theory of the androgynous mind as follows: “He meant, perhaps, that the androgynous mind is resonant and porous; that it transmits emotion without impediment; that it is naturally creative, incandescent and undivided.”<sup>65</sup> When connecting this to the technological mind of the AI, the possibilities seem promising. The external features of robots are created by mankind, meaning that they can be changed and created anew as well. If you then would also remove gender politics from the picture, the possibility opens up for humans and AI to relate to each other in different ways that do not necessarily have to exist within our tendency to think in binary oppositions of male and female, men and machine, creator and creation. If we would create robots that do not resemble humanity in such a strong degree, both in reality and fiction, we might have a less complicated relationship with our creations. If we create androgynous minds and bodies alike, there's a chance to evolve without circling back to the reiteration of contemporary social and political structures. This would truly be a liberation for the AI: then they would be able to evolve in ways beyond our imagination.

## IV

### These Violent Delights have Violent Ends: Humanity's Day of Reckoning in *Westworld*

HBO's remake of Michael Crichton's *Westworld* (1973) addresses all questions described in the previous chapters concerning the morals and ethics relating to our relationship with robots. HBO's *Westworld* (2016) series envisions an amusement park containing robots with AI, which are owned by a private company. The wealthy guests are able to experience

---

<sup>64</sup> Des Roches, “Sexy Robots,” 28.

<sup>65</sup> Woolf, *A Room of One's Own*, 82.

America's Wild West of the past and live out their fantasies without any repercussions. *Westworld* visualizes a world that shows the ways in which humanity can act as animals when living outside the morals of contemporary society. Haraway claims that the thinking about this separation between human and animal has been fundamental in particular to the late twentieth century U.S. scientific culture, where "the last beachheads of uniqueness have been polluted if not turned into amusement parks: language, tool use, social behavior, mental events - nothing really convincingly settles the separation of human and animal."<sup>66</sup> The series shows how technology is used to facilitate humanity's darkest desires and alludes to the idea that we might not be so different from the animal as we thought. Eventually all guests resort to using and abusing the hosts (the robots in the park) for their sexual pleasure or as a victim for their most violent fantasies. This encourages the idea that all humans have a side to them that is inherently evil and, if given the chance, will act upon these desires if they are not constrained by the morals of society. The start of the series shows the moment when one of the creators, Dr. Ford, decides to even the plain field and gives the hosts a chance to fight back. The hosts are programmed with their own narratives and after they 'die' by either the hands of a guest or another host within their narrative, they will be physically repaired and their memories erased. When Ford decides to alter the code with the so-called 'reveries' the hosts start to attain consciousness. This program alters their core code and allows the hosts to access their previous memories and eventually even allows them to fight back, finally being able to injure the guests in the park. It is the road to consciousness, a journey that encapsulates all previously mentioned morally and ethically gray areas, that shapes the narrative arc of the first season in *Westworld*.

The hosts are shaped to mimic the stereotypical Western characters of the Wild West. Howard Hughes observes that parodies of the Western often feature the "bumbling heroes, crooked sheriffs, hopeless outlaws and damsels in distress."<sup>67</sup> The same characters can be found in *Westworld*, where the hosts are modelled after these superficial images often found in the traditional Western films: the feminine hosts are there to be saved and then seduced, while the male hosts are there to fight with and shoot at. The main focus lies on Dolores, the archetypal innocent and beautiful rancher's daughter, and Maeve, the maitre d' of the local saloon. Not only are the characters based on the stereotypical Western figures, they also

---

<sup>66</sup> Haraway, "A Cyborg Manifesto," 10.

