Automated Technology and Man:

The Essential Cooperation of Humans and Technology

Sarah Delaney

Old Dominion University

PHIL 383T

Professor Andrew Garnar

December 2, 2022

Automated technology is one of the most fascinating and innovative developments of our society. Automation has allowed for increased precision, productivity, and efficiency in numerous fields of technology; however, limitations and boundaries are necessary in order for automation to succeed. This paper will discuss the importance of taking the users of technology into consideration when developing automated technology, specifically when the technology stands in place of a human action. This paper will also introduce two philosophers, Johannes Himmelreich and Tom Sorrell, who reinforce this argument through the presentation of technologies that are slowly being introduced into the world, the limitations of said technologies, and why the nature of humans is so important in their development. Following the presentation of the arguments of Hemmelreich and Sorrell, I will discuss my analysis of this issue in regards to said arguments.

Johannes Himmelreich, a philosopher and professor at Syracuse University, discusses the automation of the self-driving car and argues that Personal Ethics Settings (PES) are vital for self-driving cars to be successful. In regards to Personal Ethics Settings, he states his view that “Where today cars have a wheel, tomorrow they should have a dial” (Himmelreich, 2). Through his article, it is clear that Himmelreich agrees that humans need to be integrated and considered in the development of autonomous technology, and he presents this in two main arguments for PES. The first argument presented, the autonomy argument, asserts that PES are an expression of an individual's autonomy, or the right to drive. He explains that even in areas outside of automation, we use our collective autonomy to “legislate or govern in public spaces” where public safety is an issue, showing that the limitations of automation are necessary (Himmelreich, 3). The second argument, the moral proxy argument, uses the analogy that the decisions engineers make with self-driving cars are similar to that of decisions doctors make about what treatments to give to patients. As passengers have cars as moral proxies, patients have family members or trusted acquaintances. The passenger must be able to trust their car, and that can only be achieved through human governing. Put plainly, your car should

act with your interest in mind, and Personal Ethics Settings allow that to happen. Himmelreich also maintains that PES is in fact necessary for furthering the adaptation of autonomous cars. He explains that the inability to set the parameters of a car's driving makes people more anxious and hesitant to use a self-driving vehicle. Before discussing what a realistic PES system in a self-driving car might look like, Himmelreich acknowledges the counter arguments and issues that could arise from his argument for PES. He discusses one argument in particular: the best interest argument. The argument compares traffic situations to a prisoners' dilemma game, and assumes that cooperation between players is unlikely, and the only way to achieve moral equilibrium is through state regulation, meaning that in automation, Mandatory Ethics Settings (MES) would be necessary. Himmelreich argues that the traffic prisoners' dilemma analogy is not comparable. He asserts that the "game traffic is, to some extent, a matter of policy" (Himmelreich, 6). The policymakers are able to set "incentives", and the "game" that traffic is, is not black and white full cooperation or no cooperation. In fact, he explains that self-driving cars are in fact more likely to cooperate due to their ability to remember interactions with drivers and other cars, then sharing those learnings with other cars in traffic. To defend what a PES system could look like for automated cars, Himmelreich presents several different dimensions of PES in which the car would behave based on human needs, wants, and behaviors. He first emphasizes the dimensions of mobility-safety and self-interest-other-interest. The mobility-safety dimension would allow for the determination of safety decisions such as taking a different, safer route, avoiding a risky left turn, or slowing down/speeding up to avoid a cyclist. With this, self-driving cars with PES would be able to regulate "driving decisions that were impossible to regulate before" in greater detail (Himmelreich, 8). The other dimension that he argues for is one dial for both mobility-safety and self-interest-other-interest. He states that PES should be limited, and that similar to the knob to adjust temperature, there should be one dial to adjust the parameters together on mobility-safety and self-interest-other-interest, allowing the car to drive more or less cautiously to the passenger's comfort. Himmelreich's article is successful in supporting the

argument that human intervention and user consideration is of extremely great value in the development of automated technology. He highlights the wants of the users themselves and explains why incorporating a personalized, convenient factor in automation is key.

In an article titled “Cobots, “co-operation” and the replacement of human skill”, philosopher Tom Sorell introduced the idea of automated technologies, specifically robots, existing as assistants, since humans do not use robots as tools. The article discusses “cobots”: robots that will assist human operators in manufacturing. With this, he emphasizes the idea that the workers should have a say in the design process, as the technology should be catered to what will allow them to work together, and not be replaced or “put humans at the service of machines” (Sorell, 7). Sorrel starts off this article by clarifying that automation doesn’t always replace the human in the situation, but rather coexists with the human, as there are limits to both human capabilities and technological capabilities. In manufacturing, humans and robots both bring different abilities to the production process, but Sorell makes it clear that what defines a cobot is its “task-, time-, or space-sharing with humans” (Sorell, 2). Without the workers, co-bots cannot function to their highest potential and will not have the value their manufacturers desire in the production process. Sorell also addresses the shared fear that robots or AI will end up replacing humans, leaving them without jobs. This is quickly shut down with his introduction of participatory design. He highlights two academic case studies which reflect that creating a cobot that caters to “specific industrial purposes is often a matter of programming that models human skills” (Sorell, 2). This participatory approach relieves the fear of being replaced by reassuring that as long as designers make clear their intentions in working with automation, as well as communicate with the collaborating humans, robotics will be a harmonious introduction in the workplace. Robots will be fitted to a more collaborative approach, which will require workers that are familiar with the manufacturing process to participate in the development of automated robots. Sorell specifies that only humans know the nuances of the actions in manufacturing, and their knowledge is key to the success of cobots. While the article argues

