

# How mary shelley's depiction shows the threats of artificial intelligence

[Literature](#), [Russian Literature](#)



Author Mary Shelley uses the creature in her book Frankenstein to depict the dangers of humans creating artificial intelligence. Shelley presents these ideas through using Victor Frankenstein as a creator of a new species of life, and the creature as a naive, fiendish artificial intelligence. This idea is more relevant in today's world than ever before. With the advancements of computer driven artificial intelligence, the ideas Shelley formulated in Frankenstein will become a reality in the near future. While Shelley likely didn't know this would be the case, Shelley still formulated an accurate depiction of what the struggles of creating a species would be.

The creature's biggest struggle is overcoming his hamartia of lack of emotion. Throughout the novel, the creature attempts to overcome his misunderstanding of human emotions. Its inability to understand and control his emotion is demonstrated when the creature is shot after saving the little girl. Confused, the creature screamed, " This was then the reward of my benevolence!" (Shelley 279). The creature, obviously befuddled at the human's actions was furious of the actions of the shooter. Confused because of his demeanor to do good, the monster screamed with disbelief, " This was then the reward of my benevolence! I had saved a human being from destruction, and as a recompense I now writhed under the miserable pain of a wound which shattered the flesh and bone" (Shelley 280). The creature was attempting to help the innocent girl, and the man shot the creature because he misunderstood the creature's intentions. This led to a swarm of confusion, in which the creature turns into undeniable hatred stating, " I vowed eternal hatred and vengeance to all mankind" (Shelley 280). Shelley portrays these events as events of confusion, mistrust, and natural hatred

toward the creature, and the shooter. This is where Mary Shelley delineated the dangers of artificial intelligence. Humans are naturally territorial, and when a self-aware species is threatening to our safety, there will be conflict (Sack). Shelley demonstrated this by portraying the human as frightened by the sight of the creature, and shooting the creature without just reason. In the case of a Shelley created this scene in order to start the idea of humans and artificial intelligence discordant with each other.

Humanity's underlying goals and values are not the same as other species; a gorilla is not going to have the same desire to kill for food as a human, because of the difference in need and desire. The same concept is applied to the creature and forms of artificial intelligence. The creature's values are portrayed when he states, " My feelings were those of rage and revenge. I could with pleasure have destroyed the cottage and its inhabitants and have glutted myself with their shrieks and misery" (Shelley 268). The creature's prominent ideology is based on the ideas of revenge and narcissism. This ideology creates mayhem as humanity's strived for values of unity and serenity clash with the demonic values of the creature. The creature continues with his demonic values until the end of the novel when the creature's developed self-consciousness transitions his values into a higher level of tranquility. The idea of humanity's values clashing with an artificial intelligence's values is commonly warned of today. Famous entrepreneur and billionaire philanthropist, Elon Musk has warned of artificial intelligence's potentially devastating effects, stating AI could be the " biggest existential threat to mankind" (Kedmey). Musk is not talking about the threat of a

superhuman creature, but rather the threat of a machine that does not know of humanity's values, and what is morally acceptable. Elon Musk backed this up by the famous thought experiment named the 'Paperclip Maximizer'. The 'Paperclip Maximizer' experiment was an artificial intelligence experiment called a instrumental convergence. According to the Oxford Faculty of Philosophy, the instrumental convergence thesis is, "the hypothetical tendency for most sufficiently intelligent agents to pursue certain instrumental goals such as self-preservation and resource acquisition" (Bostrom 6). Throughout the book, the creature was consistently attempting to protect himself from humanity's aggressive notation. The creature's goal was not unity and serenity, but self-protection. Therefore, it can be assumed that the creature is a instrumental convergence artificial intelligence. Typically, this type of artificial intelligence is considered dangerous due to it's ability to cause damage to humanity in order to protect itself. Shelley utilized these ideas to furthermore add a complex plot to the creature's personality.

A major portion of Frankenstein is the creature's journey to understand the reasoning of his creation and progress his self-awareness. One of the early questions the creature asks itself is questioning his existence. The creature states: "My person was hideous and my stature gigantic. What did this mean? Who was I? What was I? Whence did I come? What was my destination?" (Shelley 252). These questions stated by the creature closely associate to a human when reaching consciousness. The creature continues to progress throughout the novel, becoming more aware of his place in the

world, and more dangerous in his actions. This pattern continues until the end of the novel when the creature rethinks running from Victor and decides his own fate. The quick progression of an artificial intelligence is expected when referencing a modern-day computer intelligence due to the scalability and efficiency of technology (Urban). While the reasoning behind the creature's dramatic spike in its own self-awareness is unknown in the novel, it can be assumed that the creature takes a similar path. After the creature finds the answers about the origins of his existence, the creature gets further enraged to the point of craving revenge against Victor. Shelley could be inferring of the events that may happen after creating an artificial intelligence, and what the consequences may be.

The novel Frankenstein demonstrates the possible outcomes of creating an artificial intelligence, and the dangers humanity should avoid. Shelley presented these ideas by developing the creature's social tendencies, showing the creature's awareness progress. The ideas Shelley presented in the novel refer to near future problems that humanity will need to face. Society can use the ideas Shelley wrote in 1818 to help aid decisions and increase awareness. Shelley presents these critically important themes in order to expose society to these potential critical problems.

#### Works Cited:

Bostrom, Nick. "The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents." Nick Bostrom, 2012. Web. 15 Mar. 2016.

<https://assignbuster.com/how-mary-shelleys-depiction-shows-the-threats-of-artificial-intelligence/>

Kedmey, Dan. "Elon Musk Warns Artificial Intelligence Is Like 'Summoning the Demon'" Time. Time, 27 Oct. 2014. Web. 15 Mar. 2016.

Sack, Robert David. "Human Territoriality." Google Scholar. Google, n. d. Web. 09 Mar. 2016.

Shelley, Mary Wollstonecraft, and Diana Gibson. Frankenstein. Madrid, Spain: Edimat Libros, 2000. Print.

Urban, Tim. "The Artificial Intelligence Revolution: Part 1 - Wait But Why." Wait But Why. Wait But Why, 22 Jan. 2015. Web. 15 Mar. 2016.