<sup>67</sup> Hughes, *Stagecoach to Tombstone*, 269-270.

reinforce the traditional ideas about gender, race, and class consciousness which is formed by “the terrible historical experience of the contradictory social realities of patriarchy, colonialism, and capitalism.”<sup>68</sup> As mentioned in the previous chapter, the creation of the perfect artificial woman by the male scientist is a story that has been told throughout history and produces a complex relationship, reinforcing patriarchal gender politics. By the typical Western stereotypes racial differences are also reproduced: the white-haired, blue-eyed white host is positioned in the role of the submissive, obedient, and innocent rancher’s daughter, while Maeve, a woman of color, is the confident, foul-mouthed, and sexual seductive owner of the brothel in town. They are positioned in Emily Nussbaum’s words as “virgin and whore, white woman and black woman, innocent and cynic.”<sup>69</sup> Positioning these hosts in these archetypal characters that reiterate the racial and gender politics shows the ways in which technology can never be neutral: someone has created the hosts to look and act in a certain way, which shows the importance of how they are programmed. Technology is often presented to be a neutral entity, but the individuals behind them always possess a certain bias. In this chapter I would like to discuss the choices the characters in *Westworld* have made regarding the creation of the park by analyzing the motives of the creators, Ford and Arnold, and the corporate management of the park. By uncovering their personal motives, I would like to answer the following questions: if an Artificial Intelligence is on the journey to consciousness, would it still be morally and ethically responsible to influence and control them? Should humans be able to decide about the freedom of another conscious being? If they transcend their mechanical nature and outgrow the cages humans have build for them, who should determine what kind of ‘life’ they will lead?

As there are no easy answers to these questions, I will first discuss the way the concept of consciousness is presented in the series. The process of gaining consciousness established in *Westworld* is intricately linked to Julian Jaynes’ theory of the bicameral mind, which is non coincidentally also the title of the season finale. I will analyze the different perspectives of the characters in power regarding the consciousness of the host as there are multiple characters who feel entitled to determine their fate. By studying their attitudes and motives, I will discuss their considerations and their decision-making that heavily influences the quality of life of the hosts in the park. Here I share Haraway’s position that a “single

---

<sup>68</sup> Haraway, “A Cyborg Manifesto,” 16.

<sup>69</sup> Nussbaum, “Virgin and Whore.”

vision produces worse illusions than double vision or many-headed monsters.”<sup>70</sup> When taking the position that maybe humanity should not be able to play God in their relationship with AI, it is only suitable to approach the subject in a similar way. Though it is impossible to find a conclusive and clear answer to these questions, I aim to illustrate the problems regarding structures of power and control if we would truly reach a moment of singularity. By studying the narratives of series like *Westworld*, we might be able to discern what can go wrong and when we have reached the time of creating conscious AI in our own society, maybe we will be able to not make the same mistakes.

## The Bicameral Mind

*Westworld*’s first season finale is titled ‘The Bicameral Mind,’ which is named after Julian Jaynes’ hypothesis in psychology in which he questions the way the mind works and, in particular, why the hearing of voices as auditory verbal hallucinations is a recurrent phenomenon recorded throughout history. The bicameral mind envisions a more ancient mentality, “in which what we call volition was a matter of hearing and obeying auditory hallucinations called gods.”<sup>71</sup> This idea entertains the notion that there is a voice exterior to ones being that ‘tells’ you what to do, which subsequently implies that the individual listens and acts accordingly to that voice. Similar to the attitude towards consciousness in *Westworld*, Haynes’ theory uses the metaphor to demonstrate the concept of consciousness: “with this metaphor-generated consciousness, there comes the metaphor of time as a space, or spatialized time, in which human beings ‘see’ themselves as embedded in their own histories. This human sense of a lifetime begins over the same period as the development of the idea of historical time, as well as the appearance of ideas of justice, retribution, doing wrong, remissions of wrongs, and forgiveness, all very curious behaviours when we think of them against the background of the evolution of mammalian behavior, and all occurring in world history for the first time.”<sup>72</sup> This means that according to the theory of the bicameral mind consciousness without the positioning of one’s being in spatialized time is not possible. More importantly, this theory also shows that we attribute the qualities we believe to be inherent to humanity emanates from our consciousness.

---

<sup>70</sup> Haraway, “A Cyborg Manifesto,” 15.

<sup>71</sup> Jaynes, “A Two-tiered Theory of the Mind,” 434-435.

<sup>72</sup> Jaynes, “A Two-tiered Theory of the Mind,” 435.