that the existence and success of cobots would be nonexistent without user or worker intervention, it also addresses the ways cobots benefit the workers themselves. The cobots are able to eliminate repetitive and strenuous tasks, making work a more enjoyable experience. There is also the possibility of workers acquiring new skills due to the addition of cobots into the workplace, which leads to social recognition: one of the principle things that make work good according to Gheaus and Hertzog's 2016 analysis referenced by Sorell. He explains that in the workplace, the human and the cobot will work in a way similar to an assembly line: the robot works first, then the human. Sorell highlights the downsides to the incorporation of cobots as well. He first explains that in some cases, the humans' jobs will have to be completely rethought in order to properly incorporate the cobot. While the human's job would not be taken from them, they would not have the job that they were used to; however, this again could lead to social recognition due to the acquisition of new skills. He also references a case study that reflected that the process of getting the robot to actually work "heightens appreciation of the capabilities of human workers" (Sorell, 7). As true with all technology, it will not always work as designed to or as we want it to, and human skills are not easy to replicate. When in the workplace, cobots have issues with adapting to experimental reconfigurations of work processes, which is not the same with human workers. Through his article, Sorell proves that manufacturing and incorporating cobots requires intense superintending and maintenance from humans. In order for the robots to be successful and helpful, the human workers themselves should contribute to the design and improvement process of cobots.

My position on the issue lies very similarly to that of these two philosophers. To put the argument into my own words, I believe that it is necessary for automation engineers and developers to take humans, the users of said automation, into consideration. It seems as if the technologies of AI and automation strive to move away from the role and importance of humankind; however, these technological developments cannot be done without humans. These technologies will simply not be successful or helpful to users without the intervention of

humans, whether that be users or those working in the fields of technology. As CEO of technology service company CloudFactory Mark Sears said in an interview, “Whatever you decide to automate, you will heavily rely on people in your operation to manage the process” (Staff, 2019). While it is easy for me to say I agree with both philosophers, their takes on the issue are presented through different perspectives. Himmelreich takes a wider approach to the argument, with an issue that has somewhat assimilated into our everyday lives. Self-driving cars have the possibility of becoming widespread in the future, and Himmelreich’s perspective is almost a wakeup call to readers. He seeks to educate the reader on the issue of human intervention in automation, from the perspective of the users themselves. Alternatively, Sorell presents the issue through an argument that is less relatable to the majority of people, but rather an overlooked issue. While most of society is familiar with the concepts of AI and automation, the concept of “co-bots” in the workplace and the concerns that come along with it are unpopular and overlooked. While Sorell discusses the main argument towards the significance of human intervention in automation, he also touches on the concept of humans being replaced by automation and AI, which I think is an important discussion to have within this field. While his discussion of co-bots focuses more on the automation aiding the human, Himmelreich’s discussion is focused on the user aiding the automation, which helps strengthen the argument more than Sorell. While the articles from these philosophers differ in content, their overarching ideas and views are complementary. Both philosophers touch on the idea that the introduction and transition to automation, both in everyday society and the workplace, allows humans to learn more about the field and acquire skills that strengthen their knowledge of technology. Sorrel is very deliberate and strong with this, as he emphasizes the “co” in “co-bots”: the worker and the robot collaborate, and the worker must superintend the robot. Additionally, the subject matter of the philosophers’ articles are akin to experiments in a way. Both self-driving cars and co-bots are still in the development and introduction stages of autonomous technology, and they explain that their usage in society today is somewhat of a learning

process. Himmelreich's article concludes with a proposal of what the future could look like for self-driving cars, and gives another idea for developers to work into their car designs. For Sorell, he reminds the reader that there are currently co-bots being used with students in laboratories, being altered, studied, and experimented with. Both philosophers are successful in conveying the current state of autonomous technology, and giving a look into the future of it as well; however, it is my judgment that Himmelreich's article in regards to self-driving cars is more effective in presenting the argument at hand. I think it is important that he provides a look into a technological development that we are seeing more frequently. He explains that with non-automated cars, the decision about how a car drives is in the diver's hands, but raises the question of "with self-driving cars, who decides?"

As demonstrated in this essay, philosophers Johannes Himmelreich and Tom Sorell are successful in explaining the relevance of users being taken into consideration in the field of automation. The discussion of autonomous technology, specifically self-driving cars and "co-bots", emphasizes the notion that automation cannot succeed without humans or users. While the articles further emphasized my argument that the limitations and boundaries of automation need to be recognized and intervened by humans, they were also educational in regards to issues within automation that shed light on different topics such as user autonomy, being put at the service of technology, and the concept of human adaptability and dexterity being replicated, or in this case, not being able to be replicated. With that considered, I have increased confidence in my position with the arguments presented by these two philosophers. There is a bright future for AI and autonomous technology; however, it is clear that developers are on the path to creating more technologies that will stand in place of human actions. The articles presented in this essay clarify that this will only occur when there is a clear line drawn defining that human centered automation works best when it is designed to enhance the abilities of humanity, not replace them.

Works Cited

Himmelreich, J. No wheel but a dial: why and how passengers in self-driving cars should decide how their car drives. *Ethics Inf Technol* 24, 45 (2022). https://doi-org.proxy.lib.odu.edu/10.1007/s10676-022-09668-5

Sorell, T. Cobots, “co-operation” and the replacement of human skill. *Ethics Inf Technol* 24, 44 (2022). https://doi.org/10.1007/s10676-022-09667-6

Staff, R. B. R. (2019, March 14). *Why humans will remain at center of an automated workforce*. Robotics Business Review. Retrieved November 30, 2022, from https://www.roboticsbusinessreview.com/ai/why-humans-will-remain-at-center-of-automated-workforce/