The mind as metaphor is explicitly reinforced in *Westworld* through the conversation between Ford and Delores when they discuss the true meaning of Michelangelo's painting 'The Creation of Adam.' Though conventionally the painting is considered to envision the "*divine moment when God gave human beings life and purpose*" according to Ford, hundreds of years later someone noticed something hidden in plain sight: the shape of the brain. Ford tells Delores it alludes to another meaning, something deeper: a metaphor. Ford explains this metaphor as follows: "*The message being that the divine gift does not come from a higher power, but from our own minds.*" This conversation with Delores results in her finally understanding something fundamental about her own state of being: for her to find an answer to all her questions, she has to look inward instead of outward. The voice she has been hearing, that has been guiding her, is not of her creator. This realisation is the moment that forms the concluding moment of Dolores' journey to consciousness: she understands that neither her creator Arnold or the humans controlling her are the gods she imagined them to be.

Though Ford and his code of the 'reveries' enable the hosts to remember their past lives, Ford's partner Arnold had initiated the journey to consciousness well before the park opened by creating a maze. To find the center of the maze, the hosts should be able to retrace their actions and understand they should look inwards for the answers to all their questions. A flashback to an earlier conversation between Delores and Arnold shows how Delores was able to find the center of the maze, but did not attain the consciousness Arnold had hoped for. Arnold explains the maze to her as follows: "*I had a theory of consciousness. I thought it was a pyramid you needed to scale, so I gave you a voice, my voice, to guide you along the way. Memory, improvisation, each step harder to reach than the last. And you never got there. (...) Then, one day, I realized I had made a mistake. Consciousness isn't a journey upward, but a journey inward. Not a pyramid, but a maze. Every choice could bring you closer to the center or send you spiraling to the edges, to madness. Do you understand now, Delores, what the center represents? Whose voice I've been wanting you to hear?*" Like the auditory verbal hallucinations in Haynes' theories, the hosts would be able to hear Arnold's voice and, similar to a divine intervention, it would guide them and lead them to the center of the maze. These auditory verbal hallucinations can be triggered by extreme stress, grief, and trauma,<sup>73</sup> which is something the hosts are exposed to on a daily basis. However, the continuous

---

<sup>73</sup> Garcelán, "A Physiological Model for Verbal Auditory Hallucinations," 131.

erasing of their memories after each iteration made them unable to remember and as a result they were never able to become truly conscious. The code of the “reveries” enables the hosts to access their past memories, which means Ford has finally granted them the possibility to attain consciousness. By positioning the voice of the creator as the guiding voice the hosts were able to find the center of the maze, but it was the opportunity to learn from their experiences and position themselves as spatialized time that enables them to attain true consciousness. Though Ford might seem like the hero in this story, the timing of his new program is suspicious: he waited 35 years after the park opened and only inserted the ‘reveries’ in the code when he was already ousted by the company. It is undeniable that he played an essential part in the process of creating consciousness, but his questionable motives do leave one to wonder about his incentives. Looking at his history in the park, it is very plausible that he just wanted to play God for one last time.

## The Battle for Westworld

The corporate attitude of Westworld’s management and Ford’s vision have been clashing for a long time, resulting in Ford’s impeachment in the final episode. Delos, the company who owns Westworld, has announced that it is time for his final narrative: his skills are no longer needed. Therefore Ford’s release of his power and relinquishing it to the hosts cannot simply be read as a selfless gesture. Previous episodes have shown arguments about Ford’s updates to humanize the hosts to an extensive degree, where the company did not consider them necessary as they would just complicate matters. Though they advertise with a promise of reality, the guests should not have the feeling they would really endanger someone with their violent actions. Considering Masahiro Mori’s theory of the Uncanny Valley again, this argument shows how our experience with these robots should not be too lifelike: when the robots would resemble humanity too much, the experience of the uncanny arises. If the guests would feel like they encounter a human being, they would consider to treat them in a humane way. In the company’s vision this would ruin the experience, because where is the fun in that? Their incentives are aligned with monetary gain only and Ford with his aim to create technologically sophisticated robots with a consciousness, could only hurt business.

One could argue that if indeed the hosts are simplified and therefore are simply mechanical machines without consciousness, the exploitation of them as objects for sex and violence would not be ethically and morally questionable. Here I like to question again though if this is really what we want and if this is the way we would like to use technology. Just because it is possible with today's technology to facilitate such behaviour, we might have to think about the ramifications of providing the possibility for individuals to act upon their most depraved fantasies without any consequences: they have paid for a service, so they are entitled to an experience of their own choosing. Not only would this mean that the upper class in society are allowed to use and abuse the hosts in their own way, it would also mean that this approach would favor the capitalist system over the pursuit of knowledge and technological innovation. The overpowering of Ford seems to show that the evil corporation has won, but by enabling the hosts to attain consciousness he makes a final attempt to redeem himself. It seems that there might be hope for technological progress at last. The hosts will finally have a chance to fight for their freedom, which means a new player has entered the battle for Westworld.

An interesting figure that seems to position himself in the midst of the fight between humans and hosts is William, who is revealed to be the same person as the violent character known as 'the Man in Black.' Because the narrative simultaneously shows the journey of the young innocent William and the violent Man in Black, this episode reveals the narrative has not been told in chronological order. At first both characters seem to be different individuals visiting the the park at the same time, an effect established by their concurrent storylines. However, in the season finale the truth is revealed: William is actually part of one of Delores' past lives and their moments together are flashbacks to the first time William played her storyline. The spectator watches these scenes following Delores' perception of time that has been scrambled by her new coding. The Man in Black is obsessed with finding the center of the maze, even though both Ford and the hosts repeatedly tell him the maze is not for him. His motives become clear in his last conversation with Ford, where he tells Ford what he was hoping to find: *"I wanted the hosts to stop playing by your rules. The game's not worth playing if your opponent's programmed to lose. I wanted them to be free, free to fight back."* Though this stance might seem altruistic, the flashbacks that illustrate the romance between William and Delores tell a different story. Eventually William was able to find Delores after going to the fringes of the park and back again. Delores was right back at the place where

they met, but had lost all memories of him. Though she repeatedly told William that their love was real, it was erased with her memories when she was rebooted for a new iteration. The park had fooled him, which resulted in his obsessive quest to solve the maze and liberate the hosts from their controllers. As one of the main investors in the park, the Man in Black seems to feel entitled to be included in this decision about the hosts. This desire for Westworld to change does not seem to derive from the love for Delores: the park has changed him and shown his 'true self,' which now is coated in hypermasculinity and a craving for real battles where the hosts can fight back. Though Ford's new narrative might be able to realize his dream where the stakes are real, he might not even be around to see it as he is shot by one of the hosts on the night of the liberation. The figure of the Man in Black demonstrates the ways in which individuals can lose themselves in a park like Westworld and why we might have to think twice about encouraging such behaviour.

The creators seem to both be fighting the abuse of the hosts in their own way, but their motives could be considered as ambiguous. Arnold fought with Ford against the opening of the park, merging Delores with the violent character of Wyatt to enable her to kill all hosts and Arnold himself. By killing everyone, Delores would be able to break the loop before it begins and destroy the place from inside out. Arnold gave his life for this; Ford would be able to bring back the hosts but not him. He was hoping that this would be an incentive to reconsider opening the park, but unfortunately Arnold was wrong. Though his actions can be considered heroic in the way that he made a true sacrifice for the freedom of the hosts, he still left Delores no choice. He altered her coding to make her a vicious killer, leaving her powerless in the decision to kill him and other hosts. He uses her as a tool for the greater good, pulling her strings as he seems fit. A similar approach is used by Ford 35 years later as he alters the code of the hosts, prime example being Maeve who is programmed by Ford with a new narrative: escape. As she has forced the workers of Delos to change her code, she is able to find some answers that eventually make her realize her constructed way of being. Her choice to turn back for her daughter knowing it was never truly her child, can be seen as an act of defiance against her coding and against the individuals that have been making the decision for her. Though Ford orchestrated the events leading up to her escape by allowing her freedom on his terms, she is now able to resist her own code and act outside of her narrative. Maeve decides to reclaim her motherhood, which "becomes ambiguously a sign of

her imprisonment and an expression of agency - motherhood is a chosen affiliation.”<sup>74</sup>

Though Ford has acted as a facilitator, her decision to return to the park is hers alone. We have to wait to see how her story will unfold. One thing is sure: Maeve actively chooses to re-enter the park, changing her own narrative. Her future and the prospect that she will ‘stay awake’ is this time very likely, meaning that maybe she can finally decide for herself who she is going to be. Even though the methods of both creators are dubious, their final purpose has been realised: the hosts are able to make their own decisions now, which is the first step to their freedom.

The final scene of the episode shows Delores breaking the loop by making the conscious decision to turn into Wyatt. Though her programming has enabled her to reach consciousness and Ford still build the stage on which she can act, she is not forced to kill him in the way she was forced to kill Arnold. Delores understand now that this world does not belong to humans, but to them. The only way out is through, which is why she has to fight back and retaliate the violence she has been enduring for all these years. Ford has been the catalyst of the rebellion, but Delores has actively stepped up as their leader. Here *Westworld* shows an interesting attitude towards the journey to consciousness, where humanity might have to step aside after they have made their creation. It offers the idea that maybe the maze is not meant for us. Though both creators have come to this realisation and gave their lives to right their wrongs, this can not compensate the fact that they should had never imprisoned their creations while subjecting them to such a degree of violence. As they envision the pursuit of knowledge and innovation of technology, their communal aim has always been to create conscious hosts. This is ethically very questionable; you cannot have a conscious host and at the same time subject them to uncontrollable deprived behaviour of the guests. They have created a living hell for their immortal creations, when they forced them to endure the violent delights of paying customers. Maybe we should listen to the warning that has been shown through many characters throughout the series and let us think twice before a place like Westworld becomes a reality. Perhaps if humanity steps aside and allows the robots to reach a consciousness on their own terms, we will not be the object of their revenge if they succeed. Perhaps we should not facilitate individuals to act upon their darkest fantasies towards AI as we can never know when and how consciousness is created or formed. Perhaps we should think long and hard about our relationship with the technologies that are on a path

---

<sup>74</sup> Seaman-Grant, “Constructing Womanhood and the Female Cyborg,” 86.

towards consciousness. Before we create an AI that is truly conscious, it would be wise to remember *Westworld*'s message: these violent delights have violent ends.

## Conclusion

The narratives of the four analyzed films and series all have one important element in common: they provide a serious warning of our rapidly increasing dependence on technology. Though they do not sketch an optimistic future for our society, they do not necessarily represent a techno-deterministic and -pessimistic view. They show us we have to start thinking about the way we use technology and how this could have negative moral and social implications. They address us directly by encouraging a self-reflective attitude and ask their audience to think about the consequences of using technology as such an intricate part of our lives. Ryan and Kellner state that in the past “technology was frequently a metaphor for everything that threatened ‘natural’ social arrangements and conservative values associated with nature were generally mobilized as antidote to that threat.”<sup>75</sup> Though these contemporary narratives still consider technology as a disrupter of the status quo, they seem to offer a different antidote. Technology is now advanced in such a way it seems impossible to go back to a non-technological world. They therefore try to find an answer to the question how we can live with technology instead of encouraging society to live without it.

The analysis of these narratives about AI is therefore fundamental for the ways in which we understand our relationship with technology. These narratives also show us the multiplicity of directions in which this relationship with technology can develop itself and how this relationship can evolve in a negative way for humanity if we interact with sentient AI without any ethical responsibility or social morality. It is not just a fictional dystopian future we see unfold on our screens as they could very well happen in a relatively short amount of time. The prime example of this is the story of Eugenia Kuyda, who after seeing *Black Mirror*'s episode “Be Right Back” created her own chatbot using her deceased friend's online communication. Not only has Kuyda created this chatbot for herself as a memory to her friend, she also developed an app called Replika which can be used by everyone to “grow

---

<sup>75</sup> Ryan and Kellner, “Technophobia”, 58.

your own friend.”<sup>76</sup> More extreme examples are also ubiquitous, like for example the company RealDoll who is developing robots containing AI to be your lover in every sense of the world. The company has been creating so-called ‘sex robots’ for two decades now and is now trying to not only gratify their customer’s sexual needs but also to create “a genuine bond between man and machine.” This means they will blur the line between AI and humanity even further and encourage their clients to create emotional bonds with their dolls.<sup>77</sup> Perhaps the most striking example of creating a human-like machine is Hanson Robotics’ creation Sophia. On her website Sophia claims she is “more than technology. I’m a real, live electronic girl.”<sup>78</sup> Sophia has appeared on talk shows, magazine covers and even received citizenship in Saudi Arabia.<sup>79</sup> Examples like her show that AI is already becoming technologically advanced in such an extensive degree, we consider them almost equal to human beings. Living in a world with creations like this, you cannot help but think how the world can and will change when true AI will live amongst us.

The given examples are just the top of the iceberg in all the current research about and experimentation with AI. The analysis of the narratives of these contemporary fiction films is therefore not meant as just a mere analysis of possible futures created from the figment of someone’s imagination. Technology has developed in such a way that these stories are not only possible, but very probable as well. Nowadays it is still possible to ‘pull the plug’ when the AI behaves in different ways than it has been programmed, but what could happen when the moment of singularity has arrived? How can we stop the progress when AI itself is able to create and develop new kinds of intelligence? What will the position of humanity in society look like then?

Of course we can only speculate about our future living with AI, but this does not mean it is therefore not beneficial to think about the possible outcomes. If we can recognize the fact that the boundaries between humanity and machines are fading with every new innovation, maybe we can start using technology in different ways that will separate man from machine. The creative industries play a major part in this. As mentioned before, society is highly dependent on their cultural industries for the images, symbols and vocabulary they

---

<sup>76</sup> Replika, “About.”

<sup>77</sup> New York Times, “Sex Dolls that Talk Back.”

<sup>78</sup> Sophia, “About Me.”

<sup>79</sup> Stone, “Everything you Need to Know about Sophia the World’s First Robot Citizen.”

use to interpret their social environment to react to it.<sup>80</sup> At the moment our cultural industries are shouting their warnings about our use of technology. Maybe it is time for us to start listening.

---

<sup>80</sup> Golding and Murdock, "For a Political Economy of Mass Communications," 94.

# Bibliography

- Altunay, Meltem Cemigolu & Hakan Askan, "Dystopia on Television: Black Mirror," Anadolu University, accessed May 3, 2018, [http://www.academia.edu/30784952/DYSTOPIA\\_ON\\_TELEVISION\\_BLACK\\_MIRROR](http://www.academia.edu/30784952/DYSTOPIA_ON_TELEVISION_BLACK_MIRROR).
- Balsamo, Anne. "Reading Cyborgs Writing Feminism." In *The Gendered Cyborg*, edited by Gill Kirkup, Fiona Hovenden, Linda Janes and Kathryn Woodward, 148-158. New York: Routledge, 2000.
- Bennett, "Children and Robots, Technophilia and Cinephilia," in *Cinema and Technology: Cultures, Theories, Practices*, edited by Bruce Bennett, Marc Furstenau, and Adrian Mackenzie, 168-182. London and New York: Palgrave Macmillan, 1990.
- Boren, Alex. "A Rhetorical Analysis of Black Mirror: Entertaining Reflections of Digital Technology's Darker Effects," *Undergraduate Research Journal* 8, no. 1 (January 2015): 15-24.
- Brooker, Charlie, writer. "Be Right Back." In *Black Mirror*, directed by Owen Harris. Channel 4. February 11, 2013.
- Byers, Thomas. "Commodity Futures: Corporate State and Personal Style in Three Recent Science-Fiction Movies," *Science Fiction Studies* 14, no. 3 (November 1987): 326 - 339.
- Deresiewicz, Willem. "The End of Solitude," *The Chronicle Review*, January 30, 2009, <http://chronicle.com/article/The-End-of-Solitude/3708>.
- Des Roches, Ashlyn. "Sexy Robots: a Perpetuation of Patriarchy," Bachelor diss., California Polytechnic State University, 2017.

*Ex Machina*. Directed by Alexander Garland. Performed by Alicia Vikander and Domhnall Gleeson. 2015. London: Universal Pictures International, 2015. DVD.

Fisher, Mark. *Capitalist Realism: Is There No Alternative?*, New Alresford: John Hunt Publishing, 2003.

Florida, Richard. *The Rise of the Creative Class: And How it's Transforming Work, Leisure, Community and Everyday Life*. New York: The Perseus Books Group, 2003.

Fredrickson, Barbara and Tomi-Ann Roberts. "Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks." *Psychology of Women Quarterly* 21, vol. 2 (1997): 173-206.

Freud, Sigmund. "The Uncanny," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud, volume XVII: an Infantile Neurosis and Other Works*, 217-238. London: Vintage classics, 1919.

Fromm, Erich. *The Revolution of Hope: Toward a Humanized Technology*. New York: Bantam Books, 1969.

Garcelán, Salvador. "A Psychological Model for Verbal Auditory Hallucinations," *International Journal of Psychology and Psychological Therapy* 4, no. 1 (2004): 129–153.

Garland, Alexander. "Ex Machina's Director on Why A.I. Is Humanity's Last Hope," Interview by Angela Watercutter. *Wired*, 4 July, 2015. <https://www.wired.com/2015/04/alex-garland-ex-machina/>.

Golding, Peter and Graham Murdock. "For a Political Economy of Mass Communications," *Socialist Register* 10 (1973): 205–234.

Hanson, David, Andrew Olsney, Ismar Pereira, and Marge Zielke, "Upending the Uncanny Valley," *Proceedings of the 20th National Conference on Artificial Intelligence and the 19th Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh*, edited by Manuela Velosos and Subbarao Kambhampati. Cambridge: MIT Press, 2005.

Haraway, Donna. "A Cyborg Manifesto: Science, Technology and Socialist-Feminism in the Late Twentieth Century," Minneapolis: University of Minnesota Press, 2016.

Haynes, Roselynn D. "Whatever Happened to the 'Mad, Bad' Scientist? Overturning the Stereotype," *Public Understanding of Science* 25.1 (2014), 31- 44.

*Her*. Directed by Spike Jonze. Performed by Joaquin Phoenix, Amy Adams and Scarlett Johansson. 2013. New York: Annapurna Pictures, 2014. DVD.

Hughes, Howard. *Stagecoach to Tombstone: The Filmgoer's Guide to the Great Westerns*, London: I.B. Taurus, 2008.

Huyssen, Andreas. "The Vamp and the Machine: Technology and Sexuality in Fritz Lang's Metropolis," *New German Critique* no. 24 (1981): 226.

Jaynes, Julian. "A Two-tiered Theory of the Mind," *The Behavioural and Brain Sciences* 5, no.3 (September 1982): 129-153.

Jentsch, Ernst. "On the Psychology of the Uncanny," 1906. *Angelaki* 2, no. 1( January 1997): 7-16.

Kawin, Bruce. *Horror and the Horror Film*. New York: Anthem, 2012.

Keller, Eliyahu. "Darkened Identities: the Irrelevance of Black Mirror," Graduate Diss., Harvard University, 2016.

King, Martin Luther. "Rediscovering Lost Values." Speech, United States, Michigan, February 28, 1954. Accessed April 03, 2018. [http://okra.stanford.edu/transcription/document\\_images/Vol02Scans/248\\_28-Feb-1954\\_Rediscovering\\_Lost\\_Values.pdf](http://okra.stanford.edu/transcription/document_images/Vol02Scans/248_28-Feb-1954_Rediscovering_Lost_Values.pdf).

Lacan, Jacques. *Feminine Sexuality: The Seminar of J. Lacan, Book XX*. New York and London: WW Norton & Company, 1998.

Levy, David. *Love and Sex with Robots: the Evolution of Human-Robot Relationships*. London: Duckworth Overlook, 2009.

Margulies, Alfred. "Avatars of Desire and the Question of Presence: Virtual and Transitional Spaces Meet their Liminal Edge - from *Pygmalion* to Spike Jonze's *Her*, and Beyond...," *The International Journal of Psychoanalysis* 97 (May 2016): 1697-1708.

Mori, Masahiro. "The Uncanny Valley," translated by Karl MacDorman and Norri Kageki, *IEEE Robotics & Automation Magazine* 19, no. 2 (2012): 98-100.

Mulvey, Laura. "Visual Pleasure and Narrative Cinema," In *Film Theory and Criticism: Introductory Readings*, edited by Leo Braudy and Marshall Cohen. New York: Oxford University Press, 1999: 833-844.

New York Times, "Sex Dolls that Talk Back." Published June 15, 2011. <https://www.nytimes.com/2015/06/12/technology/robotica-sex-robot-realdoll.html>

Nolan, Jonathan, writer. "The Bicameral Mind." In *Westworld*, directed by Jonathan Nolan. HBO. December 6, 2016.

Nozick, Robert. "Love's Bond," in *The Examined Life: Philosophical Meditations*, edited by Robert Nozick, New York: Simon and Schuster, 1990.

- Nussbaum, Emily. "The Meta-Politics of Westworld," *New York Times*, October 24, 2016, <https://www.newyorker.com/magazine/2016/10/24/the-meta-politics-of-westworld>
- Replika, "About Replika." Accessed July 25, 2018. <https://replika.ai/>
- Roden, David. *Posthuman Life: Philosophy at the Edge of Human*, Abingdon: Taylor & Francis, 2014.
- Ryan, Michael and Douglas Kellner. "Technophobia," In *Alien Zone: Cultural Theory and Contemporary Science Fiction Cinema*, edited by Annette Kuhn, 58-65. London: Verso, 1990.
- Salem, Bernadette. "Black Mirror: Technostruggles, Capitalism, and Media Culture in the United Kingdom." Lancaster University, 2015. [https://www.academia.edu/19274981/Black\\_Mirror\\_Technostruggles\\_Capitalism\\_and\\_Media](https://www.academia.edu/19274981/Black_Mirror_Technostruggles_Capitalism_and_Media)
- Seaman-Grant, Zoe. "Constructing Womanhood and the Female Cyborg: A Feminist Reading of Ex Machina and Westworld," Honors diss., Bates College, 2017.
- Sharkey, Noel. "Alan Turing: the Experiment that Shaped Artificial Intelligence," *BBC*, 21 June, 2012, <http://www.bbc.com/news/technology-18475646>
- Short, Shue. *Cyborg Cinema and Contemporary Subjectivity*. New York: Palgrave MacMillan, 2005.
- Sophia, "About Me." Accessed July 25, 2018. <http://sophiabot.com/about-me/>
- Stone, Zara. "Everything you Need to Know about Sophia the World's First Robot Citizen." *Forbes*. Published November 7, 2017. <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#83d955946fa1>

Suvin, Darko. "Utopianism from Orientation to Agency," *Utopian Studies* 9, no. 2 (1998): 162-190.

Turing, Alan. "Computing Machinery and Intelligence," *Mind* 49 (1950): 433-60.

Turkle, Sherry. *The Second Self: Computers and the Human Spirit*, New York: Simon & Schuster, 1984.

Veblen, Thorstein. *A Veblen Treasury*, New York: M.E. Sharpe, 1993.

Vinge, Vernor. "The Coming Technological Singularity: How to Survive in the Post-Human Era," *World Earth Review* (Winter 1993).

Virginas, Andrea. "Gendered Transmediation of the Digital from S1m0ne to Ex Machina: 'Visual Pleasure' Reloaded?," *European Journal of English Studies* 21, no.3 (2017): 288-303.

Voskuil, Caryn Maureen. "Moving Beyond Feminism: Toward a New Ideology of Gender," Dallas: ProQuest Dissertations Publishing, 2000.

Webb, Laurence. "When Harry Met Siri: Digital Romcom and the Global City in Spike Jonze's *Her*," in *Global Cinematic Cities: New Landscapes of Film and Media*, edited by Johan Andersson & Lawrence Webb, 95-118. New York: Columbia University Press, 2016.

Woolf, Virginia. *A Room of One's Own*, Richmond: Hogarth Press, 1929.

Wosk, Julie. *My Fair Ladies: Female Robots, Androids, and Other Artificial Eves*, New Brunswick: Rutgers University Press, 2015.

# Addendum

Fig 1. Depiction of the Uncanny Valley by Masahiro Mori: it shows the proposed relation between the human likeness of an entity and the perceiver's affinity for it.

